Article
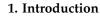
Novel Results on SNR Estimation for Bandlimited Optical Intensity Channels

Wilfried Gappmair [ID]

Institute of Communication Networks and Satellite Communications, Graz University of Technology, Inffeldgasse 12, 8010 Graz, Austria; gappmair@tugraz.at

Abstract: In a previous work of the author about non-data-aided estimation of the signal-to-noise ratio (SNR) for bandlimited optical intensity channels, a couple of limitations have been identified in terms of error performance and computational complexity. In the current paper, these deficiencies are avoided by the introduction of a second receiver filter with specific properties that is operated in parallel to the receiver filter normally used in this respect. Although not initially intended, the concept is also applied to data-aided SNR estimation by deriving a maximum likelihood algorithm and the Cramer–Rao lower bound (CRLB) as the theoretical limit of the error performance. In the next step, the dual-filter framework is used in the context of SNR estimation without knowledge about data symbols. The most significant benefit of this method is that the number of payload data employed for the estimation procedure might be selected arbitrarily long without impacting the spectral efficiency of the link. Since the computation of the true CRLB was out of scope due to complexity reasons, an asymptotic variant for very low SNR values is analyzed, which ends up in a closed-form solution. Furthermore, an algorithm based on first- and second-order moments of the samples at the dual-filter output is investigated, which turned out to be very attractive in terms of error performance and computational complexity.

Keywords: SNR estimation; optical wireless communications; intensity modulation

1. Introduction

There is no doubt that optical wireless communication (OWC) solutions have a lot of benefits compared to their radio frequency (RF) counterparts: rather inexpensive and easy to deploy, extremely high throughput, no problems with data security, no regulatory and license issues, just to mention the most significant aspects in this context [1–4]. However, when it comes to optical intensity modulation, a unipolar signal concept with respect to symbol constellation and pulse shaping is of paramount importance. Realized via pulse amplitude modulation (PAM) and root-raised cosines, normally used in the RF domain for pulse shaping, this has been investigated in [5,6] by assuming an appropriately selected bias. Although the approach is strictly bandlimited, the price to be paid is some clipping effect and less efficiency in terms of power and energy.

As an alternative, squared raised cosine and double jump functions have been suggested in [7] for pulse shaping in bandlimited optical intensity links. Clipping as well as bias problems are completely avoided by this method, with the additional benefit that the Nyquist property is still satisfied, which allows a simple detection procedure without the introduction of inter-symbol interference effects.

Nevertheless, even for OWC systems, the most important transmission parameters must be recovered at the receiver by suitable algorithms. In this context, some knowledge about the signal-to-noise ratio (SNR) is indispensable for a lot of scenarios; e.g., many adaptive systems [8] or powerful error correction schemes [9] necessitate this sort of information so that the link can be operated close to the Shannon bound. In this context, a non-data-aided (NDA) algorithm for SNR estimation has been discussed by the author

in a recently published paper [10], i.e., no a priori knowledge about data is needed in the receiver unit for this purpose, or, in other words, payload data might be used so that the spectral efficiency does not depend on the observation length required for the estimation process. This is in striking contrast to a data-aided (DA) approach, where part of the data frame is occupied by pilot sequences [11].

However, the solution developed in [10] is based on the expectation-maximization (EM) principle [12–14], which turned out to be the only way to approach the Cramer–Rao lower bound (CRLB) as the theoretical limit of the error performance [15–17]. Other algorithms for SNR estimation have been studied as well, e.g., moment-based or decision-directed methods, but they failed insofar as they exhibit a non-negligible bias and/or jitter effect. Unfortunately, the computational load of EM algorithms is considerable, so a less complex SNR estimator would be most welcome. In the current paper, this is realized by the introduction of a second filter operated in parallel to the receiver filter and using both outputs in a moment-based approach, which is characterized by its computational simplicity and the fact that the performance degradation of EM solutions in the medium SNR range is avoided.

The rest of the paper is organized as follows: In Section 2, the signal and channel model is introduced, which we use for analytical and simulation work in the sequel. In Section 3, a maximum likelihood algorithm is investigated for DA SNR estimation, complemented by the derivation of the related CRLB as the theoretical limit of the error performance. The focus of Section 4 is on the development of a method based on the first- and second-order moments of the samples available at the output of the receiver filters. Since the computation of the true CRLB is out of scope in the context of NDA SNR estimation, we concentrate on an asymptotic variant, which is available in closed form and applies to smaller SNR values. Numerical results related to bias effect and error performance are presented in Section 5. Finally, conclusions are drawn in Section 6.

2. Signal and Channel Model

The signal and channel model for the current paper are more or less the same as those in [10,11] for NDA as well as DA estimation of the SNR, respectively. Nevertheless, for convenience, the model is briefly recapitulated in the sequel. On top of that, it is also extended in a suitable manner so that new results are achievable.

In the following, it is assumed that the data symbols a_k, $k \in \mathbb{Z}$, are independent and identically distributed (i.i.d.) elements of an M-ary PAM alphabet $\mathcal{A}$. In this context, it makes sense to normalize the symbols such that $\mathbb{E}[a_k^2] = 1$, where $\mathbb{E}[\cdot]$ denotes the expectation operator. This is straightforwardly achieved by $\mathcal{A} = \frac{1}{\sqrt{\eta_M}}\{0, \ 1, \ \ldots, \ M-1\}$ and $\eta_M = \frac{1}{6}(M-1)(2M-1)$. As a consequence, the average symbol value is given by

$$\mu_a = \mathbb{E}[a_k] = \frac{1}{\sqrt{\eta_M}} \ \frac{M-1}{2} = \sqrt{\frac{3\,(M-1)}{2\,(2M-1)}} \tag{1}$$

If we adopt a pulse shape expressed by $h(t)$, the signal at the output of the opto-electrical receiver module develops as [10,11]

$$r(t) = A \sum_k a_k \, h(t - kT - \tau) + w(t) \tag{2}$$

where $A > 0$ is the channel gain, T and τ specify the symbol period as well as the propagation delay between transmitter and receiver, respectively. Of course, the signal part in (2) is impaired by $w(t)$, representing a zero-mean white Gaussian noise component with variance σ_w^2.

Furthermore, by introduction of

$$\overline{h} = \frac{1}{\sqrt{T}} \int\limits_{-\infty}^{\infty} h(t) \, dt \tag{3}$$

and the average optical power $P_0 = \mu_a \overline{h}$, the average electrical SNR is described as

$$\gamma_s = \frac{A^2 P_0^2}{\sigma_w^2} \tag{4}$$

Before the signal in (2) can be processed in further receiver stages, it has to be filtered appropriately. Denoting the corresponding impulse response by $q(t)$, the related output is obtained as $z(t) = q(t) \otimes r(t)$, where $\otimes$ characterizes the convolutional operator. This scenario is basically illustrated in Figure 1. Regarding a radio frequency (RF) system, $q(t)$ would be designed as a matched filter in order to maximize the SNR at the output. It has been shown in [18,19] that suitable pairs of transmitter-receiver filters satisfying the non-negativity and Nyquist properties are only attainable via an unconstrained min–max optimization procedure, which must be solved by rather cumbersome numerical means whenever a new filter design is required. In order to avoid this, it is suggested to implement a rectangular shape over the spectrum occupied by the user component in (2).
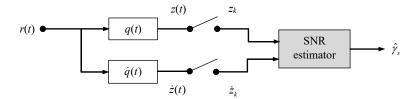

Figure 1. Signal model for SNR estimation using a dual-filter framework.

By application of the Fourier transform [20], we simply obtain $Q(f) = \mathcal{F}[q(t)] = \sqrt{T}$ for $|f| \leq (1 + \alpha)/T$ and $Q(f) = 0$ elsewhere, with α as the roll-off factor (excess bandwidth) of the selected pulse shape. As a result, the signal parts of $r(t)$ and $z(t)$ are the same, whereas the noise component is determined by $n(t) = w(t) \otimes q(t)$, representing a zero-mean non-white Gaussian process.

As already mentioned previously, a second filter will be operated in parallel to the receiver filter, as visualized in Figure 1. Performing an impulse response expressed by $\dot{q}(t)$, the corresponding output is given by $\dot{z}(t) = \dot{q}(t) \otimes r(t)$. Under the assumption that the symbol timing has been reliably recovered by properly selected estimation or synchronization algorithms [21–23], it is required that $\dot{q}(t)$ exhibits an impulse response such that the signal component of $\dot{z}(t)$ vanishes at integer multiples of the symbol period, which is satisfied only if $h(t) \otimes \dot{q}(t)\big|_{t=kT} = 0$ for all $k \in \mathbb{Z}$. In this respect, it is to be noticed that the first-order derivative of squared raised cosine or double jump functions fulfills this pre-requisite [7], i.e., $\partial h(t)/\partial t\big|_{t=kT} = 0$, $k \in \mathbb{Z}$. Hence, by shifting this problem to the frequency domain, the Fourier transform of $\dot{q}(t)$ is determined by $\dot{Q}(f) = \mathcal{F}[\dot{q}(t)] = j2\pi f T \, Q(f)$.

After having synchronized symbol timing and clocking, e.g., by implementation of one of the recovery algorithms investigated in [21–23], the T-spaced samples at the output of the dual-filter concept are furnished by

$$z_k = z(kT) = A \cdot a_k + n_k \tag{5}$$

and

$$\dot{z}_k = \dot{z}(kT) = \dot{n}_k \tag{6}$$

It is to be emphasized that the parallel filter delivers solely noise samples denoted by $\dot{n}_k$, which is explained by the filter properties introduced previously. Furthermore, both noise processes, n_k and $\dot{n}_k$, have a zero-mean non-white Gaussian character.

3. Data-Aided SNR Estimation

3.1. Log-Likelihood Function

The Cramer–Rao lower bound (CRLB) is of paramount importance when it comes to the estimation of a parameter. This is not only true for communication links, as exemplified in [16,17], but relates to any technical system [15]. The reason behind this is the fact that it represents the theoretical limit of the error performance of any estimator developed in this respect. In order to derive the CRLB of SNR estimates achieved by the dual-filter framework in Figure 1, we assume that L observables z_k and $\dot{z}_k$, $k = 0, 1, \ldots, L - 1$ are available at the output of the receiver filters. Arranging these observables in vector form, we have that $\mathbf{z} = A \cdot \mathbf{a} + \mathbf{n}$ and $\dot{\mathbf{z}} = \dot{\mathbf{n}}$, where $\mathbf{a}$, $\mathbf{n}$ and $\dot{\mathbf{n}}$ denote the corresponding data and noise sequences, each of them with L elements. For convenient reasons, we put these pieces together so that we have a single vector equation given by

$$\mathbf{y} = \begin{pmatrix} \mathbf{z} \\ \dot{\mathbf{z}} \end{pmatrix} = A \cdot \begin{pmatrix} \mathbf{a} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{n} \\ \dot{\mathbf{n}} \end{pmatrix} = A \cdot \mathbf{c} + \mathbf{v} \tag{7}$$

where $\mathbf{0}$ stands for a vector with L zero entries.

According to the channel model introduced in the previous section, the average electrical SNR in (4) is a function of the channel gain A and the standard deviation σ_w of the zero-mean white Gaussian noise process. Since P_0 depends on the PAM scheme as well as the pulse shape selected for this purpose, both of which are known in advance by the receiver unit, we better focus on the estimation of the SNR normalized by P_0^2, i.e., $\rho_s = \gamma_s / P_0^2 = A^2 / \sigma_w^2$. Therefore, the parameter vector used for the computation of the CRLB is characterized by $\mathbf{u} = (A, \sigma_w)$.

The major ingredient for the derivation of the CRLB is knowledge of the likelihood function $\Pr(\mathbf{y}|\mathbf{c}; \mathbf{u})$ describing the signal model in (7). Since both noise components, $\mathbf{n}$ and $\dot{\mathbf{n}}$, are zero-mean Gaussian variates, the likelihood function is furnished by [24,25]

$$\Pr(\mathbf{y}|\mathbf{c}; \mathbf{u}) = \frac{1}{\sqrt{(2\pi)^{2L} \det(\mathbf{R})}} e^{-\frac{1}{2}\, (\mathbf{y} - A\, \mathbf{c})^T \mathbf{R}^{-1} (\mathbf{y} - A\, \mathbf{c})} \tag{8}$$

with $\mathbf{R} = \mathbb{E}[(\mathbf{y} - A\,\mathbf{c})(\mathbf{y} - A\mathbf{c})^T]$ as the corresponding covariance matrix. This $2L \times 2L$ matrix can be partitioned into four $L \times L$ submatrices as follows:

$$\mathbf{R} = \mathbb{E}[\mathbf{v} \cdot \mathbf{v}^T] = \begin{pmatrix} \mathbb{E}[\mathbf{n} \cdot \mathbf{n}^T] & \mathbb{E}[\mathbf{n} \cdot \dot{\mathbf{n}}^T] \\ \mathbb{E}[\dot{\mathbf{n}} \cdot \mathbf{n}^T] & \mathbb{E}[\dot{\mathbf{n}} \cdot \dot{\mathbf{n}}^T] \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix} \tag{9}$$

For line $i = 0, 1, \ldots, L - 1$ and column $k = 0, 1, \ldots, L - 1$, the entries of $\mathbf{R}_{11}$ are determined by the auto-correlation of the noise samples n_i, i.e.,

$$\begin{aligned} \mathbb{E}[n_i\, n_k] &= \sigma_w^2 \int\limits_{-(1+\alpha)/T}^{(1+\alpha)/T} e^{j2\pi(i-k)fT} Q(f)Q(-f)\, df \\ &= 2(1+\alpha)\, \sigma_w^2\, \mathrm{sinc}[2(1+\alpha)(i-k)] \end{aligned} \tag{10}$$

with $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$. On the other hand, the entries of $\mathbf{R}_{12}$ characterize the cross-correlation between noise samples n_i and $\dot{n}_k$, which is given by

$$
\begin{aligned}
\mathbb{E}[n_i\,\dot{n}_k] &= \sigma_w^2 \int\limits_{-(1+\alpha)/T}^{(1+\alpha)/T} e^{j2\pi(i-k)fT} Q(f)\dot{Q}(-f)\,df \\
&= 2(1+\alpha)\sigma_w^2 \begin{cases} 0, & i = k \\ \frac{-\cos[2\pi(1+\alpha)(i-k)]+\mathrm{sinc}[2(1+\alpha)(i-k)]}{i-k}, & i \neq k \end{cases}
\end{aligned}
\tag{11}
$$

The elements of $\mathbf{R}_{21}$ are determined by $\mathbb{E}[\dot{n}_i n_k]$, which is equivalent to (11) after having swapped indexes i and k, i.e., $\mathbf{R}_{21}$ is the transpose of $\mathbf{R}_{12}$. Finally, we have that the entries of $\mathbf{R}_{22}$ are specified by the auto-correlation of $\dot{n}_i$, which yields

$$
\begin{aligned}
\mathbb{E}[\dot{n}_i\,\dot{n}_k] &= \sigma_w^2 \int\limits_{-(1+\alpha)/T}^{(1+\alpha)/T} e^{j2\pi(i-k)fT} \dot{Q}(f)\dot{Q}(-f)\,df \\
&= 2(1+\alpha)\sigma_w^2 \begin{cases} \frac{4\pi^2(1+\alpha)^2}{3}, & i = k \\ \frac{2\cos[2\pi(1+\alpha)(i-k)]+([2\pi(1+\alpha)(i-k)]^2-2)\mathrm{sinc}[2(1+\alpha)(i-k)]}{(i-k)^2}, & i \neq k \end{cases}
\end{aligned}
\tag{12}
$$

Extracting the common factor, we simply obtain

$$
\mathbf{R} = 2(1+\alpha)\,\sigma_w^2\,\mathbf{\Omega}
\tag{13}
$$

where the entries of $\mathbf{\Omega}$ are only functions of the roll-off factor α. Furthermore, it is not difficult to show that $\mathbf{\Omega} = \mathbf{\Omega}^T$. Hence, with $\mathbf{\Psi} = \mathbf{\Omega}^{-1}$, $\mathbf{\Psi} = \mathbf{\Psi}^T$, and substituting (13) into (8), the log-likelihood function (LLF) of our signal model, i.e., $\Lambda(\mathbf{y}|\mathbf{c};\mathbf{u}) = \log\Pr(\mathbf{y}|\mathbf{c};\mathbf{u})$, is expressed by

$$
\Lambda(\mathbf{y}|\mathbf{c};\mathbf{u}) = -L\log\sigma_w^2 - \frac{\mathbf{y}^T\mathbf{\Psi}\,\mathbf{y} - 2A\,\mathbf{y}^T\mathbf{\Psi}\,\mathbf{c} + A^2\mathbf{c}^T\mathbf{\Psi}\,\mathbf{c}}{4(1+\alpha)\sigma_w^2}
\tag{14}
$$

after having omitted immaterial constants and factors not depending on $\mathbf{u}$.

3.2. Modified Cramer–Rao Lower Bound

Knowing the LLF, we are able to derive the elements of the Fisher information matrix (FIM), which are determined by the second-order derivatives of the LLF with respect to A and σ_w. However, by definition of $P_n = \sigma_w^2$ and replacing the channel gain A in (14) by $\sqrt{\rho_s P_n}$, the derivatives are related to ρ_s and P_n. This approach has the advantage that the CRLB for the normalized SNR is directly obtained without the application of cumbersome transformation rules [15]. As a consequence, we have that

$$
\frac{\partial^2 \Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})}{\partial \rho_s^2} = -\frac{\mathbf{y}^T\mathbf{\Psi}\,\mathbf{c}}{8(1+\alpha)\sqrt{\rho_s^3 P_n}}
\tag{15}
$$

$$
\frac{\partial^2 \Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})}{\partial P_n^2} = \frac{L}{P_n^2} - \frac{\mathbf{y}^T\mathbf{\Psi}\,\mathbf{y}}{2(1+\alpha)P_n^3} + \frac{3\sqrt{\rho_s}\mathbf{y}^T\mathbf{\Psi}\,\mathbf{c}}{8(1+\alpha)\sqrt{P_n^5}}
\tag{16}
$$

$$
\frac{\partial^2 \Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})}{\partial \rho_s\,\partial P_n} = \frac{\partial^2 \Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})}{\partial P_n\,\partial \rho_s} = -\frac{\mathbf{y}^T\mathbf{\Psi}\,\mathbf{c}}{8(1+\alpha)\sqrt{\rho_s^3 P_n}}
\tag{17}
$$

Plugging now $\mathbf{y} = A\,\mathbf{c} + \mathbf{\nu} = \sqrt{\rho_s P_n}\,\mathbf{c} + \mathbf{\nu}$ into (15)–(17) and averaging with respect to $\mathbf{\nu}$ provides us with FIM elements depending on the selected pilot sequence $\mathbf{c}{\sim}\mathbf{a}$. In order to avoid this sort of restriction, we may extend the expected operation to $\mathbf{c}$ as well, which

constitutes the so-called modified Cramer–Rao lower bound (MCRLB) [26–28]. By doing so, we get

$$J_{11} = -\mathbb{E}\left[\frac{\partial^2 \Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})}{\partial \rho_s^2}\right] = \frac{1}{8(1+\alpha)\rho_s}\mathbb{E}_\mathbf{c}[\mathbf{c}^T\mathbf{\Psi}\,\mathbf{c}] \tag{18}$$

$$J_{22} = -\mathbb{E}\left[\frac{\partial^2 \Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})}{\partial P_n^2}\right] = -\frac{L}{P_n^2} + \frac{\rho_s}{8(1+\alpha)P_n^2}\mathbb{E}_\mathbf{c}[\mathbf{c}^T\mathbf{\Psi}\,\mathbf{c}] + \frac{1}{2(1+\alpha)P_n^3}\mathbb{E}_\mathbf{v}[\mathbf{v}^T\mathbf{\Psi}\,\mathbf{v}] \tag{19}$$

$$J_{12} = -\mathbb{E}\left[\frac{\partial^2 \Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})}{\partial \rho_s \partial P_n}\right] = \frac{1}{8(1+\alpha)P_n}\mathbb{E}_\mathbf{c}[\mathbf{c}^T\mathbf{\Psi}\,\mathbf{c}] \tag{20}$$

As a result, the MCRLB for ρ_s is achieved as

$$\begin{aligned}
\text{MCRLB}(\rho_s) &= \frac{J_{22}}{J_{11}J_{22}-J_{12}^2} \\
&= 2(1+\alpha)\rho_s^2\left(\frac{P_n}{\mathbb{E}_\mathbf{v}[\mathbf{v}^T\mathbf{\Psi}\,\mathbf{v}]-2L(1+\alpha)P_n} + \frac{4}{\rho_s\,\mathbb{E}_\mathbf{c}[\mathbf{c}^T\mathbf{\Psi}\,\mathbf{c}]}\right)
\end{aligned} \tag{21}$$

with $\mathbb{E}_\mathbf{c}[\cdot]$ and $\mathbb{E}_\mathbf{v}[\cdot]$ indicating the averaging procedures with respect to data and noise. Next, we assume that the $2L \times 2L$ matrix $\mathbf{\Psi}$ is partitioned into four $L \times L$ submatrices according to

$$\mathbf{\Psi} = \begin{pmatrix} \mathbf{\Psi}_{11} & \mathbf{\Psi}_{12} \\ \mathbf{\Psi}_{21} & \mathbf{\Psi}_{22} \end{pmatrix} \tag{22}$$

where $\psi_{mn,ik}$ denotes the entry of $\mathbf{\Psi}_{mn}$ for line i and column k. Hence, by considering the definition of $\mathbf{c}$ in (7), it is clear that

$$\mathbb{E}_\mathbf{c}[\mathbf{c}^T\mathbf{\Psi}\,\mathbf{c}] = \mathbb{E}_\mathbf{a}[\mathbf{a}^T\mathbf{\Psi}_{11}\,\mathbf{a}] = \sum_{i=0}^{L-1}\sum_{k=0}^{L-1}\mathbb{E}[a_i a_k]\,\psi_{11,ik} = L(\overline{\Psi}_0 + 2\mu_a^2\overline{\Psi}_1) \tag{23}$$

after having taken into account that $\mathbb{E}[a_i^2] = 1$ and $\mathbb{E}[a_i a_k]_{i\neq k} = \mu_a^2$ as well as using the sums

$$\overline{\Psi}_0 = \frac{1}{L}\sum_{i=0}^{L-1}\psi_{11,ii}, \quad \overline{\Psi}_1 = \frac{1}{L}\sum_{i=0}^{L-1}\sum_{k=i+1}^{L-1}\psi_{11,ik} \tag{24}$$

Furthermore, if we employ the results in (10)–(12), the averaging procedure with respect to noise yields

$$\begin{aligned}
\mathbb{E}_\mathbf{v}[\mathbf{v}^T\mathbf{\Psi}\,\mathbf{v}] &= \mathbb{E}_\mathbf{v}[\mathbf{n}^T\mathbf{\Psi}_{11}\,\mathbf{n} + \mathbf{n}^T\mathbf{\Psi}_{12}\,\dot{\mathbf{n}} + \dot{\mathbf{n}}^T\mathbf{\Psi}_{21}\,\mathbf{n} + \dot{\mathbf{n}}^T\mathbf{\Psi}_{22}\,\dot{\mathbf{n}}] \\
&= \sum_{i=0}^{L-1}\sum_{k=0}^{L-1}\left\{\mathbb{E}[n_i n_k]\,\psi_{11,ik} + \mathbb{E}[n_i \dot{n}_k]\,\psi_{12,ik} + \mathbb{E}[\dot{n}_i n_k]\,\psi_{21,ik} + \mathbb{E}[\dot{n}_i \dot{n}_k]\,\psi_{22,ik}\right\} \\
&= 4L(1+\alpha)\,P_n\,\overline{\Psi}_2
\end{aligned} \tag{25}$$

where

$$\overline{\Psi}_2 = \frac{1}{2L}\sum_{i=0}^{2L-1}\sum_{k=0}^{2L-1}\omega_{ik}\,\psi_{ik} \tag{26}$$

with ω_{ik} and ψ_{ik} as the entries of $\mathbf{\Omega}$ and $\mathbf{\Psi}$, respectively. Finally, by putting all these pieces together, the MCRLB in (21), for convenient reasons normalized by ρ_s^2, boils down to

$$\text{NMCRLB}(\rho_s) = \frac{\text{MCRLB}(\rho_s)}{\rho_s^2} = \frac{1}{L}\left(\frac{1}{2\overline{\Psi}_2 - 1} + \frac{8(1+\alpha)}{\rho_s(\overline{\Psi}_0 + 2\mu_a^2\overline{\Psi}_1)}\right) \tag{27}$$

3.3. Maximum Likelihood Estimator

The closed form of the LLF in (14) gives us the chance to develop a maximum likelihood (ML) estimator in a fairly straightforward manner by computing the first-order derivatives

of (14) with respect to channel gain A and noise power P_n, equating these results to zero, and solving both relationships for A and P_n. Specifically, we have for $\mathbf{u} = (A, P_n)$

$$\left.\frac{\partial \Lambda(\mathbf{y}|\mathbf{c}; \mathbf{u})}{\partial A}\right|_{\mathbf{u}=\hat{\mathbf{u}}} = \frac{\mathbf{y}^T \mathbf{\Psi} \, \mathbf{c} - \hat{A} \, \mathbf{c}^T \mathbf{\Psi} \, \mathbf{c}}{2(1+\alpha)\hat{P}_n} = 0 \tag{28}$$

and

$$\left.\frac{\partial \Lambda(\mathbf{y}|\mathbf{c}; \mathbf{u})}{\partial P_n}\right|_{\mathbf{u}=\hat{\mathbf{u}}} = -\frac{L}{\hat{P}_n} + \frac{\mathbf{y}^T \mathbf{\Psi} \, \mathbf{y} - 2\hat{A} \, \mathbf{y}^T \mathbf{\Psi} \, \mathbf{c} + \hat{A}^2 \mathbf{c}^T \mathbf{\Psi} \, \mathbf{c}}{4(1+\alpha)\hat{P}_n^2} = 0 \tag{29}$$

Introducing in the next step $M_{cc} = \mathbf{c}^T \mathbf{\Psi} \, \mathbf{c}$, $M_{cy} = \mathbf{y}^T \mathbf{\Psi} \, \mathbf{c}$, and $M_{yy} = \mathbf{y}^T \mathbf{\Psi} \, \mathbf{y}$, the estimates for channel gain and noise power are furnished by

$$\hat{A} = \frac{\mathbf{y}^T \mathbf{\Psi} \, \mathbf{c}}{\mathbf{c}^T \mathbf{\Psi} \, \mathbf{c}} = \frac{M_{cy}}{M_{cc}}, \quad M_{cc} > 0 \tag{30}$$

and

$$\hat{P}_n = \frac{\mathbf{y}^T \mathbf{\Psi} \, \mathbf{y} - 2\hat{A} \, \mathbf{y}^T \mathbf{\Psi} \, \mathbf{c} + \hat{A}^2 \mathbf{c}^T \mathbf{\Psi} \, \mathbf{c}}{4(1+\alpha)L} = \frac{1}{4(1+\alpha)L}\left(M_{yy} - \frac{M_{cy}^2}{M_{cc}}\right) \tag{31}$$

According to the invariance principle [29,30], the ML estimate is then given by

$$\hat{\rho}_s = \frac{\hat{A}^2}{\hat{P}_n} \tag{32}$$

Note that M_{cc} must be larger than zero; otherwise, the algorithm fails, i.e., a pilot sequence constituted by only zero elements would not work. In the following, the ML algorithm is summarized to be executed step by step (Algorithm 1):

Algorithm 1: ML estimator.

Initialization
    Pre-calculation of $\mathbf{\Psi} \leftarrow \mathbf{\Omega}^{-1}$
    Creation of pilot sequence $\mathbf{c}$ and vector $\mathbf{y}$ of observables
Computation
    Auxiliary terms: $M_{cc} \leftarrow \mathbf{c}^T \mathbf{\Psi} \, \mathbf{c}$, $M_{cy} \leftarrow \mathbf{y}^T \mathbf{\Psi} \, \mathbf{c}$, $M_{yy} \leftarrow \mathbf{y}^T \mathbf{\Psi} \, \mathbf{y}$
    Channel gain estimate: $\hat{A} \leftarrow \frac{M_{cy}}{M_{cc}}$
    Noise power estimate: $\hat{P}_n \leftarrow \frac{1}{4(1+\alpha)L}\left(M_{yy} - \frac{M_{cy}^2}{M_{cc}}\right)$
Output
    SNR estimate: $\hat{\rho}_s \leftarrow \frac{\hat{A}^2}{\hat{P}_n}$

Since $\mathbf{\Psi}$ is a $2L \times 2L$ matrix and $\mathbf{c}$ as well as $\mathbf{y}$ are vectors with $2L$ entries each, it is clear that the computational complexity of the SNR estimate is in the order of $\mathcal{O}(L^2)$ real-valued additions and multiplications.

4. Non-Data-Aided SNR Estimation

4.1. Asymptotic Cramer–Rao Lower Bound

For NDA estimation of the SNR, we cannot assume that the data sequence $\mathbf{c} \sim \mathbf{a}$ is known to the receiver. On the other hand, as already mentioned in the introductory section, via an NDA approach, we are in a position to take arbitrarily long portions of the payload data without impacting the spectral efficiency, as would be the case for a DA solution.

Regarding the computation of the CRLB, the likelihood function in (8) must be averaged first with respect to **c** before deriving the FIM elements [28]. Consequently, the related LLF is determined by

$$\Lambda(\mathbf{y};\mathbf{u}) = \log \mathbb{E}_{\mathbf{c}}[\Pr(\mathbf{y}|\mathbf{c};\mathbf{u})] = \log\left(\frac{1}{M^L}\sum_{\mathbf{c}\in\mathcal{A}^L} e^{\Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})}\right) \tag{33}$$

where $\mathcal{A}^L$ denotes the L-dimensional space spanned by the M-ary PAM alphabet. Of course, the computational complexity of (33) is in the order of $\mathcal{O}(M^L)$, i.e., even for smaller values of L, the evaluation of (33) is out of scope. Therefore, we will concentrate in the current paper on the derivation of the asymptotic Cramer–Rao lower bound (ACRLB), which applies only to lower SNR values but is achievable in closed form. To this end, the likelihood function is rewritten as

$$\Pr(\mathbf{y}\big|\mathbf{c};\mathbf{u}) = e^{\Lambda(\mathbf{y}|\mathbf{c};\mathbf{u})} = P_n^{-L}e^{-\Phi_1(\mathbf{y};\mathbf{u})+\Phi_2(\mathbf{y}|\mathbf{c};\mathbf{u})} \tag{34}$$

where

$$\Phi_1(\mathbf{y};\mathbf{u}) = \frac{\mathbf{y}^T\boldsymbol{\Psi}\,\mathbf{y}}{4(1+\alpha)P_n} \tag{35}$$

and

$$\Phi_2(\mathbf{y}|\mathbf{c};\mathbf{u}) = \frac{2\sqrt{\rho_s P_n}\,\mathbf{y}^T\boldsymbol{\Psi}\,\mathbf{c} - \rho_s P_n\,\mathbf{c}^T\boldsymbol{\Psi}\,\mathbf{c}}{4(1+\alpha)P_n} \tag{36}$$

By series expansion of (36), we obtain for smaller SNR values:

$$\Pr(\mathbf{y}|\mathbf{c};\mathbf{u}) = P_n^{-L}e^{-\Phi_1(\mathbf{y};\mathbf{u})}\sum_{k=0}^{\infty}\frac{\Phi_2^k(\mathbf{y}|\mathbf{c};\mathbf{u})}{k!}\bigg|_{\rho_s\ll 1} \approx P_n^{-L}e^{-\Phi_1(\mathbf{y};\mathbf{u})}[1+\Phi_2(\mathbf{y}|\mathbf{c};\mathbf{u})] \tag{37}$$

and averaging with respect to **c** yields

$$\Pr(\mathbf{y};\mathbf{u}) = \mathbb{E}_{\mathbf{c}}[\Pr(\mathbf{y}|\mathbf{c};\mathbf{u})]\big|_{\rho_s\ll 1} \approx P_n^{-L}e^{-\Phi_1(\mathbf{y};\mathbf{u})}[1+\Phi_2(\mathbf{y};\mathbf{u})] \tag{38}$$

where $\Phi_2(\mathbf{y};\mathbf{u}) = \mathbb{E}_{\mathbf{c}}[\Phi_2(\mathbf{y}|\mathbf{c};\mathbf{u})]$. As a result, the corresponding LLF is approximated by

$$\begin{aligned}\Lambda(\mathbf{y};\mathbf{u}) &= \log\Pr(\mathbf{y};\mathbf{u})\\ &\approx -L\log P_n - \Phi_1(\mathbf{y};\mathbf{u}) + \log[1+\Phi_2(\mathbf{y};\mathbf{u})]\\ &\approx -L\log P_n - \Phi_1(\mathbf{y};\mathbf{u}) + \Phi_2(\mathbf{y};\mathbf{u})\end{aligned} \tag{39}$$

In computing the second-order derivatives with respect to ρ_s and P_n, in the next step we arrive at

$$\frac{\partial^2\Lambda(\mathbf{y};\mathbf{u})}{\partial\rho_s^2} = \frac{\partial^2\Phi_2(\mathbf{y};\mathbf{u})}{\partial\rho_s^2} = -\frac{S_1(\mathbf{y})}{8(1+\alpha)\sqrt{\rho_s^3 P_n}} \tag{40}$$

$$\frac{\partial^2\Lambda(\mathbf{y};\mathbf{u})}{\partial P_n^2} = \frac{L}{P_n^2} - \frac{\partial^2\Phi_1(\mathbf{y};\mathbf{u})}{\partial P_n^2} + \frac{\partial^2\Phi_2(\mathbf{y};\mathbf{u})}{\partial P_n^2} = \frac{L}{P_n^2} - \frac{S_2(\mathbf{y})}{2(1+\alpha)P_n^3} + \frac{3\sqrt{\rho_s P_n}S_1(\mathbf{y})}{8(1+\alpha)P_n^3} \tag{41}$$

$$\frac{\partial^2\Lambda(\mathbf{y};\mathbf{u})}{\partial\rho_s\partial P_n} = \frac{\partial^2\Phi_2(\mathbf{y};\mathbf{u})}{\partial\rho_s\partial P_n} = -\frac{S_1(\mathbf{y})}{8(1+\alpha)\sqrt{\rho_s P_n^3}} \tag{42}$$

where $S_1(\mathbf{y}) = \mathbb{E}_{\mathbf{c}}[\mathbf{y}^T\boldsymbol{\Psi}\,\mathbf{c}]$ and $S_2(\mathbf{y}) = \mathbf{y}^T\boldsymbol{\Psi}\,\mathbf{y}$. The related FIM elements are achieved by substituting **y** into (40)–(42). The results are then averaged with respect to **y**, i.e., with respect to data and noise sequence, **c** and **v**, so that we have

$$J_{11} = -\mathbb{E}_{\mathbf{y}}\left[\frac{\partial^2\Lambda(\mathbf{y};\mathbf{u})}{\partial\rho_s^2}\right] = \frac{S_1}{8(1+\alpha)\sqrt{\rho_s^3 P_n}} \tag{43}$$

$$J_{22} = -\mathbb{E}_{\mathbf{y}}\left[\frac{\partial^2 \Lambda(\mathbf{y};\mathbf{u})}{\partial P_n^2}\right] = -\frac{L}{P_n^2} + \frac{S_2}{2(1+\alpha)P_n^3} - \frac{3\sqrt{\rho_s P_n}S_1}{8(1+\alpha)P_n^3} \tag{44}$$

$$J_{12} = -\mathbb{E}_{\mathbf{y}}\left[\frac{\partial^2 \Lambda(\mathbf{y};\mathbf{u})}{\partial \rho_s \partial P_n}\right] = \frac{S_1}{8(1+\alpha)\sqrt{\rho_s P_n^3}} \tag{45}$$

where $S_1 = \mathbb{E}_{\mathbf{y}}[S_1(\mathbf{y})]$ and $S_2 = \mathbb{E}_{\mathbf{y}}[S_2(\mathbf{y})]$ provided by (A3) and (A11) in Appendix A. With this in mind, the ACRLB is obtained after some lengthy but straightforward manipulations as

$$
\begin{aligned}
\text{ACRLB}(\rho_s) &= \frac{J_{22}}{J_{11}J_{22} - J_{12}^2} \\
&= \frac{8(1+\alpha)\rho_s}{L\mu_a^2(\overline{\Psi}_0 + 2\overline{\Psi}_1)\left(1 - \frac{\mu_a^2(\overline{\Psi}_0 + 2\overline{\Psi}_1)\rho_s}{8(1+\alpha)(2\overline{\Psi}_2 - 1) + (4\overline{\Psi}_0 - 3\mu_a^2\overline{\Psi}_0 + 2\mu_a^2\overline{\Psi}_1)\rho_s}\right)}
\end{aligned} \tag{46}
$$

Normalized by ρ_s^2, the relationship is for lower SNR values, approximated by

$$\text{NACRLB}(\rho_s) = \left.\frac{\text{ACRLB}(\rho_s)}{\rho_s^2}\right|_{\rho_s \ll 1} \approx \frac{8(1+\alpha)}{L\mu_a^2\rho_s(\overline{\Psi}_0 + 2\overline{\Psi}_1)} \tag{47}$$

4.2. Moment-Based Estimator

Regarding RF systems, moment-based (MB) estimators are well established when it comes to SNR estimation [30,31]. One of the main reasons behind this is that they are often very simple from a computational point of view, e.g., the algorithm based on second- and fourth-order moments frequently employed in the RF context, but they are also very powerful in that their error performance approaches the CRLB. However, the latter is only true for constant-modulus constellations like PSK, whereas for non-constant constellations like QAM, they exhibit a significant degradation, which could be mitigated by resorting to higher-order moments as shown in [32,33]. A particular simple solution is available if the MB principle is applied to the dual-filter framework in Figure 1, as shown in the sequel.

Computing the first-order moment of the samples at the output of $q(t)$, we have

$$M_1 = \mathbb{E}[z_k] = A\,\mathbb{E}[a_k] + \mathbb{E}[n_k] = A\,\mu_a \tag{48}$$

whereas the second-order moment of the samples at the output of $\dot{q}(t)$ is given by

$$M_2 = \mathbb{E}[\dot{z}_k^2] = \mathbb{E}[\dot{n}_k^2] = \frac{8\,\pi^2(1+\alpha)^3}{3}P_n \tag{49}$$

Of course, in a practical scenario, the moments in (48) and (49) are replaced by finite sums according to

$$\hat{M}_1 = \hat{A}\,\mu_a = \frac{1}{L}\sum_{k=0}^{L-1} z_k \tag{50}$$

and

$$\hat{M}_2 = \frac{8\,\pi^2(1+\alpha)^3}{3}\hat{P}_n = \frac{1}{L}\sum_{k=0}^{L-1}\dot{z}_k^2 \tag{51}$$

so that the related SNR estimate is furnished as

$$\hat{\rho}_s = \frac{\hat{A}^2}{\hat{P}_n} = \frac{8\,\pi^2(1+\alpha)^3}{3\mu_a^2}\frac{\hat{M}_1^2}{\hat{M}_2} \tag{52}$$

For convenient reasons, the MB algorithm is summarized as follows (Algorithm 2):

Algorithm 2: MB estimator.

Initialization

   Collection of L filter outputs: z_k, $\dot{z}_k$

Computation

   Modified first-order moment: $\hat{M}_a \leftarrow \sum\limits_{k=0}^{L-1} z_k$

   Modified second-order moment: $\hat{M}_b \leftarrow \sum\limits_{k=0}^{L-1} \dot{z}_k^2$

Output

   SNR estimate: $\hat{\rho}_s \leftarrow \dfrac{8\,\pi^2(1+\alpha)^3}{3\mu_a^2}\dfrac{\hat{M}_a^2}{\hat{M}_b}$

Since the MB algorithm is reduced to the calculation of first- and second-order moments, it is obvious that the computational complexity of the related SNR estimate is just in the order of $\mathcal{O}(L)$ real-valued additions and multiplications.

5. Numerical Results

Using a 4-PAM signal operated with no excess bandwidth, i.e., $\alpha = 0$, as well as its maximum value specified by $\alpha = 1$, Figure 2 illustrates the evolution of the error performance, normalized by ρ_s^2, as a function of $\rho_s = \gamma_s/P_0^2$ in dB for DA estimation of the SNR. For this purpose, a rather small value for the length of the pilot sequence, embodied by $L = 10$, and a larger one with $L = 100$ have been assumed. For the dual-filter solution analyzed in the current paper, we can see that the normalized MCRLB in (27)—in the diagram denoted by NMCRLB-DF and shown by solid lines—exhibits for $\rho_s \to \infty$ the same value irrespective of the selected roll-off factor, which is proportional to $1/L$, whereas for $\rho_s \ll 1$ this proportionality is given by $(1 + \alpha)/L$. For comparison purposes, the diagram also includes the normalized limit in the case of a single filter approach [11], which is denoted by NMCRLB-SF and shown in dashed-dotted style. By detailed inspection, we observe that the single-filter results are for $\rho_s \to \infty$ approximately twice as large as those achieved with the dual-filter method; this is explained by the fact that, with the output of the parallel filter, twice as many samples are available for estimation purposes.
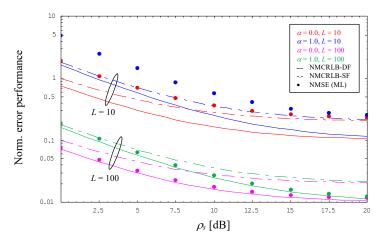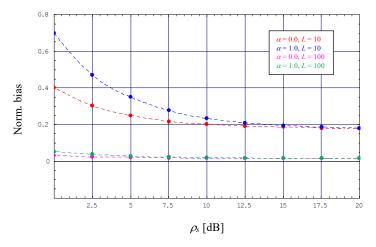

Figure 2. Normalized mean square error for data-aided SNR estimation (4-PAM).

However, with $L = 10$, we also observe that the ML estimator developed in Section 3.3 performs a normalized mean square error (NMSE) with a non-negligible amount of degradation, i.e., although evolving in parallel to the NMCRLB-DF, the difference to the theoretical limit is considerable. The reason for this deficiency is a significant bias effect, which is detailed in Figure 3 (the dotted lines are not related to any analytical work; they are just obtained by cubic interpolation of the simulation results embodied by solid dots in different styles). Nevertheless, the diagram also demonstrates that the bias might be reduced significantly when we increase the length of our pilot sequence to $L = 100$. This is confirmed

by Figure 2, where the NMSE generated by the corresponding ML estimator is now very close to NMCRLB-DF.

Figure 3. Normalized bias for data-aided ML SNR estimation (4-PAM).

Using again a 4-PAM constellation applied to $\alpha = 0$ and 1, Figure 4 visualizes the normalized error performance for NDA estimation of the SNR, but now for larger observation lengths, i.e., $L = 100$ and 1000, as is typical for an NDA situation. For comparison purposes, the diagram includes the evolution of NMCRLB-DF given by (27), in the diagram shown by solid lines, but also that of the normalized true CRLB for NDA estimation of the SNR assuming a single filter receiver (NTRCLB-SF), illustrated in dashed-dotted style, which has been derived in [10]. For $\rho_s \to \infty$, we see that the single-filter limit is again twice as large as NMCRLB-DF, irrespective of the chosen excess bandwidth.
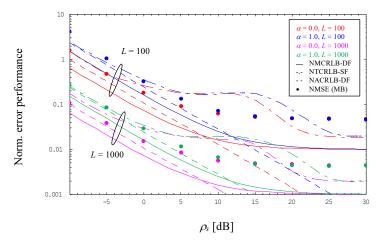

Figure 4. Normalized mean square error for non-data-aided SNR estimation (4-PAM).

The most interesting phenomenon in this respect is that NTRCLB-SF deviates significantly from NMCRLB-DF in the medium SNR range. However, the effect is well known from the open literature when it comes to non-data-aided SNR estimation of non-constant modulus constellations, like PAM or QAM schemes [10,34]. And it is in particular this SNR domain where the simple MB algorithm developed in Section 4.2 is considerably better than the single-filter bound. On the other hand, for $\rho_s \to \infty$ the normalized MSE of the MB algorithm is approximately two times larger than NTCRLB-SF, irrespective of the chosen value of α.

For $\rho_s \ll 1$ and $L = 1000$, the error performance of MB estimates is close to the normalized asymptotic CRLB for the dual-filter solution expressed by (47), in Figure 4 denoted by NACRLB-DF and depicted in dashed style. In this context, it is to be recalled that—similar to NMCRLB-DF—the asymptotic variant is also proportional to $(1 + \alpha)/L$.
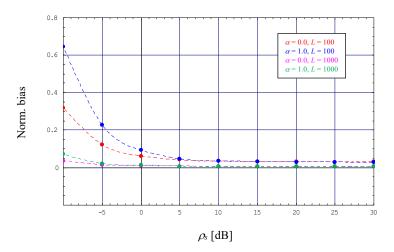
Not surprisingly, a non-negligible discrepancy is observed for smaller observation intervals, as exemplified by $L = 100$, the reason for which is a non-negligible bias effect exhibited by the MB algorithm. This sort of degradation vanishes more and more with increasing values of L, as confirmed by Figure 5.

Figure 5. Normalized bias for non-data-aided MB SNR estimation (4-PAM).

Comparing the DA and NDA methods, we can see that at least ten times the value of L is necessary for the NDA solution, i.e., $L \geq 1000$, to reduce the bias effect to an amount such that it does not have an impact on the error performance. In this respect, it is to be mentioned again that this does not affect the spectral efficiency of the communication system as such because no pilot data are needed for MB algorithms.

6. Conclusions

The major motivation of the current paper were some drawbacks and shortcomings of algorithms investigated in a previous work about NDA SNR estimation for bandlimited optical intensity links. For this reason, the implementation of a second receiver filter with particular properties has been suggested, which is operated in parallel to the receiver filter normally used in this context. Although not intended at the beginning, it could be shown that the approach might be applied to DA scenarios as well. In this respect, an ML algorithm has been developed whose error performance is for larger pilot sequences close to the CRLB as the theoretical limit; smaller pilot sequences exhibit some degradation, which is explained by a non-negligible bias effect.

For the NDA approach, it turned out that the computation of the CRLB would be out of scope even for smaller observation lengths. Therefore, the focus was on the derivation of an asymptotic version of the CRLB, which applies to lower SNR values where a closed-form solution could be achieved. In addition to this, a moment-based SNR estimator has been obtained, which is very simple from a complexity point of view and performs in the medium SNR range significantly better than the CRLB developed for a single-filter solution.

Finally, it is to be mentioned that the ML and MB algorithms have been verified with other parameter setups as well, in particular with symbol constellations other than 4-PAM, but the observations were in principle the same as those made with respect to 4-PAM.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available from the author upon mail request.

Appendix A

In this appendix, we detail the expected operations needed for the computation of the asymptotic CRLB in Section 4.1. First, we focus on the averaging of $S_1(\mathbf{y})$ with respect to $\mathbf{c} \sim \mathbf{a}$, i.e., $S_1(\mathbf{y}) = \mathbb{E}_{\mathbf{c}}[\mathbf{y}^T \mathbf{\Psi} \ \mathbf{c}]$. By detailed inspection of the signal model in (7), we have

$$\mathbf{y}^T \mathbf{\Psi} \ \mathbf{c} = \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} z_i \ \psi_{11,ik} a_k + \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} \dot{z}_i \ \psi_{21,ik} a_k \tag{A1}$$

such that the averaging with respect to a_k leads to

$$S_1(\mathbf{y}) = \mu_a \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} z_i \ \psi_{11,ik} + \mu_a \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} \dot{z}_i \ \psi_{21,ik} \tag{A2}$$

In the next step, in order to obtain $S_1 = \mathbb{E}_{\mathbf{y}}[S_1(\mathbf{y})]$, we have to perform the averaging procedure with respect to $z_i = A \ a_i + n_i = \sqrt{\rho_s P_n} \ a_i + n_i$ as well as $\dot{z}_i = \dot{n}_i$. By taking into account that these are zero-mean processes, we get

$$\begin{aligned} S_1 \ &= \mu_a \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} \mathbb{E}[z_i] \ \psi_{11,ik} + \mu_a \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} \mathbb{E}[\dot{z}_i] \ \psi_{21,ik} \\ &= \mu_a^2 \sqrt{\rho_s P_n} \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} \psi_{11,ik} \\ &= L \mu_a^2 \sqrt{\rho_s P_n} (\overline{\Psi}_0 + 2\overline{\Psi}_1) \end{aligned} \tag{A3}$$

where the last line is achieved by considering the sums in (24). Furthermore, by application of the signal model to $S_2(\mathbf{y}) = \mathbf{y}^T \mathbf{\Psi} \ \mathbf{y}$, we have

$$\mathbf{y}^T \mathbf{\Psi} \ \mathbf{y} = \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} (z_i \ \psi_{11,ik} z_k + z_i \ \psi_{12,ik} \dot{z}_k + \dot{z}_i \ \psi_{21,ik} z_k + \dot{z}_i \ \psi_{22,ik} \dot{z}_k) \tag{A4}$$

Averaging this relationship with respect to z_i and z_k yields

$$\begin{aligned} S_2 \ &= \mathbb{E}_{\mathbf{y}}[S_2(\mathbf{y})] \\ &= \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} \{\mathbb{E}[z_i \ z_k] \ \psi_{11,ik} + \mathbb{E}[z_i \ \dot{z}_k] \ \psi_{12,ik} + \mathbb{E}[\dot{z}_i \ z_k] \ \psi_{21,ik} + \mathbb{E}[\dot{z}_i \ \dot{z}_k] \ \psi_{22,ik}\} \end{aligned} \tag{A5}$$

In this context, one has to keep in mind that

$$\begin{aligned} \mathbb{E}[z_i z_k] \ &= \rho_s P_n \ \mathbb{E}[a_i a_k] + 2\sqrt{\rho_s P_n} \mathbb{E}[a_i n_k] + \mathbb{E}[n_i n_k] \\ &= \begin{cases} \rho_s P_n \ + 2(1+\alpha)\omega_{11,ii} P_n \ , & i = k \\ \mu_a^2 \rho_s P_n + 2(1+\alpha) \ \omega_{11,ik} P_n \ , & i \neq k \end{cases} \end{aligned} \tag{A6}$$

$$\mathbb{E}[z_i \dot{z}_k] = \sqrt{\rho_s P_n} \mathbb{E}[a_i \dot{n}_k] + \mathbb{E}[n_i \dot{n}_k] = 2(1+\alpha) \ \omega_{12,ik} P_n \tag{A7}$$

$$\mathbb{E}[\dot{z}_i z_k] = \sqrt{\rho_s P_n} \mathbb{E}[\dot{n}_i a_k] + \mathbb{E}[\dot{n}_i n_k] = 2(1+\alpha) \ \omega_{21,ik} P_n \tag{A8}$$

$$\mathbb{E}[\dot{z}_i \dot{z}_k] = \mathbb{E}[\dot{n}_i \dot{n}_k] = 2(1+\alpha) \ \omega_{22,ik} P_n \tag{A9}$$

Therefore,

$$\begin{aligned} S_2 \ &= \rho_s P_n \sum_{i=0}^{L-1} \psi_{11,ii} + \mu_a^2 \rho_s P_n \sum_{i=0}^{L-1} \sum_{k=0,k \neq i}^{L-1} \psi_{11,ik} + 2(1+\alpha)P_n \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} \omega_{11,ik} \ \psi_{11,ik} \\ &+ 2(1+\alpha)P_n \sum_{i=0}^{L-1} \sum_{k=0}^{L-1} (\omega_{12,ik} \ \psi_{12,ik} + \omega_{21,ik} \ \psi_{21,ik} + \omega_{22,ik} \ \psi_{22,ik}) \end{aligned} \tag{A10}$$

and by taking into account (24) and (26), we finally get

$$S_2 = L \left[\rho_s P_n \, \overline{\Psi}_0 + 2\mu_a^2 \rho_s P_n \, \overline{\Psi}_1 + 4(1+\alpha) P_n \, \overline{\Psi}_2 \right] \tag{A11}$$

References

1. Hranilovic, S. *Wireless Optical Communication Systems*; Springer: New York, NY, USA, 2004.
2. Arnon, S.; Barry, J.; Karagiannidis, G.; Schober, R.; Uysal, M. *Advanced Optical Wireless Communication Systems*; Cambridge University Press: New York, NY, USA, 2012.
3. Khalighi, M.A.; Uysal, M. Survey on free space optical communication: A communication theory perspective. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 2231–2258. [CrossRef]
4. Ghassemlooy, Z.; Arnon, S.; Uysal, M.; Xu, Z.; Cheng, J. Emerging optical wireless communications—Advances and challenges. *IEEE J. Select. Areas Commun.* **2015**, *33*, 1738–1749. [CrossRef]
5. Tavan, M.; Agrell, E.; Karout, J. Bandlimited intensity modulation. *IEEE Trans. Commun.* **2012**, *60*, 3429–3439. [CrossRef]
6. Czegledi, C.; Khanzadi, M.R.; Agrell, E. Bandlimited power-efficient signaling and pulse design for intensity modulation. *IEEE Trans. Commun.* **2014**, *62*, 3274–3284. [CrossRef]
7. Hranilovic, S. Minimum-bandwidth optical intensity Nyquist pulses. *IEEE Trans. Commun.* **2007**, *55*, 574–583. [CrossRef]
8. Chung, T.S.; Goldsmith, A.J. Degrees of freedom in adaptive modulation: A unified view. *IEEE Trans. Commun.* **2001**, *49*, 1561–1571. [CrossRef]
9. Summers, T.A.; Wilson, S.G. SNR mismatch and online estimation in turbo decoding. *IEEE Trans. Commun.* **1998**, *46*, 421–423. [CrossRef]
10. Gappmair, W. Non-data-aided SNR estimation for bandlimited optical intensity channels. *Sensors* **2023**, *23*, 802. [CrossRef]
11. Gappmair, W. Data-aided SNR estimation for bandlimited optical intensity channels. *Sensors* **2022**, *22*, 8660. [CrossRef]
12. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B* **1977**, *39*, 1–22. [CrossRef]
13. Moon, T.K. The expectation-maximization algorithm. *IEEE Signal Process. Mag.* **1996**, *13*, 47–60. [CrossRef]
14. Boujelben, M.A.; Bellili, F.; Affes, S.; Stephenne, A. EM algorithm for non-data-aided SNR estimation of linearly-modulated signals over SIMO channels. In Proceedings of the 2009 IEEE Global Telecommunications Conference (GLOBECOM), Honolulu, HI, USA, 30 November–4 December 2009.
15. Kay, S.M. *Fundamentals of Statistical Signal Processing: Estimation Theory*; Prentice Hall: Upper Saddle River, NJ, USA, 1993.
16. Mengali, U.; D'Andrea, A.N. *Synchronization Techniques for Digital Receivers*; Plenum Press: New York, NY, USA, 1997.
17. Meyr, H.; Moeneclaey, M.; Fechtel, S.A. *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*; Wiley: New York, NY, USA, 1998.
18. Liu, Y.; Bauer, P.H. Frequency domain limitations in the design of nonnegative impulse response filters. *IEEE Trans. Signal Process.* **2010**, *58*, 4535–4546.
19. Stepniak, G. Minimum bandwidth nonnegative pulses for optical transmission. *IEEE Commun. Lett.* **2019**, *23*, 438–441. [CrossRef]
20. Proakis, J.G.; Manolakis, D.G. *Digital Signal Processing: Principles, Algorithms, and Applications*; Prentice Hall: Upper Saddle River, NJ, USA, 1996.
21. Gappmair, W. On parameter estimation for bandlimited optical intensity channels. *Computation* **2019**, *7*, 11. [CrossRef]
22. Gappmair, W.; Nistazakis, H.E. Blind symbol timing estimation for bandlimited optical intensity channels. In Proceedings of the 2020 12th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), Porto, Portugal, 20–22 July 2020.
23. Gappmair, W.; Schlemmer, H. Feedback solution for symbol timing recovery in bandlimited optical intensity channels. In Proceedings of the 2022 International Conference on Broadband Communications for Next Generation Networks and Multimedia Applications (CoBCom), Graz, Austria, 12–14 July 2022.
24. Proakis, J.G. *Digital Communications*; McGraw-Hill: New York, NY, USA, 1989.
25. Papoulis, A. *Probability, Random Variables, and Stochastic Processes*; McGraw-Hill: New York, NY, USA, 1991.
26. D'Andrea, A.N.; Mengali, U.; Reggiannini, R. The modified Cramer-Rao bound and its application to synchronization problems. *IEEE Trans. Commun.* **1994**, *42*, 1391–1399. [CrossRef]
27. Gini, F.; Reggiannini, R.; Mengali, U. The modified Cramer-Rao bound in vector parameter estimation. *IEEE Trans. Commun.* **1998**, *46*, 52–60. [CrossRef]
28. Moeneclaey, M. On the true and the modified Cramer-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters. *IEEE Trans. Commun.* **1998**, *46*, 1536–1544. [CrossRef]
29. Scharf, L.L. *Statistical Signal Processing*; Prentice-Hall: Upper Saddle River, NJ, USA, 1990.
30. Pauluzzi, D.R.; Beaulieu, N.C. A comparison of SNR estimation techniques for the AWGN channel. *IEEE Trans. Commun.* **2000**, *48*, 1681–1691. [CrossRef]
31. Stephenne, A.; Bellili, F.; Affes, S. Moment-based SNR estimation over linearly-modulated wireless SIMO channels. *IEEE Trans. Wirel. Commun.* **2010**, *9*, 714–722. [CrossRef]
32. Lopez-Valcarce, R.; Mosquera, C. Sixth-order statistics-based non-data-aided SNR estimation. *IEEE Commun. Lett.* **2007**, *11*, 351–353. [CrossRef]

33. Alvarez-Diaz, M.; Lopez-Valcarce, R.; Mosquera, C. SNR estimation for multilevel constellations using higher-order moments. *IEEE Trans. Signal Process.* **2010**, *58*, 1515–1526. [CrossRef]

34. Bellili, F.; Stephenne, A.; Affes, S. Cramer-Rao lower bounds for NDA SNR estimates of square QAM modulated transmissions. *IEEE Trans. Commun.* **2010**, *55*, 3211–3218. [CrossRef]