Tradeoff between Diversity and Multiplexing Gains in Block Fading Optical Wireless Channels

Sufang Yang, Longguang Li[†], Haoyue Tang, Jintao Wang

Abstract

The diversity-multiplexing tradeoff (DMT) provides a fundamental performance metric for different multiple-input multiple-output (MIMO) schemes in wireless communications. In this paper, we explore the block fading optical wireless communication (OWC) channels and characterize the DMT in the presence of both optical peak- and average-power constraints. Three different fading distributions are considered, which reflect different channel conditions. In each channel condition, we obtain the optimal DMT when the block length is sufficiently large, and we also derive the lower and upper bounds of the DMT curve when the block length is small. These results are dramatically different from the existing DMT results in radio-frequency (RF) channels. These differences may be due to the fact that the optical input signal is real and bounded, while its RF counterpart is usually complex and unbounded.

Index terms — Peak- and average-power constraints, outage probability, average error probability, optical wireless communication, diversity-multiplexing tradeoff.

1 Introduction

As an important complement to conventional RF communication, OWC [1], [2] significantly improves the rate performance and offers an ideal solution to the spectrum scarcity in existing wireless communication systems. Recently it has been considered as a promising technique in future 6G [3]–[5]. Most current OWC systems adopt the socalled intensity-modulation and direct-detection (IM-DD) transmission scheme because of its simplicity and low-cost deployment. In such a scheme, the transmitter modulates the intensity of optical signals coming from light-emitting diodes (LEDs), and the receiver measures incoming optical intensities by photodetectors [6], [7]. As a consequence, transmit signals are proportional to optical intensities, and hence are real and nonnegative, fundamentally different from their RF counterparts. Furthermore, considering safety reasons and hardware limitations, the peak and average optical powers of transmitting signals typically have to be restricted.

The OWC channels are more sensitive to environment fluctuations than the traditional RF channels due to the IM-DD transmission scheme. Particularly in the mediumand long-range OWC, the fluctuations caused by atmospheric turbulence degrade the quality of the communication links severely. They change the temperature and pressure of the atmosphere, and eventually lead to the refractive index variations of the path links [8]–[10]. To combat the channel fading induced by atmospheric fluctuations in the OWC

This work is supported by the National Natural Science Foundation of China under Grant No. 62101192 and Shanghai Sailing Program under Grant No. 21YF1411000. (Corresponding author: L. Li)

Sufang Yang, Haoyue Tang, and Jintao Wang are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: ysf20@tsinghua.org.cn, thy17@mails.tsinghua.edu.cn, wangjintao@tsinghua.edu.cn).

Longguang Li is with Department of Communication and Electronic Engineering, East China Normal University, Shanghai 200241, China (e-mail: lgli@cee.ecnu.edu.cn).

systems, the widely adopted MIMO technique plays a pivotal role [11]–[14]. It provides many independent transmission paths available at the transmitting or receiving ends to effectively alleviate the atmospheric turbulence path loss [15]. Another benefit provided by utilizing MIMO techniques is the significant improvement on spectral efficiency when compared with single-antenna systems [16].

In this paper, we investigate the OWC-MIMO system from a DMT perspective [17]–[19]. The optimal DMT characterizes the maximal achievable diversity gain at a fixed multiplexing gain, which provides a fundamental metric for comparison between different MIMO transmission schemes, and has triggered extensive research, such as deriving the optimal DMT curves in different channel models [20]–[22], designing various optimal DMT-approaching coding schemes [23]–[25].

Most of the existing DMT results focus on the traditional RF systems [17]–[22], among which Zheng and Tse presented the classic optimal tradeoff curve for Rayleigh fading channels [17]. They showed that when the channel block length satisfies $l \ge l_{\rm th}^{\rm ZT} = n_{\rm T} + n_{\rm R} - 1$, the optimal DMT curve can be exactly characterized as

$$d_{\rm ZT}^{\star}(r) = (n_{\rm T} - r)(n_{\rm R} - r),$$
 (1)

where $d_{\text{ZT}}^{\star}(r)$ represents the maximum diversity gain achieved at the given multiplexing gain r, and n_{T} and n_{R} represent the number of transmit and receive antennas, respectively. Moreover, they proposed the bounds on DMT curve at the small block length regime $(l < l_{\text{th}}^{\text{ZT}})$, if $r \leq r_1$, with $r_1 = n_{\text{R}} - \lceil \frac{l - |n_{\text{T}} - n_{\text{R}}| - 1}{2} \rceil$, the maximum diversity gain is bounded by

$$d_{\rm ZT}^{\star}(r) \le (n_{\rm T} - r)(n_{\rm R} - r),$$
 (2)

$$d_{\rm ZT}^{\star}(r) \ge -l(r-r_1) + (n_{\rm T} - r_1)(n_{\rm R} - r_1), \tag{3}$$

otherwise the optimal DMT curve agrees with (1). This work established the theoretical framework for the DMT characterizations, and many similar results in other channels, such as Rician, Nakagami and log-normal channels are also obtained afterward [19], [21], [26], [27].

Although there has been extensive research on the DMT in RF systems, few studies tap into characterizing the optimal DMT for the OWC systems. Some recent results have shown that many derivations in RF systems are not applicable for OWC systems due to the unique IM-DD transmission scheme [27]–[29]. The existing work closely related to ours was done by Jaiswal and Bhatnagar [28]. They considered the real-valued requirements for the input, and then proposed that if the channel block length satisfies $l \geq l_{\rm th}^{\rm JB} = n_{\rm T} + n_{\rm R} - 1$, the optimal DMT curve for negative exponential channel is given by

$$d_{\rm JB}^{\star}(r) = \frac{1}{2} d_{\rm ZT}^{\star}(r).$$
 (4)

Also, under a small block length, they found analogous DMT bounds as in (2), (3).

The one-half factor in (4) reflects the loss of half of the degree of freedom compared with RF channels, which is due to the fact that optical inputs need to be real-valued. However, as mentioned before, inputs in OWC channels represent optical intensities, and hence their values must also be nonnegative. In fact, it is the nonnegativity of the optical inputs that significantly complicates the analysis of performance limits in OWC channels. Hence, directly applying half of the traditional RF MIMO capacity formula to calculate diversity gain in OWC channels cannot be theoretically justified. Also, the nonnegativity of optical inputs implies that the commonly used two-sided Gaussian random codes in RF channels are no longer admissible in OWC channels. Furthermore, due to the optical intensity inputs, the power constraints imposed on the inputs need to be described differently. All the above issues indicate the existing DMT results in RF channels cannot be directly extended here. Hence, how to characterize the optimal DMT in practical OWC channels still remains an open problem. In this paper, we first investigate the optimal DMT by fully considering practical optical input constraints in three different fading channels. We first restrict the optical inputs to be real-valued and nonnegative. Also, a peak- and an average-power constraints are imposed on the inputs. Then we use negative exponential, gamma-gamma, and log-normal distributions to model the channel with atmospheric turbulence fluctuation from strong to weak intensities. For each channel condition, we establish optimal DMT in different block length regimes. It turns out that our derived results are fundamentally different from the above existing results, which may more precisely reflect the fundamental limits of practical OWC systems.

Specifically, the main contributions in this paper are as follows.

- Bounds on Instantaneous Capacity: By using a truncated exponential random coding argument and further applying the generalized entropy power inequality (GEPI) [30], we first derive a lower bound on instantaneous capacity. Then, an upper bound is established by some algebraic manipulations on the asymptotic capacity in [31, Theorem 21]. These bounds are closed-form and proved to be optimal in terms of outage diversity gain.
- Exact Characterization of Outage Diversity Gain: With the above new instantaneous capacity bounds, we establish lower and upper bounds on outage diversity gain in general fading channels. Applying these bounds into our considered channels, we precisely characterize the outage diversity gain.
- Error Probability Bounds on Truncated Exponential Random Coding: We propose a truncated exponential random coding scheme, which helps to derive a new upper bound on average error probability. Based on this bound, we derive a tight lower bound on the optimal diversity gain in different fading channels.
- Characterization of Optimal DMT: We characterize the optimal DMT curves in different block length regimes for considered channels. Specifically, if the block length $l \ge n_{\rm T} n_{\rm R} + 1$, the optimal diversity gain $d^*(r)$ can be characterized as

$$d^{\star}(r) = (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r).$$
(5)

Otherwise under a small block length $(l < n_{\rm T} - n_{\rm R} + 1)$, the optimal diversity gain can be upper- and lower-bounded by:

$$d^{\star}(r) \le (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r),$$
 (6)

$$d^{\star}(r) \ge l(n_{\rm R} - r). \tag{7}$$

See Theorems 7, 10, and 12 in detail.

The paper is organized as follows. We end the introduction with a few notational conventions. Section 2 describes in detail the investigated channel model. In Section 3, we present new upper and lower bounds on outage diversity gain. Section 4 characterizes the average error probability with the truncated exponential random coding. The optimal DMTs of negative exponential, gamma-gamma, and log-normal channels are characterized in Sections 5, 6 and 7. Numerical examples are included in Section 8. Most of the proofs are in the appendices.

Notation: Random variables and matrices are boldfaced, e.g., h and \mathbf{X} , while their realizations are typeset in h and \mathbf{X} , respectively. \mathbf{X}_j denotes the jth column of matrix \mathbf{X} , and \mathbf{x}_{ij} or $[\mathbf{X}]_{ij}$ denotes the *i*th row and *j*th entry of \mathbf{X} . Sets are typeset in a special font, e.g., \mathcal{A} . Differential Entropy is denoted by $h(\cdot)$, and mutual information by $I(\cdot; \cdot)$. $\|\cdot\|_1$ and $\|\cdot\|_{\mathsf{F}}$ denote the \mathcal{L}_1 - and Frobenius-norm, respectively. $\log(\cdot)$ denotes the logarithm to the base of e. The expectation of a random variable is denoted by $\mathsf{E}[\cdot]$, and variance by $\mathsf{Var}[\cdot]$. $\mathcal{R}^{m \times n}$ ($\mathcal{R}^{m \times n}_+$) denotes real (nonnegative) valued set. We denote $x^+ \triangleq \max\{0, x\}$, and use symbol \doteq to denote *exponential equality*, i.e., $g(x) \doteq x^b$ indicates $\lim_{x\to\infty} \frac{\log g(x)}{\log x} = b$, and \geq, \leq are similarly defined.

2 Channel Model

Consider a MIMO channel with $n_{\rm T}$ LEDs and $n_{\rm R}$ photodetectors. The channel output is given by¹

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z},\tag{8}$$

where $\mathbf{X} \in \mathcal{R}^{n_{\mathrm{T}} \times l}_{+}$ denotes the channel input, with l being the block length; where $\mathbf{Z} \in \mathcal{R}^{n_{\mathrm{R}} \times l}$ denotes the channel noise, whose entries are independent and identically distributed (i.i.d.) inside and across blocks; and where $\mathbf{H} \in \mathcal{R}^{n_{\mathrm{R}} \times n_{\mathrm{T}}}_{+}$ denotes the channel matrix, and its entries remain constant inside one block, and i.i.d. across blocks.

The entry h_{ij} in **H** represents the nonnegative and real-valued gain from *j*th transmit antenna to *i*th receive antenna, and it depends on two factors: deterministic distance attenuation and random atmospheric turbulence loss. Hence, h_{ij} can be formulated as

$$\boldsymbol{h}_{ij} = h_{ij}^d \boldsymbol{h}_{ij}^r, \tag{9}$$

where $h_{ij}^d = e^{-\nu d_{ij}}$ represents the deterministic part with parameter ν characterizing the transmission environment and d_{ij} being the transmission distance; and where h_{ij}^r represents the random part characterizing the atmospheric turbulence intensity. In this paper, we consider three different distributions of the random atmospheric turbulence h_{ij}^r , which cover the turbulence fluctuation regimes from strong to weak intensities, and they are

• Negative exponential distribution:

$$f(h_{ij}^r) = e^{-h_{ij}^r}; (10)$$

• Gamma-gamma distribution:

$$f(h_{ij}^r) = \frac{2(\rho_1 \rho_2)^{\frac{\rho_1 + \rho_2}{2}}}{\Gamma(\rho_1)\Gamma(\rho_2)} (h_{ij}^r)^{\frac{\rho_1 + \rho_2}{2} - 1} K_{\rho_1 - \rho_2} (2\sqrt{\rho_1 \rho_2 h_{ij}^r}),$$
(11)

where ρ_1 and ρ_2 denote the irradiance fluctuation parameters with $\rho_2 < \rho_1$; where $\Gamma(\cdot)$ denotes the Gamma function with $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$; and where $K_{\tau}(\cdot)$ denotes the modified Bessel function of the second kind with τ being the order [32], [33];

• Log-normal distribution:

$$f(h_{ij}^r) = \frac{1}{h_{ij}^r \sqrt{2\pi\sigma_l^2}} \exp\left(-\frac{(\log(h_{ij}^r) - \mu_l)^2}{2\sigma_l^2}\right),$$
(12)

where μ_l and σ_l denote the expectation and variance of $\log(\mathbf{h}_{ij}^r)$, respectively.

Considering the limited dynamic working range of the light emitters and practical illumination requirements for the modulated optical sources, both peak- and averagepower constraints are imposed on the channel input, i.e.,

$$P[x_{ij} > A] = 0, \ \forall i \in \{1, \dots, n_{\rm T}\}, \ \forall j \in \{1, \dots, l\},$$
(13a)

$$\frac{1}{l} \mathsf{E} \Big[\sum_{j=1}^{l} \| \mathbf{X}_j \|_1 \Big] \le \mathsf{E}, \tag{13b}$$

where \mathbf{X}_j denotes the *j*th column input vector; where A represents the allowed maximum optical power by each antenna; and where E denotes the total average optical power

¹For simplicity, we assume the photoelectric coefficient for the photodetector is 1.

allowed across all antennas. The ratio between the allowed average power and the allowed peak power is denoted by

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}},\tag{14}$$

where $\alpha \in (0, \frac{n_{\rm T}}{2}]$, and is fixed in the paper.

Since information is carried on the intensity of the optical signal, we adopt the definition of optical signal-to-noise ratio (OSNR) [34]–[36] as follows:

$$\mathsf{OSNR} = \frac{\mathsf{E}}{\sigma_n}.\tag{15}$$

We further present some useful concepts and definitions in terms of OSNR. More details can be seen in [17].

Given a transmission scheme, an outage occurs when the mutual information of this channel can not support the target rate R, and we denote

$$\mathcal{C} \triangleq \{ \mathbf{H} : \mathbf{I}(\mathbf{X}_j; \mathbf{Y}_j | \mathbf{H} = \mathbf{H}) < R \}$$
(16)

as the outage event in terms of **H**.

Among all possible transmission schemes, the outage probability at time index j is defined as [17]

$$P_{out}(\mathsf{OSNR}) \triangleq \min_{f(\mathbf{X}_j) \text{ satisfying (13)}} P[I(\mathbf{X}_j; \mathbf{Y}_j | \mathbf{H}) \le R],$$
(17)

where \mathbf{X}_j and \mathbf{Y}_j denote the transmit and receive vectors at time j, respectively; and where $f(\mathbf{X}_j)$ denotes the input distribution of \mathbf{X}_j . In the existing literature [21], [37], the mutual information term in (17) is also called instantaneous capacity.

Now we briefly define the following diversity and multiplexing gains in terms of OSNR that will be used in the rest of the paper.

Definition 1. A transmission scheme is said to achieve multiplexing gain r, outage diversity gain $d_{out}(r)$, and diversity gain d(r) if the rate R(OSNR) satisfies

$$\lim_{\mathsf{OSNR}\to\infty} \frac{R(\mathsf{OSNR})}{\log(\mathsf{OSNR})} = r,$$
(18)

the outage probability $P_{out}(OSNR)$ satisfies

$$-\lim_{\mathsf{OSNR}\to\infty} \frac{\log \mathcal{P}_{\mathrm{out}}(\mathsf{OSNR})}{\log(\mathsf{OSNR})} = d_{\mathrm{out}}(r), \tag{19}$$

and average error probability $P_e(OSNR)$ satisfies

$$-\lim_{\mathsf{OSNR}\to\infty} \frac{\log \mathcal{P}_{e}(\mathsf{OSNR})}{\log(\mathsf{OSNR})} = d(r).$$
(20)

For each r, we define $d^{\star}(r)$ as the supremum of the diversity gain achieved over all schemes at the data rate R(OSNR).

3 Outage Probability Analysis

This section presents the new outage probability bounds, which are crucial in the following derivations of the optimal DMTs.

To estimate the outage probability in (17), we first need to characterize the instantaneous capacity. In the presence of the peak- and average-power constraints in (13a) and (13b), there are no existing results in current literature applicable here. We present new lower and upper bounds, which are closed-form and sufficiently tight at high OSNR. The lower bound is derived by using a truncated exponential random coding argument. For any time index $j \in \{1, 2, ..., l\}$, denote $\mathbf{X}_j = [\mathbf{X}_{j1}, \mathbf{X}_{j2}, ..., \mathbf{X}_{jn_T}]^{\mathsf{T}}$, and let the entries of \mathbf{X}_j be i.i.d. according to the following truncated exponential distribution:

$$f(\mathbf{X}_{ji}) = \frac{1}{\mathbf{A}} \cdot \frac{\mu}{1 - e^{-\mu}} \cdot e^{-\frac{\mu \mathbf{X}_{ji}}{\mathbf{A}}}, \quad \forall i \in \{1, 2, \dots, n_{\mathrm{T}}\},$$
(21)

where μ is a parameter satisfying

$$\frac{1}{\mu} - \frac{e^{-\mu}}{1 - e^{-\mu}} = \frac{\alpha}{n_{\rm T}}.$$
(22)

The achievable rate by this truncated exponential distribution can serve as a natural lower bound on the instantaneous capacity. Applying the GEPI in [30] we derive the following lower bound, whose rigorous proof is shown in Appendix A.

Proposition 2 (Lower Bound). Given a MIMO-OWC channel in (8), we have

$$I(\mathbf{X}_{j}; \mathbf{Y}_{j} | \mathbf{H}) \geq \frac{1}{2} \log (1 + L_{l}(\mathsf{OSNR})^{2n_{\mathrm{R}}} | \mathbf{H}\mathbf{H}^{\mathsf{T}} |),$$
(23)

where

$$L_{l} = \left(\frac{2\sigma_{n}^{2}}{\pi\alpha^{2}e}\right)^{n_{\rm R}} \left(\frac{1 - e^{-\mu}}{\mu}2^{-\frac{\mu e^{-\mu}}{1 - e^{-\mu}}}\right)^{2n_{\rm R}}$$
(24)

with μ satisfying (22).

The following upper bound is derived by first assuming the channel state information available at the transmitter, and then by some algebraic manipulations on the existing asymptotic capacity in [31, Theorem 21].

Proposition 3 (Upper Bound). Given a MIMO-OWC channel in (8), we have

$$I(\mathbf{X}_{j}; \mathbf{Y}_{j} | \mathbf{H}) \leq \frac{1}{2} \log \left(L_{u}(\mathsf{OSNR})^{2n_{\mathrm{R}}} | \mathbf{H} \mathbf{H}^{\mathsf{T}} | \right),$$
(25)

where

$$L_u = \binom{n_{\rm T}}{n_{\rm R}} \left(\frac{\sigma_n^2}{2\pi\alpha^2 e} \right)^{n_{\rm R}}.$$
 (26)

Remark 4. Note that L_u and L_l in Proposition 2 and 3 are constants independent of parameters OSNR and **H**, which play no role in the following derivations related to the OSNR exponent.

For a given channel realization H, we denote λ_i , $\forall i = \{1, \ldots, n_R\}$, as eigenvalues of HH^T with λ_i arranged in an increasing order. It is obvious they are real and nonnegative. We further express $\lambda_i = (\mathsf{OSNR})^{-a_i}$ for $i \in \{1, \cdots, n_R\}$. For the notational convenience, we denote vector $a \triangleq [a_1, \cdots, a_{n_R}]$ and vector $\lambda \triangleq [\lambda_1, \cdots, \lambda_{n_R}]$.² Decompose H as UDQ, where $U \in \mathcal{R}^{n_R \times n_R}$ and $Q \in \mathcal{R}^{n_R \times n_T}$ are orthogonal matrices,

Decompose H as UDQ, where $U \in \mathcal{R}^{n_{R} \times n_{R}}$ and $Q \in \mathcal{R}^{n_{R} \times n_{T}}$ are orthogonal matrices, and D is a diagonal matrix with $D = \text{diag}[(\mathsf{OSNR})^{-\frac{a_{1}}{2}}, \dots, (\mathsf{OSNR})^{-\frac{a_{n_{R}}}{2}}]$. More details about the matrix decomposition can be found in Appendix B.1.

For a given distribution of **H**, by [38] we prove in Appendix B.1 that the distribution of vector \boldsymbol{a} can be estimated as

$$f(a) \doteq (\mathsf{OSNR})^{-\sum_{i=1}^{n_{\mathrm{R}}} \frac{n_{\mathrm{T}} - n_{\mathrm{R}} + 2i - 1}{2} a_i} \times \iint f(\mathrm{UDQ}) d\mathrm{Q} d\mathrm{U},$$
(27)

with f(UDQ) being the distribution of **H**.

Now we are ready to present the main theorem with proof in Appendix B.2.

²It should be noted that the random forms corresponding to a and λ are denoted as a and λ .

Theorem 5 (Outage Probability Bounds). Given a channel in (8), we have

$$\int_{\mathcal{B}} f(a)da \stackrel{.}{\leqslant} \mathcal{P}_{\text{out}}(\mathsf{OSNR}) \stackrel{.}{\leqslant} \int_{\mathcal{A}} f(a)da, \tag{28}$$

where set \mathcal{A} is defined as

$$\mathcal{A} = \left\{ a : (2n_{\rm R} - \sum_{i=1}^{n_{\rm R}} a_i)^+ \le 2r \right\},\tag{29}$$

and set ${\mathcal B}$ as

$$\mathcal{B} = \left\{ a : 2n_{\mathrm{R}} - \sum_{i=1}^{n_{\mathrm{R}}} a_i \le 2r \right\};\tag{30}$$

where f(a) is defined as in (27).

4 Random Coding Error Analysis

This section presents the diversity order of truncated exponential random coding, which serves as a lower bound on the optimal diversity order.

Consider a random code with codewords i.i.d. as in (21), and the code rate is $R = r \log(\text{OSNR})$. The decoder applies the maximum-likelihood (ML) method to detect the sent message. Now consider a channel realization $\mathbf{H} = \mathbf{H}$, then the conditional error probability can be upper bounded by the following inequality, whose proof is postponed to Appendix C.

$$P(\text{error}|\mathbf{H} = \mathbf{H}) \le (\mathsf{OSNR})^{lr} \prod_{i=1}^{n_{\mathrm{R}}} \left(1 + \frac{g(\alpha, n_{\mathrm{T}})}{2} (\mathsf{OSNR})^2 \lambda_i \right)^{-\frac{l}{2}},$$
(31)

where

$$g(\alpha, n_{\rm T}) = \frac{1}{\mu^2} - \frac{e^{-\mu}}{(1 - e^{-\mu})^2}$$
(32)

with μ satisfying (22).

Now we present the main theorem on the error probability $\mathrm{P}_{\mathrm{e}}(\mathsf{OSNR})$ of this random coding scheme.

Theorem 6. Given a channel in (8), we have

$$P_{e}(\mathsf{OSNR}) \stackrel{<}{\leqslant} (\mathsf{OSNR})^{-d_{\mathrm{out}}(r)} + (\mathsf{OSNR})^{-d_{\mathrm{te}}(r)}, \tag{33}$$

where

$$(\mathsf{OSNR})^{-d_{\mathsf{te}}(r)} \doteq \int_{\mathcal{C}^c} f(a) (\mathsf{OSNR})^{-\frac{l}{2}(\sum_{i=1}^{n_{\mathsf{R}}} (2-a_i)^+ - 2r)} da, \tag{34}$$

with \mathcal{C}^c denoting the complement of outage set \mathcal{C} .

Proof: We consider the error event conditioned on the channel outage event. Specifically, we bound the error probability by

$$P_{e}(\mathsf{OSNR}) = P_{out}(\mathsf{OSNR})P_{e|out}(\mathsf{OSNR}) + P_{e,nout}(\mathsf{OSNR})$$
(35)

$$\leq P_{out}(OSNR) + P_{e,nout}(OSNR).$$
 (36)

where $P_{e|out}(OSNR)$ denotes the error probability conditioned on outage event; where $P_{e,nout}(OSNR)$ denotes the joint probability of error and no outage events.

By the definition of outage diversity, the first term in the RHS of (36) is

$$P_{out}(\mathsf{OSNR}) \doteq (\mathsf{OSNR})^{-d_{out}(r)}.$$
(37)

Substitute $\lambda_i = (\mathsf{OSNR})^{-a_i}$ into (31), we have

$$P(\text{error}|a) \stackrel{!}{\leqslant} (\mathsf{OSNR})^{-\frac{l}{2}(\sum_{i=1}^{n_{\mathrm{R}}} (2-a_i)^+ - 2r)}.$$
(38)

Multiplying with the distribution of a, and integrating over set C^c , the second term in the RHS of (36) is

$$P_{e,nout}(\mathsf{OSNR}) \doteq (\mathsf{OSNR})^{-d_{te}(r)}.$$
(39)

The proof is concluded.

5 DMT Analysis on Negative Exponential Channel

In this section we consider the OWC channel with atmospheric turbulence according to negative exponential distribution. This distribution is commonly used to model the channel with relatively high atmospheric turbulence intensity. We first present results on the outage diversity order, and based on this, we further characterize the optimal DMT.

5.1 Outage Diversity Gain Characterization

To distinguish parameters from different fading channels, here we denote $l_{\rm th}^{\rm NE}$, $d_{\rm out}^{\rm NE}(r)$, $d_{\rm te}^{\rm NE}(r)$ and $d_{\rm NE}^{\star}(r)$ as the parameters for negative exponential channel. Similar notations will also be used in the following gamma-gamma and log-normal channels.

As shown in Theorem 5, to estimate the outage diversity order, we first need to calculate the distribution of vector \boldsymbol{a} in (27). To do this, we first analyze the distribution of \mathbf{H} . Since the entries in \mathbf{H} are assumed i.i.d., and by (10) we expand the distribution of \mathbf{H} as

$$f(\mathbf{H}) = \prod_{i=1}^{n_{\mathrm{R}}} \prod_{j=1}^{n_{\mathrm{T}}} f(h_{ij})$$
(40)

$$= \prod_{i=1}^{n_{\rm R}} \prod_{j=1}^{n_{\rm T}} \exp\left(-e^{\nu d_{ij}} h_{ij} + \nu d_{ij}\right).$$
(41)

Now we decompose H = UDQ, and then for any entry $h_{ij} = [H]_{ij}$, we have

$$h_{ij} = \sum_{k=1}^{n_{\rm R}} u_{ik} \times (\mathsf{OSNR})^{-\frac{a_k}{2}} \times q_{kj}, \tag{42}$$

where $u_{ik} = [U]_{ik}$ and $q_{ik} = [Q]_{ik}$. Since a_i s are arranged in a decreasing order, at high OSNR, the last term $u_{in_{\rm R}} \times (\text{OSNR})^{-\frac{a_{n_{\rm R}}}{2}} \times q_{n_{\rm R}j}$ dominates other terms. By the fact that the entries of H are nonnegative, we have $u_{in_{\rm R}}q_{n_{\rm R}j} \ge 0$. Hence f(H) can be further simplified as

$$f(\mathbf{H}) \doteq \prod_{i=1}^{n_{\mathrm{R}}} \prod_{j=1}^{n_{\mathrm{T}}} \exp\left(-(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}\right).$$
 (43)

Substituting (43) into (27), we get

$$f(a) \doteq (\mathsf{OSNR})^{-\sum_{i=1}^{n_{\mathrm{R}}} \frac{n_{\mathrm{T}} - n_{\mathrm{R}} + 2i - 1}{2}a_i} \times \prod_{i=1}^{n_{\mathrm{R}}} \prod_{j=1}^{n_{\mathrm{T}}} \exp\left(-(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}\right).$$
(44)

Yang, Li, Tang, Wang, 10 December 2021

Now we apply Theorem 5 to bound the outage probability. Note that when $a_{n_{\rm R}} < 0$, in (44) the exponential term $\exp\left(-(\text{OSNR})^{-\frac{a_{n_{\rm R}}}{2}}\right) \to 0$, and hence $f(a) \to 0$. With this observation, we can further reduce the integral domain \mathcal{A} in (28) to be $\mathcal{A}' = \mathcal{A} \cap$ $\{a: a_{n_{\rm R}} \ge 0\}$ and \mathcal{B} in (28) to be $\mathcal{B}' = \mathcal{B} \cap \{a: a_{n_{\rm R}} \ge 0\}$. Furthermore, when $a_{n_{\rm R}} \ge 0$, $\exp\left(-(\text{OSNR})^{-\frac{a_{n_{\rm R}}}{2}}\right) \to 1$ or e^{-1} , and then we can further simplify f(a) in (44) as

$$f(a) \doteq (\mathsf{OSNR})^{-\sum_{i=1}^{n_{\mathrm{R}}} \frac{n_{\mathrm{T}} - n_{\mathrm{R}} + 2i - 1}{2}a_i}.$$
(45)

Now by definition (17), substituting (45) into (28), the outage diversity gain $d_{\text{out}}^{\text{NE}}(r)$ at multiplexing order r is bounded by

$$\inf_{a \in \mathcal{A}'} d_{\text{out}}^{\text{NE}}(r, a) \le d_{\text{out}}^{\text{NE}}(r) \le \inf_{a \in \mathcal{B}'} d_{\text{out}}^{\text{NE}}(r, a),$$
(46)

where function $d_{\text{out}}^{\text{NE}}(r, a)$ is defined as

$$d_{\rm out}^{\rm NE}(r,a) = \sum_{i=1}^{n_{\rm R}} \frac{n_{\rm T} - n_{\rm R} + 2i - 1}{2} a_i;$$
(47)

and where \mathcal{A}' is given by

$$\mathcal{A}' = \left\{ a : a_1 \ge \dots \ge a_{n_{\mathrm{R}}} \ge 0, \text{ and } (2n_{\mathrm{R}} - \sum_{i=1}^{n_{\mathrm{R}}} a_i)^+ \le 2r \right\},\tag{48}$$

and \mathcal{B}' by

$$\mathcal{B}' = \left\{ a : a_1 \ge \dots \ge a_{n_{\mathrm{R}}} \ge 0, \text{ and } 2n_{\mathrm{R}} - \sum_{i=1}^{n_{\mathrm{R}}} a_i \le 2r \right\}.$$
(49)

5.2 Optimal DMT Characterization

We now characterize the optimal DMT of negative exponential channel. In fact, our derived outage probability can serve as a lower bound on average error probability. By Fano's inequality [39], we can show

$$P_{e}(\mathsf{OSNR}) \stackrel{i}{\geq} (\mathsf{OSNR})^{-d_{out}(r)}.$$
(50)

Furthermore, note that the error probability bound on truncated exponential coding established in Section 4 can serve as an upper bound on average error probability. Hence

$$P_{e}(\mathsf{OSNR}) \leq (\mathsf{OSNR})^{-d_{out}(r)} + (\mathsf{OSNR})^{-d_{te}(r)}.$$
(51)

Comparing the bounds in (50) and (51), we have the following observations:

• If $d_{te}(r) \ge d_{out}(r)$, $d_{out}(r)$ becomes the dominant term in the RHS of (51). Combined with (50), the optimal diversity gain $d^{\star}(r)$ is equal to the outage diversity gain, i.e.,

$$d^{\star}(r) = d_{\text{out}}(r). \tag{52}$$

• If $d_{te}(r) < d_{out}(r)$, $d_{te}(r)$ turns to be the dominant term in the RHS of (51). In this case, we can simply bound $d^{\star}(r)$ as

$$d_{\rm te}(r) \le d^{\star}(r) \le d_{\rm out}(r). \tag{53}$$

With the above two observations, we only need to calculate $d_{out}(r)$ and $d_{te}(r)$. The optimal DMT of negative exponential channel is characterized in the following theorem.

Theorem 7 (DMT of Negative Exponential Channel). Given a channel with distribution in (41), if $l \ge l_{\rm th}^{\rm NE} = n_{\rm T} - n_{\rm R} + 1$, the optimal diversity order is given by

$$d_{\rm NE}^{\star}(r) = (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r), \tag{54}$$

otherwise,

$$l(n_{\rm R} - r) \le d_{\rm NE}^{\star}(r) \le (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r).$$
(55)

Proof: We first calculate $d_{\text{out}}^{\text{NE}}(r)$ and $d_{\text{te}}^{\text{NE}}(r)$, and then characterize the optimal diversity order $d_{\text{NE}}^{\star}(r)$.

By linear optimization, it is straightforward to get that the infimums of $d_{\text{out}}^{\text{NE}}(r, a)$ over feasible regimes \mathcal{A}' and \mathcal{B}' are both achieved at $a_{\text{out}}^{\star} = [2(n_{\text{R}} - r), 0, \dots, 0]$ in (46), and the infimums match. Hence the outage diversity gain is given by

$$d_{\rm out}^{\rm NE}(r) = (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r).$$
(56)

Now we calculate $d_{\text{te}}^{\text{NE}}(r)$. By reducing the integral domain \mathcal{C}^c in (34) to \mathcal{A}^c , and enlarging the integral domain \mathcal{C}^c in (34) to \mathcal{B}^c , we bound $d_{\text{te}}^{\text{NE}}(r)$ as

$$\inf_{(\mathcal{B}^c)'} d_{\mathrm{te}}^{\mathrm{NE}}(r,a) \le d_{\mathrm{te}}^{\mathrm{NE}}(r) \le \inf_{(\mathcal{A}^c)'} d_{\mathrm{te}}^{\mathrm{NE}}(r,a),$$
(57)

where function $d_{te}^{NE}(r, a)$ is given by

$$d_{\rm te}^{\rm NE}(r,a) = \sum_{i=1}^{n_{\rm R}} \frac{n_{\rm T} - n_{\rm R} + 2i - 1}{2} a_i + \frac{l}{2} \left(\sum_{i=1}^{n_{\rm R}} (2 - a_i)^+ - 2r \right);$$
(58)

and where $(\mathcal{A}^c)'$ is given by

$$(\mathcal{A}^c)' = \left\{ a : a_1 \ge \dots \ge a_{n_{\mathrm{R}}} \ge 0, \text{ and } (2n_{\mathrm{R}} - \sum_{i=1}^{n_{\mathrm{R}}} a_i)^+ \ge 2r \right\},$$
 (59)

and $(\mathcal{B}^c)'$ by

$$(\mathcal{B}^c)' = \left\{ a : a_1 \ge \dots \ge a_{n_{\mathrm{R}}} \ge 0, \text{ and } 2n_{\mathrm{R}} - \sum_{i=1}^{n_{\mathrm{R}}} a_i \ge 2r \right\}.$$
 (60)

Compared with (47), we can see $d_{te}^{NE}(r, a)$ in (58) is also linear with respect to vector a, but with an extra term $\frac{l}{2} \left(\sum_{i=1}^{n_{\rm R}} (2-a_i)^+ - 2r \right)$. Given a block length l, we can still analyze the infimum of $d_{te}^{NE}(r, a)$ over $(\mathcal{A}^c)'$ or $(\mathcal{B}^c)'$ through linear optimization. Here, we show that the infimums in (57) also match the same optimal point a_{te}^{\star} . In the following, we characterize a_{te}^{\star} under different block length l:

• when $l < n_{\rm T} - n_{\rm R} + 1$, the optimal $a_{\rm te}^{\star}$ is given by

$$a_i^{\star} = 0, \text{ for } i = 1, \cdots, n_{\mathrm{R}};$$
 (61)

• when $n_{\rm T} - n_{\rm R} + 2k - 1 \le l < n_{\rm T} - n_{\rm R} + 2(k+1) - 1$, $\forall k \in \{1, \dots, n_{\rm R} - r\}$, the optimal $a_{\rm te}^{\star}$ is given by

$$a_i^{\star} = 2, \text{ for } i = 1, \cdots, k,$$
 (62a)

$$a_i^{\star} = 0, \text{ for } i = k + 1, \cdots, n_{\rm R};$$
 (62b)

• when $l \ge n_{\rm T} - n_{\rm R} + 2((n_{\rm R} - r) + 1) - 1$, the optimal $a_{\rm te}^{\star}$ is given by

$$a_i^{\star} = 2, \text{ for } i = 1, \cdots, n_{\mathrm{R}} - r,$$
 (63a)

$$a_i^{\star} = 0, \text{ for } i = n_{\rm R} - r + 1, \cdots, n_{\rm R}.$$
 (63b)

Substituting a_{te}^{\star} into (57), we obtain that if $l < n_{T} - n_{R} + 1$, then

$$d_{\rm te}^{\rm NE}(r) = l(n_{\rm R} - r). \tag{64}$$

Otherwise if $l \ge n_{\rm T} - n_{\rm R} + 1$, we have

$$d_{\rm te}^{\rm NE}(r) = \sum_{i=1}^{k} (n_{\rm T} - n_{\rm R} + 2i - 1) + l(n_{\rm R} - k - r), \ k \in \{1, \cdots, n_{\rm R} - r\},$$
(65)

where the choice of k depends on the channel block length l (see (61), (62), and (63)).

Last, we characterize $d_{\rm NE}^{\star}(r)$. If $l \ge n_{\rm T} - n_{\rm R} + 1$ and at any k in set $\{1, \dots, n_{\rm R} - r\}$,

$$d_{\rm te}^{\rm NE}(r) \ge \sum_{i=1}^{\kappa} (n_{\rm T} - n_{\rm R} + 2 - 1) + (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - k - r) = (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r) = d_{\rm out}^{\rm NE}(r).$$
(66)

The proof in this case is concluded by combining (52) with (56) and (66).

If $l < n_{\rm T} - n_{\rm R} + 1$, by (61) and (56), we have

$$d_{\rm te}^{\rm NE}(r) < (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r) = d_{\rm out}^{\rm NE}(r).$$
(67)

Combing (53) with (56), (64) and (67), the proof is concluded in this case.

Remark 8. Notice that in (46), the difference in the upper and lower bounds on $d_{\text{out}}^{\text{NE}}(r)$ lies in the optimization domain. It is easily verified that set \mathcal{A} represents a larger set compared with the original outage set \mathcal{C} , while set \mathcal{B} is a smaller set, i.e.,

$$\mathcal{B} \subseteq \mathcal{C} \subseteq \mathcal{A}.$$
 (68)

The results above show that the infimums over different optimization domains are achieved at the same point, revealing that the bounds derived in Proposition 2 are sufficiently tight in terms of diversity gain. \triangle

Remark 9. Observing (61) and (63), we can find that a_{te}^{\star} has at most $(n_{\rm R} - r)$ entries equal to 2. Intuitively, note that $(\mathcal{B}^c)'$ denotes the feasible regime in (57), thus any point a in it satisfies $2n_{\rm R} - \sum_{i=1}^{n_{\rm R}} a_i \geq 2r$. This indicates that any a has at most $(n_{\rm R} - r)$ entries equal to 2.

6 DMT Analysis on Gamma-Gamma Channel

In this section, we consider the OWC channel with atmospheric turbulence according to gamma-gamma distribution. This distribution is commonly used to model the channel with moderate-to-strong atmospheric turbulence intensity. The outage diversity and the optimal DMT are characterized in the following.

6.1 Outage Diversity Gain Characterization

Due to the presence of $K_{\tau}(\cdot)$ in (11), it is difficult to analyze original gamma-gamma distribution directly. Instead, we choose an alternative expression in terms of power series, i.e.,

$$f(h_{ij}^r) = \sum_{n=0}^{\infty} d_n(\rho_1, \rho_2) (h_{ij}^r)^{n+\rho_2-1} + \sum_{n=0}^{\infty} d_n(\rho_2, \rho_1) (h_{ij}^r)^{n+\rho_1-1},$$
(69)

where

$$d_n(\rho_1, \rho_2) = \frac{(\rho_1 \rho_2)^{\rho_2 + n} \Gamma(\rho_1 - \rho_2)}{\Gamma(\rho_1) \Gamma(\rho_2) (1 + \rho_2 - \rho_1)_n n!}.$$
(70)

Combined with (9) and (40), we obtain the distribution of gamma-gamma channel as

$$f(\mathbf{H}) = \prod_{i=1}^{n_{\mathrm{R}}} \prod_{j=1}^{n_{\mathrm{T}}} \left\{ \sum_{n=0}^{\infty} d_n(\rho_1, \rho_2) \left(e^{\nu d_{ij}} h_{ij} \right)^{n+\rho_2 - 1} e^{\nu d_{ij}} + \sum_{n=0}^{\infty} d_n(\rho_2, \rho_1) \left(e^{\nu d_{ij}} h_{ij} \right)^{n+\rho_1 - 1} e^{\nu d_{ij}} \right\}.$$
(71)

After some algebraic manipulations, we can rewrite (71) as

$$f(\mathbf{H}) \doteq \prod_{i=1}^{n_{\mathrm{R}}} \prod_{j=1}^{n_{\mathrm{T}}} \left\{ \sum_{n=0}^{\infty} d_n(\rho_1, \rho_2) \left(e^{\nu d_{ij}} \right)^{n+\rho_2} (\mathsf{OSNR})^{-\frac{n+\rho_2-1}{2}} a_{n_{\mathrm{R}}} + \sum_{n=0}^{\infty} d_n(\rho_2, \rho_1) \left(e^{\nu d_{ij}} \right)^{n+\rho_1} (\mathsf{OSNR})^{-\frac{n+\rho_1-1}{2}} a_{n_{\mathrm{R}}} \right\}$$
(72)
$$\doteq (\mathsf{OSNR})^{-\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}(\rho-1)},$$
(73)

where ρ denotes min{ ρ_1, ρ_2 }.

Substituting (73) into (27) and (28), the outage diversity order $d_{\text{out}}^{\text{GG}}(r)$ at multiplexing order r is bounded by

$$\inf_{a \in \mathcal{A}'} d_{\text{out}}^{\text{GG}}(r, a) \le d_{\text{out}}^{\text{GG}}(r) \le \inf_{a \in \mathcal{B}'} d_{\text{out}}^{\text{GG}}(r, a),$$
(74)

where function $d_{\text{out}}^{\text{GG}}(r, a)$ is defined as

$$d_{\text{out}}^{\text{GG}}(r,a) = \sum_{i=1}^{n_{\text{R}}} \frac{(n_{\text{T}} - n_{\text{R}} + 2i - 1)}{2} a_i + \frac{n_{\text{R}} n_{\text{T}} a_{n_{\text{R}}}}{2} (\rho - 1);$$
(75)

and where sets \mathcal{A}' and \mathcal{B}' are defined as in (48) and (49), respectively.

6.2 Optimal DMT Characterization

We present the main results in the following theorem.

Theorem 10 (DMT of Gamma-Gamma Channel). Given a channel with distribution in (71), when $n_{\rm R} = 1$, if $l \ge l_{\rm th}^{\rm GG} = \rho n_{\rm T}$, the optimal diversity order is

$$d_{\rm GG}^{\star}(r) = \rho n_{\rm T} (1-r),$$
(76)

otherwise,

$$l(1-r) \le d_{\rm GG}^{\star}(r) \le \rho n_{\rm T}(1-r);$$
(77)

when $n_{\rm R} > 1$, if $l \ge l_{\rm th}^{\rm GG} = n_{\rm T} - n_{\rm R} + 1$, then

$$d_{\rm GG}^{\star}(r) = (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r), \tag{78}$$

otherwise,

$$l(n_{\rm R} - r) \le d_{\rm GG}^{\star}(r) \le (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r).$$
(79)

Proof: We follow the similar arguments as in the proof of Theorem 7. Here we mainly emphasize the differences.

Compared with the function in (47), $d_{\text{out}}^{\text{GG}}(r, a)$ in (75) contains an extra term $n_{\text{R}}n_{\text{T}}(\rho-1)a_{n_{\text{R}}}/2$. We consider the cases $n_{\text{R}} = 1$ and $n_{\text{R}} > 1$ separately to show the differences when $a_1 = a_{n_{\text{R}}}$ and $a_1 \neq a_{n_{\text{R}}}$.

When $n_{\rm R} = 1$, the infimums in (74) are both achieved at point $a_{\rm out}^{\star} = 2(n_{\rm R} - r)$. Hence, the outage diversity gain is given by

$$d_{\text{out}}^{\text{GG}}(r) = (n_{\text{T}} - n_{\text{R}} + 1 + n_{\text{R}}n_{\text{T}}(\rho - 1))(n_{\text{R}} - r)$$

= $\rho n_{\text{T}}(1 - r).$ (80)

When $n_{\rm R} > 1$, the infimums in (74) are both achieved at point $a_{\rm out}^{\star} = [2(n_{\rm R} - r), 0, \dots, 0]$, and we have

$$d_{\rm out}^{\rm GG}(r) = (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r).$$
 (81)

Now we calculate $d_{te}^{GG}(r)$. Modifying the optimization domain in (34) with \mathcal{A}^c or \mathcal{B}^c , we have

$$\inf_{(\mathcal{B}^c)'} d_{\mathrm{te}}^{\mathrm{GG}}(r,a) \le d_{\mathrm{te}}^{\mathrm{GG}}(r) \le \inf_{(\mathcal{A}^c)'} d_{\mathrm{te}}^{\mathrm{GG}}(r,a),\tag{82}$$

where function $d_{\text{te}}^{\text{GG}}(r, a)$ is defined as

$$d_{\rm te}^{\rm GG}(r,a) = \sum_{i=1}^{n_{\rm R}} \frac{n_{\rm T} - n_{\rm R} + 2i - 1}{2} a_i + \frac{n_{\rm R} n_{\rm T} a_{n_{\rm R}}}{2} (\rho - 1) + \frac{l}{2} \left(\sum_{i=1}^{n_{\rm R}} (2 - a_i)^+ - 2r \right);$$
(83)

and where sets $(\mathcal{A}^c)'$ and $(\mathcal{B}^c)'$ are defined as in (59) and (60), respectively.

Compared with the function in (58), $d_{\rm te}^{\rm GG}(r,a)$ also has an extra term $n_{\rm R}n_{\rm T}(\rho - 1)a_{n_{\rm R}}/2$. We still separately consider the cases when $n_{\rm R} = 1$ and $n_{\rm R} > 1$.

When $n_{\rm R} = 1$, if $l < n_{\rm T} - n_{\rm R} + 1 + n_{\rm R}n_{\rm T}(\rho - 1) = \rho n_{\rm T}$, the infimums in (82) are both achieved at $a_{\rm te}^{\star} = 0$, which yields

$$d_{\rm te}^{\rm GG}(r) = l(n_{\rm R} - r), \tag{84}$$

otherwise the infimums in (82) are both achieved at $a_{te}^{\star} = 2(n_{R} - r)$, thus

$$d_{\rm te}^{\rm GG}(r) = (n_{\rm T} - n_{\rm R} + 1 + n_{\rm R} n_{\rm T} (\rho - 1))(n_{\rm R} - r)$$

= $\rho n_{\rm T} (1 - r).$ (85)

When $n_{\rm R} > 1$, since sets $(\mathcal{A})'$ and $(\mathcal{B})'$ in (82) are the same as the corresponding sets in the negative exponential channel in (57), here we just need to show $d_{\rm te}^{\rm GG}(r) = d_{\rm te}^{\rm NE}(r)$. On the one hand, as $n_{\rm R}n_{\rm T}(\rho - 1)a_{n_{\rm R}}/2 \ge 0$ for any $a_{n_{\rm R}} \ge 0$, we can obtain that $d_{\rm te}^{\rm GG}(r, a) \ge d_{\rm te}^{\rm NE}(r, a)$ at any a with $a_{n_{\rm R}} \ge 0$. On the other hand, at the point $a_{\rm te}^{\star}$ defined in (61), (62), and (63), we have $d_{\rm te}^{\rm GG}(r, a_{\rm te}^{\star}) = d_{\rm te}^{\rm NE}(r, a_{\rm te}^{\star})$. The proof is concluded.

Remark 11. Notice that when $n_{\rm R} > 1$, the $n_{\rm R}$ th entry of $a_{\rm out}^{\star}$ or $a_{\rm te}^{\star}$ is always 0, which indicates the extra term $n_{\rm R}n_{\rm T}(\rho-1)a_{n_{\rm R}}/2$ plays no role on $d_{\rm out}^{\rm GG}(r,a)$ and $d_{\rm te}^{\rm GG}(r,a)$. Hence when $n_{\rm R} > 1$, we always have $d_{\rm out}^{\rm GG}(r) = d_{\rm out}^{\rm NE}(r)$ and $d_{\rm te}^{\rm GG}(r) = d_{\rm te}^{\rm NE}(r)$. Δ

7 DMT Analysis on Log-Normal Channel

In this section, we consider the OWC channel with atmospheric turbulence according to log-normal distribution. This distribution is commonly used to model the channel with weak atmospheric turbulence intensity. The outage diversity order and the optimal DMT are characterized in the following.

7.1 Outage Diversity Gain Characterization

Substituting (9) and (12) into (40), we have

$$f(\mathbf{H}) = \prod_{i=1}^{n_{\mathbf{R}}} \prod_{j=1}^{n_{\mathbf{T}}} \left\{ \frac{1}{h_{ij}\sqrt{2\pi\sigma_l^2}} \exp\left(-\frac{\left(\log(\exp(\nu d_{ij})h_{ij}) - \mu_l\right)^2}{2\sigma_l^2}\right) \right\}.$$
 (86)

Following similar arguments as in Section 5.1, we simplify f(H) as

$$f(\mathbf{H}) \doteq (\mathsf{OSNR})^{\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}} \times \prod_{i=1}^{n_{\mathrm{R}}} \prod_{j=1}^{n_{\mathrm{T}}} \left\{ \exp\left(-\frac{\left(\log\left(\exp(\nu d_{ij})(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}{2}}}\right) - \mu_{l}\right)^{2}\right)}{2\sigma_{l}^{2}}\right) \right\}$$

$$= (\mathsf{OSNR})^{\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}} \times \prod_{i=1}^{n_{\mathrm{I}}} \prod_{j=1}^{n_{\mathrm{T}}} \left\{ \exp\left(-\frac{\left(\log(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}{2}}{2}} - (\mu_{l} - \nu d_{ij})\right)^{2}\right)}{2\sigma_{l}^{2}}\right) \right\}$$

$$= (\mathsf{OSNR})^{\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}} \times \exp\left(-\log(\mathsf{OSNR}) \times \frac{n_{\mathrm{R}}n_{\mathrm{T}}}{2\sigma_{l}^{2}} \left(\frac{a_{n_{\mathrm{R}}}^{2}}{4}\log(\mathsf{OSNR}) + a_{n_{\mathrm{R}}}\beta_{1} + \frac{\beta_{2}}{\log(\mathsf{OSNR})}\right)\right)$$

$$= (\mathsf{OSNR})^{\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}} \times (\mathsf{OSNR})^{\log_{\mathsf{OSNR}}} \exp\left(-\log(\mathsf{OSNR}) \times \frac{n_{\mathrm{R}}n_{\mathrm{T}}}{2\sigma_{l}^{2}} \left(\frac{a_{n_{\mathrm{R}}}^{2}}{4}\log(\mathsf{OSNR}) + a_{n_{\mathrm{R}}}\beta_{1} + \frac{\beta_{2}}{\log(\mathsf{OSNR})}\right)\right)$$

$$= (\mathsf{OSNR})^{\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}} - \frac{n_{\mathrm{R}}n_{\mathrm{T}}}{2\sigma_{l}^{2}} \left(\frac{a_{n_{\mathrm{R}}}^{2}}{4}\log(\mathsf{OSNR}) + a_{n_{\mathrm{R}}}\beta_{1} + \frac{\beta_{2}}{\log(\mathsf{OSNR})}\right)$$

$$= (\mathsf{OSNR})^{\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}} - \frac{n_{\mathrm{R}}n_{\mathrm{T}}}{2\sigma_{l}^{2}} \left(\frac{a_{n_{\mathrm{R}}}^{2}}{4}\log(\mathsf{OSNR}) + a_{n_{\mathrm{R}}}\beta_{1} + \frac{\beta_{2}}{\log(\mathsf{OSNR})}\right)} \right)$$

$$= (\mathsf{OSNR})^{\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}} - \frac{n_{\mathrm{R}}n_{\mathrm{T}}}{2\sigma_{l}^{2}} \left(\frac{a_{n_{\mathrm{R}}}^{2}}{4}\log(\mathsf{OSNR}) + a_{n_{\mathrm{R}}}\beta_{1} + \frac{\beta_{2}}{\log(\mathsf{OSNR})}\right)} \right)$$

$$(87)$$

where $\log_{OSNR}(\cdot)$ denotes the logarithm to the base of OSNR; where β_1 and β_2 are defined as

$$\beta_1 = \frac{\sum_{i=1}^{n_{\rm R}} \sum_{j=1}^{n_{\rm T}} (\mu_l - \nu d_{ij})}{n_{\rm R} n_{\rm T}},\tag{89}$$

$$\beta_2 = \frac{\sum_{i=1}^{n_{\rm R}} \sum_{j=1}^{n_{\rm T}} (\mu_l - \nu d_{ij})^2}{n_{\rm R} n_{\rm T}};\tag{90}$$

and where (87) follows by

$$(\mathsf{OSNR})^{\log_{\mathsf{OSNR}} e^{b \cdot \log_{\mathsf{OSNR}}}} = (\mathsf{OSNR})^{b \cdot \log_{\mathsf{OSNR}} e^{\log_{\mathsf{OSNR}}}}$$
(91)

$$= (\mathsf{OSNR})^b. \tag{92}$$

Substituting (88) into (27) and (28), we bound the outage diversity gain as

$$\inf_{a \in \mathcal{A}'} d_{\text{out}}^{\text{LN}}(r, a) \le d_{\text{out}}^{\text{LN}}(r) \le \inf_{a \in \mathcal{B}'} d_{\text{out}}^{\text{LN}}(r, a),$$
(93)

where function $d_{\text{out}}^{\text{LN}}(r, a)$ is defined as

$$d_{\text{out}}^{\text{LN}}(r,a) = \sum_{i=1}^{n_{\text{R}}} \frac{(n_{\text{T}} - n_{\text{R}} + 2i - 1)}{2} a_i - \frac{n_{\text{R}} n_{\text{T}} a_{n_{\text{R}}}}{2} + \frac{n_{\text{R}} n_{\text{T}}}{2\sigma_l^2} \left(\frac{a_{n_{\text{R}}}^2}{4} \log(\text{OSNR}) + a_{n_{\text{R}}} \beta_1\right);$$
(94)

and where sets \mathcal{A}' and \mathcal{B}' are defined as in (48) and (49), respectively.

Optimal DMT Characterization 7.2

We present the main results in the following theorem.

Theorem 12 (DMT of Log-Normal Channel). Given a channel with distribution in (86), and denote $\theta \triangleq \left(\log(\text{OSNR})/2 + \beta_1 \right) / \sigma_l^2$. When $n_{\rm R} = 1$, if $l \ge l_{\rm th}^{\rm LN} = \theta n_{\rm T}$, the optimal diversity order is

$$d_{\rm LN}^{\star}(r) = \theta n_{\rm T}(1-r),\tag{95}$$

otherwise,

$$l(1-r) \le d_{\rm LN}^{\star}(r) \le \theta n_{\rm T}(1-r); \tag{96}$$

when $n_{\rm R} > 1$, if $l \ge l_{\rm th}^{\rm LN} = n_{\rm T} - n_{\rm R} + 1$, then

$$d_{\rm LN}^{\star}(r) = (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r), \qquad (97)$$

otherwise,

$$l(n_{\rm R} - r) \le d_{\rm LN}^{\star}(r) \le (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r).$$
(98)

Proof: We first calculate $d_{out}^{LN}(r)$. In (94), $d_{out}^{LN}(r,a)$ is a nonlinear function to vector a due to the extra term

$$u(a_{n_{\rm R}}) \triangleq \frac{n_{\rm R} n_{\rm T}}{2\sigma_l^2} \left(\frac{a_{n_{\rm R}}^2}{4} \log(\mathsf{OSNR}) + a_{n_{\rm R}} \beta_1 \right).$$
(99)

Notice that at high OSNR, $u(a_{n_{\rm B}})$ tends to be infinite, and dominates the sum of rest terms. For this reason, we prioritize small $a_{n_{\rm R}}$ to optimize $d_{\rm out}^{\rm LN}(r,a)$ over \mathcal{A}' and \mathcal{B}' in (93). Since this extra term $u(a_{n_{\rm R}})$ only relates to $a_{n_{\rm R}}$, we still separately consider $n_{\rm R} = 1$ and $n_{\rm R} > 1$ to distinguish $a_1 = a_{n_{\rm R}}$ and $a_1 \neq a_{n_{\rm R}}$. When $n_{\rm R} = 1$, the integral domains \mathcal{A}' and \mathcal{B}' in (28) coincide as $\{a \ge 2(n_{\rm R} - r)\}$.

We can obtain the optimal point $a_{\text{out}}^{\star} = 2(n_{\text{R}} - r)$, and we have

$$d_{\text{out}}^{\text{LN}}(r) = \left(n_{\text{T}} - n_{\text{R}} + 1 - n_{\text{R}}n_{\text{T}} + \frac{n_{\text{R}}n_{\text{T}}}{2\sigma_l^2} \left(\log(\text{OSNR}) + 2\beta_1\right)\right) (n_{\text{R}} - r) = \theta n_{\text{T}}(1 - r).$$
(100)

When $n_{\rm R} > 1$, a necessary condition to have finite infimums in (93) is the $n_{\rm R}$ th entry of a to be zero, i.e., $a_{n_{\rm R}} = 0$. Then the rest of entries in optimal point $a_{\rm out}^{\star}$ can be computed by linear optimization, and we obtain $a_{\text{out}}^{\star} = [2(n_{\text{R}} - r), 0, \cdots, 0].$ Substituting a_{out}^{\star} into (93), the outage diversity gain is given by

$$d_{\rm out}^{\rm LN}(r) = (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r).$$
 (101)

Now we calculate $d_{te}^{LN}(r)$. Modifying the integral domain in (34) as \mathcal{A}^c or \mathcal{B}^c , we have

$$\inf_{(\mathcal{B}^c)'} d_{\mathrm{te}}^{\mathrm{LN}}(r,a) \le d_{\mathrm{te}}^{\mathrm{LN}}(r) \le \inf_{(\mathcal{A}^c)'} d_{\mathrm{te}}^{\mathrm{LN}}(r,a),$$
(102)

where function $d_{te}^{LN}(r, a)$ is defined as

$$d_{\rm te}^{\rm LN}(r,a) = \sum_{i=1}^{n_{\rm R}} \frac{n_{\rm T} - n_{\rm R} + 2i - 1}{2} a_i - \frac{n_{\rm R} n_{\rm T}}{2} a_{n_{\rm R}} + \frac{n_{\rm R} n_{\rm T}}{2\sigma_l^2} \left(\frac{a_{n_{\rm R}}^2}{4} \log(\mathsf{OSNR}) + a_{n_{\rm R}}\beta_1\right) + \frac{l}{2} \left(\sum_{i=1}^{n_{\rm R}} (2 - a_i)^+ - 2r\right); \quad (103)$$

and where sets $(\mathcal{A}^c)'$ and $(\mathcal{B}^c)'$ are defined as in (59) and (60), respectively.

It is direct to see that $d_{te}^{LN}(r, a)$ is also a nonlinear function due to the extra term $u(a_{n_R})$. Using similar method argued above, we can directly show that when $n_R = 1$, if $l < \theta n_T$, both infimums of $d_{te}^{LN}(r, a)$ over $(\mathcal{A}^c)'$ and $(\mathcal{B}^c)'$ are achieved at $a_{te}^* = 0$, which yields

$$d_{\rm te}^{\rm LN}(r) = l(n_{\rm R} - r),$$
 (104)

otherwise, $a_{te}^{\star} = 2(n_{R} - r)$, and we have

$$d_{\rm te}^{\rm LN}(r) = \left(n_{\rm T} - n_{\rm R} + 1 - n_{\rm R}n_{\rm T} + \frac{n_{\rm R}n_{\rm T}}{2\sigma_l^2} \left(\log(\mathsf{OSNR}) + 2\beta_1\right)\right) (n_{\rm R} - r) = \theta n_{\rm T}(1 - r).$$
(105)

When $n_{\rm R} > 1$, following similar arguments as in the derivation of $d_{\rm out}^{\rm LN}(r)$, we obtain the optimal point $a_{\rm te}^{\star}$ has the same expressions as in (61), (62), and (63), and $d_{\rm te}^{\rm LN}(r)$ has the same expressions as in (64) and (65). Then the proof is concluded.

Remark 13. The term $u(a_{n_{\rm R}})$ causes the non-linearity of $d_{\rm out}^{\rm LN}(r, a)$ in (94). When $n_{\rm R} > 1$, a necessary condition to have a finite $d_{\rm out}^{\rm LN}(r)$ is $a_{n_{\rm R}} = 0$, and hence $u(a_{n_{\rm R}}) = 0$. Eliminating the influence of $a_{n_{\rm R}}$, the optimization problems in (93) are essentially equivalent to the linear problems in (46).

One issue in Theorem 12 is when $n_{\rm R} = 1$, at high OSNR, the optimal diversity gain $d_{\rm LN}^*(r)$ tends to be infinite, which fails to reflect the performance limits of practical OWC systems. In fact, similar phenomena are also observed in the existing literature [21], [27], [40]. This issue is due to the fact that $d_{\rm LN}^*(r)$ contains the OSNR related term, while the diversity gain essentially measures the system performance under high OSNR. To deal with this problem, we adopt the definition of asymptotically relative diversity order (ARDO) proposed in [27], and then analyze the optimal diversity gain of log-normal channel. The ARDO of a transmission scheme is given by

$$\hat{d}^{\star}(r) = \lim_{\mathsf{OSNR}\to\infty} \frac{d^{\star}(r)}{d^{\star}_{\mathrm{BM}}(r)},\tag{106}$$

where $d^{\star}(r)$ denotes the optimal diversity order, and where $d^{\star}_{BM}(r)$ denotes the optimal diversity order of a benchmark scheme. In the following, we use the outage diversity order at r = 0 of the SISO system as the benchmark³, such that

$$\hat{d}_{\rm LN}^{\star}(r) = \frac{d_{\rm LN}^{\star}(r)}{d_{\rm out}^{\rm LN}(0)}.$$
(107)

Substituting $d_{LN}^{\star}(r)$ in (96) and (97) into (107), we have the following proposition:

Proposition 14. Given a channel with distribution in (86), when $n_{\rm R} = 1$, the ARDO under a large block length is given by

$$\hat{d}_{\rm LN}^{\star}(r) = n_{\rm T}(1-r);$$
 (108)

under a small block length we have

$$0 \le \hat{d}_{\rm LN}^{\star}(r) \le n_{\rm T}(1-r).$$
 (109)

³The benchmark is chosen by considering the fact that outage diversity order at r = 0 is the maximal achievable diversity order for any transmission scheme.



Figure 1: Outage diversity gain for SISO channel.

8 Numerical Results and Discussions

8.1 Optimal DMT Analysis

In this section, we present numerical examples of the derived DMT results. For the notational convenience, a channel with $n_{\rm R}$ receive, $n_{\rm T}$ transmit antennas, and l block length is simply denoted by $n_{\rm T} \times n_{\rm R} \times l$.

Figures 1, 2, and 3 depict the outage diversity gain $d_{\text{out}}(r)$ at the multiplexing gain r under different channel conditions for SISO, MISO, and MIMO channels. These curves can also represent the optimal spatial diversity gain $d^{\star}(r)$ when $l \geq l_{\text{th}}$, where l_{th} represents the block length threshold of corresponding channels, such as $l_{\text{th}}^{\text{NE}}$, $l_{\text{th}}^{\text{GG}}$, $l_{\text{th}}^{\text{LN}}$ (see Theorems 7, 10, and 12).

Figure 1 shows the derived outage diversity gain $d_{out}(r)$ under three different channel distributions, i.e., negative exponential, gamma-gamma, and log-normal channels. Besides, we choose two different values of ρ given in [41] to show their influence on the gamma-gamma channel, i.e., (1) $\rho = 10.1614$, (2) $\rho = 17.4369$. For negative exponential and gamma-gamma channels, $d_{out}(r)$ linearly decreases with the increasing r. For the log-normal channel, $d_{out}(r)$ tends to be infinite when r = 0 and returns to 0 when r takes the maximum value of 1. Since log-normal, gamma-gamma, and negative exponential distributions model the channel with weak, moderate, and strong atmospheric turbulence intensities, respectively, this figure also reflects the fact that increasing turbulence intensity may hamper the diversity gain. As shown in Figure 2, the above conclusions also hold in the MISO channels.

From Figure 3, we observe that the maximum outage and multiplexing gains are given by $d_{\max}^{\star} = (n_{\mathrm{T}} - n_{\mathrm{R}} + 1)n_{\mathrm{R}}$ and $r_{\max}^{\star} = n_{\mathrm{R}}$, and the outage diversity gains are equal under these three different fading MIMO channels.

Remark 15. Although different distributions may lead to different outage probabilities, the result in Figure 3 reveals that the exponents of outage probabilities are identical regardless of the fading properties. In fact, this phenomenon has been explained in [19, Theorem 21]. Specifically, as long as some relevant attributes by different channels are identical, the derived outage diversity gain will be the same. Therefore, we could



Figure 2: Outage diversity gain for $n_{\rm T} \times 1$ MISO channel.



Figure 3: Outage diversity gain for $n_{\rm T} \times n_{\rm R}$ MIMO channel.



Figure 4: Bounds on optimal diversity gain for 16×4 MIMO channel.

infer that the result shown in Figure 3 is also caused by some internal attributes which behave the same among these three fading channels. \triangle

When $l < l_{\rm th}$, Figure 4 depicts the bounds on optimal $d^{\star}(r)$ for a 16 × 4 MIMO channel with l = 7, and l = 10, respectively. Here, $l_{\rm th} = 13$. It is direct to see $d^{\star}(r)$ is upper-bounded by $d_{\rm out}(r)$ and lower-bounded by $d_{\rm te}(r)$ for the above channels. For any $l < l_{\rm th}$ and $r < n_{\rm R}$, the lower bound on diversity gain is strictly below the upper bound, implying that there is always uncertainty on the characterization of optimal DMT curve. The uncertainty tends to decrease as the block length or the multiplexing gain increases.

We also depict the optimal $d^*(r)$ as a function of the number of transmit antenna $n_{\rm T}$ or receive antenna $n_{\rm R}$ in Figure 5. The left curve of Figure 5 depicts the optimal diversity gain $d^*(r)$ as a function of $n_{\rm T}$ with fixed parameters $n_{\rm R} = 4$ and r = 2, while the right curve depicts $d^*(r)$ as a function of $n_{\rm R}$ with fixed parameters $n_{\rm T} = 16$ and r = 2. We observe that $d^*(r)$ monotonically increases as $n_{\rm T}$ increases, while increasing $n_{\rm R}$ does not always guarantee larger diversity gain. In fact, our derived $d^*(r)$ is a linear function with $n_{\rm T}$, but a second-order function of $n_{\rm R}$, which is fundamentally different from the RF DMT result established before [17].

Remark 16. Since the optimal diversity order $d^{\star}(r)$ is a second-order function of $n_{\rm R}$, hence, installing $n_{\rm R}^{\star} = \lfloor \frac{n_{\rm T}+1+r}{2} \rfloor$ antennas at the receiving end may achieve the best performance.

8.2 DMT Comparison with Existing Results

Table 1 compares our derived DMT results with existing results in RF and OWC channels. Note that the slope of $d_{out}(r)$ or $d_{te}(r)$ in [17] is two times large as in [28], and the threshold of block length in [17] is the same as in [28], which are different from our derived results. In the following, to avoid ambiguity, we denote l_{th} in [17] or [28] as l_{th}^{ZJ} .

To numerically compare the optimal DMT curves for different channels in Table 1, Figures 6 and 7 depict these DMT curves of 16×4 and 16×11 MIMO channels when $l \geq l_{\rm th}$. For any multiplexing gain r, we can see that $d_{\rm ZT}^{\star}(r)$ is always large than $d_{\rm JB}^{\star}(r)$



Figure 5: Optimal diversity gain at fixed $n_{\rm R}$ or $n_{\rm T}$.

Table 1: Results of different MIMO channels

	$d_{ m out}(r)$	$d_{\rm te}(r) \ (l \le l_{\rm th})$	$l_{ m th}$
Rayleigh [17]	$(n_{\mathrm{T}} - r)(n_{\mathrm{R}} - r)$	$-l(r-r_1) + (n_{\rm T} - r_1)(n_{\rm R} - r_1)$	$n_{\rm T} + n_{\rm R} - 1$
Negative Exponential [28]	$\frac{1}{2}(n_{\rm T}-r)(n_{\rm R}-r)$	$-\frac{l}{2}(r-r_1)+\frac{1}{2}(n_{\rm T}-r_1)(n_{\rm R}-r_1)$	$n_{\rm T} + n_{\rm R} - 1$
Negative Exponential (us)	$(n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r)$	$l(n_{ m R}-r)$	$n_{\rm T} - n_{\rm R} + 1$

where $r_1 = n_{\rm R} - [(l - |n_{\rm T} - n_{\rm R}| - 1)/2].$



Figure 6: Comparisons of the optimal DMT curve for $16 \times 4 \times \infty$ MIMO channel.



Figure 7: Comparisons of the optimal DMT curve for $16 \times 11 \times \infty$ MIMO channel.

and $d_{NE}^{\star}(r)$. This implies that the optical signals usually suffer more severe fading than their traditional RF counterparts.

In Figures 6 and 7 we also observe that, even though $d_{\rm NE}^{\star}(r)$ and $d_{\rm JB}^{\star}(r)$ both characterize optimal diversity gain, these two curves are fundamentally different. $d_{\rm JB}^{\star}(r)$ is derived by only considering the fact that input signal is real, and then simply putting a half factor into the traditional RF capacity expression. Here, our derived $d_{\rm NE}^{\star}(r)$ is obtained by taking the unique inputs constraints into consideration. Comparing these two figures, $d_{\rm NE}^{\star}(r)$ tends to match $d_{\rm ZT}^{\star}(r)$ as $n_{\rm R}$ decreases, which is caused by the fact that $d_{\rm NE}^{\star}(r)$ approaches $(n_{\rm T} - r)(n_{\rm R} - r)$ as $n_{\rm R}$ decreases. For specific $n_{\rm R}$, $d_{\rm NE}^{\star}(r)$ may intersect with $d_{\rm JB}^{\star}(r)$ at some point (r_2, d_2) , and when $r < r_2$, $d_{\rm NE}^{\star}(r)$ is strictly less than $d_{\rm JB}^{\star}(r)$, while $d_{\rm NE}^{\star}(r)$ is strictly larger than $d_{\rm JB}^{\star}(r)$ when $r > r_2$.

Moreover, we numerically study the variation of $d^*(r)$ when l decreases for a 16×4 MIMO channel. Recall that $l_{\rm th}^{\rm ZJ}$ denotes the block length threshold for the previous literature [17], [28]. Hence, we have $l_{\rm th}^{\rm NE} = 13$ and $l_{\rm th}^{\rm ZJ} = 19$. We select three different block lengths: (1) $l = \infty$, (2) l = 15, (3) l = 10, as shown in Figures 6, 8 and 9. Note that as l decreases, the gap between the upper and lower bounds on $d^*(r)$ increases. However, our derived $d_{\rm NE}^*(r)$ has the slowest uncertainty growth rate compared with the existing results. Since when l decreases from ∞ to 15, the bounds $d_{\rm ZT}^*(r)$ and $d_{\rm JB}^*(r)$ become very loose. In contrast, our derived $d_{\rm NE}^*(r)$ is still tight until l = 10.

9 Conclusion Remarks

This paper investigates three different block fading optical wireless communication channels and derives the optimal DMT under both optical peak- and average-power input constraints. We first establish an upper bound on the optimal diversity gain by characterizing the outage diversity gain. Then by analyzing the error probability of a random coding scheme, we establish a new lower bound. It turns out that these two bounds are close, hence give good approximations on the optimal diversity gain. In fact, at a large block length regime, these bounds match, thus we can precisely characterize the optimal diversity gain. Our derived DMT results are fundamentally different from their



Figure 8: Comparisons of the optimal DMT curve for $16\times4\times15$ MIMO channel.



Figure 9: Comparisons of the optimal DMT curve for $16\times4\times10$ MIMO channel.

counterparts in traditional RF channels. These differences are due to the unique input constraints in the optical wireless channels. From the perspective of a practical $n_{\rm T} \times n_{\rm R}$ OWC system design, our results imply at a specific multiplexing order r, letting the numbers of transmit and receive antennas satisfy $\lfloor n_{\rm T} + 1 + r \rfloor = 2n_{\rm R}$ may achieve the best performance in terms of diversity gain.

A Derivation of Instantaneous Capacity Bounds

A.1 Proof of Proposition 2

Recall $\mathbf{X}_j = [\mathbf{X}_{j1}, \mathbf{X}_{j2}, \dots, \mathbf{X}_{jn_T}]^{\mathsf{T}}$, it is straightforward to get

$$\mathsf{E}[\mathbf{X}_{ji}] = \mathsf{A}\left(\frac{1}{\mu} - \frac{e^{-\mu}}{1 - e^{-\mu}}\right), \quad \forall i \in \{1, 2, \dots, n_{\mathrm{T}}\}$$
(110)

and

$$h(\mathbf{X}_{ji}) = \log A + 1 - \frac{\mu e^{-\mu}}{1 - e^{-\mu}} - \log \frac{\mu}{1 - e^{-\mu}}.$$
(111)

Next, we define a new random vector $\widetilde{\mathbf{X}}_j = [\widetilde{\mathbf{X}}_{j1}, \widetilde{\mathbf{X}}_{j2}, \dots, \widetilde{\mathbf{X}}_{jn_T}]^{\mathsf{T}}$, whose entries are i.i.d. according to Gaussian distribution with expectation 0, and variance

$$\operatorname{Var}\left[\widetilde{\mathbf{X}}_{ji}\right] = \frac{1}{2\pi e} 2^{2\operatorname{h}(\mathbf{X}_{ji})}.$$
(112)

By the Gaussian differential entropy formula, we can immediately show

$$h(\mathbf{\tilde{X}}_{ji}) = h(\mathbf{X}_{ji}), \ \forall i \in \{1, \dots, n_{\mathrm{T}}\}.$$
(113)

For a given channel realization $\mathbf{H} = \mathbf{H}$, apply the generalized EPI in [30], and we have

$$\mathbf{H}\mathbf{X}_j) \ge \mathbf{h}(\mathbf{H}\widetilde{\mathbf{X}}_j) \tag{114}$$

$$= \frac{n_{\rm R}}{2} \log \left(2\pi e \left| {\rm HK}_{\tilde{\mathbf{X}}_j \tilde{\mathbf{X}}_j} {\rm H}^{\mathsf{T}} \right|^{\frac{1}{n_{\rm R}}} \right), \tag{115}$$

where $K_{\widetilde{\mathbf{X}}_{i}\widetilde{\mathbf{X}}_{j}}$ denotes the covariance matrix of $\widetilde{\mathbf{X}}_{j}$ which is diagonal with its *i*th entry

$$\left[\mathbf{K}_{\widetilde{\mathbf{X}}_{j}\widetilde{\mathbf{X}}_{j}}\right]_{ii} = \mathsf{Var}\left[\widetilde{\mathbf{X}}_{ji}\right], \quad \forall i \in \{1, 2, \dots, n_{\mathrm{T}}\}.$$
(116)

Substituting (111) into (112), (116) and (115), we have

h(

$$h(\mathrm{H}\mathbf{X}_j) \ge \frac{n_{\mathrm{R}}}{2} \log\left(2\pi e \left|\frac{\sigma_n^2}{2\pi\alpha^2 e} \mathrm{T}(\alpha, n_{\mathrm{T}})^2 (\mathsf{OSNR})^2 \mathrm{H}\mathrm{H}^{\mathsf{T}}\right|^{\frac{1}{n_{\mathrm{R}}}}\right),\tag{117}$$

where

$$T(\alpha, n_{\rm T}) = 2 \frac{1 - e^{-\mu}}{\mu} 2^{-\frac{\mu e^{-\mu}}{1 - e^{-\mu}}},$$
(118)

with μ satisfying (22).

Now we lower-bound the instantaneous capacity as

$$_{j};\mathbf{Y}_{j}|\mathbf{H}) = h(\mathbf{H}\mathbf{X}_{j} + \mathbf{Z}_{j}|\mathbf{H}) - h(\mathbf{Z}_{j})$$

$$(119)$$

$$\geq \frac{1}{2} \log \left(2^{2 \operatorname{h}(\mathbf{H}\mathbf{X}_j)} + 2^{2 \operatorname{h}(\mathbf{Z}_j)} \right) - \operatorname{h}(\mathbf{Z}_j) \tag{120}$$

$$= \frac{1}{2} \log \left(1 + \frac{2^{2\pi (\Pi \Lambda)}}{(2\pi e \sigma_n)^{n_{\mathrm{R}}}} \right)$$
(121)

$$\geq \frac{1}{2} \log \left(1 + L_l (\mathsf{OSNR})^{2n_{\mathrm{R}}} | \mathbf{H} \mathbf{H}^{\mathsf{T}} | \right), \tag{122}$$

where (119) follows by the independence between **H** and \mathbf{Z}_j , (120) by applying the EPI [42], and (122) by substituting (117) into (121). The proof is concluded.

 $I(\mathbf{X}$

A.2 Proof of Proposition 3

At high OSNR, by the asymptotic upper bound in [31, Theorem 21], we have

$$I(\mathbf{X}_j; \mathbf{Y}_j | \mathbf{H}) \le n_{\mathrm{R}} \log \mathsf{A} + \frac{1}{2} \log \left(\frac{\mathsf{V}_{\mathbf{H}}^2}{(2\pi e \sigma_n)^{n_{\mathrm{R}}}} \right), \tag{123}$$

where $V_{\mathbf{H}} = \sum_{\mathcal{U}} |\det \mathbf{H}_{\mathcal{U}}|$ with $H_{\mathcal{U}}$ denoting the $n_{\mathrm{R}} \times n_{\mathrm{R}}$ submatrix of H indexed by set \mathcal{U} .

We further upper bound the RHS of (123) as

$$n_{\rm R}\log \mathsf{A} + \frac{1}{2}\log\left(\frac{\mathsf{V}_{\mathbf{H}}^2}{(2\pi e\sigma_n)^{n_{\rm R}}}\right) \le n_{\rm R}\log \mathsf{A} + \frac{1}{2}\log\left(\frac{\binom{n_{\rm T}}{n_{\rm R}}\left(\sum_{\mathcal{U}}|\mathbf{H}_{\mathcal{U}}|^2\right)}{(2\pi e\sigma_n)^{n_{\rm R}}}\right)$$
(124)

$$= n_{\rm R} \log \mathsf{A} + \frac{1}{2} \log \left(\frac{\binom{n_{\rm T}}{n_{\rm R}} |\mathbf{H}\mathbf{H}^{\mathsf{T}}|}{(2\pi e \sigma_n)^{n_{\rm R}}} \right), \tag{125}$$

where (124) holds by Cauchy–Schwarz inequality, and (125) by Cauchy–Binet formula: $\sum_{\mathcal{U}} |\mathbf{H}_{\mathcal{U}}|^2 = |\mathbf{H}\mathbf{H}^{\mathsf{T}}|$. The proof is concluded.

B Derivation of Outage Probability Bounds

B.1 Preliminary

B.1.1 Decomposition on Channel Matrix

Given the channel matrix $\mathbf{H} \in \mathcal{R}^{n_{\mathbf{R}} \times n_{\mathrm{T}}}_+$, we first decompose it by LQ factorization

$$\mathbf{H} = \mathbf{L}\mathbf{Q},\tag{126}$$

where $L \in \mathcal{R}^{n_R \times n_R}$ is lower-triangular and $Q \in \mathcal{R}^{n_R \times n_T}$ is orthogonal satisfying $QQ^{\mathsf{T}} = \mathbf{I}_{n_R \times n_R}$. Combined with (126), we can rewrite HH^{T} as

$$HH^{\mathsf{T}} = LL^{\mathsf{T}}.$$
 (127)

Then perform eigenvalue decomposition on HH^{T} , we obtain

$$HH^{\mathsf{T}} = U\Lambda U^{\mathsf{T}},\tag{128}$$

where $U \in \mathcal{R}^{n_{\mathrm{R}} \times n_{\mathrm{R}}}$ is orthogonal satisfying $UU^{\mathsf{T}} = \mathbf{I}_{n_{\mathrm{R}} \times n_{\mathrm{R}}}$ and $\Lambda \in \mathcal{R}^{n_{\mathrm{R}} \times n_{\mathrm{R}}}$ is diagonal with $[\Lambda]_{ii} = \lambda_i, \ \forall i = \{1, \cdots, n_{\mathrm{R}}\}$. Combining (127) with (128), we can decompose LL^{T} as

$$LL^{\mathsf{T}} = \left(U\Lambda^{\frac{1}{2}}\right) \left(U\Lambda^{\frac{1}{2}}\right)^{\mathsf{T}}.$$
(129)

Then we have $L = U\Lambda^{\frac{1}{2}}$. Further substitute L into (126), and we have

$$\mathbf{H} = \mathbf{U}\Lambda^{\frac{1}{2}}\mathbf{Q}.\tag{130}$$

Denote $D \triangleq \Lambda^{\frac{1}{2}}$, and rewrite λ_i in terms of a_i , $\forall i = \{1, \dots, n_R\}$, then we get

$$\mathbf{D} = \operatorname{diag}[(\mathsf{OSNR})^{-\frac{a_1}{2}}, \dots, (\mathsf{OSNR})^{-\frac{a_{n_{\mathbf{R}}}}{2}}].$$
(131)

B.1.2 Derivation of Ep. (27)

By [38], we have

$$f(a) = \xi \; [\log(\mathsf{OSNR})]^{n_{\mathrm{R}}} \times \left(\prod_{i=1}^{n_{\mathrm{R}}} (\mathsf{OSNR})^{-\frac{(n_{\mathrm{T}}-n_{\mathrm{R}}+1)}{2}a_i}\right)$$

$$\times \prod_{i < j} \left| (\mathsf{OSNR})^{-a_i} - (\mathsf{OSNR})^{-a_j} \right| \times \int_{\mathbf{V}_{n_{\mathbf{R}}, n_{\mathbf{R}}}} \int_{\mathbf{V}_{n_{\mathbf{R}}, n_{\mathbf{T}}}} f(\mathbf{UDQ}) d\mathbf{Q} d\mathbf{U}, \quad (132)$$

where ξ is a normalization constant.

Form the definition in (19), we observe that the outage diversity order $d_{out}(r)$ is fully determined by the corresponding OSNR exponent of outage probability. Note that

$$\lim_{\mathsf{OSNR}\to\infty} \frac{\log(\xi[\log(\mathsf{OSNR})]^{n_{\mathrm{R}}})}{\log(\mathsf{OSNR})} = 0.$$
(133)

Thus, at high OSNR, we have

$$\xi[\log(\mathsf{OSNR})]^{n_{\mathrm{R}}} \doteq \mathrm{OSNR}^{0}. \tag{134}$$

Second, note that the term $|(OSNR)^{-a_i} - (OSNR)^{-a_j}|$ is only determined by the smaller exponent at high OSNR, and thus we can simplify the following product term as

$$\prod_{i < j} \left| (\mathsf{OSNR})^{-a_i} - (\mathsf{OSNR})^{-a_j} \right| \doteq \prod_{i=1}^{n_{\mathrm{R}}} (\mathsf{OSNR})^{-(i-1)a_i}.$$
(135)

Substituting (133) and (135) into (132), the proof is concluded.

B.2 Proof of Theorem 5

B.2.1 Upper Bound

To prove the new bounds on outage probability in Theorem 5, we employ the bounds obtained in Proposition 2. We first prove the upper bound on outage probability, and then prove the lower bound.

Denote $R = r \log \text{OSNR}$. Substitute (23) into the RHS of (17), and we obtain

$$P_{\text{out}}(\text{OSNR}) \leq P\left[\frac{1}{2}\log(1 + L_l(\text{OSNR})^{2n_{\text{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|) \leq r\log(\text{OSNR})\right]$$
(136)
$$= P\left[1 + L_l(\text{OSNR})^{2n_{\text{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}| \leq \text{OSNR}^{2r}\right].$$
(137)

Since $\lambda_i = (\mathsf{OSNR})^{-a_i}, \forall i \in \{1, \dots, n_R\}$, are the the eigenvalues of \mathbf{HH}^{T} , we have

$$|\mathbf{H}\mathbf{H}^{\mathsf{T}}| = \prod_{i=1}^{n_{\mathrm{R}}} \lambda_i = \mathsf{OSNR}^{-\sum_{i=1}^{n_{\mathrm{R}}} a_i}.$$
(138)

Substituting (138) into (137), we get

$$P_{\text{out}}(\mathsf{OSNR}) \le P\left[1 + L_l(\mathsf{OSNR})^{2n_{\text{R}} - \sum_{i=1}^{n_{\text{R}}} a_i} \le (\mathsf{OSNR})^{2r}\right]$$
(139)

$$\doteq \mathbf{P}\left[\left(\mathsf{OSNR}\right)^{\left(2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_{i}\right)^{+}} \le \left(\mathsf{OSNR}\right)^{2r}\right] \tag{140}$$

$$= \mathbf{P}\left[(2n_{\mathbf{R}} - \sum_{i=1}^{n_{\mathbf{R}}} a_i)^+ \le 2r \right]$$
(141)

$$= \int_{\mathcal{A}} f(a)da, \tag{142}$$

where (140) follows by the fact that $1 + c \cdot \mathsf{OSNR}^b \doteq \mathsf{OSNR}^{b^+}$, and (142) by the definition in (29). The proof is concluded.

B.2.2 Lower Bound

Substituting (25) into the RHS of (17), and following the similar arguments as in Section B.2.1, we have

$$P_{out}(\mathsf{OSNR}) \ge P\left[\frac{1}{2}\log(L_u(\mathsf{OSNR})^{2n_{\mathrm{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|) \le r\log(\mathsf{OSNR})\right]$$
(143)

$$= \mathbf{P} \left[L_u(\mathsf{OSNR})^{2n_{\mathbf{R}} - \sum_{i=1}^{n_{\mathbf{R}}} a_i} \le (\mathsf{OSNR})^{2r} \right]$$
(144)

$$\doteq \mathbf{P}\left[(\mathsf{OSNR})^{\left(2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_{i}\right)} \le (\mathsf{OSNR})^{2r}\right]$$
(145)

$$\doteq \mathbf{P}\left[2n_{\mathbf{R}} - \sum_{i=1}^{n_{\mathbf{R}}} a_i \le 2r\right] \tag{146}$$

$$= \int_{\mathcal{B}} f(a)da. \tag{147}$$

The proof is concluded.

Proof of Eq. (31) \mathbf{C}

Denote the codebook as $\{X(0), X(1), \ldots, X(\mathsf{OSNR}^{lr} - 1)\}$. Suppose the sent codeword is X(0), and we consider the event when the ML decoder decides erroneously in favor of X(1). This event occurs only if the projection distance of Y(0) - HX(0) on the direction of HX(1) - HX(0) is larger than $d_{01}/2$, where $d_{01} = || HX(1) - HX(0) ||_{\mathsf{F}}$ is the distance between HX(1) and HX(0). The probability on the occurrence of this event is⁴

$$P(X(0) \to X(1) | \mathbf{H} = H) = P\left(\frac{\langle \mathbf{Y}(\mathbf{0}) - HX(0), HX(1) - HX(0) \rangle}{\| HX(1) - HX(0) \|_{\mathsf{F}}} \ge \frac{1}{2} \| HX(1) - HX(0) \|_{\mathsf{F}}\right).$$
(148)

Since X(0) is sent, and $\mathbf{Z}(\mathbf{0}) = \mathbf{Y}(\mathbf{0}) - \mathrm{HX}(0)$, then the projection of $\mathbf{Y}(\mathbf{0}) - \mathrm{HX}(0)$ on HX(1) - HX(0) is still a Gaussian variable. In the following, we use the term $\frac{z}{\|HX(1)-HX(0)\|_{\mathsf{F}}}$ to denote the projection, where z is a Gaussian variable with expectation 0 and variance $\|HX(1) - HX(0)\|_{\mathsf{F}}^2 \sigma_n^2$. Thus we further upper-bound the error probability as

$$P(X(0) \to X(1) | \mathbf{H} = \mathbf{H}) = P\left(\frac{z}{\| HX(1) - HX(0) \|_{\mathsf{F}}} \ge \frac{1}{2} \| HX(1) - HX(0) \|_{\mathsf{F}}\right)$$
(149)

$$=Q\left(\sqrt{\frac{\parallel \mathrm{HX}(1) - \mathrm{HX}(0) \parallel_{\mathsf{F}}^{2}}{4\sigma_{n}^{2}}}\right)$$
(150)

$$\leq \exp\left(-\frac{\parallel \mathrm{HX}(1) - \mathrm{HX}(0) \parallel_{\mathsf{F}}^{2}}{8\sigma_{n}^{2}}\right),\tag{151}$$

where (151) holds because $Q(t) \leq 1/2 \exp(-t^2/2)$.

Note that

$$\| \operatorname{HX}(1) - \operatorname{HX}(0) \|_{\mathsf{F}}^{2} = \sum_{k=1}^{n_{\mathrm{R}}} \sum_{j=1}^{l} |\sum_{i=1}^{n_{\mathrm{T}}} [\operatorname{X}_{ij}(1) - \operatorname{X}_{ij}(0)] h_{ki}|^{2},$$
(152)

⁴Without loss of generality, for any $X \in \mathcal{R}^{m \times n}$ and $Y \in \mathcal{R}^{m \times n}$, $\langle X, Y \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} y_{ij}$.

where $X_{ij}(\cdot) = [X(\cdot)]_{ij}$, and $h_{ki} = [H]_{ki}$. Setting $\Omega_j = [X_{1j}(1) - X_{1j}(0), \cdots, X_{n_T j}(1) - X_{n_T j}(0)]^{\mathsf{T}}$, we rewrite (152) as

$$\| \operatorname{HX}(1) - \operatorname{HX}(0) \|_{\mathsf{F}}^{2} = \sum_{j=1}^{l} \Omega_{j}^{\mathsf{T}} \operatorname{H}^{\mathsf{T}} \operatorname{H}\Omega_{j}.$$
 (153)

By SVD, we decompose $\mathrm{H}^{\mathsf{T}}\mathrm{H} = \mathrm{V}\Lambda\mathrm{V}^{\mathsf{T}}$, where $\Lambda = \mathrm{diag}[\lambda_1, ..., \lambda_{n_{\mathrm{T}}}]$. Define $\beta_j = [\beta_{j1}, \cdots, \beta_{jn_{\mathrm{T}}}]$, and let $\beta_j = \Omega_j^{\mathsf{T}}\mathrm{V}$. We can rewrite (153) as

$$\| \operatorname{HX}(1) - \operatorname{HX}(0) \|_{\mathrm{F}}^{2} = \sum_{j=1}^{l} \sum_{i=1}^{n_{\mathrm{T}}} \lambda_{i} \beta_{ji}^{2}.$$
(154)

Substituting (154) into (151), we have

$$P(X(0) \to X(1) | \mathbf{H} = \mathbf{H}) \le \exp\left(-\frac{1}{8\sigma_n^2} \sum_{j=1}^l \sum_{i=1}^{n_T} \lambda_i \beta_{ji}^2\right).$$
(155)

Averaging over the ensemble of truncated exponential random codes, we get

$$P(\mathbf{X}(\mathbf{0}) \to \mathbf{X}(\mathbf{1}) | \mathbf{H} = \mathbf{H}) \le \mathsf{E}_{\boldsymbol{\beta}_{ji}} \left[\exp\left(-\frac{1}{8\sigma_n^2} \sum_{j=1}^l \sum_{i=1}^{n_{\mathrm{T}}} \lambda_i \boldsymbol{\beta}_{ji}^2\right) \right]$$
(156)

$$=\prod_{j=1}^{l}\mathsf{E}_{\boldsymbol{\beta}_{j}}\left[\exp(-\frac{1}{8\sigma_{n}^{2}}\sum_{i=1}^{n_{\mathrm{T}}}\lambda_{i}\boldsymbol{\beta}_{ji}^{2})\right]$$
(157)

$$= \det\left(\mathbf{I} + \frac{g(\alpha, n_{\mathrm{T}})}{2} (\mathsf{OSNR})^2 \mathrm{HH}^{\mathsf{T}}\right)^{-\frac{1}{2}}, \qquad (158)$$

where (157) follows by the independence between vectors β_p and β_q , $\forall p \neq q$, and (158) by the derivation in the following Appendix C.0.1.

Note that at transmit rate $R = r \log(\text{OSNR})$, we have in total $(\text{OSNR})^{lr}$ codewords. Then applying the union bound, the decoded error probability can be upper-bounded as

$$P(\text{error}|\mathbf{H} = \mathbf{H}) \le (\mathsf{OSNR})^{lr} \det \left(\mathbf{I} + \frac{g(\alpha, n_{\mathrm{T}})}{2} (\mathsf{OSNR})^2 \mathbf{H} \mathbf{H}^{\mathsf{T}} \right)^{-\frac{t}{2}}$$
(159)

$$= (\mathsf{OSNR})^{lr} \prod_{i=1}^{n_{\mathrm{R}}} \left(1 + \frac{g(\alpha, n_{\mathrm{T}})}{2} (\mathsf{OSNR})^2 \lambda_i \right)^{-\frac{l}{2}}.$$
 (160)

C.0.1 Proof of Eq. (158)

Denote the matrix $\mathbf{V} = [\mathbf{V}_1, \cdots, \mathbf{V}_{n_T}]$, where $\mathbf{V}_p = [v_{1p}, \cdots, v_{n_Tp}]^{\mathsf{T}}$ is the *p*th column of V. Recall that V is an orthogonal matrix, i.e.,

$$\sum_{i=1}^{n_{\rm T}} v_{ip} v_{ip} = 1, \quad \forall p \in \{1, \cdots, n_{\rm T}\},$$
(161)

$$\sum_{i=1}^{n_{\mathrm{T}}} v_{ip} v_{iq} = 0, \quad \forall p \neq q.$$

$$(162)$$

Note that

$$\boldsymbol{\beta}_{jp} = \Delta \mathbf{X}_{1j} v_{1p} + \dots + \Delta \mathbf{X}_{n_{\mathrm{T}}j} v_{n_{\mathrm{T}}p}$$
(163)

$$\boldsymbol{\beta}_{jq} = \Delta \mathbf{X}_{1j} v_{1q} + \dots + \Delta \mathbf{X}_{n_{\mathrm{T}}j} v_{n_{\mathrm{T}}q}, \tag{164}$$

where $\Delta \mathbf{X}_{ij} = \mathbf{X}_{ij}(\mathbf{1}) - \mathbf{X}_{ij}(\mathbf{0})$.

Recall that \mathbf{X}_{ij} , $\forall i \in \{1, \dots, n_{\mathrm{T}}\}$ and $j \in \{1, \dots, l\}$, follows i.i.d. truncated exponential distribution in (21), and hence it is directly to verify

$$\mathsf{E}[\boldsymbol{\beta}_{ji}] = 0, \tag{165}$$

$$\operatorname{Var}[\beta_{ji}] = 2\mathsf{A}^2 \left(\frac{1}{\mu^2} - \frac{e^{-\mu}}{(1 - e^{-\mu})^2} \right)$$
(166)

$$\triangleq 2\mathsf{A}^2 g(\alpha, n_{\mathrm{T}}). \tag{167}$$

By the central limit theorem, we can approximate β_{ji} as a Gaussian random variable with expectation (165) and variance (167), and hence by averaging over the distribution of β_{ji} in (157), we have

$$\prod_{j=1}^{l} \mathsf{E}_{\boldsymbol{\beta}_{j}} \left[\exp\left(-\frac{1}{8\sigma_{n}^{2}} \sum_{i=1}^{n_{\mathrm{T}}} \lambda_{i} \boldsymbol{\beta}_{ji}^{2}\right) \right] = \det\left(\mathbf{I} + \frac{g(\alpha, n_{\mathrm{T}})}{2} (\mathsf{OSNR})^{2} \mathrm{HH}^{\mathsf{T}}\right)^{-\frac{1}{2}}.$$
 (168)

The proof is concluded.

References

- V. W. S. Chan, "Free-Space Optical Communications," J. Lightw. Technol., vol. 24, no. 12, pp. 4750–4762, Dec. 2006.
- [2] L. Zeng, D. C. O'Brien, H. L. Minh, G. E. Faulkner, K. Lee, D. Jung, Y. Oh, and E. T. Won, "High Data Rate Multiple Input Multiple Output (MIMO) Optical Wireless Communications Using White LED Lighting," *IEEE J. Sel. Area. Comm.*, vol. 27, no. 9, pp. 1654–1662, Dec. 2009.
- [3] J. Song, W. Ding, F. Yang, H. Yang, B. Yu, and H. Zhang, "An Indoor Broadband Broadcasting System Based on PLC and VLC," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 299–308, Mar. 2015.
- [4] M. Kashef, M. Ismail, M. Abdallah, K. A. Qaraqe, and E. Serpedin, "Energy Efficient Resource Allocation for Mixed RF/VLC Heterogeneous Wireless Networks," *IEEE J. Sel. Area. Comm.*, vol. 34, no. 4, pp. 883–893, Mar. 2016.
- [5] H. Kaushal and G. Kaddoum, "Optical Communication in Space: Challenges and Mitigation Techniques," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 57–96, Aug. 2017.
- [6] C. Cox, E. Ackerman, R. Helkey, and G. E. Betts, "Techniques and Performance of Intensity-Modulation Direct-Detection Analog Optical Links," *IEEE Trans. Microw. Theory Tech.*, vol. 45, no. 8, pp. 1375–1383, Aug. 1997.
- [7] S. C. J. Lee, S. Randel, F. Breyer, and A. M. J. Koonen, "Pam-DMT for Intensity-Modulated and Direct-Detection Optical Communication Systems," *IEEE Photon. Technol. Lett.*, vol. 21, no. 23, pp. 1749–1751, Sep. 2009.
- [8] X. Zhu and J. M. Kahn, "Free-Space Optical Communication Through Atmospheric Turbulence Channels," *IEEE Trans. Commun.*, vol. 50, no. 8, pp. 1293–1300, Aug. 2002.
- [9] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation through Random Media*. 2nd ed. SPIE Press: Bellingham, WA, USA, 2005.
- [10] O. Korotkova and I. Toselli, "Non-Classic Atmospheric Optical Turbulence," Appl. Sci., vol. 11, no. 18, p. 8487, Sep. 2021.

- [11] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An Overview of MIMO Communications - A Key to Gigabit Wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Nov. 2004.
- [12] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE J. Sel. Top. Sign. Proces.*, vol. 8, no. 5, pp. 742–758, Apr. 2014.
- [13] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for Next Generation Wireless Systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186– 195, Feb. 2014.
- [14] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems," *IEEE J. Sel. Top. Sign. Proces.*, vol. 10, no. 3, pp. 436–453, Feb. 2016.
- [15] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE J. Sel. Area. Comm.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
- [16] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.
- [17] L. Zheng and D. N. C. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073– 1096, May 2003.
- [18] R. Narasimhan, "Finite-SNR Diversity-Multiplexing Tradeoff for Correlated Rayleigh and Rician MIMO Channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3965–3979, Aug. 2006.
- [19] L. Zhao, W. Mo, Y. Ma, and Z. Wang, "Diversity and Multiplexing Tradeoff in General Fading Channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1549–1557, Mar. 2007.
- [20] A. R. Heidarpour, G. K. Kurt, and M. Uysal, "Finite-SNR Diversity-Multiplexing Tradeoff for Network Coded Cooperative OFDMA Systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1385–1396, Nov. 2017.
- [21] H. Nouri, F. Touati, and M. Uysal, "Diversity-Multiplexing Tradeoff for Log-Normal Fading Channels," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 3119–3129, Jun. 2016.
- [22] D. Zahavi, L. Zhang, I. Maric, R. Dabora, A. J. Goldsmith, and S. Cui, "Diversity-Multiplexing Tradeoff for the Interference Channel With a Relay," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 963–982, Dec. 2015.
- [23] U. Habib, A. E. Aighobahi, T. Quinlan, S. D. Walker, and N. J. Gomes, "Analog Radio-Over-Fiber Supported Increased RAU Spacing for 60 GHz Distributed MIMO Employing Spatial Diversity and Multiplexing," J. Lightw. Technol., vol. 36, no. 19, pp. 4354–4360, Apr. 2018.
- [24] T. Yang and Z. Xi, "Incremental Decode-and-Forward Protocols for D2D Communication: A Diversity-Multiplexing-Tradeoff Perspective," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13927–13931, Aug. 2020.
- [25] C. G. Tsinos and K. Berberidis, "A Cooperative Uplink Transmission Technique With Improved Diversity–Multiplexing Tradeoff," *IEEE Trans. Veh. Technol.*, vol. 64, no. 7, pp. 2883–2896, Aug. 2015.

- [26] W. Shin, S. Chung, and Y. H. Lee, "Diversity–Multiplexing Tradeoff and Outage Performance for Rician MIMO Channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1186–1196, Mar. 2008.
- [27] M. Safari and M. Uysal, "Cooperative Diversity over Log-Normal Fading Channels: Performance Analysis and Optimization," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1963–1972, May 2008.
- [28] A. Jaiswal and M. R. Bhatnagar, "Free-Space Optical Communication: A Diversity-Multiplexing Tradeoff Perspective," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 1113–1125, Jul. 2018.
- [29] P. Sharda and M. R. Bhatnagar, "Diversity-multiplexing Tradeoff for Indoor Visible Light Communication," in Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob), Thessaloniki, Greece, Oct. 2020, pp. 1–6.
- [30] R. Zamir and M. Feder, "A Generalization of the Entropy Power Inequality with Applications," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1723–1728, Sep. 1993.
- [31] L. Li, S. M. Moser, L. Wang, and M. Wigger, "On the Capacity of MIMO Optical Wireless Channels," *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5660–5682, Sep. 2020.
- [32] E. W. Weisstein, "Modified Bessel Function of the Second Kind," 2002. [Online]. Available: https://mathworld.wolfram.com/
- [33] M. R. Bhatnagar, "A One Bit Feedback Based Beamforming Scheme for FSO MISO System Over Gamma-Gamma Fading," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1306–1318, Jan. 2015.
- [34] M. Riedl, Optical Design Fundamentals for Infrared Systems. 2nd ed. SPIE Press: Bellingham, WA, USA, 2001.
- [35] M. S. Faruk, Y. Mori, and K. Kikuchi, "In-Band Estimation of Optical Signal-to-Noise Ratio From Equalized Signals in Digital Coherent Receivers," *IEEE Photon.* J., vol. 6, no. 1, pp. 1–9, Feb. 2014.
- [36] K. Keykhosravi, E. Agrell, M. Secondini, and M. Karlsson, "When to Use Optical Amplification in Noncoherent Transmission: An Information-Theoretic Approach," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2438–2445, Apr. 2020.
- [37] A. A. Farid and S. Hranilovic, "Outage Capacity Optimization for Free-Space Optical Links With Pointing Errors," J. Lightw. Technol., vol. 25, no. 7, pp. 1702–1710, Jul. 2007.
- [38] A. Edelman, "Eigenvalues and Condition Numbers of Random Matrices," SIAM J. Matrix Anal. Appl., vol. 9, no. 4, p. 543–560, Dec. 1988.
- [39] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.
- [40] G. Pan, E. Ekici, and Q. Feng, "Performance Analysis of Multi-Branch Multi-Hop Wireless Relay Systems over Log-Normal Channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 223–233, Jan. 2014.
- [41] J. Parikh and V. K. Jain, "Study on Statistical Models of Atmospheric Channel for FSO communication Link," in *Proc. Nirma Univ. Int. Conf. Eng. (NUiCONE)*, Ahmedabad, India, Dec. 2011, pp. 1–7.
- [42] R. König and G. Smith, "The Entropy Power Inequality for Quantum Systems," *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1536–1548, Jan. 2014.