Free-Space Optical Communications Over Lognormal Fading Channels Using OOK With Finite Extinction Ratios

Luanxia Yang, Xuegui Song, Julian Cheng, *Senior Member, IEEE*, and Jonathan F. Holzman, *Member, IEEE*

***Abstract*—Free-space optical communication links operating over lognormal turbulence channels using on-off keying (OOK) are studied in this work. Such systems can suffer from irreducible error floors that result from the use of demodulation with fixed and unoptimized detection thresholds. The resulting error floors are analyzed for the general case of low and high state offsets (i.e., finite extinction ratios). An electrical signal-to-noise ratio (SNR) optimized detection system is applied. The system uses the electrical SNRs to implement adaptive detection thresholds and eliminate the error floors. The system can accommodate operation with finite extinction ratios, as it uses the method of moments and maximum likelihhod estimation techniques to estimate the low and high state offsets and electrical SNR. Numerical results show that the SNR gap between the electrical-SNR-optimized detection system and the adaptive detection system is 2.3 dB at a bit-error rate of 10^{-5} without a state offset. The SNR gap increases to 4.5 dB with a state offset of $\xi = 0.2$.**

***Index Terms*—Atmospheric turbulence, optical wireless communications, on-off keying.**

I. Introduction

Free-space optical (FSO) communication links have important advantages over radio frequency links. Such FSO systems offer broadband operation, high link security, and freedom from spectral license regulations. But optical signals that are transmitted over free-space are subject to amplitude and phase distortion due to transient inhomogeneities of atmospheric temperature and pressure [1], [2]. The resulting scintillation or fading is a major cause of performance degradation for FSO systems. The performance degradation is especially pronounced for FSO systems using irradiance modulation and direct detection (IM/DD) with on-off keying (OOK) and fixed detection thresholds that are non-adaptive and inherently unoptimized [3], [4]. This can produce irreducible error floors when operation is extended to high signal-to-noise ratios (SNRs) [5], [6].

Attempts to overcome the irreducible error floors of OOK IM/DD systems have focused on the application of adaptive detection thresholds. Adaptive detection of OOK signal can be classified into three categories. The first category, idealized adaptive detection, applies bit-by-bit adaptations to the detection threshold on the (typically) nanosecond timescale of the bit interval [2], [7]. The second category, quasi-static adaptive detection, applies adaptations to the detection threshold on the (typically) millisecond timescale of the turbulence coherence time [8], [9], [10]. The third category, electrical-SNR-optimized detection [11], [12], [13], unlike the prior two categories, does not require knowledge of the instantaneous channel state information (CSI), and it applies adaptations to the detection threshold on the second- or minute-long timescale over which the turbulence exhibits stationary statistics [14]. Operation with electrical-SNR-optimized detection offers the practical advantages of operation with fixed detection thresholds, as only slow adaptations are needed to define the detection thresholds[1]. Operation with electrical-SNR-optimized detection also offers the performance advantages of operation with idealized adaptive detection thresholds, as it avoids irreducible error floors. Unfortunately, existing electrical-SNR-optimized systems must assume perfect knowledge of the electrical SNR and turbulence probability distribution function (pdf).

While perfect knowledge of the turbulence distribution can be difficult to realize, it is possible to determine and make use of the statistical moments of the turbulence. The authors in [15]- [18] have, for example, made use of moments with Pearson curves, John curves and saddlepoints to approximate pdfs, albeit with somewhat restrictive conditions [19]. Similarly, the authors in [20] have recently derived unified pdf formulas based on the generalized Laguerre polynomial series expansion, although its parameters depend on the type of fading distribution.

In light of above systems and limitations, the contributions of this paper are as follows:

1) Novel analytical error floor expressions are derived for a general representation of practical FSO links with finite low and high state offsets (i.e., finite extinction ratios) operating over various turbulence channels.
2) The electrical-SNR-optimized system is implemented without requiring perfect knowledge of the instantaneous CSI and turbulence pdf. The turbulence pdf is approximated by a sum of Laguerre polynomials. With perfect knowledge of the turbulence pdf and no state offsets, the optimum detection rule reduces to that shown in [11], [12].
3) Method of moments estimation (MoME) and maximum likelihhood estimation (MLE) are used to estimate the

L. Yang, X. Song, J. Cheng and J. F. Holzman are with the School of Engineering, The University of British Columbia, Kelowna, BC, Canada. (e-mail: {luanxia.yang, xuegui.song, julian.cheng, jonathan.holzman}@ubc.ca)

This work was supported by an NSERC Discovery Grant.

[1]On the timescale of stationary statistics, the electrical SNR is constant, and the detection threshold is a fixed detection threshold.

state offsets and electrical SNR, and the electrical-SNR-optimized system is then employed to operate without irreducible error floors.

The reminder paper is organized as follows. Section II describes the system and channel models (for a finite extinction ratio). Section III derives the irreducible error floors of OOK with fixed detection thresholds. Section IV introduces our electrical-SNR-optimized detection system for operation with unknown turbulence model. Section V addresses the estimation of unknown state offsets for operation with finite extinction ratios and electrical SNRs. Section VI presents numerical results and discussions. Section VII makes some concluding remarks.

II. System and Channel Models

In an OOK IM/DD system, the transmitted intensity is a positive quantity that can be expressed as

$$\hat{s}(t) = 1 + \xi + \sum_i a_i g(t - iT_p) \tag{1}$$

where $a_i \in \{-1, 1\}$ is the data, and T_p is the symbol duration. In (1), pulse shaping is defined as $g(t) = 1$ for $0 < t < T_p$, and $g(t) = 0$ otherwise, and the positive parameter ξ is the low and high state offset that results from operation with a finite extinction ratio[2] [21], i.e., extinction ratio $= (2 + \xi)/\xi$. Finite extinction ratios are due to practical considerations for semiconductor laser transmitters, which often operate with finite power levels for the low and high states. Typical values of ξ are between 0.1053 and 0.2857 [22]. When $\xi \neq 0$, the low and high states of the received electrical signal are affected by turbulence. When $\xi = 0$, the received electrical signal specializes to the classical model discussed in [5].

The signal $\hat{s}(t)$ is transmitted through an atmospheric turbulence channel and is distorted by a multiplicative intensity process $I(u, t)$. The received electrical signal after photodetection can be written as

$$r(t) = R[(1 + \xi)I(u, t) + \sum_i I(u, t)a_i g(t - iT_p)] + n(u, t). \tag{2}$$

The photodetector responsivity, without loss of generality, is $R = 1$, and $I(u, t)$ is assumed to be a normalized stationary random process for signal scintillation caused by atmospheric turbulence and is modeled as lognormal distribution in this work, where u is an event in the sample space. The term $n(u, t)$ is additive white Gaussian noise process due to thermal noise and/or ambient shot noise. Using a p-i-n photodiode and following [11], the shot noise is assumed to be dominated by ambient shot noise. (Both ambient shot noise and thermal noise are statistically independent of the desired signal.) The total noise power is $\sigma_g^2 = \sigma_s^2 + \sigma_T^2$, where σ_s^2 and σ_T^2 denote the respective ambient shot noise power and the thermal noise power.

The received signal is sampled at time T_p. The sample $I(u, t = T_p)$ is a random variable (RV) I, and the sample $n(u, t = T_p)$ is a RV N having zero mean and variance $\sigma_g^2 = N_0/2$, where N_0 is the noise power spectral density. If "0" is transmitted, s_0 is true and the laser is in the low state, so the sample for demodulation is $r|_{s_0} = \xi I + N$. If "1" is transmitted, s_1 is true and the laser is in the high state, so the sample for demodulation is $r|_{s_1} = (2 + \xi)I + N$. It is important to note that the nonzero state offset ξ leads to turbulence dependence for the received signal when s_0 or s_1 is true.

The common statistical models that are used to characterize atmospheric turbulence channels are the lognormal, K, negative exponential, and Gamma-Gamma models [23]. The lognormal distribution characterizes weak turbulence and is suitable for characterizing FSO communications in clear sky links over several hundred meters [24]. The K-distribution is suitable for describing strong turbulence over links that are several kilometers in length [25]. The negative exponential distribution describes the limiting case of saturated scintillation [26]. The Gamma-Gamma distribution is a generalized model that can be applied to a wide range of turbulence conditions [27]. In this work, a lognormal turbulence channel is emphasized, but the developed detection algorithm is sufficiently general and can be applied to any turbulence models.

For the lognormal channel model, the optical irradiance I is given by

$$I = \exp(X) \tag{3}$$

where X is a Gaussian RV with mean μ and variance σ^2. Consequently, I follows a lognormal distribution with a pdf given by [24]

$$f_I(I) = \frac{1}{\sqrt{2\pi}\sigma I} \exp\left(-\frac{(\ln I - \mu)^2}{2\sigma^2}\right). \tag{4}$$

Normalizing the mean, i.e., $E[I] = 1$, where $E[\cdot]$ is the expectation operation, the pdf of I can be written as

$$f_I(I) = \frac{1}{\sqrt{2\pi}\sigma I} \exp\left(-\frac{(\ln I + \sigma^2/2)^2}{2\sigma^2}\right). \tag{5}$$

The parameter σ is the scintillation level [28]. Turbulence effects on the performance are minimal when scintillation levels are below 0.1, so the electrical-SNR-optimized detection system is characterized for scintillation levels ranging from 0.1 to 0.5. This is the typical range for scintillation levels [11], [29]. In this paper, the case for which the statistical details of (4) or (5) are completely unknown at the receiver is considered.

III. OOK with Fixed and Unoptimized Detection Thresholds

In the low state, the received signal ($r = \xi I + N$) is a sum of two RVs, N and I_s, where $I_s = \xi I$. Since N and I_s are assumed to be independent, the pdf of the received low state signal is the convolution of the marginal pdfs of I_s and N

[2]Extinction ratio, when used to describe the performance of an optical transmitter used in digital communications, is simply the ratio of the power used to transmit a logic state "1", to the power used to transmit a logic state "0".

according to

$$f(r|s_0) = \frac{1}{\xi} f_I\left(\frac{r}{\xi}\right) * f_N(r) = \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{\left(\ln\frac{x}{\xi} + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma_g} \exp\left(-\frac{(r-x)^2}{2\sigma_g^2}\right) dx \quad (6)$$

where $*$ denotes the convolution operation, and $f_N(r) = \frac{1}{\sqrt{2\pi}\sigma_g} \exp\left(-\frac{r^2}{2\sigma_g^2}\right)$ denotes the noise pdf.

In the high state, the pdf of the received signal ($r = (2+\xi)I + N$) can be defined in a similar manner according to

$$f(r|s_1) = \frac{1}{2+\xi} f_I\left(\frac{r}{2+\xi}\right) * f_N(r) = \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{\left(\ln\frac{x}{2+\xi} + \frac{\sigma^2}{2}\right)^2}{2\sigma^2}\right) \times \frac{1}{\sqrt{2\pi}\sigma_g} \exp\left(-\frac{(r-x)^2}{2\sigma_g^2}\right) dx. \quad (7)$$

For a given fixed detection threshold T_{th}, the probability of false alarm P_F and probability of miss P_M can be respectively written as

$$P_F = \int_{T_{th}}^\infty f(r|s_0)\, dr = \int_{T_{th}}^\infty \frac{1}{\xi} f_I\left(\frac{r}{\xi}\right) * f_N(r)\, dr \quad (8)$$

and

$$P_M = \int_{-\infty}^{T_{th}} f(r|s_1)\, dr = \int_0^{T_{th}} \frac{1}{2+\xi} f_I\left(\frac{r}{2+\xi}\right) * f_N(r)\, dr. \quad (9)$$

Assuming that p_1 represents the *a priori* probability that "1" is sent, i.e., $p_1 = \frac{1}{2}$ means "0"s and "1"s are equally likely to be sent, one can write the bit-error rate (BER) for OOK using a fixed detection threshold T_{th} as

$$\begin{aligned} P_e &= (1-p_1)P_F + p_1 P_M \\ &= \frac{(1-p_1)\exp\left(-\frac{\sigma^2}{8}\right)}{\sqrt{2\pi}\sigma} \int_0^\infty \frac{\sqrt{\xi}}{x^{3/2}} \exp\left(-\frac{\ln^2\frac{x}{\xi}}{2\sigma^2}\right) \times Q\left(\frac{T_{th}-x}{\sigma_g}\right) dx + \frac{p_1 \exp\left(-\frac{\sigma^2}{8}\right)}{\sqrt{2\pi}\sigma} \int_0^\infty \frac{\sqrt{2+\xi}}{x^{3/2}} \\ &\quad \times \exp\left(-\frac{\ln^2\frac{x}{2+\xi}}{2\sigma^2}\right) Q\left(\frac{x-T_{th}}{\sigma_g}\right) dx \\ &= \frac{(1-p_1)\exp\left(-\frac{\sigma^2}{8}\right)}{\sqrt{2\pi}\sigma} \int_0^\infty \frac{\sqrt{\xi}}{x^{3/2}} \exp\left(-\frac{\ln^2\frac{x}{\xi}}{2\sigma^2}\right) \times Q\left(\sqrt{\gamma}(T_{th}-x)\right) dx + \frac{p_1 \exp\left(-\frac{\sigma^2}{8}\right)}{\sqrt{2\pi}\sigma} \int_0^\infty \frac{\sqrt{2+\xi}}{x^{3/2}} \\ &\quad \times \exp\left(-\frac{\ln^2\frac{x}{2+\xi}}{2\sigma^2}\right) Q\left(\sqrt{\gamma}(x-T_{th})\right) dx \end{aligned} \quad (10)$$

TABLE I
ERROR FLOOR EXPRESSIONS FOR FSO SYSTEMS EMPLOYING FIXED DETECTION THRESHOLDS OF $T_{th} = (1+\xi)E[I]$ OVER A LOGNORMAL FADING CHANNEL WITH $\sigma = 0.25$

ξ	0.15	0.18	0.2	0.25
Theoretical error floor	0.0044	0.0049	0.0054	0.0065
Simulated error floor	0.0045	0.0049	0.0054	0.0066

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}}\, dt$ is the Gaussian Q-function, and we have denoted the electrical SNR by $\gamma = (E[I])^2/N_0$ [11], or simply $\gamma = 1/N_0$ under a normalized mean assumption.

In large SNR regimes, when γ approaches infinity or equivalently when $\sigma_g^2 = N_0/2$ approaches zero, the Gaussian distribution approaches a Dirac delta function $\delta(\cdot)$. Hence, one can have

$$\lim_{\gamma\to\infty} f_N(r) = \delta(r) \quad (11)$$

and

$$\lim_{\gamma\to\infty} \frac{1}{a} f_I\left(\frac{r}{a}\right) * f_N(r) = \frac{1}{a} f_I\left(\frac{r}{a}\right) \quad (12)$$

where a is a constant taking either ξ or $2+\xi$. When the electrical SNR is asymptotically large (i.e., $\gamma \to \infty$), using (8), (9), and (12), one obtains

$$\lim_{\gamma\to\infty} P_F = \int_{T_{th}}^\infty \frac{1}{\xi} f_I\left(\frac{r}{\xi}\right) dr = 1 - F_I\left(\frac{T_{th}}{\xi}\right) \quad (13)$$

and

$$\lim_{\gamma\to\infty} P_M = \int_0^{T_{th}} \frac{1}{2+\xi} f_I\left(\frac{r}{2+\xi}\right) dr = F_I\left(\frac{T_{th}}{2+\xi}\right) \quad (14)$$

where $F_I(\cdot)$ represents the cumulative distribution function (CDF) of the irradiance I. Therefore, the false alarm probability and miss probability in large SNR regimes are determined by the CDF of the irradiance evaluated at T_{th}/ξ and $T_{th}/(2+\xi)$, respectively. Substituting (13) and (14) into (10) gives

$$\begin{aligned} \lim_{\gamma\to\infty} P_e &= \lim_{\gamma\to\infty} (1-p_1)P_F + p_1 P_M \\ &= (1-p_1) Q\left(\frac{\ln T_{th} - \ln\xi + \sigma^2/2}{\sigma}\right) + p_1 Q\left(\frac{\ln(2+\xi) - \ln T_{th} - \sigma^2/2}{\sigma}\right) \end{aligned} \quad (15)$$

which is the error floor for an OOK IM/DD system with a fixed detection threshold through lognormal turbulence channels. As shown from (15), the error floor depends on both T_{th} and ξ, and typically one chooses the fixed detection threshold as $T_{th} = (E[r|s_1] + E[r|s_0])/2$. When $\xi = 0$, $T_{th} = E[I]$, and the analytical error floor expression in (15) is equivalent to [5, eq. (20)], which was derived under an assumption of normalized second moment, i.e., $E[I^2] = 1$. When $\xi \neq 0$, it is simple to show that $T_{th} = (1+\xi)E[I]$.

It is important to note that the error floor varies with the offset ξ. For a lognormal turbulence channel with $\sigma = 0.25$ and equal *a priori* data symbol probability, the predicted error floors are shown in Table I for different values of ξ. It is

seen that an increase of ξ results in a higher error floor. The theoretical error floors are verified with simulated BER limits in Table I.

Following the same approach, one can predict the error floors for different turbulence channel models based on the corresponding CDFs. The resulting error floors are summarized in Table II, where $K_{\alpha-\beta}(\cdot)$ denotes the modified Bessel function of the second kind with order $\alpha-\beta$, and the function $h(x,y,z,w)$ is defined as

$$h(x,y,z,w) = \left(\frac{xyz}{w}\right)^{x} \frac{\Gamma(y-x)}{\Gamma(x+1)\Gamma(y)} {}_1F_2\left(x; x+1, x-y+1; \frac{xyz}{w}\right) \tag{16}$$

where $\Gamma(\cdot)$ is the Gamma function and ${}_1F_2(\cdot;\cdot,\cdot;\cdot)$ is the generalized hypergeometric function [30].

IV. OOK with Electrical-SNR-Optimized Detection Thresholds

A performance trade-off can be established between operation with fixed detection thresholds (which can suffer from irreducible error floors) and adaptive detection thresholds (which require knowledge of the instantaneous SNR for each data symbol). This is done with an electrical-SNR-optimized detection system [11], [12]. The approach considers the optimization problem

$$\arg\min_{T_{th}} P_e = \arg\min_{T_{th}} [(1-p_1)P_F + p_1 P_M]. \tag{17}$$

From (10) and (17), it is clear that electrical-SNR-optimized detection requires knowledge of T_{th}, ξ, and the underlying turbulence model. To find the detection threshold that minimizes the BER at a given electrical SNR, one can take the derivative of (10) with respect to T_{th} and set it to zero, i.e., $\frac{\partial P_e}{\partial T_{th}} = 0$. This gives

$$-(1-p_1) f(T_{th}|s_0) + p_1 f(T_{th}|s_1) = 0 \tag{18}$$

where $f(T_{th}|s_0)$ and $f(T_{th}|s_1)$ are the likelihood functions evaluated at T_{th}. Assuming perfect knowledge of the pdf for the turbulence model and $\xi = 0$, the optimum detection rule reduces to the rule discussed in [11], [12]. The location of the electrical-SNR-optimized detection threshold lies at the intersection of two scaled likelihood functions: $(1-p_1)f(r|s_0)$ and $p_1 f(r|s_1)$. As shown in Fig. 1, when the electrical SNR approaches infinity, the total area underneath the intersected pdfs, i.e., $(1-p_1)P_F + p_1 P_M$, will become infinitely small. The electrical-SNR-optimized detection can therefore be used to eliminate the error floors caused by a receiver using fixed detection thresholds [13], [31].

To accommodate the fact that the FSO receiver may not always know the underlying turbulence model, the turbulence distribution can be approximated by sample moments. The approximated turbulence distribution can then be used to derive the electrical-SNR-optimized detection threshold.

The density functions of numerous statistical models on the positive half-line can be approximated by a sum of

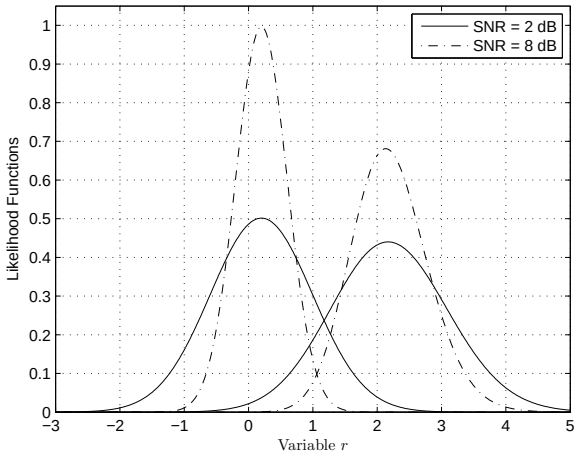

Fig. 1. The likelihood functions $f(r|s_0)$ and $f(r|s_1)$ with $\sigma = 0.25$ and $\xi = 0.2$ when $\gamma = 2$ dB and $\gamma = 8$ dB. The likelihood functions are a result of the convolution of the lognormal pdf and Gaussian pdf.

Laguerre polynomials [19], [32]. Using this approach, one can approximate the pdf of I as [33]

$$f_I(I) \approx \frac{I^v \exp(-I/c)}{c^{v+1}\Gamma(v+1)} \sum_{j=0}^{\infty} \delta_j L_j(v, I/c) \tag{19}$$

where $L_j(v, I/c)$ is a Laguerre polynomial of order j in I/c and is written as

$$L_j\left(v, \frac{I}{c}\right) = \sum_{k=0}^{j} \frac{(-1)^k \Gamma(v+j+1)}{k!(j-k)!\Gamma(v+j-k+1)} \left(\frac{I}{c}\right)^{j-k} \tag{20}$$

and

$$\delta_j = \sum_{k=0}^{j} (-1)^k \frac{j!\Gamma(v+1)}{k!(j-k)!\Gamma(v+j-k+1)} \mu_{\frac{I}{c}}[j-k] \tag{21}$$

where the jth moment of I is denoted by $\mu_I[j]$. In (19), the parameters $c = \frac{\mu_I[2]-\mu_I^2[1]}{\mu_I[1]}$ and $v = \frac{\mu_I[1]}{c} - 1$ are chosen to have the mean and variance of the Gamma RV I_0, whose pdf is $f_{I_0}(I) = \frac{I^v \exp(-I/c)}{c^{v+1}\Gamma(v+1)}$, match those of RV I. From (19), the corresponding characteristic function (CF) and moment generating function (MGF) for RV I can also be obtained. The detailed derivations are given in the Appendix. These analytical expressions can be used to estimate the performance of an FSO system over the lognormal fading. Substituting (19) into (6) and (7) yields the likelihood functions

$$f(r|s_0) = \frac{1}{\xi\sqrt{2\pi}\sigma_g c^{v+1}} \sum_{j=0}^{\infty} \delta_j \int_0^{\infty} \left(\frac{x}{\xi}\right)^{v-1} \exp\left(-\frac{x}{\xi c}\right) \times \exp\left(-\gamma(r-x)^2\right) L_j\left(v, \frac{x}{\xi}\right) dx \tag{22}$$

TABLE II
PDF AND ERROR FLOOR EXPRESSIONS FOR VARIOUS TURBULENCE CHANNEL MODELS

Turbulence Models	PDF	Error Floors
Lognormal	$\frac{1}{\sqrt{2\pi}\sigma I}$ $\times \exp\{-\frac{(\ln I+\sigma^2)^2}{2\sigma^2}\}$	$(1-p_1)Q\left(\frac{\ln T_{th}-\ln\xi+\sigma^2/2}{\sigma}\right)$ $+p_1 Q\left(\frac{\ln(2+\xi)-\ln T_{th}-\sigma^2/2}{\sigma}\right)$
K-distribution	$\frac{2\alpha^{\frac{\alpha+1}{2}}}{\Gamma(\alpha)} I^{\frac{\alpha-1}{2}}$ $\times K_{\alpha-1}(2\sqrt{\alpha I})$	$(1-p_1)\{1-\frac{1}{2}[h(1,\alpha,T_{th},\xi)+h(\alpha,1,T_{th},\xi)]\}$ $+p_1\{\frac{1}{2}[h(1,\alpha,T_{th},2+\xi)+h(\alpha,1,T_{th},2+\xi)]\}$
Gamma-Gamma	$\frac{2(\alpha\beta)^{\frac{\alpha+\beta}{2}}}{\Gamma(\alpha)\Gamma(\beta)} I^{\frac{\alpha+\beta}{2}-1}$ $\times K_{\alpha-\beta}(2\sqrt{\alpha\beta I})$	$(1-p_1)\{1-\frac{1}{2}[h(\beta,\alpha,T_{th},\xi)+h(\alpha,\beta,T_{th},\xi)]\}$ $+p_1\{\frac{1}{2}[h(\beta,\alpha,T_{th},2+\xi)+h(\alpha,\beta,T_{th},2+\xi)]\}$
Negative Exponential	$\frac{1}{\mu}\exp\{-\frac{I}{\mu}\}$	$(1-p_1)\exp\left(\frac{T_{th}}{\xi\mu}\right)+p_1\left\{1-\exp\left[\frac{T_{th}}{(2+\xi)\mu}\right]\right\}$

and

$$f(r|s_1)=\frac{1}{(2+\xi)\sqrt{2\pi}\sigma_g c^{v+1}}\sum_{j=0}^{\infty}\delta_j\int_0^{\infty}\left(\frac{x}{2+\xi}\right)^{v-1} \times\exp\left(-\frac{x}{(2+\xi)c}\right)\exp\left(-\gamma(r-x)^2\right)L_j\left(v,\frac{x}{2+\xi}\right)dx. \tag{23}$$

Substituting (22) and (23) into (18) yields

$$-\frac{1-p_1}{\xi}\sum_{j=0}^{\infty}\delta_j\int_0^{\infty}\left(\frac{x}{\xi}\right)^{v-1}\exp\left(-\gamma(x^2-2xT_{op})\right) \times\exp\left(-\frac{x}{\xi c}\right)L_j\left(v,\frac{x}{\xi}\right)dx+\frac{p_1}{2+\xi}\sum_{j=0}^{\infty}\delta_j \times\int_0^{\infty}\left(\frac{x}{2+\xi}\right)^{v-1}\exp\left(-\gamma(x^2-2xT_{op})\right) \times\exp\left(-\frac{x}{(2+\xi)c}\right)L_j\left(v,\frac{x}{2+\xi}\right)dx=0. \tag{24}$$

The detection threshold can be obtained numerically with respect to a given offset ξ and electrical SNR from (24). On the timescale of stationary statistics, the electrical SNR is constant, and the detection threshold is a fixed detection threshold. A comparison of the electrical-SNR-optimized detection thresholds, acquired by the approximated and exact lognormal pdfs, are presented in Table III. The thresholds are obtained by averaging 10 calculated detection thresholds. As shown from Table III, the approximated pdf can be used to calculate the detection threshold with high accuracy when the electrical SNR is less than 16 dB. For higher values of SNR, the calculated detection thresholds lose accuracy, and the corresponding BER curve deviates from the BER curve obtained with perfect knowledge of the lognormal pdf. This discrepancy occurs because the Laguerre-polynomial-based pdf approximation can not accurately describe the behaviors of the lognormal pdf near the origin. Fortunately, this inaccuracy does not concern most practical FSO systems, as they typically operate at relatively low SNR values [34].

V. PARAMETER ESTIMATION

As the electrical-SNR-optimized detection threshold introduced in Section IV requires knowledge of the state offset ξ and electrical SNR γ, it is necessary to estimate ξ and γ. MoME and MLE are used for this estimation in this section.

TABLE III
COMPARISON OF DETECTION THRESHOLDS BY USING AN EXACT AND APPROXIMATED LOGNORMAL PDF WITH $\sigma=0.25$

SNR (dB)	Thresholds with exact PDF	Thresholds with approximated PDF	Sample variance
0	0.9497	0.9505	3.11×10^{-8}
4	0.8633	0.8637	2.27×10^{-8}
8	0.7528	0.7496	1.89×10^{-8}
12	0.6302	0.6214	2.56×10^{-8}
16	0.5087	0.4984	1.08×10^{-7}
20	0.3981	0.4697	8.62×10^{-6}
24	0.3036	0.5239	1.03×10^{-5}

With bit-by-bit interleaved fading channels[3] [2], [11], it is assumed that there are $2L$ sampled signals during the observation interval. The vectors $\mathbf{R}=[r[0] \quad ... \quad r[2L-1]]^T$, $\mathbf{I}_f=[I[0] \quad ... \quad I[2L-1]]^T$, and $\mathbf{N}=[n[0] \quad ... \quad n[2L-1]]^T$ represent the received signal vector, fading coefficient vector, and noise vector, respectively. Assuming that a training sequence of length $2L$ is transmitted with L consecutive 1's followed by L consecutive 0's, one can write the received signal at the lth bit interval when bit 1 is transmitted as

$$r[l]|s_1=(2+\xi)I[l]+n[l], \quad l=0,1,...,L-1 \tag{25}$$

where $I[l]$ and $n[l]$ represent the fading coefficient and noise during the lth bit interval, respectively. Similarly, if L 0's are transmitted, the received signal at the kth bit interval can be written as

$$r[k]|s_0=\xi I[k]+n[k], \quad k=L,L+1,...,2L-1. \tag{26}$$

A. Method of Moments Estimation

Using (25) and (26), one can obtain the estimation of ξ as

$$\hat{\xi}=\frac{\frac{1}{L}\sum_{k=L}^{2L-1}r[k]|s_0}{\frac{1}{2L}\sum_{l=0}^{L-1}r[l]|s_1-\frac{1}{2L}\sum_{k=L}^{2L-1}r[k]|s_0}. \tag{27}$$

[3]A typical Gbps FSO system operates in a quasi-static atmospheric turbulence channel with a coherence time on the order of milliseconds. This coherence time is much longer than the nanosecond bit interval. Thus, the same fading coefficient affects a block of information bits, and the system performance suffers from correlation. However, we can transform a quasi-static channel into a block fading channel by way of block interleaving [35]. In such a system, we place each information bit in different blocks, such that each block (i.e., information bit) experiences independent fading from that of neighbouring blocks.

To assess the performance of the moment estimator $\hat{\xi}$, approximate expressions can be derived for the mean and variance of $\hat{\xi}$ when the sample size is asymptotically large. Assuming the statistics $\mathbf{T} = [T_1 \quad T_2]^T$, where $T_1 = \frac{1}{L}\sum_{l=0}^{L-1} r[l]|s_1$ and $T_2 = \frac{1}{L}\sum_{k=L}^{2L-1} r[k]|s_0$, one can obtain the covariance matrix as

$$\begin{aligned}\mathbf{C}_T &= \begin{pmatrix} \text{Var}[T_1] & \text{Cov}[T_1, T_2] \\ \text{Cov}[T_2, T_1] & \text{Var}[T_2] \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{L}[\sigma_g^2 + (2+\xi)^2\text{Var}[I]] & 0 \\ 0 & \frac{1}{L}(\sigma_g^2 + \xi^2\text{Var}[I]) \end{pmatrix}.\end{aligned} \tag{28}$$

Here, Var$[\cdot]$ denotes the variance, and Cov$[\cdot,\cdot]$ denotes covariance of two RVs. The estimator $\hat{\xi}$ can be rewritten as

$$\hat{\xi} \triangleq \varphi(\mathbf{T}) = \frac{2T_2}{T_1 - T_2}. \tag{29}$$

The estimator in (27) is consistent, i.e., $\hat{\xi} \xrightarrow{Pr} \xi$ as $L \longrightarrow \infty$, and is asymptotically Gaussian distributed, i.e., $\sqrt{L}(\hat{\xi}-\xi) \xrightarrow{L} \mathcal{N}(0, \sigma_{\hat{\xi}}^2)$. Performing a first-order Taylor expansion of $\varphi(\cdot)$ about the point $\mathbf{T} = E[\mathbf{T}]$ gives [36]

$$\hat{\xi} \approx \varphi(\mathbf{T})\Big|_{\mathbf{T}=E[\mathbf{T}]} + \sum_{i=1}^{2} \frac{\partial\varphi}{\partial T_i}\Big|_{\mathbf{T}=E[\mathbf{T}]} (T_i - E[T_i]) \tag{30}$$

where $E[\mathbf{T}] = [(2+\xi)E[I] \quad \xi E[I]]^T$. Taking the expectation of (30), one has

$$E[\hat{\xi}] \approx \varphi(\mathbf{T})\Big|_{\mathbf{T}=E[\mathbf{T}]} = \xi \tag{31}$$

and the asymptotic variance of $\hat{\xi}$ can be expressed as [37]

$$\begin{aligned}\sigma_{\hat{\xi}}^2 =& \text{Var}[\hat{\xi}] = \frac{\partial\varphi}{\partial T_i}\Big|^T_{\mathbf{T}=E[\mathbf{T}]} \mathbf{C_T} \frac{\partial\varphi}{\partial T_i}\Big|_{\mathbf{T}=E[\mathbf{T}]} \\ =& \frac{\sigma_g^2[\xi^2 + (2+\xi)^2] + 2\xi^2(2+\xi)^2\text{Var}[I]}{4L(E[I])^2}.\end{aligned} \tag{32}$$

Using (25) and (26), one can obtain the estimation of the turbulence mean $m = \exp(\mu + \sigma^2/2)$ and N_0, respectively, as

$$\hat{m} = \frac{1}{2L}\sum_{l=0}^{L-1} r[l]|s_1 - \frac{1}{2L}\sum_{k=L}^{2L-1} r[k]|s_0 \tag{33}$$

and

$$\hat{N}_0 = 2\frac{(\hat{m}|s_1)^2\,\hat{\mu}_r[2]|s_0 - (\hat{m}|s_0)^2\,\hat{\mu}_r[2]|s_1}{(\hat{m}|s_1)^2 - (\hat{m}|s_0)^2} \tag{34}$$

where $\hat{m}|s_1 = \frac{1}{L}\sum_{l=0}^{L-1} r[l]|s_1$, $\hat{\mu}_r[2]|s_0 = \frac{1}{L}\sum_{k=L}^{2L-1} r^2[k]|s_0$, $\hat{m}|s_0 = \frac{1}{L}\sum_{k=L}^{2L-1} r[k]|s_0$, and $\hat{\mu}_r[2]|s_1 = \frac{1}{L}\sum_{l=0}^{L-1} r^2[l]|s_1$. Using (33) and (34), one can obtain the estimation of γ as

$$\begin{aligned}\hat{\gamma} &= \frac{\hat{m}^2}{\hat{N}_0} \\ &= \frac{(\hat{m}|s_1 - \hat{m}|s_0)^2\left[(\hat{m}|s_1)^2 - (\hat{m}|s_0)^2\right]}{8\left[(\hat{m}|s_1)^2\,\hat{\mu}_r[2]|s_0 - (\hat{m}|s_0)^2\,\hat{\mu}_r[2]|s_1\right]}.\end{aligned} \tag{35}$$

B. Maximum Likelihood Estimation

The estimator based on the maximum-likelihood principle is the most popular approach to obtaining practical estimators. Additionally, for most cases of practical interest, its performance is optimal for large data records and is approximately the minimum variance unbiased estimator due to its approximate efficiency.

For the MLE, with bit-by-bit interleaved fading channels [2], [11], we transmit a training sequence consisting of $2L$ consecutive 1's. Assuming the received signal model is the same as (25), one can write the pdf of the received signal as [13]

$$\begin{aligned}f(r[k]|s_1;\boldsymbol{\theta}) =& f_I(r[k]) * f_N(r[k]) \\ =& \int_0^\infty \exp\left(-\frac{(\ln x - \ln(2+\xi) - \mu)^2}{2\sigma^2}\right) \\ &\times \frac{1}{\sqrt{2N_0}\pi\sigma x}\exp\left(-\frac{(r[k]-x)^2}{N_0}\right) dx\end{aligned} \tag{36}$$

where $\boldsymbol{\theta} = [\mu \quad \sigma^2 \quad N_0 \quad \xi]^T$ denotes the unknown vector, and $f_N(r[k]) = \frac{1}{\sqrt{\pi N_0}}\exp\left(-\frac{r^2[k]}{N_0}\right)$ is the noise pdf. Assuming that the components of the received signal vector $\mathbf{R}$ are independent, we can write the pdf of the received signal when s_1 is true as

$$\begin{aligned}f(\mathbf{R};\boldsymbol{\theta}) =& \prod_{k=0}^{2L-1} f(r[k]|s_1;\boldsymbol{\theta}) \\ =& \prod_{k=0}^{2L-1}\int_0^\infty \exp\left(-\frac{(\ln x - \ln(2+\xi) - \mu)^2}{2\sigma^2}\right) \\ &\times \frac{1}{\sqrt{2N_0}\pi\sigma x}\exp\left(-\frac{(r[k]-x)^2}{N_0}\right) dx.\end{aligned} \tag{37}$$

The MLE of the unknown vector $\boldsymbol{\theta}$ is obtained by maximizing the log-likelihood function

$$\begin{aligned}L(\mathbf{R};\boldsymbol{\theta}) =& \ln f(\mathbf{R};\boldsymbol{\theta}) \\ =& \ln\prod_{k=0}^{2L-1} f_I(r[k]) * f_N(r[k]) \\ =& \sum_{k=0}^{2L-1}\ln\int_0^\infty \exp\left(-\frac{(\ln x - \ln(2+\xi) - \mu)^2}{2\sigma^2}\right) \\ &\times \frac{1}{\sqrt{2N_0}\pi\sigma x}\exp\left(-\frac{(r[k]-x)^2}{N_0}\right) dx.\end{aligned} \tag{38}$$

Taking the derivative of (38) with respect to the unknown parameter and setting it equal to zero, we can obtain the MLE of the unknown vector $\boldsymbol{\theta}$. As it is difficult to obtain a closed-form expression for each unknown parameter, the expectation-maximization (EM) algorithm can be implemented numerically to determine the MLE. This method, although iterative in nature, is guaranteed under mild conditions to converge and produce a local maximum [37].

In order to simplify the problem, we decompose the original data sets into the independent data sets $y_1[k] = I[k]$ and

$y_2[k] = n[k]$, where $y_1[k]$ and $y_2[k]$ are the complete data, and they are related to the original data as $r[k] = y_1[k] + y_2[k]$. Instead of maximizing $\ln f(\mathbf{R};\boldsymbol{\theta})$, we can maximize $\ln f(\mathbf{Y};\boldsymbol{\theta})$, where $\mathbf{Y} = [\boldsymbol{y}_1 \quad \boldsymbol{y}_2]^T$, $\boldsymbol{y}_1 = [y_1[0] \quad ... \quad y_1[2L-1]]^T$ and $\boldsymbol{y}_2 = [y_2[0] \quad ... \quad y_2[2L-1]]^T$. Since $y_1[k] = I[k]$, we have

$$\begin{aligned}
&\ln f(y_1[k];\boldsymbol{\theta}) \\
&= \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}y_1[k]}\exp\left(-\frac{(\ln y_1[k] - \ln(2+\xi) - \mu)^2}{2\sigma^2}\right)\right) \\
&= -\ln\sqrt{2\pi\sigma^2} - \ln y_1[k] - \frac{(\ln y_1[k] - \ln(2+\xi) - \mu)^2}{2\sigma^2}.
\end{aligned} \tag{39}$$

Similarly, we have

$$\begin{aligned}
&\ln f(y_2[k];\boldsymbol{\theta}) \\
&= \ln\left(\frac{1}{\sqrt{\pi N_0}}\exp\left(-\frac{y_2^2[k]}{N_0}\right)\right) \\
&= -\ln\sqrt{\pi N_0} - \frac{y_2^2[k]}{N_0}.
\end{aligned} \tag{40}$$

Assuming $\hat{\boldsymbol{\theta}}^{(j)} = [\hat{\mu}^{(j)} \quad (\hat{\sigma}^2)^{(j)} \quad (\hat{N_0})^{(j)} \quad (\xi)^{(j)}]^T$ is an estimate of $\boldsymbol{\theta}$ in the jth iteration, each iteration of the EM algorithm can be written as follows.

E-step: This step determines the conditional expectation of the complete data

$$\begin{aligned}
U(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(j)}) &= E_{\mathbf{Y}|\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)}}[\ln f(\mathbf{Y};\boldsymbol{\theta})] \\
&= E_{\boldsymbol{y}_1|\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)}}[\ln f(\boldsymbol{y}_1;\boldsymbol{\theta})] + E_{\boldsymbol{y}_2|\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)}}[\ln f(\boldsymbol{y}_2;\boldsymbol{\theta})] \\
&= \int \ln f(\boldsymbol{y}_1;\boldsymbol{\theta}) f(\boldsymbol{y}_1|\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)}) d\boldsymbol{y}_1 \\
&\quad + \int \ln f(\boldsymbol{y}_2;\boldsymbol{\theta}) f(\boldsymbol{y}_2|\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)}) d\boldsymbol{y}_2.
\end{aligned} \tag{41}$$

where we have

$$f(\boldsymbol{y}_1|\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)}) = \frac{f(\mathbf{R}|\boldsymbol{y}_1;\hat{\boldsymbol{\theta}}^{(j)}) f(\boldsymbol{y}_1;\hat{\boldsymbol{\theta}}^{(j)})}{f(\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)})} \tag{42}$$

and

$$f(\boldsymbol{y}_2|\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)}) = \frac{f(\mathbf{R}|\boldsymbol{y}_2;\hat{\boldsymbol{\theta}}^{(j)}) f(\boldsymbol{y}_2;\hat{\boldsymbol{\theta}}^{(j)})}{f(\mathbf{R};\hat{\boldsymbol{\theta}}^{(j)})} \tag{43}$$

and where

$$\begin{aligned}
f(\mathbf{R}|\boldsymbol{y}_1;\hat{\boldsymbol{\theta}}^{(j)}) &= \prod_{k=0}^{2L-1} \frac{1}{\sqrt{2\pi(\hat{\sigma}^2)^{(j)}}(r[k] - y_1[k])} \\
&\times \exp\left(-\frac{(\ln(r[k] - y_1[k]) - \ln(2 + (\xi)^{(j)}) - \hat{\mu}^{(j)})^2}{2(\hat{\sigma}^2)^{(j)}}\right)
\end{aligned} \tag{44}$$

and

$$f(\mathbf{R}|\boldsymbol{y}_2;\hat{\boldsymbol{\theta}}^{(j)}) = \prod_{k=0}^{2L-1} \frac{1}{\sqrt{\pi(N_0)^{(j)}}} \exp\left(-\frac{(r[k] - y_2[k])^2}{(N_0)^{(j)}}\right). \tag{45}$$

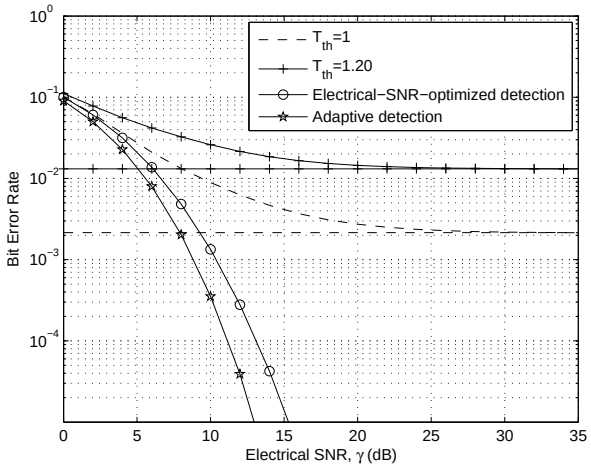

Fig. 2. BERs of OOK modulated systems using fixed detection thresholds T_{th}, electrical-SNR-optimized detection thresholds and adaptive detection thresholds over a lognormal turbulence channel with $\sigma = 0.25$ and $\xi = 0$.

M-step: This step maximizes (41) with respect to $\boldsymbol{\theta}$

$$\boldsymbol{\theta}^{(j+1)} = \arg\max_{\boldsymbol{\theta}} U(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{(j)}) \tag{46}$$

where $\hat{\boldsymbol{\theta}}^{(j+1)}$ is the new estimate of $\boldsymbol{\theta}$. For the EM algorithm, the conditional expectation of the complete data is nondecreasing until it reaches a fixed point. This fixed point is the MLE of $\boldsymbol{\theta}$, i.e., $\hat{\boldsymbol{\theta}}_{ML} = [\hat{\mu}_{ML} \quad \hat{\sigma}^2_{ML} \quad \hat{N}_{0,ML} \quad \hat{\xi}_{ML}]^T$. Based on the invariance property of the MLE, we obtain the MLE of μ_I as $\hat{\mu}_{I,ML} = \exp\left(\hat{\mu}_{ML} + \frac{\hat{\sigma}^2_{ML}}{2}\right)$. The MLE of γ can be obtained as

$$\hat{\gamma}_{ML} = \frac{(\hat{\mu}_{I,ML})^2}{\hat{N}_{0,ML}} = \frac{\left(\exp\left(\hat{\mu}_{ML} + \frac{\hat{\sigma}^2_{ML}}{2}\right)\right)^2}{\hat{N}_{0,ML}}. \tag{47}$$

The Cramér-Rao lower bound (CRLB) of $\hat{\gamma}$ can be calculated using [37]

$$\begin{aligned}
&\text{Var}[\hat{\gamma}] \\
&\geq \left[\frac{\partial\gamma}{\partial\mu} \quad \frac{\partial\gamma}{\partial\sigma^2} \quad \frac{\partial\gamma}{\partial N_0} \quad \frac{\partial\gamma}{\partial\xi}\right] \mathbf{I}^{-1}(\boldsymbol{\theta}) \left[\frac{\partial\gamma}{\partial\mu} \quad \frac{\partial\gamma}{\partial\sigma^2} \quad \frac{\partial\gamma}{\partial N_0} \quad \frac{\partial\gamma}{\partial\xi}\right]^T
\end{aligned} \tag{48}$$

where $\mathbf{I}(\boldsymbol{\theta})$ is the Fisher information matrix.

VI. NUMERICAL RESULTS

Figures 2 and 3 show the BERs versus electrical SNR when the OOK modulated system uses fixed detection thresholds of $T_{th} = 1$ and $T_{th} = 1.2$ with $\sigma = 0.25$ for the lognormal channel. For expository purposes, the parameter ξ is set to be $\xi = 0$ for Fig. 2 and $\xi = 0.2$ for Fig. 3. It is observed that the BER curves obtained by using Monte Carlo simulation show excellent agreement with the derived error floors in large SNR regimes and the error floors decrease for lower fixed detection thresholds.

To eliminate the error floors and improve the performance, the system with electrical-SNR-optimized detection thresholds

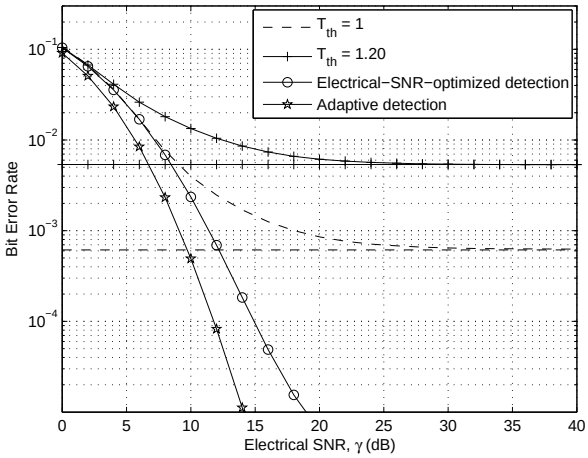

Fig. 3. BERs of OOK modulated systems using fixed detection thresholds T_{th}, electrical-SNR-optimized detection thresholds and adaptive detection thresholds over a lognormal turbulence channel with $\sigma = 0.25$ and $\xi = 0.2$.

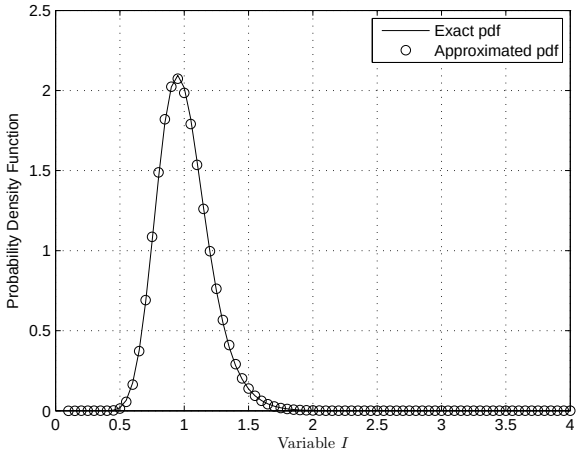

Fig. 4. Comparison of an approximated pdf using $J = 3$ sample moments and an exact pdf for a lognormal fading channel with $\sigma = 0.25$.

is used. The BERs for the system with the electrical-SNR-optimized detection thresholds are shown in Figs. 2 and 3, along with the BERs for the system with the adaptive detection thresholds. Both electrical-SNR-optimized detection thresholds are obtained by using the approximated lognormal pdf with $J = 3$ sample moments. Both the electrical-SNR-optimized and adaptive detection threshold results exhibit no error floors for increasing electrical SNR values. As expected, the system with electrical-SNR-optimized detection thresholds does not perform as well as the optimum OOK system using adaptive detection thresholds. For example, in Fig. 2, the OOK modulated system using adaptive detection thresholds requires an SNR of 13 dB to attain a BER of 10^{-5}, while the system using electrical-SNR-optimized detection thresholds requires an SNR of 15.3 dB to achieve the same BER performance. The corresponding SNR penalty factor in Fig. 2 for the system using an electrical-SNR-optimized detection threshold, is 2.3 dB at BER of 10^{-5}. The corresponding SNR penalty factor in Fig. 3 for the system using an electrical-SNR-optimized detection threshold, increases to 4.5 dB when $\xi = 0.2$. This performance difference can be factored into the ultimate FSO system design to offset the complexity of implementing systems with adaptive detection thresholds (and their need for knowledge of the instantaneous SNR). It is also important to point out that the BER performance achieved by the electrical-SNR-optimized system does not require rapid adjustment of the detection threshold. Since practical FSO systems typically operate at constant transmit power, the detection threshold only needs to be calculated once over durations of seconds or even minutes. The electrical-SNR-optimized system can therefore reduce the implementation complexity, compared to that of the idealized system using adaptive threshold detection.

In Fig. 4, the approximated lognormal pdf using $J = 3$ sample moments is compared with the exact lognormal pdf, for $\sigma = 0.25$. The absolute errors between these pdfs are shown explicitly in Fig. 5. The approximated lognormal pdf shows good agreement with the exact lognormal pdf when $\sigma = 0.25$. However, for higher σ values ($\sigma > 0.75$), the approximation of the lognormal pdf becomes inaccurate as integer moments can not uniquely determine the lognormal pdf. Fortunately, such scintillation levels are not encounted in practice [38]. A comparison of absolute errors from the pdf approximations using different numbers of sample moments is also given in Fig. 5. Clearly, larger numbers of sample moments can reduce the absolute error, but this comes at the cost of higher computational complexity. In general, a higher scintillation level σ will require higher order sample moments and the resulting approximation can become increasingly inaccurate. The Laguerre-polynomial-based approximation is accurate for the 0.1 to 0.5 range of σ values that is of interest to FSO applications [11], [29].

In Fig. 6, we compare the BER performance between the approximated lognormal pdf, for different values of σ with $J = 3$ sample moments, and the exact lognormal pdf. The two simulated error rate curves show good agreement over a wide range of SNR values. For the large SNR regime, the BER results from the approximate pdf have reduced accuracy, because the approximated pdf based on Laguerre polynomials is unable to characterize the behaviors of the lognormal pdf near the origin.

In order to evaluate the estimator performance, the sample variance of the electrical SNR estimator is compared with the CRLB. The variance of the electrical SNR estimator is given by

$$\hat{\sigma}^2_{\hat{\gamma}} = \frac{1}{M-1} \sum_{i=0}^{M-1} (\hat{\gamma}_i - \bar{\hat{\gamma}})^2 \tag{49}$$

where $\hat{\gamma}_i$ is the estimation by using MoME or MSE at the ith trail, M represents the total number of trails, and $\bar{\hat{\gamma}}$ is the sample mean of the electrical SNR estimator. In order to assess the estimator, Monte Carlo simulations are used to obtain $\hat{\sigma}^2_{\hat{\gamma}}$. In the simulation, different training sequence lengths are used to estimate the mean and noise variance, $M = 1 \times 10^4$ trials are used to calculate the variance of the electrical SNR estimator, and ξ is set as 0.2. Figure 7 plots the

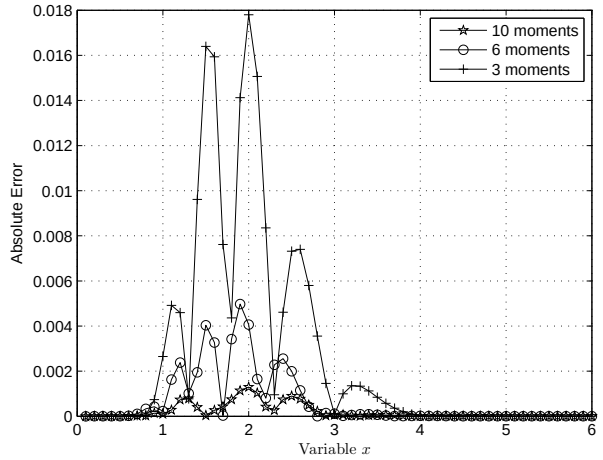

Fig. 5. The absolute error between the approximated pdf and the exact lognormal pdf with $\sigma = 0.25$ and $J = 3, 6, 10$ sample moments.

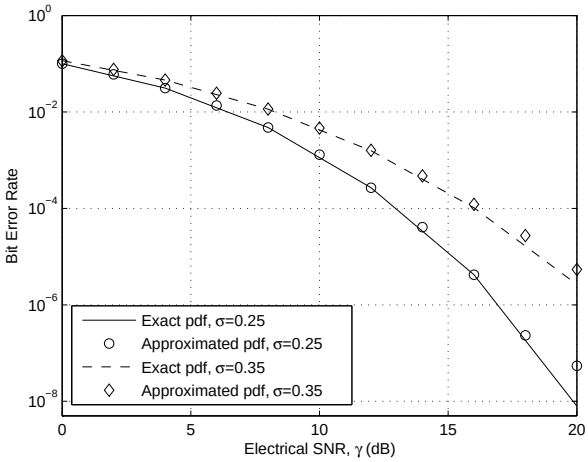

Fig. 6. Comparison of BERs obtained by the approximated pdf and the exact lognormal pdf with $\xi = 0$ and $J = 3$ sample moments.

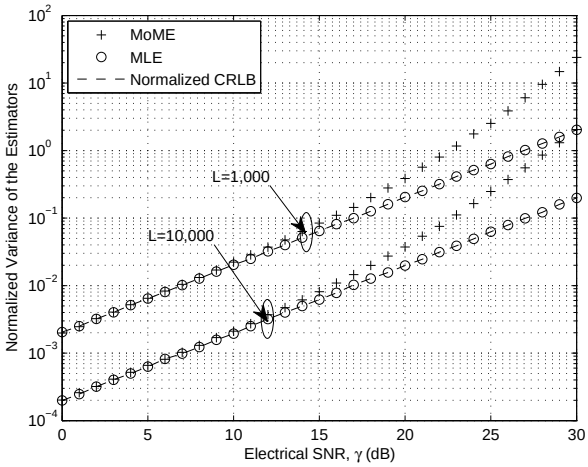

Fig. 7. Comparison of MoME and MLE normalized sample variance for different training sequence lengths over a lognormal turbulence channel with $\sigma = 0.25$. The normalized MSE is computed over $M = 1 \times 10^4$ trials.

normalized sample variance of the electrical SNR estimator, which is defined as the sample variance scaled by γ, versus the average electrical SNR. It is shown that the normalized sample variance for MLE attains the normalized CRLB, which is obtained by scaling the CRLB by γ. However, there is a discrepency between the normalized sample variance for MoME and the normalized CRLB for SNR values greater than 12 dB due to the inaccurate estimation of the noise variance. It is shown that the discrepency between the normalized sample variance for MoME and the normalized CRLB will disapear when $\xi = 0$. In this case, the received signal is the noise when 0 is transmitted. Thus, the noise variance can be accurately estimated by transmitting a training sequence with consecutive 0's. When $\xi \neq 0$ and 0 is transmitted, the received signal is the noise as well as the fading coefficient term. This leads to inaccurate estimation of the noise variance if a training sequence transmitted with consecutive 0's.

VII. CONCLUSION

It is known that FSO systems operating with OOK and fixed detection thresholds can suffer from irreducible error floors and power inefficiency. With this in mind, the resulting error floors are analyzed here (and validated with simulations) for lognormal turbulence channels and quantified for the general case having low and high state offsets, i.e., with finite extinction ratios. It is shown that the error floors can be eliminated by using electrical-SNR-optimized detection thresholds that minimize the average BER. The electrical-SNR-optimized system with the Laguerre-polynomials-based approximate pdf for the turbulence is found to be effective for typical FSO systems, which operate at relatively low SNR values, as it yields near-optimal BER performance. It is concluded that MLE is the preferred estimation technique for electrical-SNR-optimized detection with a finite extinction ratio, although such estimation comes at the cost of higher computational complexity.

APPENDIX

The CF of a RV I is the Fourier transform of its pdf, $f_I(I)$, and it is defined by

$$\Phi_I(\omega) = \int_{-\infty}^{\infty} f_I(I) \exp(j\omega I)\, dI \tag{50}$$

or

$$\Phi_I(\omega) = \mathrm{Re}[\Phi_I(\omega)] + j\mathrm{Im}[\Phi_I(\omega)] \tag{51}$$

where $j^2 = -1$. In (51), $\mathrm{Re}[\cdot]$ and $\mathrm{Im}[\cdot]$ denote the real and imaginary parts, respectively. Both can be written, respectively, as

$$\mathrm{Re}[\Phi_I(\omega)] = \int_{0}^{\infty} f_I(I) \cos(\omega I)\, dI \tag{52}$$

and

$$\mathrm{Im}[\Phi_I(\omega)] = \int_{0}^{\infty} f_I(I) \sin(\omega I)\, dI. \tag{53}$$

Using (19), one can approximate (52) as

$$\begin{aligned}&\mathrm{Re}[\Phi_I(\omega)]\\&\approx \int_0^\infty \frac{I^v \exp(-I/c)}{c^{v+1}} \sum_{n=0}^{\infty} \delta_n L_n\left(v, \frac{I}{c}\right) \cos(\omega I)\, dI\\&= \frac{1}{c^{v+1}} \sum_{n=0}^{\infty} \delta_n \int_0^\infty I^v \exp(-I/c) L_n\left(v, \frac{I}{c}\right) \cos(\omega I)\, dI.\end{aligned} \tag{54}$$

Substituting (20) into (54), one has

$$\begin{aligned}\mathrm{Re}[\Phi_I(\omega)] \approx& \frac{1}{c^{v+1}} \sum_{n=0}^{\infty} \delta_n \sum_{k=0}^{n} \frac{(-1)^k \Gamma(\alpha)}{k!(n-k)!\Gamma(\alpha-k)}\\&\times \int_0^\infty I^v \exp(-I/c) \left(\frac{I}{c}\right)^{n-k} \cos(\omega I)\, dI\\=& \sum_{n=0}^{\infty} \delta_n \sum_{k=0}^{n} \frac{(-1)^k \Gamma(\alpha)}{k!(n-k)!\Gamma(\alpha-k)(1+c^2\omega^2)^{\frac{\alpha-k}{2}}}\\&\times \cos((\alpha-k)\arctan(c\omega))\end{aligned} \tag{55}$$

where $\alpha = v + n + 1$. In deriving the last equality of (55), an integral identity [30, eq. 3.944(6)] has been used.

Similarly, substituting (19) and (20) into (53) and using an integral identity [30, eq. 3.944(5)], one obtains

$$\begin{aligned}\mathrm{Im}[\Phi_I(\omega)] \approx& \sum_{n=0}^{\infty} \delta_n \sum_{k=0}^{n} \frac{(-1)^k \Gamma(\alpha)}{k!(n-k)!\Gamma(\alpha-k)(1+c^2\omega^2)^{\frac{\alpha-k}{2}}}\\&\times \sin((\alpha-k)\arctan(c\omega)).\end{aligned} \tag{56}$$

The approximate CF is then found to be

$$\begin{aligned}\Phi_I(\omega) \approx& \sum_{n=0}^{\infty} \delta_n \sum_{k=0}^{n} \frac{(-1)^k \Gamma(\alpha)}{k!(n-k)!\Gamma(\alpha-k)(1+c^2\omega^2)^{\frac{\alpha-k}{2}}}\\&\times [\cos((\alpha-k)\arctan(c\omega)) + j\sin((\alpha-k)\arctan(c\omega))].\end{aligned} \tag{57}$$

Using an integral identity [30, eq. 3.326(2)], one can obtain the MGF as

$$M_I(s) \approx \sum_{n=0}^{\infty} \delta_n \sum_{k=0}^{n} \frac{(-1)^k \Gamma(v+n+1)}{k!(n-k)!(1-sc)^{v+n-k+1}}. \tag{58}$$

REFERENCES

[1] V. W. S. Chan, "Free-space optical communications," *IEEE/OSA J. Lightwave Technol.*, vol. 24, pp. 4750-4762, Dec. 2006.

[2] J. H. Shapiro and R. C. Harney, "Burst-mode atmospheric optical communication," in *Proc.1980 Nat. Telecommun. Conf.*, 1980, pp. 27.5.1-27.5.7.

[3] N. Chand, J. J. Loriz, A. J. Hunton, and B. M. Eteson, "Performance comparison of NRZ and RZ modulations with and without forward error corrections for free-space optical communication," in *Proc. SPIE*, vol. 5892, pp. 58920U-1-58920U-8, Sept. 2005.

[4] N. Liu, W. D. Zhong, Y. He, K. H. Heng, and T. H. Cheng, "Comparison of NRZ and RZ modulations in laser intersatellite communication systems," in *Proc. 2008 Int. Conf. Advanced Infocomm Tech.*, Shenzhen, 2008, pp. 29-32.

[5] J. Li, J. Q. Liu, and D. P. Taylor, "Optical communication using subcarrier PSK intensity modulation through atmospheric turbulence channels," *IEEE Trans. Commun.*, vol. 55, pp. 1598-1606, Aug. 2007.

[6] A. Jurado-Navas, J.M. Garrido-Balsells, M. Castillo-Vázquez, and A. Puerta-Notario, "Closed-form expressions for the lower-bound performance of variable weight multiple pulse-position modulation optical links through turbulent atmospheric channels," *IET Commun.*, vol. 6, pp. 390397, Apr. 2012.

[7] M. Tycz, M. W. Fitzmaurice, and D. A. Premo, "Optical communication system performance with tracking error induced signal fading," *IEEE Trans. Commun.*, vol. 21, pp. 1069-1072, Sept. 1973.

[8] M. L. B. Riediger, R. Schober, and L. Lampe, "Fast multiple-symbol detection for free-space optical communications," *IEEE Trans. Commun.*, vol. 57, pp. 1119-1128, Apr. 2009.

[9] M. L. B. Riediger, R. Schober, and L. Lampe, "Blind detection of on-off keying for free-space optical communications." in *Proc. CCECE*, Niagara Falls, Canada, 2008, pp. 1361-1364.

[10] H. R. Burris, *et al.*, "Laboratory implementation of an adaptive thresholding system for free-space optical communication receivers with signal dependent noise," in *Proc. SPIE*, 2005, vol. 5892, pp. 1-20.

[11] X. Zhu and J. M. Kahn, "Free-space optical communication through atmospheric turbulence channels," *IEEE Trans. Commun.*, vol. 50, pp. 1293-1300, Oct. 2002.

[12] H. Moradi, H. H. Refai, and P. G. LoPresti, "Thresholding-Based Optimal Detection of Wireless Optical Signals," *J. Opt. Comm. Net.*, vol. 2, pp. 689-700, Sept. 2010.

[13] L.Yang, J. Cheng and J. F. Holzman, "Electrical-SNR-Optimized detection threshold for OOK IM/DD optical wireless communications," in *Proc. Canadian Workshop Inform. Theory*, Toronto, ON, Jun. 18-21, 2013, pp. 186-189.

[14] D. A. DeWolf, "Are strong irradiance fluctuations log normal or Rayleigh distributed," *J. Opt. Soc. Amer.*, vol. 57, pp. 787-797, Jun. 1967.

[15] H. Solomon and M. Stephens, "Approximations to density functions using Pearson curves," *J. Amer. Stat. Assoc.*, vol. 73, pp. 153-160, Mar. 1978.

[16] W. P. Elderton and N. L. Johnson, *Systems of Frequency Curves*, 1st ed. Oxford: Cambridge University Press, 1969.

[17] C. Rose and M. D. Smith, *Mathematical Statistics with Mathematica*, 1st ed. New York: Springer-Verlag, 2002.

[18] N. Reid, "Saddlepoint methods and statistical inference," *Stat. Sci.*, vol. 3, pp. 213-238, May 1988.

[19] S. B. Provost, "Moment-based density approximants," *Math. J.*, vol. 9, pp. 727-756, May 2005.

[20] C. C. Chai and T. T. Tjhung, "Unified Laguerre polynomial series based distribution of small scale fading envelopes," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 3988-3999, Oct. 2009.

[21] G. P. Agrawal, *Fiber-Optic Communication Systems*, 3rd ed. New York: John Wiley & Sons, 2002.

[22] *Measuring extinction ratio of optical transmitters*, Agilent Tech., Santa Clara, CA, Applicat. Note 1550-8, 2001, pp. 5.

[23] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation Through Random Media*. Bellingham, WA: SPIE Press, 1998.

[24] E. J. Lee and V. W. S. Chan, "Part 1: Optical communication over the clear turbulent atmospheric channel using diversity," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1896-1906, Nov. 2004.

[25] M. Niu, J. Cheng, and J. F. Holzman, "Exact error rate analysis of equal gain and selection diversity for coherent free-space optical systems on strong turbulence channels," *Opt. Express*, vol. 18, pp. 13915-13926, Jun. 2010.

[26] W. O. Popoola and Z. Ghassemlooy, "BPSK subcarrier intensity modulated free-space optical communications in atmospheric turbulence," *IEEE/OSA J. Lightwave Technol.*, vol. 27, pp. 967-973, Apr. 2009.

[27] A. Al-Habash, L. C. Andrews, and R. L. Phillips, "Mathematical model for the irradiance probability density function of a laser beam propagating through turbulent media," *Opt. Eng.*, vol. 40, pp. 1554-1562, Aug. 2001.

[28] J. C. Brandenburg and J. Q. Liu, "Signal detection in optical communications through the atmospheric turbulence channel," in *Proc. IEEE Global Telecommun. Conf.*, New Orleans, LA, Nov. 30-Dec. 4, 2008, pp. 2741-2745.

[29] M. Uysal, J. Li, and M. Yu, "Error rate performance analysis of coded free-space optical links over Gamma-Gamma atmospheric turbulence channels," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1229-1233, Jun. 2006.

[30] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 4th ed, San Diego: Academic Press, 2000.

[31] L.Yang, J. Cheng and J. F. Holzman, "Optical communications over lognormal fading channels using OOK," in *Proc. 2nd Int. Workshop Opt. Wireless Commun.*, pp. 99-103, Newcastle Upon Tyne, Oct. 21, 2013.

[32] R. Price, "An orthonormal Laguerre expansion yielding Rices envelope density function for two sine waves in noise," *IEEE Trans. Inf. Theory*, vol. 34, pp. 1375-1382, Nov. 1998.

[33] L. Devroye, "On random variate generation when only moments or fourier coefficients are known," *Math. Comput. Simul.*, vol. 31, pp. 71-89, Feb. 1989.

[34] A. A. Farid and S. Hranilovic, "Capacity bounds for wireless optical intensity channels with Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 56, pp. 6066-6077, Dec. 2010.

[35] H. E. Gamal and A. R. Hammons, "On the design of algebraic space-time codes for MIMO block-fading channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 151-163, Nov. 2003.

[36] E. L. Lehmann, *Elements of Large-Sample Theory*, 1st ed. New York: Springer-Verlag, 1999.

[37] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, 1st ed. Upper Saddle River: Prentice Hall, 1993.

[38] I. I. Kim, E. L. Woodbridge, V. J. Chan, and B. R. Strickland,, "Scintillation measurements performed during the limited-visibility lasercom experiment," in *Proc. SPIE*, San Jose, CA, May, 1998, pp. 209-220.