FSO Channel Estimation for OOK Modulation with APD Receiver over Atmospheric Turbulence and Pointing Errors

Mohammad Taghi Dabiri,[a], Seyed Mohammad Sajad Sadough[a,*], and Mohammad Ali Khalighi[b]

[a] *Department of Electrical Engineering, Shahid Beheshti University G. C., 1983963113, Tehran, Iran.*
[b] *Aix Marseille University, CNRS, Centrale Marseille, Institut Fresnel, Marseille France.*

Abstract

In the free-space optical (FSO) links, atmospheric turbulence and pointing errors lead to scintillation in the received signal. Due to its ease of implementation, intensity modulation with direct detection (IM/DD) based on ON-OFF-keying (OOK) is a popular signaling scheme in these systems. For long-haul FSO links, avalanche photo diodes (APDs) are commonly used, which provide an internal gain in photo-detection, allowing larger transmission ranges, as compared with PIN photo-detector (PD) counterparts. Since optimal OOK detection at the receiver requires the knowledge of the instantaneous channel fading coefficient, channel estimation is an important task that can considerably impact the link performance. In this paper, we investigate the channel estimation issue when using an APD at the receiver. Here, optimal signal detection is quite more delicate than in the case of using a PIN PD. In fact, given that APD-based receivers are usually shot-noise limited, the receiver noise will have a different distribution depending on whether the transmitted bit is '0' or '1', and moreover, its statistics are further affected by the scintillation. To deal with this, we first consider minimum mean-square-error (MMSE), maximum *a posteriori* probability (MAP) and maximum likelihood (ML) channel estimation over an observation

*Corresponding author
Email address: s_sadough@sbu.ac.ir (Seyed Mohammad Sajad Sadough)

window encompassing several consecutive received OOK symbols. Due to the high computational complexity of these methods, in a second step, we propose an ML channel estimator based on the expectation-maximization (EM) algorithm which has a low implementation complexity, making it suitable for high data-rate FSO communications. Numerical results show that for a sufficiently large observation window, by using the proposed EM channel estimator, we can achieve bit error rate performance very close to that with perfect channel state information. We also derive the Cramer-Rao lower bound (CRLB) of MSE of estimation errors and show that for a large enough observation window, this CRLB can be adequately tight.

Keywords: Free-space optics (FSO), atmospheric turbulence, ON-OFF keying, avalanche photo-detector (APD), channel estimation, Cramer-Rao lower bound.

---

1. Introduction

1.1. Background

Under clear sky conditions, the reliability and performance of free space optical (FSO) links can be severely affected by atmospheric conditions and pointing errors [1]. Due to the inherent complexity of phase modulation and the related high implementation complexity, most current commercial FSO systems use intensity modulation with direct detection (IM/DD) based on ON-OFF keying (OOK) [2]. This way, at the receiver, the optical signal is converted to an electrical one by a photo-detector (PD). While PIN PDs are typically suitable for ranges up to several hundred meters, for long-haul links, avalanche PDs (APDs) are the preferred solution, despite their higher cost [3]. Thanks to their high internal gain, they can provide improved signal-to-noise ratio (SNR) capability, as compared with PIN-based receivers. In such receivers, shot noise is mostly dominant [3], whose distribution can be well approximated by a Gaussian [4]. The mean and the variance of this random process will depend on the received signal intensity, thus on the transmitted symbol (i.e., whether the transmitted

bit is '0' or '1') as well as on the actual channel fading coefficient. Note that for OOK demodulation, the receiver requires the knowledge of the channel state information (CSI) to adjust the detection threshold [5, 6]. Since the channel fading affects both the received signal intensity and the receiver noise parameters, the CSI should be estimated of enough accuracy. It is worth mentioning that the coherence time of FSO channels is usually very large (typically on the order of ms), and hence, the channel fading coefficient remains constant over a large number of consecutive bits for typical transmission rates of FSO communications [1].

1.2. Related Works

Channel estimation has been extensively investigated in the context of radio-frequency (RF) networks (see [7], and the references therein). However, due to the particularities of OOK modulation and APD-based receivers, such channel estimation techniques and results are not directly applicable to FSO systems. So, it is important to develop appropriate channel identification solutions and decision metrics for optimal signal detection. A number of previous works have studied this issue. In [8, 9], the authors investigated channel estimation over atmospheric turbulence for the case of pulse position modulation (PPM). Note that PPM has a lower bandwidth efficiency compared to OOK [1]. In [10, 11] for an FSO system using OOK modulation, the channel is estimated using some pilot symbols. Also, in [12], the estimated channel was exploited to adjust the detection threshold at the receiver. However, it is well known that the insertion of pilot bits inside each data frame, incurs a signaling overhead, i.e., a loss in the effective data throughput. Obviously, it is highly preferable to avoid using a pilot overhead while ensuring the good receiver performance, i.e., accurate data detection.

1.3. Contributions

In this paper, to increase the bandwidth efficiency of FSO links and to avoid any pilot overhead at the transmitter, we propose efficient data-aided channel estimation methods. We firstly consider the minimum mean-square-error

(MMSE), the maximum *a posteriori* probability (MAP) and the maximum likelihood (ML) criteria to develop channel estimators. We show that the MMSE estimator requires evaluating complex integrals whereas the MAP and ML estimators need complex numerical computations to find the instantaneous channel attenuation coefficient. Hence, from a practical (real-time) implementation point of view, the computational complexity of these estimators may not be suitable for an FSO system working at very high data-rates. To reduce the complexity of the channel estimator, in a second step, we propose an iterative ML estimator based on the expectation-maximization (EM) algorithm. Nevertheless, as known from the general convergence property of the EM algorithm, there is no guarantee that the iterative steps of EM converge to the global maximum unless an accurate initial estimate is provided [13]. In practice, several initial estimates are used to initialize the EM algorithm in order to guarantee its convergence toward a global maximum. However, obtaining these initial estimates requires sending several training sequences which leads to a loss of the bandwidth efficiency. To solve this problem, we use a blind averaging scheme to calculate the initial channel estimate, without requiring any training symbol. The important point is that the proposed EM-based channel estimator incurs a negligible increase in the receiver's computational complexity and processing delay since it requires only one iteration to converge, making it particularly suitable for practical implementations. We will show that the proposed estimator can achieve performance very close to the perfect CSI case, provided that the observation window is sufficiently large. We also derive the Cramer-Rao lower bound (CRLB) that we use as a benchmark and show that for a large enough observation window, the CRLB becomes a quite tight bound for the mean-square-error (MSE) of the proposed channel estimator.

1.4. Paper Structure

In Section 2, we describe our system model along with our main assumptions. In Section 3, we present the different proposed channel estimators for the FSO receiver and derive the expression of the CRLB. Next, in Section 4, we present

our numerical results to study the performance of the proposed method, and lastly in Section 5, we draw our conclusions.

2. System Model

We assume an IM/DD FSO link with non-return-to-zero (NRZ) OOK modulation over an atmospheric turbulence channel in the presence of pointing errors. In the sequel, we first introduce the received signal model and data detection under consideration and then, summarize the channel model that we consider in this study.

2.1. Signal Model and Data Detection

As mentioned previously, we consider the use of an APD at the receiver. The exact distribution of APD output electrons in response to the mean of absorbed photons is rather complex [4], but it can accurately be approximated by a Gaussian, provided that the mean of absorbed photons is sufficiently large, what is usually the case in practice [3]. This simplifies the derivation of closed form analytical expressions for evaluating the system performance. This way, the mean and the variance of this Gaussian distribution will be $m\,G$ and $m\,G^2F$, respectively, where m denotes the average number of the absorbed photons, G is the average APD gain and F is its excess noise factor [4]. The APD output photocurrent corresponding to the k-th symbol interval, i.e., $[(k-1)T_b, kT_b)$ with T_b being the symbol duration, can be written as [14]:

$$r_k = \mu h s_k + n_k, \tag{1}$$

where h denotes the channel attenuation coefficient, incorporating the channel loss and the effects of atmospheric turbulence and pointing errors, assumed to be constant over a large number of transmitted bits. Also, s_k denotes the transmitted symbol with transmitted optical power P_t, which takes the values of P_1 or P_0 for the cases of the transmission of a bit '1' or '0', respectively, for the considered NRZ OOK signaling scheme. In the sequel, without loss of

generality, we assume that $P_0 = \alpha_e P_1$ where α_e is the optical source extinction ratio and has the range $0 \leq \alpha_e < 1$. Furthermore in (1), the parameter μ equals $\frac{eG\eta}{\tilde{h}_p \nu}$, where e denotes the electron charge, η is the APD quantum efficiency, ν is the optical frequency and $\tilde{h}_p$ stands for the Planck constant. Also, n_k is the photo-current noise, including thermal noise, dark current, as well as the shot noise arising from the received signal and the background radiations. While dark current noise can practically be neglected, a Gaussian distribution can accurately model the sum of other noise sources with the variance given as follows

$$\sigma_{tot}^2 = \sigma_{s,i}^2 h + \sigma_0^2; \quad \text{for } i \in \{0, 1\}, \tag{2}$$

where $\sigma_{s,1}^2 = 2eGF\mu BP_1$ and $\sigma_{s,0}^2 = \alpha_e \sigma_{s,1}^2$ are the variances of the shot noise for the cases of the transmission of a bit '1' and a bit '0'. Also, B is the bandwidth of the receiver low-pass filter, which is placed at the transimpedance amplifier (TIA) output, and is set approximately to $1/T_b$ [3]. Also, $\sigma_0^2 = \sigma_b^2 + \sigma_{th}^2$ where $\sigma_b^2 = 2eGF\mu BP_b$ is the variance of shot noise due to background power P_b and $\sigma_{th}^2 = \frac{4K_b T_r B}{R_l}$ is the variance of thermal noise with K_b being the Boltzmann constant, T_r the receiver's equivalent temperature, and R_l the resistance of the TIA. Note that the presented formulation can be simplified to the case of PIN PD by setting $F = 1$ and $G = 1$. According to (2), the variance of the signal-induced shot noise depends on the transmit optical power (here, P_0 or P_1) and also on the channel coefficient h. Then, for optimal signal detection (i.e., OOK demodulation) the received signal r_k should be compared with a threshold γ_{th}.

The average link bit error rate (BER) is given by

$$P(e) = \int_0^\infty P(e|h) f_h(h) dh, \tag{3}$$

where $P(e|h)$ is the BER conditioned on the channel coefficient h and $f_h(h)$ is the probability density function (PDF) of h. According to (1) and (2), it can be easily verified that $P(e|h)$ is equal to

$$P(e|h) = \frac{1}{4} \text{erfc}\left(\frac{\gamma_{th} - \alpha_e \mu h P_1}{\sqrt{2\left(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2\right)}} \right)$$

$$+\frac{1}{4}\text{erfc}\left(\frac{\mu h P_1-\gamma_{th}}{\sqrt{2\left(\sigma_{s,1}^2 h+\sigma_0^2\right)}}\right) \tag{4}$$

where erfc(.) is the well known complementary error function [15].

Under perfect CSI conditions, the optimal ML decision threshold that minimizes the BER is obtained by differentiating (4), and setting the result to zero. By doing so, the optimum threshold $\gamma_{th,opt}$ is obtained as in (5).

$$\gamma_{th,opt}=-\frac{\mu P_1\sigma_0^2}{\sigma_{s,1}^2}+\left(\frac{\mu^2 P_1^2\sigma_0^4}{\sigma_{s,1}^4}+\frac{\alpha_e\mu^2 P_1^2\sigma_s^2 h^2+(1+\alpha_e)\mu^2 P_1^2\sigma_0^2 h}{\sigma_{s,1}^2}\right.$$
$$\left.+\frac{(\sigma_{s,1}^2 h+\sigma_0^2)(\alpha_e\sigma_{s,1}^2 h+\sigma_0^2)}{(\alpha_e-1)\sigma_{s,1}^2 h}\ln\left(\frac{\alpha_e\sigma_{s,1}^2 h+\sigma_0^2}{\sigma_{s,1}^2 h+\sigma_0^2}\right)\right)^{0.5}. \tag{5}$$

In the special case of signal-independent noise (e.g. when using a PIN PD), the optimal ML decision threshold is simplified as

$$\gamma_{th,opt}=\frac{(\alpha_e+1)\,\mu P_1 h}{2}. \tag{6}$$

Notice that according to (5), the optimum threshold depends on the channel coefficient h, which means that for optimal ML detection, the receiver needs to estimate continuously the variations of h and to adjust the threshold accordingly.

2.2. FSO Channel Model

In FSO links, in addition to the atmospheric turbulence which is commonly modeled by the Gamma-Gamma distribution, pointing errors cause further fluctuation of the received signal intensity. This latter can be due to beam wandering, building sway, thermal expansion and weak earthquakes of tall buildings, etc. We model the channel coefficient as the product of three components, i.e., $h{=}h_l h_a h_p$. Here, h_l is the (deterministic) propagation loss, and h_a and h_p represent the effects of atmospheric turbulence, and pointing errors, respectively [16]. Modeling h_a by the Gamma-Gamma distribution, we denote by α and β the effective numbers of large-scale and small-scale turbulence eddies, respectively. Under the assumption of plane wave propagation, α and β are given by

[17],

$$\alpha = \left[\exp\left(0.49\sigma_R^2 \Big/ \left(1+1.11\sigma_R^{\frac{12}{5}}\right)^{\frac{7}{6}}\right)-1\right]^{-1}, \tag{7}$$

and

$$\beta = \left[\exp\left(0.51\sigma_R^2 \Big/ \left(1+0.69\sigma_R^{\frac{12}{5}}\right)^{\frac{5}{6}}\right)-1\right]^{-1}, \tag{8}$$

where σ_R^2 denotes the Rytov variance. The PDF of h is then given by [16, Eq. (12)]

$$\begin{aligned} f_h(h) =& \frac{\alpha\beta\gamma^2}{A_0 h_l \Gamma(\alpha)\Gamma(\beta)} \\ & \times G_{1,3}^{3,0}\left(\frac{\alpha\beta}{A_0 h_l}h \;\middle|\; \begin{matrix} \gamma^2 \\ \gamma^2-1, \alpha-1, \beta-1 \end{matrix}\right), \end{aligned} \tag{9}$$

where $G_{1,3}^{3,0}(.)$ is the Meijer's G function, $\Gamma(.)$ is the well-known Gamma function. Also, $\gamma = w_{L_{eq}}/2\sigma_j$ is the ratio between the equivalent beam radius at the receiver $w_{L_{eq}}$ and the pointing jitter standard deviation σ_j [16]. We have $w_{L_{eq}}^2 = w_L^2\sqrt{\pi}.\big(1-\text{erfc}(v)\big)/\left(2v\exp\left(-v^2\right)\right)$, with w_L being the beam spot radius at the receive plane (at distance d_0), $v = \sqrt{\pi}r/\left(\sqrt{2}w_L\right)$ and r the radius of a circular detector aperture. Moreover, the parameter $A_0 = [erf(v)]^2$ represents the geometric loss, i.e., the fraction of the collected power (when no pointing error or turbulence occurs).

3. Channel Estimation

As stated previously, for optimal signal detection at the receiver, we need to estimate the channel coefficient h. Here, we propose three approaches for channel estimation, i.e., MMSE, MAP, and ML estimators and an efficient algorithm for implementing the last one based on the EM algorithm with no need to any training symbol. We assume that we receive data in an observation window of length L_s, $\underline{r} = \{r_1, r_2, ..., r_{L_s}\}$, related to L_s transmitted signals, $\underline{s} = \{s_1, s_2, ..., s_{L_s}\}$ during which the channel is assumed to remain unchanged (the quasi-static or frozen channel model). By assuming that the two OOK

symbols are equally likely and independent, the PDF of $\underline{r}$ conditioned on h, can be written as

$$\begin{aligned} p(\underline{r}|h) &= \prod_{k=1}^{L_s} p(r_k|h) \\ &= \frac{1}{2}\prod_{k=1}^{L_s}\Big[p(r_k|s_k = P_1, h) + p(r_k|s_k = P_0, h)\Big]. \end{aligned} \tag{10}$$

According to the system model in Section 2, $p(r_k|s_k, h)$ for $s_k = P_1$ and $s_k = P_0 = \alpha_e P_1$ are given by

$$\begin{aligned} p(r_k|s_k = P_1, h) =& \frac{1}{\sqrt{2\pi(\sigma_{s,1}^2 h + \sigma_0^2)}} \\ & \times \exp\bigg(-\frac{|r_k - \mu P_1 h|^2}{2\big(\sigma_{s,1}^2 h + \sigma_0^2\big)}\bigg), \end{aligned} \tag{11}$$

and

$$\begin{aligned} p(r_k|s_k = P_0, h) =& \frac{1}{\sqrt{2\pi(\alpha_e\sigma_{s,1}^2 h + \sigma_0^2)}} \\ & \times \exp\bigg(-\frac{|r_k - \alpha_e\mu P_1 h|^2}{2\big(\alpha_e\sigma_{s,1}^2 h + \sigma_0^2\big)}\bigg). \end{aligned} \tag{12}$$

3.1. MMSE Channel Estimation

As its name indicates, the MMSE estimation method minimizes the MSE, which is a common measure of estimator quality. The estimation error vector is given by $h - \hat{h}$ and its MSE is written as

$$\text{MSE} = E\Big\{(h - \hat{h})^2|\underline{r}\Big\}. \tag{13}$$

Differentiating (13) with respect to $\hat{h}$ gives the MMSE estimate of h, denoted by $\hat{h}_{MMSE}$ as

$$\begin{aligned} \hat{h}_{MMSE} &= E\{h|\underline{r}\} \\ &= \int_0^\infty h p(h|\underline{r})dh \\ &= \frac{\int_0^\infty h p(\underline{r}|h) f_h(h)dh}{p(\underline{r})} \\ &= \frac{\int_0^\infty h p(\underline{r}|h) f_h(h)dh}{\int_0^\infty p(\underline{r}|h) f_h(h)dh}. \end{aligned} \tag{14}$$

Substituting (10) to (14), $\hat{h}_{MMSE}$ can be obtained as in (15), where $f_h(h)$ is given by (9).

$$\hat{h}_{MMSE} = \frac{\int_0^\infty \frac{h}{2} \prod_{k=1}^{L_s} \left[\frac{1}{\sqrt{2\pi(\sigma_{s,1}^2 h + \sigma_0^2)}} \exp\left(- \frac{|r_k - \mu P_1 h|^2}{2(\sigma_{s,1}^2 h + \sigma_0^2)} \right) + \frac{1}{\sqrt{2\pi(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2)}} \exp\left(- \frac{|r_k - \alpha_e \mu P_1 h|^2}{2(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2)} \right) \right] f_h(h) dh}{\int_0^\infty \frac{1}{2} \prod_{k=1}^{L_s} \left[\frac{1}{\sqrt{2\pi(\sigma_{s,1}^2 h + \sigma_0^2)}} \exp\left(- \frac{|r_k - \mu P_1 h|^2}{2(\sigma_{s,1}^2 h + \sigma_0^2)} \right) + \frac{1}{\sqrt{2\pi(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2)}} \exp\left(- \frac{|r_k - \alpha_e \mu P_1 h|^2}{2(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2)} \right) \right] f_h(h) dh}. \tag{15}$$

Given the relatively high computational complexity of this estimator, it is not well suitable for real-time implementation in a typically high data-rate FSO system. Note that in addition, this estimator needs the knowledge of the channel statistical distribution. For theses reasons, we consider in the following channel estimation based on the MAP and ML criteria.

3.2. MAP and ML Channel Estimation

The MAP channel estimate is obtained by maximizing the logarithm of the posterior function, $\ln p(h|\underline{r})$, as

$$\begin{aligned}
\hat{h}_{MAP} &= \arg \max_h \ln p(h|\underline{r}) \\
&= \arg \max_h \ln p(\underline{r}|h) + \ln\left(f_h(h)\right) \\
&= \arg \max_h \ln \prod_{k=1}^{L_s} \Big[p(r_k|s_k = P_1, h) + p(r_k|s_k = P_0, h) \Big] \\
&\quad + \ln\left(f_h(h)\right) \\
&= \arg \max_h \sum_{k=1}^{L_s} \ln \Big[p(r_k|s_k = P_1, h) + p(r_k|s_k = P_0, h) \Big] \\
&\quad + \ln\left(f_h(h)\right) \\
&= \arg \max_h \sum_{k=1}^{L_s} \ln \Bigg[\frac{1}{\sqrt{2\pi(\sigma_{s,1}^2 h + \sigma_0^2)}} \\
&\quad \times \exp\left(- \frac{|r_k - \mu P_1 h|^2}{2(\sigma_{s,1}^2 h + \sigma_0^2)} \right) \\
&\quad + \frac{1}{\sqrt{2\pi(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2)}}
\end{aligned}$$

$$\times \exp\left(-\frac{|r_k - \alpha_e \mu P_1 h|^2}{2\big(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2\big)} \right) \Bigg] + \ln\left(f_h(h)\right) \ , \tag{16}$$

where the last equality is according to (9), (11) and (12). Similarly, the ML channel estimate is obtained by maximizing the log-likelihood function $\ln p(\underline{r}|h)$, as

$$\begin{aligned} \hat{h}_{ML} &= \arg\max_h \ln p(\underline{r}|h) \\ &= \arg\max_h \sum_{k=1}^{L_s} \ln \Bigg[\frac{1}{\sqrt{2\pi(\sigma_{s,1}^2 h + \sigma_0^2)}} \\ &\quad \times \exp\left(-\frac{|r_k - \mu P_1 h|^2}{2\big(\sigma_{s,1}^2 h + \sigma_0^2\big)} \right) \\ &\quad + \frac{1}{\sqrt{2\pi(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2)}} \\ &\quad \times \exp\left(-\frac{|r_k - \alpha_e \mu P_1 h|^2}{2\big(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2\big)} \right) \Bigg] \ , \end{aligned} \tag{17}$$

where the last equality is according to (11) and (12). The advantage of MAP estimation compared to ML estimation is the exploitation of the knowledge of fading statistics (if available). The common way for deriving the optimum h from (16) or (17), is to differentiate and set the result to zero. However, as we can see form (16) and (17), such an operation is not straightforward. Hence, to obtain the optimum h based on MAP or ML criterion, we have to resort to numerical methods which are however complex to implement, especially, for high data rate FSO communications.

3.3. EM-Based Channel Estimation

Given the high computational complexity of MMSE, MAP and ML estimators, here, we propose to use the EM algorithm which is an iterative procedure to estimate the channel state based on its previous estimate [13]. According to the terminology of the EM algorithm, the received sequence $\underline{r}$ and $Y = (\underline{r}, \underline{s})$ are referred to as incomplete and complete data sets, respectively. Each iteration of the EM algorithm is composed of two steps: the E-step and the M-step.

3.3.1. E-step

The E-step finds the so-called *auxiliary* function $\Theta(h|h^i)$, which is defined as the expectation of the log-likelihood of h, $\log p(Y|h)$, where the expectation is taken with respect to $\underline{s}$ conditioned on $\underline{r}$ and the latest i-th estimate of h, denoted by h^i [13]. We have

$$\Theta(h|h^i) = E_{\underline{s}}\Big\{ \log\big(p(\underline{r},\underline{s}|h)\ \big|\ \underline{r}, h^i\Big\}. \tag{18}$$

The joint PDF of $\underline{r}$ and $\underline{s}$ conditioned on h, can be obtained as

$$\begin{aligned} p(Y|h) = p(\underline{r},\underline{s}|h) &= \prod_{k=1}^{L_s} p(r_k, s_k|h) \\ &= \prod_{k=1}^{L_s} p(r_k|s_k,h)p(s_k). \end{aligned} \tag{19}$$

Based on (19) and by assuming that the two OOK symbols are equally likely, $\Theta(h|h^i)$ can be obtained after straightforward calculation as given in (22), where

$$\mathcal{A}_k = \frac{p(r_k|s_k = P_1, h^i)}{2p(r_k|h^i)}, \tag{20}$$

and

$$\mathcal{B}_k = \frac{p(r_k|s_k = P_0, h^i)}{2p(r_k|h^i)}. \tag{21}$$

$$\begin{aligned} \Theta(h|h^i) &= \sum_{k=1}^{L_s} \Big[\ln\Big(p\big(r_k|s_k=P_1,h\big)p\big(s_k=P_1\big)\Big)p\big(s_k=P_1|r_k,h^i\big) \\ &\quad + \ln\Big(p\big(r_k|s_k=P_0,h\big)p\big(s_k=P_0\big)\Big)p\big(s_k=P_0|r_k,h^i\big)\Big] \\ &= \sum_{k=1}^{L_s}\Bigg[\ln\left(\frac{p(r_k|s_k=P_1,h)}{2}\right)\frac{p(r_k|s_k=P_1,h^i)}{2p(r_k|h^i)} \\ &\quad + \ln\left(\frac{p(r_k|s_k=P_0,h)}{2}\right)\frac{p(r_k|s_k=P_0,h^i)}{2p(r_k|h^i)}\Bigg] \\ &= -\sum_{k=1}^{L_s}\Bigg[\left(\frac{|r_k - h\mu P_1|^2}{2\big(\sigma_{s,1}^2 h + \sigma_0^2\big)} + \frac{1}{2}\ln\Big(8\pi\big(\sigma_{s,1}^2 h + \sigma_0^2\big)\Big)\right)\mathcal{A}_k \\ &\quad + \left(\frac{|r_k - \alpha_e h\mu P_1|^2}{2\big(\alpha_e\sigma_{s,1}^2 h + \sigma_0^2\big)} + \frac{1}{2}\ln\Big(8\pi\big(\alpha_e\sigma_{s,1}^2 h + \sigma_0^2 + \big)\Big)\right)\mathcal{B}_k\Bigg]. \end{aligned} \tag{22}$$

3.3.2. M-step

The M-step finds h^{i+1}, i.e., the next value of h that maximizes $\Theta(h|h^i)$. More precisely, we have

$$h^{i+1} = \arg\max_{h} \Theta(h|h^i). \tag{23}$$

Given (22), we can rewrite (23) as in (24).

$$\begin{aligned} h^{i+1} = \arg\min_{h} & \sum_{k=1}^{L_s} \left(\frac{|r_k - h\mu P_1|^2}{\sigma_{s,1}^2 h + \sigma_0^2} + \log\left(\sigma_{s,1}^2 h + \sigma_0^2\right) \right) \mathcal{A}_k \\ & + \left(\frac{|r_k - \alpha_e h\mu P_1|^2}{\alpha_e \sigma_{s,1}^2 h + \sigma_0^2} + \log\left(\alpha_e \sigma_{s,1}^2 h + \sigma_0^2\right) \right) \mathcal{B}_k. \end{aligned} \tag{24}$$

To solve (24), the optimum h can be obtained by differentiating the argument with respect to h and setting the result equal to zero. By doing so, we obtain

$$\Re_4 h^4 + \Re_3 h^3 + \Re_2 h^2 + \Re_1 h + \Re_0 = 0, \tag{25}$$

where

$$\Re_0 = \sum_{k=1}^{L_s} \left(\sigma_0^6 \sigma_{s,1}^2 - \sigma_0^4 \sigma_{s,1}^2 r_k^2 - 2\mu P_1 \sigma_0^6 r_k\right) \left(\mathcal{A}_k + \alpha_e \mathcal{B}_k\right), \tag{26}$$

$$\begin{aligned} \Re_1 = & \sum_{k=1}^{L_s} \left(2(\mu P_1)^2 \sigma_0^6 + \sigma_0^4 \sigma_{s,1}^4\right) \left(\mathcal{A}_k + \alpha_e^2 \mathcal{B}_k\right) \\ & + \left(2\sigma_0^4 \sigma_{s,1}^4 - 2\sigma_0^2 \sigma_{s,1}^4 r_k^2 - 4\mu P_1 \sigma_0^4 \sigma_{s,1}^2 r_k\right) \left(\alpha_e \mathcal{A}_k + \alpha_e \mathcal{B}_k\right), \end{aligned} \tag{27}$$

$$\begin{aligned} \Re_2 = & \sum_{k=1}^{L_s} \left(\sigma_0^2 \sigma_{s,1}^6 - \sigma_{s,1}^6 r_k^2 - 2\mu P_1 \sigma_0^2 \sigma_{s,1}^4 r_k\right) \left(\alpha_e^2 \mathcal{A}_k + \alpha_e \mathcal{B}_k\right) \\ & + \left(4(\mu P_1)^2 \sigma_0^4 \sigma_{s,1}^2 + 2\sigma_0^2 \sigma_{s,1}^6\right) \left(\alpha_e \mathcal{A}_k + \alpha_e^2 \mathcal{B}_k\right) \\ & + (\mu P_1)^2 \sigma_0^4 \sigma_{s,1}^2 \left(\mathcal{A}_k + \alpha_e^3 \mathcal{B}_k\right), \end{aligned} \tag{28}$$

$$\begin{aligned} \Re_3 = & \sum_{k=1}^{L_s} \left(2(\mu P_1)^2 \sigma_0^2 \sigma_{s,1}^4 \left(\alpha_e + \alpha_e^2\right) + \alpha_e^2 \sigma_{s,1}^2\right) \mathcal{A}_k \\ & + \left(2(\mu P_1)^2 \sigma_0^2 \sigma_{s,1}^4 \left(\alpha_e^2 + \alpha_e^3\right) + \alpha_e^2 \sigma_{s,1}^2\right) \mathcal{B}_k, \end{aligned} \tag{29}$$

$$\Re_4 = \sum_{k=1}^{L_s} \alpha_e^2 (\mu P_1)^2 \sigma_{s,1}^6 \mathcal{A}_k + \alpha_e^3 (\mu P_1)^2 \sigma_{s,1}^6 \mathcal{B}_k. \tag{30}$$

After solving (25) using [18], the EM estimate of h at the $(i+1)$-th iteration is obtained as

$$h^{i+1} = \frac{-\Re_3 + 2\Re_4\sqrt{\frac{2\Re_r}{3} + \mathcal{Q}} + 2\Re_4\sqrt{\frac{4\Re_r}{3} - \mathcal{Q} + \frac{2\Re_s}{\sqrt{\frac{2\Re_r}{3}+\mathcal{Q}}}}}{4\Re_4}, \tag{31}$$

where

$$\mathcal{Q} = 2\sqrt[3]{\mathcal{Q}_1 + \sqrt{\mathcal{Q}_2}} + 2\sqrt[3]{\mathcal{Q}_1 - \sqrt{\mathcal{Q}_2}}, \tag{32}$$

$$\mathcal{Q}_1 = \frac{2\Re_r^3 + 27\Re_s^2}{432}, \tag{33}$$

$$\mathcal{Q}_2 = \Big(4\Re_r^6 + 108\Re_r^5 + 729\Re_s^4 - 144\Re_r^4\Re_t - 1728\Re_r^2\Re_t^2 - 6912\Re_t^3\Big)\Big/186624, \tag{34}$$

$$\Re_r = \frac{3\Re_3^2 - 8\Re_4\Re_2}{8\Re_4^2}, \tag{35}$$

$$\Re_s = \frac{\Re_3^3 - 4\Re_4\Re_3\Re_2 + 8\Re_4^2\Re_1}{8\Re_4^3}, \tag{36}$$

$$\Re_t = \frac{-3\Re_3^4 + 16\Re_4\Re_3^2\Re_2 - 64\Re_4^2\Re_3\Re_1 + 256\Re_4^3\Re_0}{256\Re_4^4}. \tag{37}$$

Equation (31) consists of a series of simple additions and multiplications, whose number increases linearly with L_s, unlike MMSE and ML estimators as in (15) and (17), respectively. In the special case of signal-independent noise (e.g. by using a PIN PD), the EM-based estimate of h at the $(i+1)$-th iteration can be simplified as

$$h^{i+1} = \frac{\sum_{k=1}^{L_s}(\mathcal{A}_k + \alpha_e\mathcal{B}_k)r_k}{\mu P_1 \sum_{k=1}^{L_s}\mathcal{A}_k + \alpha_e^2\mathcal{B}_k}. \tag{38}$$

As known from the general convergence property of the EM algorithm, there is no guarantee that the iterative steps converge to the global maximum unless an accurate initial estimate is made available for h. Hence, to initialize the EM algorithm, one requires an initial estimate h^0, which is usually acquired by means of some pilot symbols. However, to ensure an accurate initial estimate, we need to send multiple pilot symbols, which reduces the spectral efficiency. Let us denote by M the number of bits '1' in the observation window of length L_s. Obviously, M is unknown to the receiver. However, the receiver knows that M

belongs to the set $\{0, 1, ..., L_s\}$ and distributed as a Binomial random variable. For large values of L_s, M tends to its expected value, i.e., $M \approx E[M] = \frac{L_s}{2}$. To obtain an initial estimate for h, without requiring any pilot symbol, we use an unbiased initial estimate as

$$\begin{aligned} h^0 &= \frac{2}{L_s \mu P_1 (1+\alpha_e)} \sum_{k=1}^{L_s} r_k \\ &= \frac{2\mu P_1 \big(M + (L_s - M)\alpha_e\big)}{L_s \mu P_1 (1+\alpha_e)} h + Z', \end{aligned} \tag{39}$$

where $Z' = \sum_{k=1}^{L_s} n_k$, is the additive white Gaussian noise with zero mean and variance

$$\sigma^2_{Z'} = 4 \frac{L_s(\sigma_0^2 + \alpha_e \sigma^2_{s,1}) + M\sigma^2_{s,1}(1-\alpha_e)}{\big(L_s \mu P_1 (1+\alpha_e)\big)^2}, \tag{40}$$

which is a decreasing function of L_s. Hence, it is clear that when L_s is chosen sufficiently large, M becomes very close to $L_s/2$ and h^0 becomes close to h. We will later refer to this initial estimate calculation as "Blind" estimator.

3.4. Cramer-Rao Lower Bound

The CRLB on the variance of an unbiased estimate of h, is given by [13],

$$CRLB(h) = \frac{1}{I_{\text{fish}}}, \tag{41}$$

where I_{fish} is referred to as the Fisher's information. In our considered system model, for a received sequence $\underline{r}$, I_{fish} is defined as

$$\begin{aligned} I_{\text{fish}} &= -E_{\underline{r},h}\left\{ \frac{\partial^2}{\partial h^2} \ln p(\underline{r}, h) \right\} \\ &= -E_{\underline{r},h}\left\{ \frac{\partial^2}{\partial h^2} \ln \Big(p(\underline{r}|h) f_h(h) \Big) \right\} \\ &= -E_{\underline{r},h}\left\{ \frac{\partial^2}{\partial h^2} \ln p(\underline{r}|h) \right\} - E_h\left\{ \frac{\partial^2}{\partial h^2} \ln f_h(h) \right\}. \end{aligned} \tag{42}$$

We define the vector $\underline{r}' = \Big[r'_1, r'_2, ..., r'_{L_s}\Big]$, where the first M elements of $\underline{r}'$ correspond to the transmitted bits '1' and the remaining $L_s - M$ elements correspond to the transmitted bits '0'. For simplicity, we assume the transmitted

sequence $\underline{s}$, is known at the receiver. According to this assumption, $p(r_k|h)$ is equal to $p(r|s_k,h)$ and then we have

$$\begin{aligned}
I_{\text{fish}} = &-\frac{1}{2^{L_s}}\sum_{M=1}^{L_s}\binom{L_s}{M}\times E_{\underline{r}',h}\Bigg\{\frac{\partial^2}{\partial h^2}\Bigg(\sum_{j=1}^{M}\ln p(r'_j|h) \\
&+\sum_{j=M+1}^{L_s}\ln p(r'_j|h)\Bigg)\Bigg\} - E_h\Bigg\{\frac{\partial^2}{\partial h^2}\ln f_h(h)\Bigg\} \\
= &\frac{1}{2^{L_s}}\sum_{M=1}^{L_s}\binom{L_s}{M}E_{\underline{r}',h}\Bigg\{\frac{\partial^2}{\partial h^2}\Bigg(\frac{M}{2}\ln\left(\sigma_{s,1}^2h+\sigma_0^2\right) \\
&+\frac{L_s-M}{2}\ln\left(\alpha_e\sigma_{s,1}^2h+\sigma_0^2\right)+\sum_{j=1}^{M}\frac{(r'_j-\mu P_1h)^2}{2(\sigma_{s,1}^2h+\sigma_0^2)} \\
&+\sum_{j=M+1}^{L_s}\frac{(r'_j-\alpha_e\mu P_1h)^2}{2(\alpha_e\sigma_{s,1}^2h+\sigma_0^2)}\Bigg)\Bigg\} \\
&-E_h\Bigg\{\frac{\partial^2}{\partial h^2}\ln f_h(h)\Bigg\},
\end{aligned} \tag{43}$$

where $\binom{n}{m}$ is the number of combinations of m items out of n. Given that $\frac{1}{2^{L_s}}\sum_{M=1}^{L_s}\binom{L_s}{M}M=\frac{L_s}{2}$ and after straightforward algebra, we obtain

$$\begin{aligned}
I_{\text{fish}} = &\frac{L_s}{4}\int_0^{\infty}\Bigg\{\frac{3\sigma_{s,1}^4}{(\sigma_{s,1}^2h+\sigma_0^2)^2}+\frac{3\alpha_e^2\sigma_{s,1}^4}{(\alpha_e\sigma_{s,1}^2h+\sigma_0^2)^2} \\
&+\frac{2(\mu P_1)^2}{\sigma_{s,1}^2h+\sigma_0^2}+\frac{2\alpha_e^2(\mu P_1)^2}{\alpha_e\sigma_{s,1}^2h+\sigma_0^2} \\
&+\frac{\partial^2}{\partial h^2}\ln f_h(h)\Bigg\}f_h(h)dh.
\end{aligned} \tag{44}$$

4. Performance Study of the Proposed Estimators

In this section, we provide numerical results in terms of BER and MSE to evaluate the performance of the proposed MMSE, MAP, ML, blind and EM-based channel estimators while considering the case of perfect CSI and the CRLB as benchmarks. We consider uncoded NRZ-OOK modulation and set the system parameters as specified in Table 1, following our parameter definition in Section II.

Table 1: System Parameters for Simulations

Description	**Parameter**	**Setting**
APD Gain	G	100
APD Quantum Efficiency	η	0.9
Avalanche Unization Factor	k_{eff}	0.028
Plank's Constant	$\tilde{h}_p$	6.6×10^{-34}
Wavelength	λ	1550 nm
Boltzmann's Constant	K_B	1.38×10^{-23} J.s
Receiver Load	R_l	1 kΩ
Receiver Temperature	T_r	300° K
Symbol Duration	T_b	10^{-9}
Modulation Extinction Ratio	α_e	0.2
Aperture Radius	r	5 cm
Normalized Beam Width	w_L/r	6
Normalized Jitter	σ_j/r	2
Background Power	P_b	10 nW
Link Range	d_0	1 km
Rytov Variance	σ_R^2	0.2

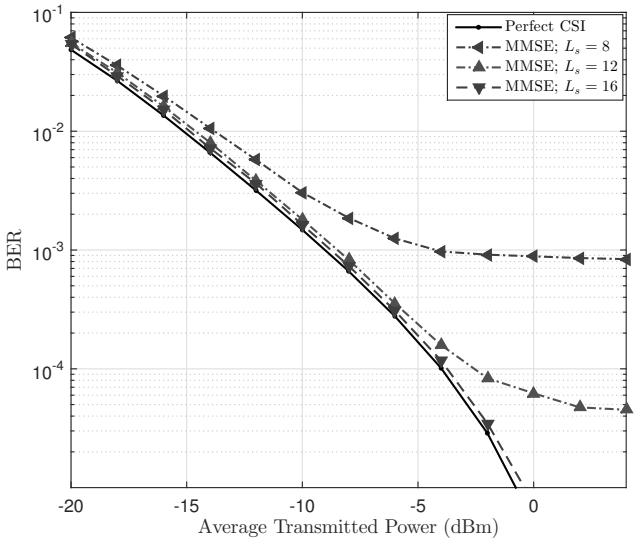


Figure 1: BER performance of MMSE estimator for different length of the observation window $L_s \in \{8, 12, 16\}$.

4.1. BER Analysis

In Figs. 1 and 2, we have presented the simulated BER as a function of the average transmit power $P_t = (P_1 + P_0)/2$ for the cases of MMSE, MAP and ML channel estimation, respectively. Results show that (as expected) with increasing the length of the observation window, i.e., L_s, the performance of the estimators become closer to that with perfect CSI. For instance, by setting L_s to 16 for MMSE, MAP and ML estimators, their BER performance is very close to the perfect CSI case.

In Fig. 3, we have shown plots of BER for the proposed EM-based channel estimation after one and ten iterations. We have also presented the corresponding BER plots for the blind method that we use for deriving an initial channel estimate, see (39). Firstly, we notice that with increasing the length of the observation window, the performance of proposed EM-based method becomes quite close to the perfect CSI case. More specifically, to achieve average target BERs of 10^{-4} for instance, L_s should be larger than 24. We also notice that

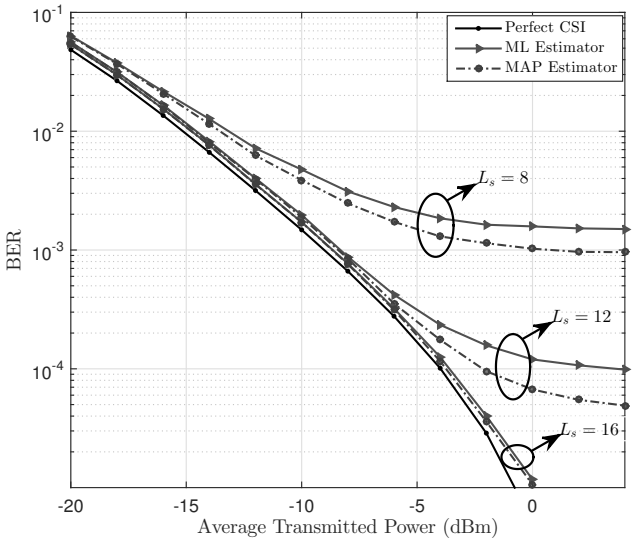


Figure 2: BER performance of MAP and ML estimators for different length of the observation window $L_s \in \{8, 12, 16\}$.

after a single iteration, the receiver converges (i.e., attains the global maximum) and further iterations result in a negligible performance improvement. This is a clear advantage of the proposed method, since it implies a relatively low computational complexity and also a low and fixed processing delay. As noticed, in order to achieve the BER performance close to the receiver with perfect CSI, a large enough L_s should be selected. However, a large L_s increases the processing load and delay. To show the optimum L_s for each of the proposed channel estimators, we have plotted the BER curves as a function of L_s in Fig. 4, for a transmitted power of -9 dBm (corresponding to an average target BER of $\approx 10^{-3}$ for perfect CSI), and in Fig. 5, for a transmitted power of -3.8 dBm (for an average target BER of $\approx 10^{-4}$ for the perfect CSI case). As expected, the best performance is achieved by the MMSE estimator, followed by MAP, ML, and EM-based estimators, and the simple blind estimator has the worst performance. For instance, from Fig. 4, we notice that the receiver with EM-based estimator can achieve an error probability close to the receiver with perfect CSI

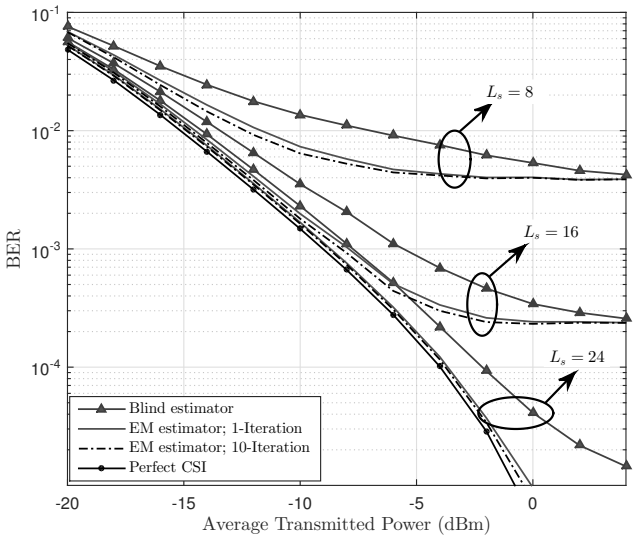


Figure 3: BER performance of EM-based and blind (Eq. (39)) estimators for different length of the observation window $L_s \in \{8, 16, 24\}$.

for $L_s = 20$, while, for MMSE, MAP, ML and blind estimators, L_s is equal to 16, 17, 18 and 120, (not shown in the figure), respectively. The receiver with EM-based estimator can still achieve an error probability close to the receiver with perfect CSI for $L_s = 20$, while, for MMSE, MAP, ML and blind estimators, L_s is equal to 16, 18, 19 and 210, respectively. From Figs. 4 and 5 we can conclude that the required observation window length does not really depend on the transmit power P_t.

Note that although the EM-based estimator needs a larger L_s compared to MMSE, MAP and ML estimators to achieve the same estimation accuracy, according to Equations (15), (16), (17) and (31), the computational complexity of the EM-based estimator is significantly lower than the three other methods.

4.2. MSE Analysis

In Fig. 6, we have shown the normalized MSE of the proposed MMSE, MAP, ML, blind and EM-based estimators and compared them with the CRLB for

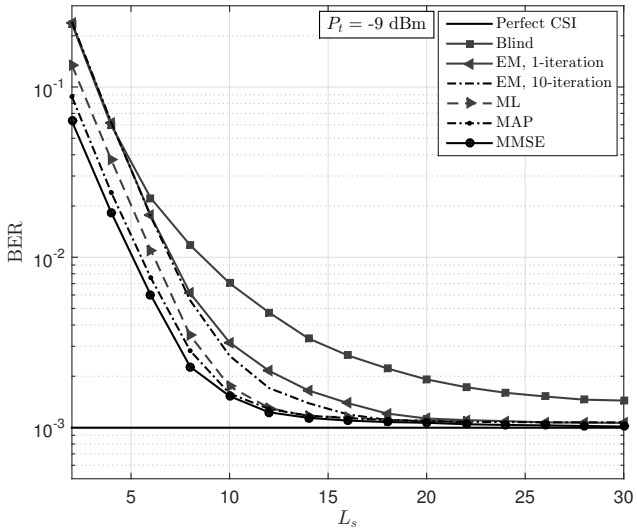


Figure 4: BER performance of EM-based, MAP, ML and MMSE estimators versus the length of the observation window L_s; $P_t = -9$ dBm and the average target BER is equal to 10^{-3}.

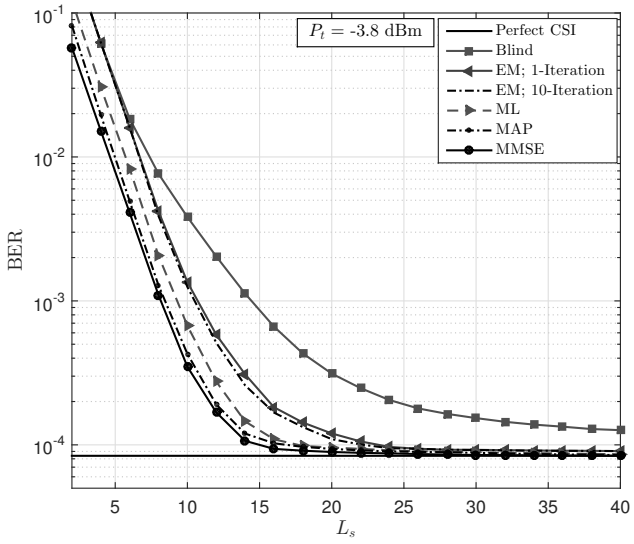


Figure 5: BER performance of EM-based, MAP, ML and MMSE estimators versus the length of the observation window L_s; $P_t = -3.8$ dBm and the average target BER is equal to 10^{-4}.

Table 2: Comparison of Complexity of the Proposed Channel Estimators

Channel estimator	**MMSE**	**MAP**	**ML**	**EM** (with one iteration)
Requirement to the channel PDF	YES	YES	NO	NO
Minimum required L_s to attain 0.95 BER performance with perfect CSI at P_t=-9 dBm	L_s=16	L_s=17	L_s=18	L_s=20
Minimum required L_s to attain 0.95 BER performance with perfect CSI at P_t=-3.8 dBm	L_s=16	L_s=18	L_s=19	L_s=26
Overall complexity	High due to complex integral calculations, see (15)	High due to high computational load, see (16)	High due to high computational load, see (17)	Relatively low; only requires the simple addition and multiplication operations, see (24)

different L_s values. We can see from these figures that, as expected, the ML and the EM-based estimators have the same MSE, which approaches that of the MMSE estimator by increasing L_s. Also, by increasing L_s, the MSE of the proposed methods become closer to the CRLB as it can be seen from Fig. 6c, for L_s=24.

4.3. Complexity and Latency Comparison and Discussion

Due to typically high date rate of FSO links, computational complexity is an important issue for the implementation of these systems [1]. Besides the presented simulation results, in order to draw general conclusions on the advantage of the proposed EM-based estimator, we compare here the processing load of the four considered estimators. According to (15), (16), (17), (31) and (39), the computational complexities of all proposed methods increase linearity with L_s. In other words, their complexity is in order of $O(L_s)$. As we noticed from simulation results, the proposed EM-based estimator has a performance close to the MMSE and the MAP estimators for large enough L_s while benefiting from a significantly lower computational complexity. We have summarized our comments on the implementation complexity of the proposed estimators in Table II.

Lastly, note that typical L_s of about 24 is quite acceptable, given the quasi-

static nature of the FSO channel and the fact that it is equivalent to an estimation delay on the order of ns for a Gbps FSO link.

5. Conclusions

We considered in this work the case of long-range FSO links with APD-based receivers and investigated optimal signal demodulation for the case of NRZ OOK signaling, in the presence of signal-dependent shot noise. For optimal OOK signal demodulation, the receiver requires the knowledge of the instantaneous channel attenuation coefficient. A channel estimation step is hence unavoidable prior to signal detection. For this purpose, we first studied three estimators based on MMSE, MAP and ML criteria, that calculate the channel coefficient over an observation window encompassing several consecutive received symbols. Due to the computational complexity of these estimators, we then proposed an ML channel estimator based on the iterative EM algorithm. We investigated the performance of the proposed EM-based estimator through numerical simulations, which alleviated the advantage of this estimator compared to the three other methods. It is worth mentioning that this estimator converges after a single iteration, and in addition, does not require the knowledge of the channel statistical distribution. The relatively low computational complexity and the low estimation delay of the proposed method and the fact that it does not rely on the transmission of pilot sequences, makes it particularly suitable for implementation in FSO communication links.

References

[1] M. A. Khalighi, M. Uysal, Survey on free space optical communication: A communication theory perspective, Communications Surveys & Tutorials, IEEE 16 (4) (2014) 2231–2258.

[2] M. R. Bhatnagar, Z. Ghassemlooy, S. Zvanovec, M.-A. Khalighi, M. M. Abadi, Quantized feedback based differential signaling for free-space optical

communication system, IEEE Transactions on Communications 64 (12) (2016) 5176–5188.

[3] F. Xu, M.-A. Khalighi, S. Bourennane, Impact of different noise sources on the performance of PIN-and APD-based FSO receivers, in: Telecommunications (ConTEL), Proceedings of the 2011 11th International Conference on, IEEE, 2011, pp. 211–218.

[4] P. Webb, R. McIntyre, J. Conradi, Properties of avalanche photodiodes, RCA review 35 (1974) 234–278.

[5] L. Yang, X. Song, J. Cheng, J. F. Holzman, Free-space optical communications over lognormal fading channels using ook with finite extinction ratios, IEEE Access 4 (2016) 574–584.

[6] L. Yang, B. Zhu, J. Cheng, J. F. Holzman, Free-space optical communications using on–off keying and source information transformation, Journal of Lightwave Technology 34 (11) (2016) 2601–2609.

[7] S. M.-S. Sadough, M.-A. Khalighi, Recent Developments in Channel Estimation and Detection for MIMO Systems, in RADIOI COMMUNICATION, INTECH Press, Apr. 2010.

[8] M. Cole, K. Kiasaleh, Signal intensity estimators for free-space optical communications through turbulent atmosphere, Photonics Technology Letters, IEEE 16 (10) (2004) 2395–2397.

[9] M. Cole, K. Kiasaleh, Signal intensity estimators for free-space optical communication with array detectors, Communications, IEEE Transactions on 55 (12) (2007) 2341–2350.

[10] K. Kiasaleh, Channel estimation for FSO channels subject to Gamma-Gamma turbulence, in: Proc. International Conference on Space Optical Systems and Applications (ICSOS) 2012, 2012.

[11] F. Xu, A. Khalighi, P. Caussé, S. Bourennane, Channel coding and time-diversity for optical wireless links, Optics express 17 (2) (2009) 872–887.

[12] K. Kiasaleh, Receiver architecture for channel-aided, OOK, APD-based FSO communications through turbulent atmosphere, IEEE Transactions on Communications 63 (1) (2015) 186–194.

[13] S. M. Kay, Fundamentals of Statistical Signal Processing, volume I: Estimation Theory, Prentice Hall, 1993.

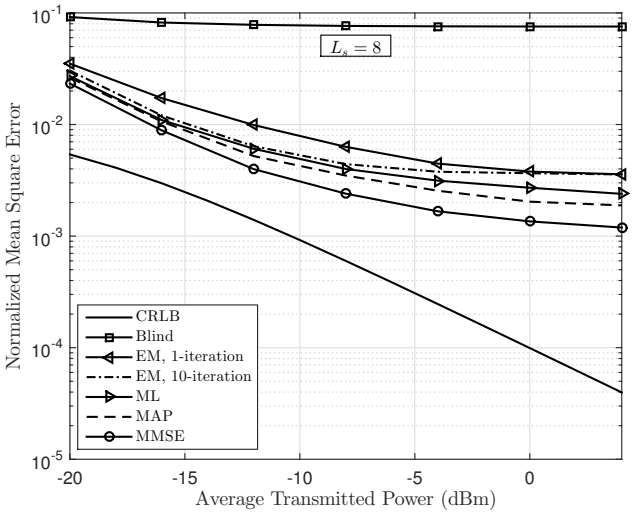
[14] R. Gagliardi, S. Karp, Optical Communications, Wiley, 1995.

[15] J. G. Proakis, Intersymbol Interference in Digital Communication Systems, Wiley Online Library, 2003.

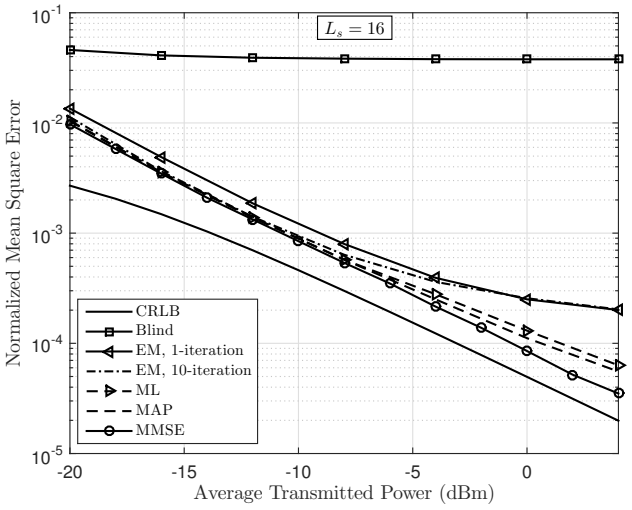
[16] H. G. Sandalidis, T. A. Tsiftsis, G. K. Karagiannidis, Optical wireless communications with heterodyne detection over turbulence channels with pointing errors, Journal of Lightwave Technology 27 (20) (2009) 4440–4445.

[17] L. C. Andrews, R. L. Phillips, Laser Beam Propagation Through Random Media, Vol. 1, SPIE press Bellingham, WA, 2005.

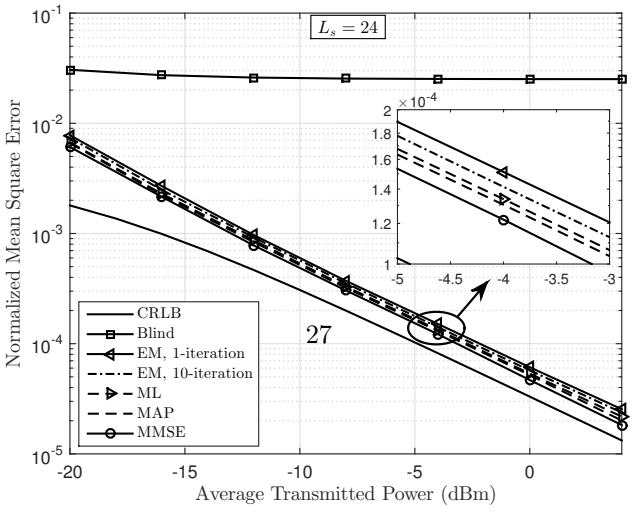
[18] E. W. Weisstein, Quartic Equation (2002).


$L_s = 8$
Normalized Mean Square Error
Average Transmitted Power (dBm)
CRLB
Blind
EM, 1-iteration
EM, 10-iteration
ML
MAP
MMSE

(a)


$L_s = 16$
Normalized Mean Square Error
Average Transmitted Power (dBm)
CRLB
Blind
EM, 1-iteration
EM, 10-iteration
ML
MAP
MMSE

(b)


$L_s = 24$
Normalized Mean Square Error
Average Transmitted Power (dBm)
CRLB
Blind
EM, 1-iteration
EM, 10-iteration
ML
MAP
MMSE

(c)