Free-Space Optical MIMO Transmission with APD Receivers in Atmospheric Turbulence

Neda Cvijetic, Stephen G. Wilson, Senior Member, IEEE,
and Maïté Brandt-Pearce, Senior Member, IEEE

Abstract

In this paper we analyze multiple-input multiple-output (MIMO) free-space optical (FSO) communication systems with $Q$-ary pulse position modulation (PPM) and avalanche photodiodes (APDs) in the receiver array in the presence of atmospheric turbulence, and derive the general maximum-likelihood (ML) detection rule and the Chernoff bound on error probability ($P_e$) for the MIMO setup. The obtained analytical results are evaluated under both the Webb-Gaussian and jointly-Gaussian models for receiver outputs. It is shown that accuracy of jointly-Gaussian modelling increases with receiver-array size. Moreover, MIMO systems with APD receivers are shown to be markedly superior to analogous non-diverse systems, particularly in strong turbulence, with full spatial diversity achieved in a Rayleigh channel. Finally, it is show that for a given thermal noise level, selection of APD gain based on a non-fading SNR criterion for a single APD receiver results in nearly-optimal $P_e$ in MIMO systems at arbitrary levels of atmospheric turbulence.

Index Terms

Free-space optical (FSO) communication, multiple-input/multiple-output (MIMO) processing, avalanche photodiode (APD), atmospheric turbulence, pulse-position modulation (PPM)

I. INTRODUCTION

FREE-space optical (FSO) technology may become prominent in next generation broadband networks. Multi-gigabit potential data rates, unlicensed spectrum, excellent security and quick and inexpensive setup are among its most attractive features [1]. Current link speeds and lengths vary, but it is envisioned that FSO may be best suited for multi Gb/s rates over distances up to a few kilometers. Yet FSO links, assumed to be line-of-sight, suffer from atmospheric turbulence from inhomogeneities in the air’s index of refraction, as well as thermally-noise-limited receivers [1]. Desire to overcome these limitations motivates this study of MIMO FSO systems with avalanche photodiodes (APDs) in the receiver array. The key idea is that FSO MIMO techniques can improve performance in a manner similar to RF MIMO systems [2], in this case via spatial diversity [1], while APDs help approach shot-noise-limited operation.
Since lasers are normally intensity modulated and detection is noncoherent (direct), previous FSO MIMO studies have often employed $Q$-ary pulse position modulation (PPM) as an energy-efficient transmission method [1], [3], [4]. Moreover, PPM avoids adaptive threshold adjustment required in on-off keying (OOK) [5]. In [1] it is shown that full transmitter diversity is obtained with repetition PPM across the laser array, while a performance analysis of multi-pulse PPM is provided in [3]. Additional codes for optical MIMO are studied in [4], [7]–[9].

The studies in [1], [3]–[6] treated ideal photon-counting receivers for differing fading scenarios and background radiation levels. Receivers with simple p-i-n detectors would likely be thermally-noise-limited in practice, so it is of great interest to consider receivers with APD front-ends due to the internal current gain they provide. Yet APDs themselves are noisy in that the number of secondary photoelectrons (PEs) produced for each primary PE is a random variable, adding an excess noise to the process over what an ideal optical amplifier with the same gain would achieve. These effects have been studied in [10]–[12] and applied to optical receiver performance in [13]. The accuracy of a Gaussian model for the relevant random variables in OOK and PPM was assessed in [13]–[15] for a single APD receiver. In [16] performance analysis was extended to turbulent optical channels for the single APD case and PPM signalling. In [17], work from [6] was extended to an APD array for an air-to-ground system in the single-transmitter case.

However, to the best of our knowledge, an error probability analysis of FSO systems in optical fading for multiple transmitters and multiple APD receivers has not yet appeared in the literature. We endeavor to cover this scenario. Section II introduces the FSO MIMO system with APD receivers as well as the channel fading models. In Section III, we derive the maximum-likelihood (ML) detection rule and Chernoff bound on error probability and apply them to the Webb-Gaussian and jointly-Gaussian models for APD receiver outputs. We also highlight some subtleties of the jointly-Gaussian approximation. Section IV presents analytical and simulation results for error probability in FSO MIMO channels with atmospheric turbulence. Section V concludes the paper.

II. SYSTEM MODEL

Let us assume that $M$ lasers are pointed at a receiving array of $N$ APDs as shown in Figure 1. The beamwidth of each laser is wide enough to illuminate the entire receiver array and the laser array produces the same total optical power irrespective of $M$ for fairness of comparison. Letting $\gamma_{nm}$ represent the optical power multiplier due to fading between transmitter $m$ and receiver $n$, we have that the $n^{th}$ APD receives an optical power

$$P_n = \frac{M}{M} \sum_{m=1}^{M} \gamma_{nm} P/M$$  \hspace{1cm} (1)
where \( P \) is the received power contributed by all lasers to a single APD in the absence of atmospheric fading. (This assumption parallels means of comparison in RF MIMO systems as well.) The spacing between adjacent lasers and photodetectors is assumed sufficient to make \( \{\gamma_{nm}\} \) reasonably independent [18]. The adopted signaling method is \( Q \)-ary PPM with all lasers repeating the same signal (repetition coding). The symbol duration is \( T_s \) seconds with PPM slot time \( T = T_s/Q \). The slot duration is related to the bit interval \( T_b \) by

\[
T = T_b \frac{\log_2 Q}{Q} \tag{2}
\]

Each detector may also intercept a background radiation \( P_0 \) that merely adds to the total irradiance and is determined by the optical field of view and optical bandwidth prior to the detector as [19]:

\[
P_0 = [W(\lambda)\Omega_{fv}][\Delta \lambda][A_l] \tag{3}
\]

where \( P_0 \) is in Watts, \( W(\lambda) \) is the source radiance (Watts/cm\(^2\)-\(\mu\)m-srad), \( \Omega_{fv} \) is the receiver’s field of view (srad), \( \Delta \lambda \) is the optical bandwidth of the detector and \( A_l \) is the aperture (lens) area. If background radiation from a diffuse sky is assumed, \( W(\lambda) \approx 10^{-4} \) (Watts/cm\(^2\)-\(\mu\)m-srad) [19]; combined with a field of view \( \Omega_{fv} \approx 2 \times 10^{-3} \) srad for a lens with diameter of 2 cm, \( \Delta \lambda = 0.01 \lambda = 0.015 \mu m, A_l \approx 3.14 \) cm\(^2\), slot time \( T = 0.2 \) ns corresponding to \( R_b = 2.5 \) Gbps and 2-PPM, this yields \( P_0 T_b \approx 10^{-17} \) Joules, or -170 dBj, which is the value assumed throughout this analysis.

We define \( \Gamma = \{\gamma_{nm}, n = 1, ..., N, m = 1, ..., M\} \) as the channel optical power fading matrix and employ two statistical models for its p.d.f., \( f_\Gamma(\gamma) \). In lognormal fading [18], [20], the p.d.f. for the optical amplitude path gain, \( A \), is

\[
f_A(a) = \frac{1}{(2\pi \sigma_X^2)^{1/2}a} \exp\left(-(\ln a - \mu_X)^2/2\sigma_X^2\right), \quad a > 0 \tag{4}
\]

where \( A = e^X \) and \( X \) is normal with mean \( \mu_X \) and variance \( \sigma_X^2 \). Since \( \Gamma = A^2 \), \( \Gamma \) is also lognormal. It is assumed that \( E[\Gamma] = 1 \) so that, on average, the turbulence neither amplifies nor attenuates the optical power; this requires \( \mu_X = -\sigma_X^2/2 \) in (4). To characterize the degree of fading, we use the ‘scintillation index’ \( (\Psi) \) defined as

\[
\Psi = \frac{E[\Gamma^2]}{E^2[\Gamma]} - 1 \tag{5}
\]

and related to the log-amplitude variance \( \sigma_X^2 \) by \( \Psi = e^{4\sigma_X^2} - 1 \). In Rayleigh fading, the p.d.f. of the amplitude path gain \( A \) is

\[
f_A(a) = 2ae^{-a^2}, \quad a > 0 \tag{6}
\]

where the p.d.f. is again normalized so that \( E[\Gamma] = E[A^2] = 1 \). \( \Gamma \) is now exponentially distributed [22], [23]. Finally, the channel is modelled as slow-fading and non-frequency-selective since the turbulence
correlation times are assumed to be much longer than the PPM symbol duration, while the channel delay spread is much shorter than $T_s$.

We model each receiver as an APD with mean gain $G$ and excess noise figure $F$ (16) followed by a transimpedance amplifier with transimpedance $R$ (Figure 2) that adds electronic thermal and shot noise, which is modelled as white, Gaussian, and independent of the optical shot noise process [21]. The amplifier output is integrated over each PPM slot and for each detector, producing a matrix of observations $Z = \{Z_{nq}\}$, where $n$ indexes APDs and $q$ indexes slots. $Z$ then forms a sufficient statistic for optimal detection.

III. OPTIMAL DETECTION RULE

Given that $Z_{nq}$ is the integrated voltage at detector $n$ and slot $q$, and $Z = \{Z_{nq}, n = 1, ..., N, q = 1, ..., Q\}$ represents the set of received data, the ML decision rule is

$$
\hat{q} = \arg \max_i f_Z(Z|i)
$$

where $f_Z(Z|i)$ is the p.d.f. for $Z$ conditioned on PPM slot $i$ being sent and is understood to be conditioned on $\Gamma$ as well. By independence among slots and across detectors, we observe that

$$
f_Z(Z|i) = \prod_{q=1}^{Q} \prod_{n=1}^{N} f(Z_{nq}|i)
$$

$$
= \prod_{n=1}^{N} \left[ \frac{f_{on_n}(Z_{ni})}{f_{off_n}(Z_{ni})} \prod_{q=1}^{Q} f_{off}(Z_{nq}) \right]
$$

where $f_{on_n}$ and $f_{off}$ denote the p.d.f.s for ‘on’ and ‘off’ PPM slots respectively. (It is noted that only the ‘on’ slot statistics vary across detectors as influenced by $\Gamma$ through (1)). This produces

$$
\hat{q} = \arg \max_i \sum_{n=1}^{N} \ln \left[ \frac{f_{on_n}(Z_{ni})}{f_{off}(Z_{ni})} \right] = \arg \max_i \sum_{n=1}^{N} \lambda(Z_{ni})
$$

which amounts to forming column sums of metrics $\lambda(Z_{ni})$ for each slot and deciding in favor of the largest.

In the two-signal hypothesis test (2-PPM) with $\Lambda_i = \sum_{n=1}^{N} \lambda(Z_{ni})$, we are interested in $P_{e|\Gamma,i} = P[\Lambda_j > \Lambda_i|i$ sent], where $i$ indicates the ‘on’ slot and $j$ the ‘off’ slot without loss of generality. The Chernoff bound is invoked to determine $P_{e|\Gamma,i}$ for cases where the exact analytical approach becomes intractable and/or simulation may be very time-consuming at low error probabilities:

$$
P_{e|\Gamma,i} \leq \min_{s \geq 0} E_i[e^{s(\Lambda_j - \Lambda_i)}]
$$

$$
\leq \min_{s \geq 0} \prod_{n} E_i[e^{s\lambda(Z_{ni})}]E_i[e^{-s\lambda(Z_{ni})}]
$$

which follows from definition of the $\Lambda$ metrics and conditional independence of observations. Here $E_i[\ ]$ denotes expectation of the enclosed under the condition that light is sent in the $i^{th}$ PPM slot. Using
moment generating functions (m.g.f), (10) can be expressed as

\[ P_{e|\Gamma,i} \leq \min_{s \geq 0} \prod_n \Phi_{off}(s) \Phi_{on}(s) \]  

(11)

where

\[ \Phi_{on}(s) = \int f_{on}(x) e^{-s\lambda(x)} dx \]
\[ \Phi_{off}(s) = \int f_{off}(x) e^{s\lambda(x)} dx \]

Noting that \( \lambda(x) = \ln[f_{on}(x)/f_{off}(x)] \) leads to

\[ \Phi_{on}(s) = \int [f_{on}(Z_{ni})]^{1-s}[f_{off}(Z_{ni})]^s dZ_{ni} \]  

(12)
\[ \Phi_{off}(s) = \int [f_{on}(Z_{nj})]^s[f_{off}(Z_{ni})]^{1-s} dZ_{nj} \]  

(13)

To generalize the binary PPM result to the Q-ary case with \( N \) detectors, we can apply the union bound and note that all pairwise error probabilities are equal, yielding the result

\[ P_{e|\Gamma,i} \leq (Q - 1) \min_{s \geq 0} \prod_n \Phi_{off}(s) \Phi_{on}(s) \]  

(14)

which is \( Q - 1 \) times the binary Chernoff bound. Since the metrics \( \Lambda_i \) are sums of random variables over \( N \), increasing \( N \) will give an increasingly tighter Chernoff bound on \( P_e \). In the following sections, we will develop probability models for \( Z_{nq} \) and apply them to both the ML rule and the resulting Chernoff bound on \( P_e \).

A. Webb-Gaussian Model

In an APD, primary photoelectrons (PEs) create secondary ones via an avalanche mechanism such that the number of secondary PEs, \( m \), per primary PE is random; we let \( P_n[m] \) be the probability of observing \( m \) secondary PEs at the \( n^{th} \) APD output over a PPM slot time \( T \). The Webb model [12] given by (15) is an excellent approximation for the exact \( P_n[m] \) distribution [11]:

\[ P_n[m] = \frac{1}{(2\pi\alpha_n G^2 F)^{1/2}} \left[ 1 + \frac{m - G\alpha_n}{\alpha_n G F (F - 1)} \right]^{3/2} \]
\[ \times \exp\left[ -\frac{(m - G\alpha_n)^2}{2\alpha_n G^2 F (1 + \frac{m - G\alpha_n}{G\alpha_n F (F - 1)})} \right] \]

(15)

for \( m = 0,1,2,\ldots; G \) is mean APD gain typically in the range of 10-200, and \( F = F(G) \) is the APD excess noise factor given by

\[ F(G) = \xi G + (2 - \frac{1}{G})(1 - \xi) \]  

(16)
where $\xi$ is the ionization coefficient ratio of the APD semiconductor material. The average number of primary PEs during ‘on’ slots is $\alpha_n = \eta(\sum_{m=1}^{M} \gamma_{nm}P/M + P_0)/h\nu$, where $\eta$ is the APD quantum efficiency, $h$ is Planck’s constant, $\nu$ is the optical frequency, and $P_0$ is the incident background power (W); for ‘off’ slots, $\alpha_n = \alpha = \eta P_0/h\nu$. Added to this source of randomness is the post-amplifier electronic noise that is modelled as white and Gaussian with zero mean and variance $\sigma^2 = 2kT T_e R V^2$ at the output of the slot integrator (Figure 2), where $T_e$ denotes the post-amplifier’s equivalent electrical noise temperature (K) and $k$ is Boltzmann’s constant. Deriving the model for $Z_{nq}$ via the Webb model with additive Gaussian noise entails convolving (16) with a Gaussian random variable (r.v.) to obtain a mixture of Gaussians for $f_Z(z)$, the p.d.f. of $Z_{nq}$. If shot-noise were to heavily dominate, the distribution for $Z_{nq}$ would approach $P_n[m]$, i.e. $f_Z(z) = \sum m P_n[m] \delta(z - meR)$, where $e$ is electronic charge and $R$ is the post-amplifier trans-impedance. If additive Gaussian noise were dominant, as in the thermal noise limit, $f_Z(z)$ would be well-approximated by a single Gaussian p.d.f.:

### B. Jointly Gaussian Model

Earlier work on APDs [10]–[14] indicates that a Gaussian model for $P_n[m]$ is accurate for most practical scenarios, including moderate incident optical power levels and/or APD gains. When combined with additive Gaussian noise from the post-amplifier, this assumption yields a Gaussian distribution for $f_Z(z)$. The Gaussian model for $f_Z(z)$ has been shown to be valid in receiver threshold tests for almost all operating conditions [14]. Under this approximation, (9) becomes

$$\hat{q} = \arg \min_i \sum_{n=1}^{N} \left\{ \frac{(Z_{ni} - \mu_{on})^2}{\sigma_{on}^2} - \frac{(Z_{ni} - \mu_{off})^2}{\sigma_{off}^2} \right\}$$  \hspace{1cm} (17)

where $\mu_{on}$, $\mu_{off}$, $\sigma_{on}^2$, $\sigma_{off}^2$ are the means and variances for p.d.f.s of ‘on’ and ‘off’ PPM slots respectively. Thus, under jointly-Gaussian modelling for receiver outputs, the ML rule produces a quadratic detector (17) which formally reduces to a linear one in special cases where $\sigma_{off}^2 = \sigma_{on}^2$. For the ‘on’ slots:

$$\mu_{on} = G Re \eta \left( \sum_{m=1}^{M} \gamma_{nm} \frac{P}{M} + P_0 \right) T/h\nu$$ \hspace{1cm} (18)

$$\sigma_{on}^2 = G^2 (Re)^2 F(G) \eta \left( \sum_{m=1}^{M} \gamma_{nm} \frac{P}{M} + P_0 \right) T/h\nu + 2kT T_e R$$ \hspace{1cm} (19)

in $V$ and $V^2$ respectively. To obtain the statistics for the ‘off’ PPM slots, we set $P = 0$ in (18) and (19). Because of the Gaussian model and unequal variances for the ‘on’ and ‘off’ slots, the metrics $\Lambda_i = \sum_{n=1}^{N} \lambda(Z_{ni})$ are sums of non-central chi-squared r.v.s, with non-centrality parameters conditioned on the
fading variables. This is more readily observed if we express (17) as

$$\Lambda_i = \sum_{n=1}^{N} (a_n Z_{ni}^2 - 2b_n Z_{ni} + c_n)$$  \hspace{1cm} (20)

$$= \sum_{n=1}^{N} a_n(Z_{ni} - b_n/a_n)^2 + d_n$$  \hspace{1cm} (21)

where

$$a_n = \frac{\sigma_{off}^2 - \sigma_{on}^2}{\sigma_{off}^2 \sigma_{on}^2}$$
$$b_n = \frac{\mu_{on} \sigma_{off}^2 - \mu_{off} \sigma_{on}^2}{\sigma_{off}^2 \sigma_{on}^2}$$
$$c_n = \frac{\mu_{on} \sigma_{off}^2 - \mu_{off} \sigma_{on}^2}{\sigma_{off}^2 \sigma_{on}^2}$$
$$d_n = c_n/a_n - (b_n/a_n)^2$$

It is then noted that \((Z_{ni} - b_n/a_n)^2\) is a non-central \(\chi^2(\delta)\) r.v. with one degree of freedom and non-centrality parameter \(\delta_n = \mu_{on_{(off)}} - b_n/a_n\), depending on whether \(Z_{ni}\) is an ‘on’ or ‘off’ slot. The m.g.f for a non-central \(\chi^2(\delta)\) r.v. can be readily derived in closed form as:

$$\Phi_n(s) = E[e^{sX}] = \exp(\delta_n^2 s / (1 - 2s)) / \sqrt{1 - 2s}$$  \hspace{1cm} (22)

By substituting (21) into (10) and performing the needed translation, scaling, and cancellation, the following Chernoff bound is obtained:

$$P_{\epsilon|\Gamma} \leq (Q-1) \min_{s \geq 0} \prod_{n=1}^{N} \Phi_{on_n}(a_n \sigma_{on_n}^2 s) \Phi_{off}(-a_n \sigma_{off}^2 s)$$  \hspace{1cm} (23)

where \(\Phi_{on_n}\) and \(\Phi_{off}\) are given by (22); the dependence on \(n\) is due to conditioning on the fading \(\Gamma\) matrix.

C. Model Comparison

Figure 3 shows the evolution of the metric \(\lambda(Z_{nq})\) defined in (9) under both Webb-Gaussian and jointly-Gaussian models for several \(T_e\) values and \(N = 1\). Under both formulations, the resulting \(\lambda(Z_{nq})\) metrics would generally be well-approximated by second-order polynomials. A linear fit to \(\lambda(Z_{nq})\) curves is also plausible, particularly for higher \(T_e\) values. If we make the affine approximation \(\lambda(Z_{nq}) \approx \rho_n Z_{nq} + \beta\), the ML rule of (9) reduces to the suboptimal linear decision rule

$$\hat{q} = \arg \max_i \sum_{n=1}^{N} \rho_n Z_{ni}$$  \hspace{1cm} (24)

which may also be derived through (20) by setting \(a_n = 0\). An equal gain combining (EGC) rule that does not have channel gain knowledge may also be obtained either from (21) by setting \(a_n = 0\) and \(b_n = 1\), or
from (24) by setting $\rho_n = 1$. However, Figure 3 reveals a subtle difference between the two models: the $\lambda(Z_{nq})$ curves are always monotone-increasing [24] under the Webb-Gaussian formulation, while this is not the case under jointly-Gaussian modelling. The monotonicity of Webb-Gaussian $\lambda(Z_{nq})$ curves proves that a linear decision rule is in fact optimal for the $N = 1$ case and binary PPM. Therefore, since the exact p.d.f.s for the detector output $Z_{nq}$ are strictly speaking non-Gaussian, the quadratic decision rule in (17) and the non-monotonic nature of the resulting $\lambda(Z_{nq})$ curves are in fact artifacts of the jointly-Gaussian assumption; the Gaussian model places too much probability density on the negatively-valued outcomes. For $N = 1$ and $Q = 2$, for example, (17) divides the decision plane by two linear boundaries that intersect at a right angle. The Webb-Gaussian model, however, would give no such intersection.

Nonetheless, we consider both models in subsequent analysis for two important reasons. First, the jointly-Gaussian model for $Z_{nq}$ has clear analytical advantages over the Webb-Gaussian formulation; Chernoff bounding with Webb-Gaussian statistics, for example, does not yield a closed-form expression for $P_e|\Gamma,i$, requiring numerical methods for calculating the bound for $P_e$. Secondly, as we will go on to show, the discrepancies between the Webb-Gaussian and jointly-Gaussian decision metrics are small and decrease with receiver array size, $N$, making jointly-Gaussian modelling of APD receiver statistics particularly well-suited to FSO MIMO systems.

IV. RESULTS AND ANALYSIS

We present analytical bounding and simulation of $P_e$ for FSO MIMO systems, emphasizing the role of $M$, $N$ and APD gain. Non-fading as well as log-normal and Rayleigh fading channels are treated. Unless otherwise noted, the physical parameters are as listed in Table 1, corresponding to a wavelength of 1.55 $\mu$m and a data rate of 2.5 Gb/s. It is assumed here that lasers operate on a peak-power constraint so the emphasis is on the $Q = 2$ case. Under an average-energy constraint $Q > 2$ becomes preferred.

A. MIMO Diversity Performance

Figure 4 shows the Chernoff bound on $P_e$ versus the optical energy parameter $PT_b$ in a non-fading channel under both Webb-Gaussian and jointly-Gaussian statistics, with each model using its respective optimal decision rule. The bounds were evaluated numerically using (11) – (13) for Webb-Gaussian and (23) for jointly-Gaussian statistics. $N = 1$, 2, and 4 APD receivers are considered. (The reader may note that increasing the number of transmitters offers no benefit in the non-fading case since the laser array produces the same total optical power irrespective of $M$ and all path gains $\gamma_{nm} = 1$ in (1).) The $P_e$ curves for analogous p-i-n-based receiver arrays are also shown to highlight the APD advantage. In Figure 4, the $PT_b$ improvement due to APD over p-i-n receivers at $P_e = 10^{-12}$ is over 10 dB for all $N$ considered since APDs are able to move the receivers out of the thermal-noise-limited regime toward
the shot-noise-limited case. Chernoff bounds indicate that the agreement between the Webb-Gaussian and jointly-Gaussian models for APD receiver statistics is best for high $P_e$, with the jointly-Gaussian model overestimating $P_e$ at lower values. For $N = 1$, for example, the jointly-Gaussian model overestimates the required $PT_b$ for $P_e = 10^{-12}$ by about 1.5 dB. However, Figure 4 reveals that this discrepancy significantly decreases as $N$ increases. At $N = 4$, the difference in $P_e$ between the two models is negligible throughout. This improvement should be attributed to an averaging effect. Since the decision metrics $\lambda(Z_{ni})$ built from the Webb-Gaussian and Gaussian models respectively are similar nonlinear functions of the same random variables, their p.d.fs are inherently similar. Summing $N$ such metrics to form $\Lambda_i$, as in (9), enhances this similarity due to the convolution of underlying densities. Thus, the true error probabilities for the two rules approach each other as $N$ increases; this is reflected by their Chernoff bounds. The accuracy of modelling APD receiver statistics as jointly-Gaussian increases with detector-array size and is therefore particularly appropriate for FSO systems with receiver diversity. We exploit this fact and the analytical convenience of the jointly-Gaussian model to perform the subsequent numerical analysis of $P_e$.

Figure 4 also illustrates that for APD receivers, the required $PT_b$ for $P_e = 10^{-12}$ is reduced by about 2.5 dB by increasing $N = 1$ to $N = 2$; the same is true for the increase from $N = 2$ to $N = 4$. Most of this improvement is due to the increase in effective receiver aperture. For strictly quantum-limited APD receivers, the effective SNR is proportional to $PT_b$ rather than $(PT_b)^2$ as in the thermal noise limit, and doubling $N$ would yield a 3 dB improvement in $PT_b$. However the APD receivers here are not fully quantum-noise-limited under the optimal $G$, which explains the 2.5 dB gain in $PT_b$ for each doubling of $N$. P-i-n detectors, on the other hand, are decidedly thermally-noise-limited, and doubling $N$ saves only 1.5 dB in $PT_b$. It it also noted that fixing the total receiver aperture area as $N$ is increased would actually degrade the error performance in the thermal-noise limit.

Figure 5 considers the error performance of 1x1, 2x2 and 4x4 MIMO systems in Rayleigh amplitude fading. (It is noted that the results presented here are for uncoded systems; sophisticated error-control codes would provide additional coding gain.) Given the jointly-Gaussian statistics, the quadratic detector (17) that employs known channel gains is optimal. Likewise shown in Figure 3 is the performance of a linear combining rule (24) that employs known channel gains, and that of an equal-gain combining (EGC) rule that does not have channel gain knowledge. (The EGC rule is readily obtained from (24) by setting $\rho_n = 1$.) The improvement in $PT_b$ for the 2x2 system over its non-diverse counterpart at $P_e = 10^{-6}$ is about 30 dB (by extrapolation.) The analogous improvement for the 4x4 setup over the 2x2 system is about 10 dB. The slope in log-log coordinates of the $P_e$ curves at high energy levels is $-MN$, indicating full diversity is obtained despite the simple coding at the transmitter. (In terms of diversity, 1x$M$ and $N$x1 systems achieve diversities $M$, $N$ respectively.) Finally, since the marginal improvement in $PT_b$ decreases
as $M$ and $N$ increase, moderately-sized transmitter and receiver arrays would be the most judicious design choice. It is also noteworthy from Figure 5 that the three decision rules perform equally for $M = N = 1$, within limits of simulation error, and are essentially indistinguishable for large MIMO designs. Therefore, although the quadratic detector of (17) is formally optimal due to jointly-Gaussian data modelling, the linear rule that forms a different partition of observation space is practically equivalent because negligible probability density resides in the decision zones where these two rules differ. This will in fact remain true for all parameter selections of interest for well-designed FSO MIMO systems. The strong performance of the equal-gain combiner is of even greater practical importance as it avoids the need for channel gain measurement or estimation.

Thus far, we have considered $P_e$ performance of non-fading and severely turbulent (Rayleigh) channels. Lognormal fading is the most commonly-accepted model for atmospheric fading due to scintillation, and Figure 6 shows error performance of 1x1 and 2x2 systems in a lognormal channel with $\Psi = 0.15$ and $\Psi = 0.7$. We observe that the MIMO gains are smaller here, since the channel model presents smaller probability of deep fades than the Rayleigh case, yet they remain significant for a moderately strong scintillation index. For $\Psi = 0.7$, for example, the $PT_b$ gain of the 2x2 setup is about 8 dB at $P_e = 10^{-6}$.

### B. Optimal APD Gain Selection

In the remainder of the section we focus on the choice of APD gain $G$ that optimizes system performance. Intuitively, increasing $G$ moves receiver operation away from the thermal-noise limit, yet excessive $G$ eventually leads to excess shot noise dominated performance through the function $F(G)$ (16).

Figure 7 plots average $P_e$ versus mean APD gain, $G$, for various FSO MIMO systems in a lognormal fading channel with $\Psi = 0.25$ and $PT_b = -160$ dBJ. We note that receiver diversity alone improves $P_e$ significantly more than does transmitter diversity alone. At $G = 30$, for example, $P_e \simeq 10^{-4}$ in the $M = 1, N = 2$ system, while $P_e \simeq 10^{-2}$ for $M = N = 1$, an improvement of two orders of magnitude. The analogous reduction in $P_e$ in the $M = 2, N = 1$ system over the non-diverse case is significantly lower. This occurs because the total transmitted power is fixed independently of $M$, whereas the received power increases with $N$. The diversity gains would in fact be equivalent under a receiver power constraint. (A similar effect holds for RF MIMO systems, although for slightly different reasons — namely, contributions from multiple RF transmitters do not add coherently at a single receiver, whereas fields from a single transmitter are combined coherently across a receiver array. Thus, larger receiver arrays are preferred to large transmitter arrays when feasible, although both would achieve diversity order $MN$ with proper space-time coding.) The optimal APD gain in Figure 7, in terms of yielding the lowest $P_e$, remains approximately constant as the diversity order of the system increases. For the cases considered here at $\Psi$
= 0.25, the optimal $G$ remains between 30 and 35 throughout. We may thus select the optimal $G$ for the MIMO system by selecting the optimal $G$ of its non-diverse counterpart ($M = N = 1$).

Since turbulence in FSO channels can be highly variable, we also study the impact of $G$ on error probability as fading severity changes. The optimal $G$ depends on received optical power and since the latter is highly variable in presence of atmospheric turbulence, it suggests that adaptive APD gain optimization might be necessary. This, however, might be prohibitively complex to implement. Figure 8 shows average $P_e$ versus mean APD gain for a $M = N = 2$ system in both lognormal and Rayleigh fading. We observe that the optimal gain remains largely unchanged at $G = 30$ over the range of conditions tested and that it is very close to $G = 27$ chosen in Table 1 since it maximizes the non-fading electrical SNR of a single APD receiver for the underlying parameter set. From this we conclude that $G$ can be essentially optimized by maximizing SNR on a non-fading MIMO channel as done in [25]. This finding also extends conclusions for a single APD receiver in [16].

Figure 9 illustrates that as thermal noise varies, the optimal $G$ changes significantly. For small $T_e$, thermal noise is diminishing, indicating that large APD gain is not warranted; APD excess noise dominates even for small $G$. As effective input noise temperature $T_e$ increases from 70 K to 1000 K, the optimal gain increases from $G = 15$ to $G = 50$. Since this range of APD gains is relatively large and it is quite difficult to adaptively alter $G$ in practical systems, the most effective design approach is selecting the optimal non-fading APD gain for the underlying set of physical parameters and taking appropriate measures to keep $T_e$ relatively constant.

V. CONCLUSIONS

The performance of FSO MIMO communication systems in turbulent atmospheric channels with $Q$-ary PPM signalling and APDs in the receiver array has been analyzed. We established the general ML detection rule and derived the Chernoff bound on error probability for the MIMO setup, which was applied it to both the Webb-Gaussian and jointly-Gaussian models for receiver outputs. Subtleties of the jointly-Gaussian model versus the Webb-Gaussian formulation have been highlighted. Analytical results indicate that the accuracy of modelling APD receiver statistics as jointly-Gaussian increases with detector-array size and is thus particularly appropriate for FSO systems with receiver diversity. Numerical results for error probability have been presented for lognormal and Rayleigh fading channels in binary PPM, indicating that MIMO systems with APD receivers are superior to analogous non-diverse systems in all fading environments. Full spatial diversity is achieved in a Rayleigh channel with independent path gains. The performance of a linear equal-gain-combiner is shown to be virtually identical to that of the ML detector even in severe turbulence. Finally, it is shown that selecting the APD gain that maximizes the
non-fading electrical SNR of a single APD receiver results in nearly optimal $P_e$ in FSO MIMO systems for arbitrary levels of atmospheric turbulence and a constant thermal noise level.

ACKNOWLEDGMENT

The authors would like to thank Dr. Ting Wang of NEC Laboratories America for technical support on this research.

REFERENCES


Fig. 1. FSO MIMO setup with APD receivers in atmospheric turbulence.

Fig. 2. Detailed view of APD-based receiver.
TABLE I

PHYSICAL PARAMETER SET.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>0.75</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.55 $\mu$m</td>
</tr>
<tr>
<td>$P_0$</td>
<td>-170 dB</td>
</tr>
<tr>
<td>$G$</td>
<td>27</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.4 (InGaAs APD)</td>
</tr>
<tr>
<td>$T_e$</td>
<td>400 K</td>
</tr>
<tr>
<td>$R$</td>
<td>100 $\Omega$</td>
</tr>
<tr>
<td>$T_b$</td>
<td>0.4 nsec</td>
</tr>
<tr>
<td>$Q$</td>
<td>2</td>
</tr>
</tbody>
</table>

Fig. 3. Webb-Gaussian and jointly-Gaussian metrics for several values of effective input noise temperature, $T_e$ (K); $PT_b = -155$ dB

Fig. 4. Chernoff bound on $P_e$ for 1x1, 2x2 and 4x4 non-fading channels (left); $P_e$ for equivalent p-i-n receivers (right)
Fig. 5. Average error probability for 1x1, 2x2 and 4x4 FSO MIMO systems with APD receivers in Rayleigh fading

Fig. 6. Average error probability for 1x1 and 2x2 FSO MIMO systems with APD receivers in lognormal fading
Fig. 7. Average error probability versus mean APD gain for various MIMO systems; EGC, lognormal channel, $\Psi = 0.25, PT_b = -160$ dB

Fig. 8. Average error probability versus mean APD gain for $M = N = 2$ system under several $\Psi$ values; EGC, lognormal channel, $T_e = 400$ K, $PT_b = -160$ dB
Fig. 9. Average error probability versus mean APD gain for $M = N = 2$ system under several $T_e$ values; EGC, lognormal channel, $\Psi = 0.25$, $PT_b = -160$ dBm