MCRB for Timing and Phase Offset for Low-Rate Optical Communication with Self-Phase Modulation

MCRB for Timing and Phase Offset for Low-Rate Optical Communication with Self-Phase Modulation

Debarati Sen, *Member, IEEE*, Henk Wymeersch, *Member, IEEE*, Naga V. Irukulapati, *Student Member, IEEE*, Erik Agrell, *Senior Member, IEEE*, Pontus Johannisson, Magnus Karlsson, Peter A. Andrekson, *Fellow, IEEE*

***Abstract*—We derive the modified Cramér–Rao bound (MCRB) for symbol timing and phase offset estimation in the presence of nonlinear self-phase modulation (SPM) in a dispersion compensated long-haul optical fiber link with coherent detection at data rates below 10 Gigabaud. In the presence of a low-pass filter at the receiver front-end, we find that SPM degrades the MCRB. Moreover, depending on the pulse shape, SPM induces underdamped oscillation on the bounds.**

***Index Terms*—modified Cramér–Rao bound, self-phase modulation, timing and phase estimation.**

I. Introduction

SELF-phase modulation (SPM) is an important impairment in fiber-optic communication systems using dual-polarization multilevel quadrature amplitude modulation (M-QAM) formats. SPM induces a non-linear phase shift proportional to signal power, leading to spectral broadening. However, most of the literature neglects the impact of SPM on one of the first tasks in a coherent receiver: synchronization [1]. Existing studies on synchronization such as [2], [3] have focused exclusively on linear impairments. Several synchronization algorithms have been proposed for coherent receivers [4], [5]. In terms of bounds, a Cramér–Rao bound for timing offset estimation in presence of SPM was derived in [6], [7]. However, a detailed analysis in presence of other synchronization parameters such as phase and frequency is missing. In this work, we extend [6], [7] and derive a modified Cramér–Rao bound (MCRB) for joint symbol timing and phase offset estimation in the presence of SPM for low-rate (i.e., below 10 Gigabaud) dual polarized M-QAM transmission.

II. Model

A. Signal Model for Low Baud Rates

The transmitted signal over two polarizations is given by

$$\mathbf{x}_0(t) = \sum_{n=1}^{N} \mathbf{s}_n \sqrt{A}\, p\,(t - nT_s - \tau)\, e^{j\theta}, \tag{1}$$

where $\mathbf{s}_n$ is a vector of 2 M-QAM symbols drawn independent and identically distributed (i.i.d.) from a unit-energy constellation, A is the power in each polarization, τ is a timing offset, θ is a phase offset, T_s is the symbol duration, N is the number of transmitted data symbols, and $p(t)$ is a real, unit-energy return-to-zero (RZ) pulse, time limited to $[-T_s/2, T_s/2]$ [8].

We assume a system consisting of N_a spans each of length L with fiber amplifiers. The attenuation factor and non-linearity parameter associated with the fiber are denoted by α and γ respectively. It is assumed that in each span the dispersion in the single-mode fiber (SMF) is perfectly compensated by dispersion compensating fiber and that the attenuation loss is compensated perfectly by Erbium doped fiber amplifiers. Each amplifier generates complex circularly symmetric Gaussian amplified spontaneous emission (ASE) noise in each polarization with power spectral density $N_0 = h\nu n_{\mathrm{sp}}(G-1)$, where G is the gain of the amplifier, $h\nu$ is the energy of the photon, and n_{sp} is the spontaneous emission factor. Overall, the received signal after the q-th span is given by [9]

$$\mathbf{x}_q(t) = \mathbf{U}_q \mathbf{x}_{q-1}(t)\, \exp\left(j\gamma L_{\mathrm{eff}} \|\mathbf{x}_{q-1}(t)\|^2\right) + \mathbf{n}_q(t), \tag{2}$$

where $L_{\mathrm{eff}} = [1 - \exp(-\alpha L)]/\alpha$ is the effective length of the SMF, $\mathbf{U}_q$ is a unitary mixing matrix, and $\mathbf{n}_q(t)$ is complex Gaussian noise with double-sided power spectral density N_0. Note that no dispersive effects are considered. The final received signal is

$$\tilde{\mathbf{r}}(t) = \mathbf{U}\,\mathbf{x}_0(t)\, \exp\left(j\gamma L_{\mathrm{eff}} \sum_{q=0}^{N_a-1} \|\mathbf{x}_q(t)\|^2\right) + \mathbf{n}(t), \tag{3}$$

where $\mathbf{n}(t)$ is the aggregate complex Gaussian noise with double-sided power spectral density $N_a N_0$, and $\mathbf{U} = \prod_q \mathbf{U}_q$. The signal $\tilde{\mathbf{r}}(t)$ is filtered with an anti-aliasing filter with frequency response $H(f) = 1$, for $|f| < 1/(2T)$, and $H(f) = 0$ elsewhere, where T is the sampling time of the filter.

B. A Simplified Signal Model

The exponent in (3) can be expanded into three terms: a signal term $j\gamma L_{\mathrm{eff}} N_a \|\mathbf{x}_0(t)\|^2$, a signal–noise interaction term (for $N_a > 1$), and a noise–noise interaction term (which is negligible in practice). For mathematical convenience, we will discard the latter two terms in several of the derivations. This leads to the simplified model, where (3) is replaced by

$$\tilde{\mathbf{r}}(t) = \mathbf{U}\,\mathbf{x}_0(t)\, \exp\left(j\gamma L_{\mathrm{eff}} N_a \|\mathbf{x}_0(t)\|^2\right) + \mathbf{n}(t). \tag{4}$$

Note that (3) and (4) are equivalent when $N_a = 1$.

Debarati Sen was with the Communication Systems Group, Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden, and is now with Samsung Electronics, Bangalore, India. Henk Wymeersch, Naga V. Irukulapati, and Erik Agrell are with the Communication Systems Group, Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden. Pontus Johannisson, Peter A. Andrekson, and Magnus Karlsson are with the Photonics Laboratory, Department of Microtechnology and Nanoscience, Chalmers University of Technology, Gothenburg, Sweden. Email: debarati@gssst.iitkgp.ernet.in, {henkw, vnaga, agrell, pontus.johannisson, magnus.karlsson, peter.andrekson}@chalmers.se.

C. Problem Formulation

Our aim is to compute lower bounds on the performance of timing and phase estimators for the model from Section II-A. We will denote these bounds by $\mathrm{MCRB}(\tau)$ and $\mathrm{MCRB}(\theta)$, respectively. We will consider two distinct cases: (i) $T \ll T_s$, so that the filter does not affect the signal component in (3) or (4). We will call this scenario "without low-pass filter (LPF)"; (ii) $T \approx T_s$, so that the filter serves as a true low-pass filter.

Comment on the validity of the model from Section II-A: This idealized model is motivated in [7, Section II-B], and is only justified for low baud rates (e.g., 10 Gbd). Our study aims to investigate the isolated effect of SPM on sychronization. For higher baud rates, dispersive effects (e.g., chromatic dispersion) should be accounted for, and we expect the impact of non-linearity to be less significant.

III. Modified Cramér–Rao Bound

A. Description

The modified Cramér–Rao bound (MCRB) is a lower bound on the error variance of any unbiased estimator. Given an unknown deterministic vector parameter $\mathbf{\Theta} = [\theta, \tau]$, the unknown symbols $\mathbf{S}$, and a vector representation of received signal $\mathbf{r}_{\mathrm{vec}}$, we introduce the 2×2 Fisher information matrix (FIM) [1] $\mathbf{F}(\mathbf{\Theta}) = \begin{bmatrix} F_{\theta\theta} & F_{\theta\tau} \\ F_{\tau\theta} & F_{\tau\tau} \end{bmatrix}$, with

$$[F(\mathbf{\Theta})]_{ij} = -\mathbb{E}\left[\frac{\partial^2 \ln p(\mathbf{r}_{\mathrm{vec}}|\mathbf{\Theta},\mathbf{S})}{\partial \Theta_i \, \partial \Theta_j}\right], \tag{5}$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator and $\ln p(\mathbf{r}_{\mathrm{vec}}|\mathbf{\Theta},\mathbf{S})$ is the log-likelihood function. The expectation is taken over the noise and unknown data symbols. Then

$$\mathrm{var}(\theta - \hat{\theta}) \geq \mathrm{MCRB}(\theta) = \frac{1}{F_{\theta\theta} - F_{\tau\theta}^2 / F_{\tau\tau}} \tag{6}$$

$$\mathrm{var}(\tau - \hat{\tau}) \geq \mathrm{MCRB}(\tau) = \frac{1}{F_{\tau\tau} - F_{\tau\theta}^2 / F_{\theta\theta}}\,. \tag{7}$$

We will derive the MCRB based on the simplified model from Section II-B, which will result in approximate lower bounds for the orginal model from Section II-A.

B. Derivation of MCRB without Low Pass Filter

The received signal vector $\mathbf{r}_{\mathrm{vec}}$ consists of samples $\mathbf{r}_k = \tilde{\mathbf{x}}_k + \mathbf{n}_k$ at rate $1/T$, comprising a signal term $\tilde{\mathbf{x}}_k = \mathbf{U}\mathbf{x}_k \exp\left(j\beta \|\mathbf{x}_k\|^2\right)$, where we have introduced $\beta = \gamma L_{\mathrm{eff}} N_a$, $\mathbf{x}_k = \mathbf{x}_0(kT)$, and an i.i.d. noise term $\mathbf{n}_k$. The first step towards computing the MCRB is to derive the expression for the likelihood function:

$$\begin{aligned} p(\mathbf{r}_{\mathrm{vec}}|\mathbf{\Theta},\mathbf{S}) &= \prod_{k=-\infty}^{+\infty} p(\mathbf{r}_k|\mathbf{\Theta},\mathbf{S}) \\ &\propto \prod_{k=-\infty}^{+\infty} \exp\left(-\frac{\|\mathbf{r}_k - \tilde{\mathbf{x}}_k\|^2}{N_a N_0 / T}\right). \end{aligned} \tag{8}$$

Taking the logarithm and substituting into (5) yields

$$[F(\mathbf{\Theta})]_{ij} = \frac{2}{N_a N_0 / T} \sum_{k=-\infty}^{+\infty} \mathbb{E}\left[\Re\left\{\frac{\partial \tilde{\mathbf{x}}_k^{\mathrm{H}}}{\partial \Theta_i}\frac{\partial \tilde{\mathbf{x}}_k}{\partial \Theta_j}\right\}\right]. \tag{9}$$

Substituting (1) into $\mathbf{x}_0(kT)$ and introducing $\xi_{n,k} = kT - nT_s - \tau$, we find that

$$\begin{aligned} \tilde{\mathbf{x}}_k &= \sqrt{A} e^{j\theta} \mathbf{U} \sum_{n=1}^{N} \mathbf{s}_n p(\xi_{n,k}) \exp\left(j\beta \left\|\sum_{m=1}^{N} \mathbf{s}_n \sqrt{A} p(\xi_{m,k}) e^{j\theta}\right\|^2\right) \\ &= \sqrt{A} e^{j\theta} \mathbf{U} \sum_{n=1}^{N} \mathbf{s}_n p(\xi_{n,k}) \exp\left(j\beta A \|\mathbf{s}_n\|^2 |p(\xi_{n,k})|^2\right), \end{aligned} \tag{10}$$

where the last equality is due to the finite duration of $p(t)$. It will be useful to introduce $I_{kl} = \int_{-\infty}^{+\infty} |p(t)|^k |\dot{p}(t)|^l \,\mathrm{d}t$ and $J_{kl} = \int_{-\infty}^{+\infty} |p(t)|^k (\dot{p}(t))^l \,\mathrm{d}t$.

Derivation of $F_{\theta\tau} = F_{\tau\theta}$: Since

$$\begin{aligned} \frac{\partial \tilde{\mathbf{x}}_k^H}{\partial \theta} &= -j e^{-j\theta} \sqrt{A} \mathbf{U}^{\mathrm{H}} \sum_{n=1}^{N} \mathbf{s}_n^{\mathrm{H}} p(\xi_{n,k}) \\ &\quad \times \exp\left(-j\beta A \|\mathbf{s}_n\|^2 |p(\xi_{n,k})|^2\right) \end{aligned} \tag{11}$$

and

$$\begin{aligned} \frac{\partial \tilde{\mathbf{x}}_k}{\partial \tau} &= -e^{j\theta} \sqrt{A} \mathbf{U} \sum_{m=1}^{N} \mathbf{s}_m \dot{p}(\xi_{m,k}) \exp\left(j\beta A \|\mathbf{s}_m\|^2 |p(\xi_{m,k})|^2\right) \\ &\quad - 2j\beta A^{3/2} e^{j\theta} \sum_{m=1}^{N} \mathbf{s}_m \|\mathbf{s}_m\|^2 |p(\xi_{m,k})|^2 \dot{p}(\xi_{m,k}) \\ &\quad \times \exp\left(j\beta A \|\mathbf{s}_m\|^2 |p(\xi_{m,k})|^2\right) \end{aligned} \tag{12}$$

we see that

$$\mathbb{E}\left[\Re\left\{\frac{\partial \tilde{\mathbf{x}}_k^{\mathrm{H}}}{\partial \theta}\frac{\partial \tilde{\mathbf{x}}_k}{\partial \tau}\right\}\right] = -2\beta A^2 \sum_{n=1}^{N} E_4 |p(\xi_{n,k})|^3 \dot{p}(\xi_{n,k}), \tag{13}$$

where $\dot{p}(t)$ is the derivative of $p(t)$. In (13) we used the fact that data symbols are i.i.d. and introduced $E_l = \mathbb{E}\left\{\|\mathbf{s}_n\|^l\right\}$. Therefore,[1] $F_{\theta\tau}$ is given as

$$\begin{aligned} F_{\theta\tau} &= \frac{-4\beta A^2}{N_a N_0 / T} \sum_{k=-\infty}^{+\infty} \sum_{n=1}^{N} E_4 |p(\xi_{n,k})|^3 \dot{p}(\xi_{n,k}) \\ &= \frac{-4\gamma L_{\mathrm{eff}} A^2 E_4 N J_{31}}{N_0}. \end{aligned} \tag{14}$$

Note that $F_{\theta\tau} = 0$ whenever the system is linear (i.e., $\gamma = 0$) or whenever the pulse $p(t)$ is even.[2]

[1] Using $\sum_{k=-\infty}^{+\infty} |p(\xi_{n,k})|^3 \dot{p}(\xi_{n,k}) = 1/T \int_{-\infty}^{+\infty} |p(t)|^3 \dot{p}(t) \mathrm{d}t$, for sufficiently small T.

[2] Interestingly, $F_{\theta\tau} \neq 0$ for non-even pulses.

Derivation of $F_{\theta\theta}$ and $F_{\tau\tau}$: Using similar reasoning, $F_{\theta\theta}$ is given as

$$F_{\theta\theta} = \frac{2A}{N_a N_0/T}\sum_{k=-\infty}^{+\infty}\sum_{n=1}^{N} E_2 \left|p(kT-nT_s-\tau)\right|^2 = \frac{4AN}{N_a N_0}, \quad (15)$$

where we have used $E_2 = 2$ and $\int |p(t)|^2 \mathrm{d}t = 1$. Note that $F_{\theta\theta}$ does not depend on the nonlinearity of the channel. Finally, it is easily shown that

$$\mathbb{E}\left\{\frac{\partial \tilde{\mathbf{x}}_k^{\mathrm{H}}}{\partial \tau}\frac{\partial \tilde{\mathbf{x}}_k}{\partial \tau}\right\} = A\sum_{n=1}^{N}\left|\dot{p}(\xi_{n,k})\right|^2\left(2+4\beta^2A^2E_6\left|p(\xi_{n,k})\right|^4\right). \quad (16)$$

Then, $F_{\tau\tau}$ is given as

$$F_{\tau\tau} = \frac{4AN}{N_a N_0} I_{02} + 8\frac{\gamma^2 L_{\mathrm{eff}}^2 N_a A^3 E_6 N}{N_0} I_{42}. \quad (17)$$

Note that the first term is the conventional Fisher information without nonlinearity.

MCRB: Substitution of $F_{\theta\tau}$, $F_{\tau\tau}$, and $F_{\theta\theta}$ into (6) leads to

$$\mathrm{MCRB}\left(\theta\right) = \frac{\frac{N_a N_0}{4AN}}{1-\frac{\gamma^2 L_{\mathrm{eff}}^2 A^2 E_4^2 J_{31}^2}{\frac{I_{02}}{N_a^2}+2\gamma L_{\mathrm{eff}}^2 A^2 E_6 I_{42}}}, \quad (18)$$

where the numerator is the conventional MCRB without nonlinearity. It is easily verified that the second term in the denominator is nonnegative, so that the nonlinearity always increases the MCRB. Similarly

$$\mathrm{MCRB}\left(\tau\right) = \frac{\frac{N_a N_0}{4ANI_{02}}}{1+\gamma^2 L_{\mathrm{eff}}^2 N_a^2 A^2 \frac{2E_6 I_{42} - E_4^2 J_{31}^2}{I_{02}}}, \quad (19)$$

where the numerator is the conventional MCRB without nonlinearity. Contrary to the phase, the second term in the denominator of $\mathrm{MCRB}\left(\tau\right)$ can be positive or negative (depending on the sign of $2E_6 I_{42} - E_4^2 J_{31}^2$), so that the nonlinearity can reduce or increase the MCRB. In particular, for even pulses $J_{31}^2 = 0$, so that nonlinearity can only reduce the MCRB.

C. Derivation of MCRB with Low-Pass Filter

Here, $\mathbf{r}_{\mathrm{vec}}$ denotes the vector representation of the signal after filtering $\tilde{\mathbf{r}}(t)$ by the filter $h(t)$ (the inverse Fourier transform of $H(f)$, introduced in Section II-A). We define the filtered signal as (with $\otimes$ denoting convolution)

$$\mathbf{r}_F(t) = \left(\mathbf{U}\mathbf{x}_0(t)\exp\left(j\beta\left\|\mathbf{x}_0(t)\right\|^2\right)\right)\otimes h(t) + \mathbf{w}(t) = \sqrt{A}e^{j\theta}\mathbf{U}\sum_{n=1}^{N}\mathbf{s}_n z_n(t-\tau;\mathbf{s}_n) + \mathbf{w}(t), \quad (20)$$

where we have introduced $\mathbf{w}(t)$ as the noise at the output of the filter, and

$$z_n(t;\mathbf{s}_n) = \left(p(t-nT_s)\exp\left(j\beta A\left\|\mathbf{s}_n\right\|^2\left|p(t-nT_s)e^{j\theta}\right|^2\right)\right)\otimes h(t). \quad (21)$$

Table I

SYSTEM AND CHANNEL PARAMETER VALUES

Parameters	Symbol	Value, unit
Nonlinearity parameter	γ	1.2 $\mathrm{W}^{-1}\mathrm{km}^{-1}$
Attenuation	α	0.25 dB/km
Length/span	L	80 km
Spontaneous emission factor	n_{sp}	1.5
Bandwidth	B, $1/T_s$	14 GHz
Number of spans	N_a	22
Wavelength	λ	1.55 μm

From $\mathbf{r}_F(t)$ we can create a vector representation $\mathbf{r}_{\mathrm{vec}}$ and again determine the Fisher information matrix. Following a line of reasoning similar to section III-B, we easily find that

$$[\mathbf{F}(\mathbf{\Theta})]_{ij} = \frac{2AN}{N_a N_0/T}\mathbb{E}\left[\left\|\mathbf{s}_n\right\|^2 G(\Theta_i,\Theta_j;\mathbf{s}_n)\right], \quad (22)$$

where

$$G(\Theta_i,\Theta_j;\mathbf{s}_n) = \sum_{k=-\infty}^{+\infty}\Re\left\{\frac{\partial\left\{z_n^{\mathrm{H}}(kT-\tau;\mathbf{s}_n)e^{-j\theta}\right\}}{\partial\Theta_i}\frac{\partial\left\{z_n(kT-\tau;\mathbf{s}_n)e^{j\theta}\right\}}{\partial\Theta_j}\right\}. \quad (23)$$

Derivation of $F_{\theta\theta}$, $F_{\tau\tau}$, and $F_{\theta\tau}$: It is readily shown that

$$G(\theta,\theta;\mathbf{s}_n) = \frac{1}{T}\int_{-1/T}^{1/T}\left|Z_n(f;\mathbf{s}_n)\right|^2\mathrm{d}f, \quad (24)$$

$$G(\tau,\tau;\mathbf{s}_n) = \frac{4\pi^2}{T}\int_{-1/T}^{1/T}f^2\left|Z_n(f;\mathbf{s}_n)\right|^2\mathrm{d}f, \quad (25)$$

$$G(\theta,\tau;\mathbf{s}_n) = -\frac{2\pi}{T}\int_{-1/T}^{1/T}f\left|Z_n(f;\mathbf{s}_n)\right|^2\mathrm{d}f. \quad (26)$$

Substitution into (22) then yields $F_{\theta\theta}$, $F_{\tau\tau}$, and $F_{\theta\tau}$, respectively. The expectation in (22) over the data symbols $\mathbf{s}_n$ can be carried out numerically for any pulse. For even pulses, $G(\theta,\tau;\mathbf{s}_n) = 0$, so that $F_{\theta\tau} = 0$. From $F_{\theta\theta}$, $F_{\tau\tau}$, and $F_{\theta\tau}$, we can then compute $\mathrm{MCRB}\left(\theta\right)$ and $\mathrm{MCRB}\left(\tau\right)$.

IV. NUMERICAL RESULTS

A. Scenario

We will now investigate a 112 Gbit/s 16-QAM dual polarization system using the system parameters given in Table I. The pulse $p(t)$ is an RZ pulse with duty cycles of 33% or 67% from [8], and $\mathbf{U} = \mathbf{I}_2$. We will consider two low-pass filters: one for $T = T_s$ (labeled LPF1) and one for $T = T_s/2$ (labeled LPF2). The MCRBs derived for the simplified model from Section II-B will be complemented with performance results of practical estimators, applied to the model from Section II-A, extended to dual polarization: a feed-forward (FF) timing estimator from [1, pp. 433-437] and the well-known Viterbi and Viterbi (V-V) phase estimator [1, pp. 280-281].

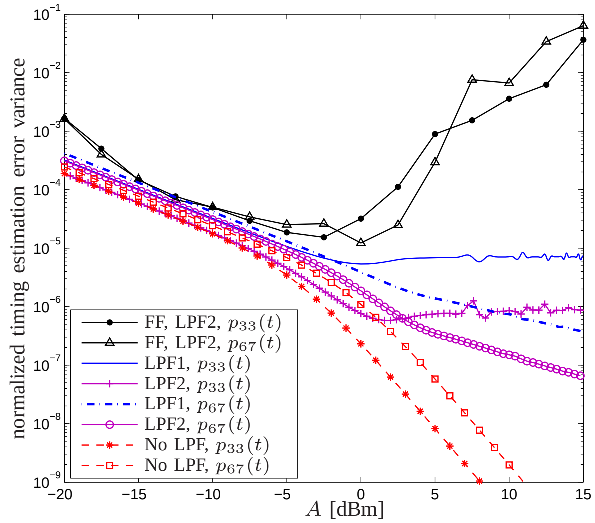

Figure 1. MCRB for timing without and two cases with LPF (LPF1: $T = T_s$ and LPF2: $T_s = 0.5T$) for even pulses with duty cycles of 33% and 67% along with FF timing estimation.

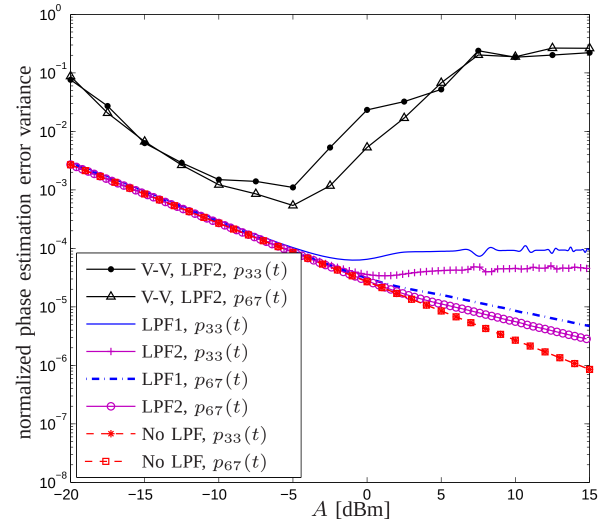

Figure 2. MCRB for phase without and two cases with LPF (LPF1: $T = T_s$ and LPF2: $T_s = 0.5T$) for even pulses with duty cycles of 33% and 67% along with phase estimation by V-V algorithm.

B. Results

Timing: For timing, we observe that without low-pass filter, the slope of the MCRB increases for input powers above -5 dBm. The pulse with the shortest duty cycle has the lowest MCRB. For the pulse with 33% duty cycle, with LPF1 (resp. LPF2), $\mathrm{MCRB}(\tau)$ stops decreasing monotonously after -1.5 dBm (resp. $+1.5$ dBm), and exhibits underdamped oscillations. This is due to the interaction between spectral broadening (energy leaking outside the filter bandwidth) and increased received power. For the pulse with 67% duty cycle, we observe a reduced slope of the MCRB with higher input powers. The pulse with the shortest duty cycle now has the highest MCRB. This can be explained by the fact that the shorter the duty cycle, the higher the peak power, thus the more susceptible the pulse is to SPM and spectral broadening. The practical FF estimator follows the MCRB up to around -5 dBm, and then degrades significantly. We see that the MCRB is quite loose at high transmit power, but still provides insight as to when algorithms may fail.

Phase: For phase, we see from Fig. 2 that contrary to $\mathrm{MCRB}(\tau)$, $\mathrm{MCRB}(\theta)$ without LPF does not depend on the pulse shape. In the case with low-pass filter, the conclusions from $\mathrm{MCRB}(\tau)$ carry over. As V-V operates after FF timing recovery, we expect a large gap between the V-V estimator and the MCRB, which can be attributed to the use of the simplified model.

Relation to BER: Finally, we note that to achieve a pre-decoding BER of 10^{-3} with 16-QAM, the required input power is approximately -2 dBm. Both $\mathrm{MCRB}(\tau)$ and $\mathrm{MCRB}(\theta)$ are in the nonlinear regime at this input power level. Hence, synchronization algorithms may fail unless we use a sufficiently broad low-pass filter.

V. Conclusions

We have derived approximate MCRBs for timing and phase estimation in presence of nonlinear self-phase modulation for single-channel optical communication at data rates below 10 Gigabaud with dispersion compensated fibers. We found that the MCRB for timing estimation is reduced due to the presence of SPM, but only when a prefilter with sufficiently large bandwidth is present. The MCRB for phase estimation is highly dependent on the bandwidth of the prefilter, with a too narrow prefilter leading to a flooring effect of the MCRB. Finally, we observed that the derived MCRBs give a good indication of the performance of practical estimators.

References

[1] U. Mengali and A. D'Andrea, *Synchronization techniques for digital receivers*. Springer, 1997.

[2] D. Zibar, A. Bianciotto, Z. Wang, A. Napoli, and B. Spinnler, "Analysis and dimensioning of fully digital clock recovery for 112 Gb/s coherent polmux QPSK systems," *Proceedings ECOC*, 2009.

[3] F. Hauske, C. Xie, N. Stojanovic, and M. Chen, "Analysis of polarization effects to digital timing recovery in coherent receivers of optical communication systems," *ITG-Fachbericht-Photonische Netze*, 2010.

[4] E. Ip and J. Kahn, "Feedforward carrier recovery for coherent optical communications," *Journal of Lightwave Technology*, vol. 25, no. 9, pp. 2675–2692, 2007.

[5] T. Nakagawa, M. Matsui, T. Kobayashi, K. Ishihara, R. Kudo, M. Mizoguchi, and Y. Miyamoto, "Non-data-aided wide-range frequency offset estimator for QAM optical coherent receivers," in *Optical Fiber Communication Conference*, Optical Society of America, 2011.

[6] 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.

[7] A. S. Tan, H. Wymeersch, P. Johannisson, M. Sjödin, E. Agrell, M. Karlsson, and P. A. Andrekson, "Modified Cramér–Rao bound for clock recovery in the presence of self-phase modulation," *Journal of Lightwave Technology*, vol. 30, no. 16, pp. 2556–2562, 2012.

[8] 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.

[9] 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.