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

Sufang Yang, *Graduate Student Member, IEEE,* Longguang Li, *Member, IEEE,* Haoyue Tang, *Student Member, IEEE,* and Jintao Wang, *Senior Member, IEEE*

***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.**

I. Introduction

As an essential complement to radio-frequency (RF) communication, optical wireless communication (OWC) significantly improves the rate performance and offers an ideal solution to spectrum scarcity. It has recently been considered as a promising technique in future 6G [1], [2]. For a general OWC system, it covers unguided visible, infrared (IR), or ultraviolet (UV) light bands [3], [4]. In this paper, we mainly focus on the free-space terrestrial OWC system, operated nearly at the visible and IR frequencies.

Modern OWC system typically uses either light-emitting diodes (LEDs) or lasers as the transmitter and photodetectors as the receiver [3], [4], which is widely known as the intensity-modulation and direct-detection (IM-DD) transmission scheme. As a consequence, transmitting signals are in the form of optical intensities and hence are real and nonnegative, fundamentally different from their RF counterparts. Furthermore, the peak and average optical power of transmitting signals need to be constrained by safety reasons and hardware limitations.

The OWC system is more sensitive to channel attenuation than the traditional RF system due to the IM-DD transmission scheme, particularly in medium- and long-range communication. In a free-space OWC system, the fluctuations caused by atmospheric turbulence severely degrade the quality of communication links. They change the temperature and pressure of the atmosphere and eventually lead to refractive index variations of the path links [5], [6]. To combat the channel fading induced by atmospheric turbulence, the widely adopted multiple-input multiple-output (MIMO) technique plays a pivotal role, which provides many independent transmission paths available at the transmitting or receiving ends [7]. Another benefit provided by utilizing MIMO technique is the significant improvement in spectral efficiency when compared with single-antenna systems [8].

In this paper, we investigate the optimal diversity-multiplexing tradeoff (DMT) for the MIMO-OWC system. 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. It has triggered extensive research deriving the optimal DMTs for different fading channels, such as Rayleigh, Rician, Nakagami and log-normal fading channels [9]–[16].

Most of the existing DMT results focus on the traditional RF system [9]–[14], among which Zheng and Tse presented the classic optimal tradeoff curve for Rayleigh fading channels [9]. They showed that when the channel block length satisfies $l \geq l_{\text{ra}} = n_{\text{T}} + n_{\text{R}} - 1$, the optimal DMT curve can be exactly characterized as

$$d^{\star}_{\text{ra}}(r) = (n_{\text{T}} - r)(n_{\text{R}} - r), \tag{1}$$

where $d^{\star}_{\text{ra}}(r)$ denotes the maximum diversity gain achieved at the given multiplexing gain r, and n_{T} and n_{R} denote the number of transmit and receive antennas, respectively. Moreover, they proposed the bounds on DMT curve at the small block length regime ($l < l_{\text{ra}}$): if $r < r_{\text{ra}} = \min\{n_{\text{R}}, n_{\text{R}} - \lceil \frac{l-(n_{\text{T}}-n_{\text{R}})-1}{2} \rceil\}$, the maximum diversity gain can be upper- and lower-bounded by:

$$d^{\star}_{\text{ra}}(r) \leq (n_{\text{T}} - r)(n_{\text{R}} - r), \tag{2}$$

This work was supported in part by the National Natural Science Foundation of China under Grant No. 62101192, in part by Shanghai Sailing Program under Grant No. 21YF1411000, and in part by Tsinghua University-China Mobile Research Institute Joint Innovation Center. This article was presented in part at the 2022 IEEE International Symposium on Information Theory. *(Corresponding author: Longguang Li.)*

Sufang Yang is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: ysf20@mails.tsinghua.edu.cn).

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

Haoyue Tang was with the Department of Electronic Engineering, Tsinghua University, Beijing 100190, China. She is now with the Institute of Network Science, Yale University, New Haven, CT 06520 USA (e-mail: haoyue.tang@yale.edu).

Jintao Wang is with the Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, Beijing 100084, China, also with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the Research Institute, Tsinghua University in Shenzhen, Shenzhen 518057, China (e-mail: wangjintao@tsinghua.edu.cn).

$$d^{\star}_{\rm ra}(r) \geq l(n_{\rm ra}-r), \tag{3}$$

where $n_{\rm ra}=r^2_{\rm ra}/l+\big(1-(n_{\rm T}+n_{\rm R})/l\big)r_{\rm ra}+n_{\rm T}n_{\rm R}/l$; otherwise, the optimal DMT curve agrees with (1).

Some existing works considered the unique characteristics of the OWC system and derived some DMT results [15]–[19]. Among them, [15] and [16] investigated outdoor and indoor OWC systems, respectively. For the indoor OWC system, [15] assumed that the user/receiver position is uniformly distributed over a circular area. Then it derived a closed-form DMT for the SISO system and a numerical DMT for the MIMO system. For the outdoor OWC system, [16] assumed that the channel is influenced by atmospheric turbulence. Then it proposed that for an exponential fading channel, the maximal diversity gain at multiplexing gain r is $d^{\star}_{\rm ra}(r)/2$ under a large block length. Also, under a small block length, they also found analogous DMT bounds as in (2) and (3), except for an extra factor of $1/2$.

The extra factor of $1/2$ reflects the loss of half of the degree of freedom, which is due to the fact that inputs need to be real-valued in the OWC system, while the counterparts in the RF system are complex. However, as mentioned before, inputs in the OWC system represent optical intensities. Hence, their values must also be nonnegative. Furthermore, since the optical intensity exactly represents power, the power constraints imposed on the inputs also need to be described differently. Therefore, 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 exponential, gamma-gamma, and log-normal distributions to model the channel with atmospheric turbulence 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 the practical OWC system.

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) [23], we first derive a lower bound on instantaneous capacity. Then, an upper bound is established by some algebraic manipulations on the asymptotic capacity in [24, Thm. 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 $l \geq n_{\rm T}-n_{\rm R}+1$, the optimal diversity gain $d^{\star}(r)$ can be characterized as

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

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:

$$d^{\star}(r) \leq (n_{\rm T}-n_{\rm R}+1)(n_{\rm R}-r),$$
$$d^{\star}(r) \geq l(n_{\rm R}-r).$$

See Theorems 3, 4, and 5 in detail.

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

Notation: Random variables, vectors and matrices are boldfaced, e.g., $\boldsymbol{h}$, $\boldsymbol{a}$ and $\mathbf{X}$, while their realizations are typeset in h, a and $\rm X$, respectively. $\mathbf{X}_j$ denotes the jth column of matrix $\mathbf{X}$, and $\boldsymbol{x}_{ij}$ or $[\mathbf{X}]_{ij}$ denotes the ith row and jth entry of $\mathbf{X}$. Sets are typeset in a special font, e.g., $\mathcal{A}$. Differential Entropy is denoted by $\mathsf{h}(\cdot)$, and mutual information by $\mathrm{I}(\cdot;\cdot)$. $\|\cdot\|_1$ 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 $\mathbb{E}[\cdot]$, and variance by $\mathbb{V}[\cdot]$. $\mathbb{R}^{m\times n}$ ($\mathbb{R}^{m\times n}_{+}$) 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 $\dot{\geqslant}$,$\dot{\leqslant}$ are similarly defined.

II. Channel Model

Consider a MIMO channel with $n_{\rm T}$ LEDs/lasers and $n_{\rm R}$ photodetectors. The channel output is given by[1]

$$\mathbf{Y}=\mathbf{H}\mathbf{X}+\mathbf{Z}, \tag{4}$$

where l denotes the block length, the channel input $\mathbf{X}\in\mathbb{R}^{n_{\rm T}\times l}_{+}$, the channel noise $\mathbf{Z}\in\mathbb{R}^{n_{\rm R}\times l}$, whose entries are independent and identically distributed (i.i.d.) inside and across blocks, and the channel matrix $\mathbf{H}\in\mathbb{R}^{n_{\rm R}\times n_{\rm T}}_{+}$, whose entries remain constant inside one block and i.i.d. across blocks. For simplicity, we assume $n_{\rm T}\geq n_{\rm R}$ and $\mathbf{H}$ is of full rank.

Each entry $\boldsymbol{h}_{ij}$ in $\mathbf{H}$ represents the nonnegative and real-valued channel gain from jth transmit antenna to ith receive

[1]Without loss of generality, we assume the photoelectric coefficient for the photodetector is 1.

antenna, and depends on two factors: deterministic distance attenuation and random atmospheric turbulence loss [16], [25], [26]. As a consequence, $\boldsymbol{h}_{ij}$ can be formulated as

$$\boldsymbol{h}_{ij} = h_{ij}^{d}\boldsymbol{h}_{ij}^{r}, \tag{5}$$

where $h_{ij}^{d} = e^{-\nu d_{ij}}$ denotes the deterministic part with parameter ν characterizing the transmission environment and d_{ij} being the transmission distance, $\boldsymbol{h}_{ij}^{r}$ denotes the random part characterizing the atmospheric turbulence intensity[2]. In this paper, we consider three different distributions of the random atmospheric turbulence $\boldsymbol{h}_{ij}^{r}$, which cover the turbulence fluctuation regimes from strong to weak intensities, and they are

- *Exponential distribution:*

$$f(h_{ij}^{r}) = e^{-h_{ij}^{r}}; \tag{6}$$

- *Gamma-gamma distribution:*

$$f(h_{ij}^{r}) = \frac{2(\rho_1\rho_2)^{\frac{\rho_1+\rho_2}{2}}}{\Gamma(\rho_1)\Gamma(\rho_2)}(h_{ij}^{r})^{\frac{\rho_1+\rho_2}{2}-1} \times K_{\rho_1-\rho_2}(2\sqrt{\rho_1\rho_2 h_{ij}^{r}}), \tag{7}$$

  where ρ_1 and ρ_2 denote the irradiance fluctuation parameters with $\rho_2 < \rho_1$, $\Gamma(\cdot)$ denotes the Gamma function with $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$, and $K_\tau(\cdot)$ denotes the modified Bessel function of the second kind with τ being the order [27], [28];
- *Log-normal distribution:*

$$f(h_{ij}^{r}) = \frac{1}{h_{ij}^{r}\sqrt{2\pi\sigma_l^2}}\exp\left(-\frac{(\log(h_{ij}^{r})-\mu_l)^2}{2\sigma_l^2}\right), \tag{8}$$

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

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

$$\mathrm{P}[\boldsymbol{x}_{ij} > \mathsf{A}] = 0,\ \forall i \in \{1,\dots,n_{\mathrm{T}}\},\ \forall j \in \{1,\dots,l\}, \tag{9a}$$

$$\frac{1}{l}\mathbb{E}[\parallel \mathbf{X} \parallel_1] \le \mathsf{E}, \tag{9b}$$

where A denotes the allowed maximum optical power by each antenna, and E denotes the total average optical power allowed across all antennas. The ratio between the allowed average power and the allowed peak power is denoted by

$$\alpha \triangleq \frac{\mathsf{E}}{\mathsf{A}}, \tag{10}$$

where $\alpha \in (0, \frac{n_{\mathrm{T}}}{2}]$ and is fixed in the paper [24], [32].

The additive noise $\mathbf{Z}$ is induced by the influence of thermal noise and ambient light [29], and each entry is modeled as i.i.d. Gaussian distribution, i.e., $\boldsymbol{z}_{ij} \sim \mathcal{N}(0,\sigma^2)$, where $\boldsymbol{z}_{ij}$ denotes the ith row and jth column entry of $\mathbf{Z}$.

[2]Some existing works also consider the channel fading caused by pointing errors [25], [26]. In this work, we assume the transmit and receive antennas are perfectly aligned and then ignore the effect of pointing errors. Nevertheless, exploring the DMT results while including the fading caused by pointing errors is still meaningful in the future.

Since information is carried on the intensity of the optical signal, we adopt the definition of optical signal-to-noise ratio (OSNR) [29]–[31] as follows:

$$\mathsf{OSNR} = \frac{\mathsf{E}}{\sigma}. \tag{11}$$

We further present some useful concepts and definitions in terms of OSNR. More details can be seen in [9], [13], [16], [25]. Given a transmission scheme, an outage occurs when the instantaneous (conditioned) channel capacity of this channel can not support the target rate R. We denote

$$\mathcal{O} \triangleq \{\mathrm{H} : \mathrm{C}(\mathbf{X}_j;\mathbf{Y}_j|\mathbf{H}=\mathrm{H}) < R\} \tag{12}$$

as the set of outage event, where $\mathbf{X}_j$ and $\mathbf{Y}_j$ denote the transmit and receive vectors at time j, and $\mathrm{C}(\mathbf{X}_j;\mathbf{Y}_j|\mathbf{H}=\mathrm{H})$ denotes the instantaneous capacity at the channel state $\mathbf{H}=\mathrm{H}$. Therefore, the outage probability is defined as

$$\mathrm{P}_{\text{out}}(\mathsf{OSNR}) \triangleq \mathrm{P}\big\{\mathrm{C}(\mathbf{X}_j;\mathbf{Y}_j|\mathbf{H}) \le R\big\}. \tag{13}$$

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 $d_{\text{out}}(r)$, and diversity gain $d(r)$ if the rate $R(\mathsf{OSNR})$ satisfies

$$\lim_{\mathsf{OSNR}\to\infty}\frac{R(\mathsf{OSNR})}{\log(\mathsf{OSNR})} = r, \tag{14}$$

the outage probability $\mathrm{P}_{\text{out}}(\mathsf{OSNR})$ satisfies

$$-\lim_{\mathsf{OSNR}\to\infty}\frac{\log \mathrm{P}_{\text{out}}(\mathsf{OSNR})}{\log(\mathsf{OSNR})} = d_{\text{out}}(r), \tag{15}$$

and average error probability $\mathrm{P}_{\mathrm{e}}(\mathsf{OSNR})$ satisfies

$$-\lim_{\mathsf{OSNR}\to\infty}\frac{\log \mathrm{P}_{\mathrm{e}}(\mathsf{OSNR})}{\log(\mathsf{OSNR})} = d(r). \tag{16}$$

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

III. 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 (13), we first need to characterize the instantaneous capacity. In the presence of the peak- and average-power constraints in (9a) and (9b), 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 [32], [33]. For any time index $j \in \{1,2,\dots,l\}$, denote $\mathbf{X}_j = [\mathbf{X}_{j1},\mathbf{X}_{j2},\dots,\mathbf{X}_{jn_{\mathrm{T}}}]^{\intercal}$, and let the entries of $\mathbf{X}_j$ be i.i.d. according to the following truncated exponential distribution:

$$f(\mathrm{X}_{ij}) = \frac{1}{\mathsf{A}}\cdot\frac{\mu}{1-e^{-\mu}}\cdot e^{-\frac{\mu \mathrm{X}_{ij}}{\mathsf{A}}}, \quad \mathrm{X}_{ij} \in [0,\mathsf{A}],\ i \in \{1,2,\dots,n_{\mathrm{T}}\}, \tag{17}$$

where μ is a parameter satisfying

$$\frac{1}{\mu}-\frac{e^{-\mu}}{1-e^{-\mu}}=\frac{\alpha}{n_{\mathrm{T}}}. \tag{18}$$

The achievable rate by this truncated exponential distribution can serve as a natural lower bound on the instantaneous capacity. Applying the GEPI in [23] we derive the following lower bound, whose rigorous proof is shown in Appendix A.

Proposition 1 (Lower Bound)**.** Given a MIMO-OWC channel in (4), we have

$$\mathsf{C}(\mathbf{X}_j;\mathbf{Y}_j|\mathbf{H})\geq\frac{1}{2}\log\big(1+L_l(\mathsf{OSNR})^{2n_{\mathrm{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|\big), \tag{19}$$

where

$$L_l=\Big(\frac{e}{2\pi\alpha^2}\Big)^{n_{\mathrm{R}}}\bigg(\frac{1-e^{-\mu}}{\mu}e^{-\frac{\mu e^{-\mu}}{1-e^{-\mu}}}\bigg)^{2n_{\mathrm{R}}} \tag{20}$$

with μ satisfying (18).

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 [24, Thm. 21].

Proposition 2 (Upper Bound)**.** Given a MIMO-OWC channel in (4), we have

$$\mathsf{C}(\mathbf{X}_j;\mathbf{Y}_j|\mathbf{H})\leq\frac{1}{2}\log\big(L_u(\mathsf{OSNR})^{2n_{\mathrm{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|\big), \tag{21}$$

where

$$L_u=\binom{n_{\mathrm{T}}}{n_{\mathrm{R}}}\Big(\frac{1}{2\pi\alpha^2 e}\Big)^{n_{\mathrm{R}}}. \tag{22}$$

Remark 1. Note that L_u and L_l in Proposition 1 and 2 are constants independent of parameters OSNR and $\mathbf{H}$, which play no role in the following derivations related to the OSNR exponent.

Given a channel realization H, we denote the eigenvalue of $\mathrm{H}\mathrm{H}^{\mathsf{T}}$ by $\lambda_i,\ \forall i\in\{1,\ldots,n_{\mathrm{R}}\}$, with λ_i arranged in increasing order. Recall that H is of full rank, we can obtain that λ_i, $\forall i$, is real and positive. We define a new vector a (with the random form $\boldsymbol{a}$), which is one-to-one correspondence with the set of eigenvalues of H. i.e.,

$$\mathsf{OSNR}^{-a_i}=\lambda_i,\quad\forall i\in\{1,\cdots,n_{\mathrm{R}}\}. \tag{23}$$

where a_i denotes the ith entry of a. It is directly to obtain that

$$a_i\in\mathbb{R},\quad\forall i\in\{1,\cdots,n_{\mathrm{R}}\}. \tag{24}$$

Besides, since $\lambda_1\leq\cdots\leq\lambda_{n_{\mathrm{R}}}$, we have

$$a_1\geq\cdots\geq a_{n_{\mathrm{R}}}. \tag{25}$$

We decompose H as UDQ, where $\mathrm{U}\in\mathbb{R}^{n_{\mathrm{R}}\times n_{\mathrm{R}}}$ and $\mathrm{Q}\in\mathbb{R}^{n_{\mathrm{R}}\times n_{\mathrm{T}}}$ are orthogonal matrices, and D is a diagonal matrix with $\mathrm{D}=\mathrm{diag}[(\mathsf{OSNR})^{-\frac{a_1}{2}},\ldots,(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}]$. More details about the matrix decomposition can be found in Appendix B-A1. Combined with the relationship between a and eigenvalues of H in (23), we can characterize the distribution of $\boldsymbol{a}$ with the representation of the distribution of $\mathbf{H}$, i.e.,

$$\begin{aligned}f(a)\doteq(\mathsf{OSNR})^{-\sum_{i=1}^{n_{\mathrm{R}}}\frac{n_{\mathrm{T}}-n_{\mathrm{R}}+2i-1}{2}a_i}\\ \times\iint f(\mathrm{UDQ})d\mathrm{Q}d\mathrm{U},\end{aligned} \tag{26}$$

with $f(\mathrm{UDQ})$ being the distribution of $\mathbf{H}$. The proof of (26) can be found in Appendix B-A2.

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

Theorem 1 (Outage Probability Bounds)**.** Given a channel in (4), we have

$$\int_{\mathcal{B}}f(a)da\ \dot{\leqslant}\ \mathrm{P}_{\text{out}}(\mathsf{OSNR})\ \dot{\leqslant}\int_{\mathcal{A}}f(a)da, \tag{27}$$

with

$$\mathrm{P}_{\text{out}}(\mathsf{OSNR})\doteq\int_{\mathcal{O}}f(a)da, \tag{28}$$

where $\mathcal{O}$ denotes the outage set, $f(a)$ is defined as in (26), and set $\mathcal{A}$ is defined as

$$\mathcal{A}\triangleq\bigg\{a=[a_1,a_2,\cdots,a_{n_{\mathrm{R}}}]:(2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_i)^+\leq 2r\bigg\}, \tag{29}$$

and set $\mathcal{B}$ is defined as

$$\mathcal{B}\triangleq\bigg\{a=[a_1,a_2,\cdots,a_{n_{\mathrm{R}}}]:2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_i\leq 2r\bigg\}. \tag{30}$$

Remark 2. Since there is no closed-form expression for the instantaneous capacity, we do not know the explicit outage event set $\mathcal{O}$. Hence, it is difficult to derive the outage probability in (13). However, from (27) and (28), we can obtain

$$\mathcal{B}\subseteq\mathcal{O}\subseteq\mathcal{A}. \tag{31}$$

It indicates that we can utilize the new sets $\mathcal{A}$ and $\mathcal{B}$ to calculate the bounds on outage probability as in (27), thereby estimating the properties of the outage diversity gain.

IV. Random Coding Error Analysis

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

Consider a random code with codewords i.i.d. as in (17), and the code rate is $R=r\log(\mathsf{OSNR})$. The decoder applies the maximum-likelihood (ML) method to detect the sent message. Now consider a channel realization $\mathbf{H}=\mathrm{H}$, then the conditional error probability can be upper bounded by the following inequality, whose proof is postponed to Appendix C.

$$\begin{aligned}&\mathrm{P}(\text{error}|\mathbf{H}=\mathrm{H})\\ &\quad\leq(\mathsf{OSNR})^{lr}\prod_{i=1}^{n_{\mathrm{R}}}\bigg(1+\frac{g(\alpha,n_{\mathrm{T}})}{2}(\mathsf{OSNR})^2\lambda_i\bigg)^{-\frac{l}{2}},\end{aligned} \tag{32}$$

where

$$g(\alpha,n_{\mathrm{T}})=\frac{1}{\mu^2}-\frac{e^{-\mu}}{(1-e^{-\mu})^2}, \tag{33}$$

with μ satisfying (18).

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

Theorem 2. Given a channel in (4), we have

$$\mathrm{P_e}(\mathsf{OSNR}) \dot{\leqslant} (\mathsf{OSNR})^{-d_{\text{out}}(r)} + (\mathsf{OSNR})^{-d_{\text{te}}(r)}, \quad (34)$$

where

$$(\mathsf{OSNR})^{-d_{\text{te}}(r)} \doteq \int_{\mathcal{O}^c} f(a)(\mathsf{OSNR})^{-\frac{l}{2}(\sum_{i=1}^{n_{\mathrm{R}}}(2-a_i)^+-2r)}da, \quad (35)$$

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

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

$$\begin{aligned}&\mathrm{P_e}(\mathsf{OSNR})\\ &\quad= \mathrm{P_{out}}(\mathsf{OSNR})\mathrm{P_{elout}}(\mathsf{OSNR}) + \mathrm{P_{e,nout}}(\mathsf{OSNR}) \quad (36)\\ &\quad\le \mathrm{P_{out}}(\mathsf{OSNR}) + \mathrm{P_{e,nout}}(\mathsf{OSNR}). \quad (37)\end{aligned}$$

where $\mathrm{P_{elout}}(\mathsf{OSNR})$ denotes the error probability conditioned on outage event, $\mathrm{P_{e,nout}}(\mathsf{OSNR})$ denotes the joint probability of error and no outage events.

By the definition of outage diversity, the first term in the RHS of (37) is

$$\mathrm{P_{out}}(\mathsf{OSNR}) \doteq (\mathsf{OSNR})^{-d_{\text{out}}(r)}. \quad (38)$$

Substitute $\lambda_i = (\mathsf{OSNR})^{-a_i}$ into (32), we have

$$\mathrm{P}(\text{error}|a) \dot{\leqslant} (\mathsf{OSNR})^{-\frac{l}{2}(\sum_{i=1}^{n_{\mathrm{R}}}(2-a_i)^+-2r)}. \quad (39)$$

Multiplying with the distribution of a, and integrating over set $\mathcal{O}^c$, the second term in the RHS of (37) is

$$\mathrm{P_{e,nout}}(\mathsf{OSNR}) \doteq (\mathsf{OSNR})^{-d_{\text{te}}(r)}. \quad (40)$$

The proof is concluded. ∎

V. DMT Analysis on Exponential Channel

In this section we consider the OWC channel with atmospheric turbulence according to 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 gain, and based on this, we further characterize the optimal DMT.

A. Outage Diversity Gain Characterization

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

As shown in Theorem 1, to estimate the outage diversity gain, we first need to calculate the distribution of vector $\boldsymbol{a}$ in (26). To do this, we first analyze the distribution of $\mathbf{H}$. Since the entries in $\mathbf{H}$ are assumed i.i.d., and by (6) we expand the distribution of $\mathbf{H}$ as

$$f(\mathrm{H}) = \prod_{i=1}^{n_{\mathrm{R}}}\prod_{j=1}^{n_{\mathrm{T}}} f(h_{ij}) \quad (41)$$

$$= \prod_{i=1}^{n_{\mathrm{R}}}\prod_{j=1}^{n_{\mathrm{T}}} \exp\big(-e^{\nu d_{ij}}h_{ij} + \nu d_{ij}\big). \quad (42)$$

By Appendix B-A1, we can decompose $\mathrm{H} = \mathrm{UDQ}$, then we have

$$h_{ij} = \sum_{k=1}^{n_{\mathrm{R}}} u_{ik} \times (\mathsf{OSNR})^{-\frac{a_k}{2}} \times q_{kj}, \quad (43)$$

where $h_{ij} = [\mathrm{H}]_{ij}$, $u_{ik} = [\mathrm{U}]_{ik}$ and $q_{ik} = [\mathrm{Q}]_{ik}$. By $a_1 \ge \cdots \ge a_{n_{\mathrm{R}}}$, we obtain the last term $u_{in_{\mathrm{R}}} \times (\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}} \times q_{n_{\mathrm{R}}j}$ dominates other terms in (43) at high OSNR, i.e.,

$$h_{ij} \doteq u_{in_{\mathrm{R}}} \times (\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}} \times q_{n_{\mathrm{R}}j}. \quad (44)$$

Since the entries of H are nonnegative, we have $u_{in_{\mathrm{R}}}q_{n_{\mathrm{R}}j} \ge 0$, thereby

$$h_{ij} \doteq (\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}. \quad (45)$$

Substituting (45) into (42), we obtain

$$\begin{aligned}f(\mathrm{H}) &\doteq \prod_{i=1}^{n_{\mathrm{R}}}\prod_{j=1}^{n_{\mathrm{T}}} \exp\Big\{-e^{\nu d_{ij}}(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}} + \nu d_{ij}\Big\}\\ &\doteq \prod_{i=1}^{n_{\mathrm{R}}}\prod_{j=1}^{n_{\mathrm{T}}} \exp\Big\{-(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}\Big\}\\ &= \exp\Big\{-n_{\mathrm{R}}n_{\mathrm{T}}(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}\Big\}. \quad (46)\end{aligned}$$

Therefore, the distribution $f(a)$ for the exponential channel can be determined by

$$\begin{aligned}f(a) &\doteq (\mathsf{OSNR})^{-\sum_{i=1}^{n_{\mathrm{R}}}\frac{n_{\mathrm{T}}-n_{\mathrm{R}}+2i-1}{2}a_i}\\ &\quad\times \exp\Big\{-n_{\mathrm{R}}n_{\mathrm{T}}(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}\Big\}. \quad (47)\end{aligned}$$

To further simplify (47), we observe that if $a_{n_{\mathrm{R}}} < 0$, the term $\exp\Big\{-n_{\mathrm{R}}n_{\mathrm{T}}(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}\Big\}$ approaches 0 at high OSNR, hence (47) can be simplified as

$$f(a) \doteq 0, \text{ if } a_{n_{\mathrm{R}}} < 0. \quad (48)$$

Otherwise, the term $\exp\Big\{-n_{\mathrm{R}}n_{\mathrm{T}}(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}\Big\}$ approaches $e^{-n_{\mathrm{R}}n_{\mathrm{T}}}$ if $a_{n_{\mathrm{R}}} = 0$, or approaches 1 if $a_{n_{\mathrm{R}}} > 0$. Hence,

$$f(a) \doteq (\mathsf{OSNR})^{-\sum_{i=1}^{n_{\mathrm{R}}}\frac{n_{\mathrm{T}}-n_{\mathrm{R}}+2i-1}{2}a_i}, \text{ if } a_{n_{\mathrm{R}}} \ge 0. \quad (49)$$

Now we apply Theorem 1 to derive the outage probability for the exponential channel. Substituting (48) and (49) into (28), we obtain

$$\begin{aligned}\mathrm{P_{out}}(\mathsf{OSNR}) &\doteq \int_{\mathcal{O}} f(a)da\\ &\doteq \int_{\mathcal{O}\cap\{a_{n_{\mathrm{R}}}\ge 0\}} f(a)da + \int_{\mathcal{O}\cap\{a_{n_{\mathrm{R}}}<0\}} f(a)da\\ &\doteq \int_{\mathcal{O}\cap\{a_{n_{\mathrm{R}}}\ge 0\}} (\mathsf{OSNR})^{-\sum_{i=1}^{n_{\mathrm{R}}}\frac{n_{\mathrm{T}}-n_{\mathrm{R}}+2i-1}{2}a_i}da\\ &\quad + \int_{\mathcal{O}\cap\{a_{n_{\mathrm{R}}}<0\}} 0da\\ &= \int_{\mathcal{O}'} (\mathsf{OSNR})^{-\sum_{i=1}^{n_{\mathrm{R}}}\frac{n_{\mathrm{T}}-n_{\mathrm{R}}+2i-1}{2}a_i}da, \quad (50)\end{aligned}$$

where $\mathcal{O}' = \mathcal{O}\cap\{a_{n_{\mathrm{R}}} \ge 0\}$. Then, combined with relationship

between sets $\mathcal{O}$, $\mathcal{A}$ and $\mathcal{B}$ in (31), we can obtain that the outage diversity gain $d_{\text{out}}^{\text{exp}}(r)$ at multiplexing gain r is bounded by

$$\inf_{a\in\mathcal{A}'} d_{\text{out}}^{\text{exp}}(r,a) \le d_{\text{out}}^{\text{exp}}(r) \le \inf_{a\in\mathcal{B}'} d_{\text{out}}^{\text{exp}}(r,a), \tag{51}$$

where function $d_{\text{out}}^{\text{exp}}(r,a)$ is given by

$$d_{\text{out}}^{\text{exp}}(r,a) = \sum_{i=1}^{n_{\text{R}}} \frac{n_{\text{T}}-n_{\text{R}}+2i-1}{2} a_i, \tag{52}$$

and $\mathcal{A}'$ is given by

$$\mathcal{A}' = \left\{ a : a_1 \ge \cdots \ge a_{n_{\text{R}}} \ge 0, \text{ and } (2n_{\text{R}} - \sum_{i=1}^{n_{\text{R}}} a_i)^+ \le 2r \right\}, \tag{53}$$

and $\mathcal{B}'$ by

$$\mathcal{B}' = \left\{ a : a_1 \ge \cdots \ge a_{n_{\text{R}}} \ge 0, \text{ and } 2n_{\text{R}} - \sum_{i=1}^{n_{\text{R}}} a_i \le 2r \right\}. \tag{54}$$

B. Optimal DMT Characterization

We now characterize the optimal DMT of exponential channel. In fact, our derived outage probability can serve as a lower bound on average error probability. By Fano's inequality [35], we can show

$$\text{P}_{\text{e}}(\mathsf{OSNR}) \dot{\geqslant} (\mathsf{OSNR})^{-d_{\text{out}}(r)}. \tag{55}$$

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

$$\text{P}_{\text{e}}(\mathsf{OSNR}) \dot{\leqslant} (\mathsf{OSNR})^{-d_{\text{out}}(r)} + (\mathsf{OSNR})^{-d_{\text{te}}(r)}. \tag{56}$$

Comparing the bounds in (55) and (56), we have the following observations:

- If $d_{\text{te}}(r) \ge d_{\text{out}}(r)$, $d_{\text{out}}(r)$ becomes the dominant term in the RHS of (56). Combined with (55), the optimal diversity gain $d^\star(r)$ is equal to the outage diversity gain, i.e.,
$$d^\star(r) = d_{\text{out}}(r). \tag{57}$$
- If $d_{\text{te}}(r) < d_{\text{out}}(r)$, $d_{\text{te}}(r)$ turns to be the dominant term in the RHS of (56). In this case, we can simply bound $d^\star(r)$ as
$$d_{\text{te}}(r) \le d^\star(r) \le d_{\text{out}}(r). \tag{58}$$

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

Theorem 3 (DMT of Exponential Channel)**.** Given a channel with distribution in (42), if $l \ge l_{\text{exp}} = n_{\text{T}} - n_{\text{R}} + 1$, the optimal diversity gain is given by

$$d_{\text{exp}}^\star(r) = (n_{\text{T}}-n_{\text{R}}+1)(n_{\text{R}}-r); \tag{59}$$

otherwise,

$$l(n_{\text{R}}-r) \le d_{\text{exp}}^\star(r) \le (n_{\text{T}}-n_{\text{R}}+1)(n_{\text{R}}-r). \tag{60}$$

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

By linear optimization, it is straightforward to obtain that the infimums of $d_{\text{out}}^{\text{exp}}(r,a)$ over feasible regimes $\mathcal{A}'$ and $\mathcal{B}'$ are both achieved at $a_{\text{out}}^\star = [2(n_{\text{R}}-r), 0, \ldots, 0]$ in (51), and the infimums match. Hence the outage diversity gain is given by

$$d_{\text{out}}^{\text{exp}}(r) = (n_{\text{T}}-n_{\text{R}}+1)(n_{\text{R}}-r). \tag{61}$$

Now we calculate $d_{\text{te}}^{\text{exp}}(r)$. By reducing the integral domain $\mathcal{O}^c$ in (35) to $\mathcal{A}^c$ and enlarging the integral domain $\mathcal{O}^c$ in (35) to $\mathcal{B}^c$, and combined with the expression of $f(a)$ in (48). we bound $d_{\text{te}}^{\text{exp}}(r)$ as

$$\inf_{(\mathcal{B}^c)'} d_{\text{te}}^{\text{exp}}(r,a) \le d_{\text{te}}^{\text{exp}}(r) \le \inf_{(\mathcal{A}^c)'} d_{\text{te}}^{\text{exp}}(r,a), \tag{62}$$

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

$$\begin{aligned} d_{\text{te}}^{\text{exp}}(r,a) = &\sum_{i=1}^{n_{\text{R}}} \frac{n_{\text{T}}-n_{\text{R}}+2i-1}{2} a_i \\ &+ \frac{l}{2}\left(\sum_{i=1}^{n_{\text{R}}} (2-a_i)^+ - 2r \right), \end{aligned} \tag{63}$$

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

$$(\mathcal{A}^c)' = \left\{ a : a_1 \ge \cdots \ge a_{n_{\text{R}}} \ge 0, \text{ and } (2n_{\text{R}} - \sum_{i=1}^{n_{\text{R}}} a_i)^+ \ge 2r \right\}, \tag{64}$$

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

$$(\mathcal{B}^c)' = \left\{ a : a_1 \ge \cdots \ge a_{n_{\text{R}}} \ge 0, \text{ and } 2n_{\text{R}} - \sum_{i=1}^{n_{\text{R}}} a_i \ge 2r \right\}. \tag{65}$$

Compared with (52), we can see $d_{\text{te}}^{\text{exp}}(r,a)$ in (63) is also linear with respect to vector a, but with an extra term $\frac{l}{2}\left(\sum_{i=1}^{n_{\text{R}}}(2-a_i)^+ - 2r\right)$. Given a block length l, we can still analyze the infimum of $d_{\text{te}}^{\text{exp}}(r,a)$ over $(\mathcal{A}^c)'$ or $(\mathcal{B}^c)'$ through linear optimization. Here, we show that the infimums in (62) also match the same optimal point $a_{\text{te}}^\star$. In the following, we characterize $a_{\text{te}}^\star$ under different block length l:

- when $l < n_{\text{T}} - n_{\text{R}} + 1$, the optimal $a_{\text{te}}^\star$ is given by
$$a_i^\star = 0, \text{ for } i = 1, \cdots, n_{\text{R}}; \tag{66}$$
- when $n_{\text{T}} - n_{\text{R}} + 2k - 1 \le l < n_{\text{T}} - n_{\text{R}} + 2(k+1) - 1$, $\forall k \in \{1, \cdots, n_{\text{R}} - r\}$, the optimal $a_{\text{te}}^\star$ is given by
$$a_i^\star = 2, \text{ for } i = 1, \cdots, k, \tag{67a}$$
$$a_i^\star = 0, \text{ for } i = k+1, \cdots, n_{\text{R}}; \tag{67b}$$
- when $l \ge n_{\text{T}} - n_{\text{R}} + 2((n_{\text{R}}-r)+1) - 1$, the optimal $a_{\text{te}}^\star$ is given by
$$a_i^\star = 2, \text{ for } i = 1, \cdots, n_{\text{R}} - r, \tag{68a}$$
$$a_i^\star = 0, \text{ for } i = n_{\text{R}} - r + 1, \cdots, n_{\text{R}}. \tag{68b}$$

Substituting $a_{\text{te}}^\star$ into (62), we obtain that if $l < n_{\text{T}} - n_{\text{R}} + 1$, then

$$d_{\text{te}}^{\text{exp}}(r) = l(n_{\text{R}}-r). \tag{69}$$

This article has been accepted for publication in IEEE Transactions on Information Theory. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIT.2023.3244567

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

$$d_{\mathrm{te}}^{\mathrm{exp}}(r) = \sum_{i=1}^{k}(n_{\mathrm{T}} - n_{\mathrm{R}} + 2i - 1) + l(n_{\mathrm{R}} - k - r),\ k \in \{1, \cdots, n_{\mathrm{R}} - r\}, \quad (70)$$

where the choice of k depends on the channel block length l (see (66), (67), and (68)).

Last, we characterize $d_{\mathrm{exp}}^{\star}(r)$. If $l \geq n_{\mathrm{T}} - n_{\mathrm{R}} + 1$ and at any k in set $\{1, \cdots, n_{\mathrm{R}} - r\}$,

$$\begin{aligned} d_{\mathrm{te}}^{\mathrm{exp}}(r) &\geq \sum_{i=1}^{k}(n_{\mathrm{T}} - n_{\mathrm{R}} + 2 - 1) \\ &\quad + (n_{\mathrm{T}} - n_{\mathrm{R}} + 1)(n_{\mathrm{R}} - k - r) \\ &= (n_{\mathrm{T}} - n_{\mathrm{R}} + 1)(n_{\mathrm{R}} - r) \\ &= d_{\mathrm{out}}^{\mathrm{exp}}(r). \end{aligned} \quad (71)$$

The proof in this case is concluded by combining (57) with (61) and (71).

If $l < n_{\mathrm{T}} - n_{\mathrm{R}} + 1$, by (66) and (61), we have

$$\begin{aligned} d_{\mathrm{te}}^{\mathrm{exp}}(r) &< (n_{\mathrm{T}} - n_{\mathrm{R}} + 1)(n_{\mathrm{R}} - r) \\ &= d_{\mathrm{out}}^{\mathrm{exp}}(r). \end{aligned} \quad (72)$$

Combing (58) with (61), (69) and (72), the proof is concluded in this case. ∎

Remark 3. Notice that in (51), the difference in the upper and lower bounds on $d_{\mathrm{out}}^{\mathrm{exp}}(r)$ lies in the optimization domain. Since $\mathcal{B} \subseteq \mathcal{O} \subseteq \mathcal{A}$, the results above show that the infimums over different optimization domains are achieved at the same point, revealing that the bounds derived in Proposition 1 are sufficiently tight in terms of diversity gain.

VI. 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.

A. Outage Diversity Gain Characterization

Due to the presence of $K_{\tau}(\cdot)$ in (7), it is difficult to analyze original gamma-gamma distribution directly. Instead, we choose an alternative expression in terms of power series, i.e.,

$$\begin{aligned} f(h_{ij}^{r}) &= \sum_{n=0}^{\infty} d_n(\rho_1, \rho_2)(h_{ij}^{r})^{n+\rho_2-1} \\ &\quad + \sum_{n=0}^{\infty} d_n(\rho_2, \rho_1)(h_{ij}^{r})^{n+\rho_1-1}, \end{aligned} \quad (73)$$

where

$$d_n(\rho_1, \rho_2) = \frac{(\rho_1\rho_2)^{\rho_2+n}\Gamma(\rho_1 - \rho_2)}{\Gamma(\rho_1)\Gamma(\rho_2)(1 + \rho_2 - \rho_1)_n n!}. \quad (74)$$

Combined with (5) and (41), we obtain the distribution of gamma-gamma channel as

$$\begin{aligned} f(\mathrm{H}) = \prod_{i=1}^{n_{\mathrm{R}}}\prod_{j=1}^{n_{\mathrm{T}}}\Bigg\{ &\sum_{n=0}^{\infty} d_n(\rho_1, \rho_2)\left(e^{\nu d_{ij}} h_{ij}\right)^{n+\rho_2-1} e^{\nu d_{ij}} \\ &+ \sum_{n=0}^{\infty} d_n(\rho_2, \rho_1)\left(e^{\nu d_{ij}} h_{ij}\right)^{n+\rho_1-1} e^{\nu d_{ij}} \Bigg\}. \end{aligned} \quad (75)$$

Recall that $h_{ij} \doteq (\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}$, thus we can rewrite (75) as

$$\begin{aligned} &f(\mathrm{H}) \doteq \\ &\prod_{i=1}^{n_{\mathrm{R}}}\prod_{j=1}^{n_{\mathrm{T}}}\Bigg\{ \sum_{n=0}^{\infty} d_n(\rho_1, \rho_2)\left(e^{\nu d_{ij}}\right)^{n+\rho_2}(\mathsf{OSNR})^{-\frac{n+\rho_2-1}{2}a_{n_{\mathrm{R}}}} \\ &\quad + \sum_{n=0}^{\infty} d_n(\rho_2, \rho_1)\left(e^{\nu d_{ij}}\right)^{n+\rho_1}(\mathsf{OSNR})^{-\frac{n+\rho_1-1}{2}a_{n_{\mathrm{R}}}} \Bigg\}. \end{aligned} \quad (76)$$

Note that if $a_{n_{\mathrm{R}}} < 0$, then $f(\mathrm{H}) \to \infty$, which is impossible. Hence we can constrain $a_{n_{\mathrm{R}}}$ to be $a_{n_{\mathrm{R}}} \geq 0$. Hence, after some algebraic manipulations, we can simplify $f(\mathrm{H})$ to

$$f(\mathrm{H}) \doteq (\mathsf{OSNR})^{-\frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}(\rho-1)}, \quad a_{n_{\mathrm{R}}} \geq 0. \quad (77)$$

where ρ denotes $\min\{\rho_1, \rho_2\}$.

Substituting (77) into (26) and (27), the outage diversity gain $d_{\mathrm{out}}^{\mathrm{gg}}(r)$ at multiplexing gain r is bounded by

$$\inf_{a \in \mathcal{A}'} d_{\mathrm{out}}^{\mathrm{gg}}(r, a) \leq d_{\mathrm{out}}^{\mathrm{gg}}(r) \leq \inf_{a \in \mathcal{B}'} d_{\mathrm{out}}^{\mathrm{gg}}(r, a), \quad (78)$$

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

$$\begin{aligned} d_{\mathrm{out}}^{\mathrm{gg}}(r, a) &= \sum_{i=1}^{n_{\mathrm{R}}} \frac{(n_{\mathrm{T}} - n_{\mathrm{R}} + 2i - 1)}{2} a_i \\ &\quad + \frac{n_{\mathrm{R}}n_{\mathrm{T}}a_{n_{\mathrm{R}}}}{2}(\rho - 1), \end{aligned} \quad (79)$$

and sets $\mathcal{A}'$ and $\mathcal{B}'$ are defined as in (53) and (54), respectively.

B. Optimal DMT Characterization

We present the main results in the following theorem.

Theorem 4 (DMT of Gamma-Gamma Channel)**.** Given a channel with distribution in (75), when $n_{\mathrm{R}} = 1$, if $l \geq l_{\mathrm{gg}} = \rho n_{\mathrm{T}}$, the optimal diversity gain is

$$d_{\mathrm{gg}}^{\star}(r) = \rho n_{\mathrm{T}}(1 - r); \quad (80)$$

otherwise,

$$l(1 - r) \leq d_{\mathrm{gg}}^{\star}(r) \leq \rho n_{\mathrm{T}}(1 - r). \quad (81)$$

When $n_{\mathrm{R}} > 1$, if $l \geq l_{\mathrm{gg}} = n_{\mathrm{T}} - n_{\mathrm{R}} + 1$, then

$$d_{\mathrm{gg}}^{\star}(r) = (n_{\mathrm{T}} - n_{\mathrm{R}} + 1)(n_{\mathrm{R}} - r); \quad (82)$$

otherwise,

$$l(n_{\mathrm{R}} - r) \leq d_{\mathrm{gg}}^{\star}(r) \leq (n_{\mathrm{T}} - n_{\mathrm{R}} + 1)(n_{\mathrm{R}} - r). \quad (83)$$

Proof: We follow the similar arguments as in the proof of Theorem 3. Here we mainly emphasize the differences.

Compared with the function in (52), $d_{\text{out}}^{\text{gg}}(r,a)$ in (79) 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 $n_{\text{R}}=1$, the infimums in (78) are both achieved at point $a_{\text{out}}^{\star}=2(n_{\text{R}}-r)$. Hence, the outage diversity gain is given by

$$\begin{aligned} d_{\text{out}}^{\text{gg}}(r) &= (n_{\text{T}}-n_{\text{R}}+1+n_{\text{R}}n_{\text{T}}(\rho-1))(n_{\text{R}}-r) \\ &= \rho n_{\text{T}}(1-r). \end{aligned} \tag{84}$$

When $n_{\text{R}}>1$, the infimums in (78) are both achieved at point $a_{\text{out}}^{\star}=[2(n_{\text{R}}-r),0,\ldots,0]$, and we have

$$d_{\text{out}}^{\text{gg}}(r)=(n_{\text{T}}-n_{\text{R}}+1)(n_{\text{R}}-r). \tag{85}$$

Now we calculate $d_{\text{te}}^{\text{gg}}(r)$. Modifying the optimization domain in (35) with $\mathcal{A}^c$ or $\mathcal{B}^c$, we have

$$\inf_{(\mathcal{B}^c)'} d_{\text{te}}^{\text{gg}}(r,a) \le d_{\text{te}}^{\text{gg}}(r) \le \inf_{(\mathcal{A}^c)'} d_{\text{te}}^{\text{gg}}(r,a), \tag{86}$$

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

$$\begin{aligned} d_{\text{te}}^{\text{gg}}(r,a) = &\sum_{i=1}^{n_{\text{R}}} \frac{n_{\text{T}}-n_{\text{R}}+2i-1}{2}a_i + \frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}}{2}(\rho-1) \\ &+ \frac{l}{2}\left(\sum_{i=1}^{n_{\text{R}}}(2-a_i)^+ - 2r\right), \end{aligned} \tag{87}$$

and sets $(\mathcal{A}^c)'$ and $(\mathcal{B}^c)'$ are defined as in (64) and (65), respectively.

Compared with the function in (63), $d_{\text{te}}^{\text{gg}}(r,a)$ also has an extra term $n_{\text{R}}n_{\text{T}}(\rho-1)a_{n_{\text{R}}}/2$. We still separately consider the cases when $n_{\text{R}}=1$ and $n_{\text{R}}>1$.

When $n_{\text{R}}=1$, if $l<n_{\text{T}}-n_{\text{R}}+1+n_{\text{R}}n_{\text{T}}(\rho-1)=\rho n_{\text{T}}$, the infimums in (86) are both achieved at $a_{\text{te}}^{\star}=0$, which yields

$$d_{\text{te}}^{\text{gg}}(r)=l(n_{\text{R}}-r); \tag{88}$$

otherwise, the infimums in (86) are both achieved at $a_{\text{te}}^{\star}=2(n_{\text{R}}-r)$, thus

$$\begin{aligned} d_{\text{te}}^{\text{gg}}(r) &= (n_{\text{T}}-n_{\text{R}}+1+n_{\text{R}}n_{\text{T}}(\rho-1))(n_{\text{R}}-r) \\ &= \rho n_{\text{T}}(1-r). \end{aligned} \tag{89}$$

When $n_{\text{R}}>1$, since sets $(\mathcal{A})'$ and $(\mathcal{B})'$ in (86) are the same as the corresponding sets in the exponential channel in (62), here we just need to show $d_{\text{te}}^{\text{gg}}(r)=d_{\text{te}}^{\text{exp}}(r)$. On the one hand, as $n_{\text{R}}n_{\text{T}}(\rho-1)a_{n_{\text{R}}}/2\ge 0$ for any $a_{n_{\text{R}}}\ge 0$, we can obtain that $d_{\text{te}}^{\text{gg}}(r,a)\ge d_{\text{te}}^{\text{exp}}(r,a)$ at any a with $a_{n_{\text{R}}}\ge 0$. On the other hand, at the point $a_{\text{te}}^{\star}$ defined in (66), (67), and (68), we have $d_{\text{te}}^{\text{gg}}(r,a_{\text{te}}^{\star})=d_{\text{te}}^{\text{exp}}(r,a_{\text{te}}^{\star})$. The proof is concluded. ∎

Remark 4. Notice that when $n_{\text{R}}>1$, the n_{R}th entry of $a_{\text{out}}^{\star}$ or $a_{\text{te}}^{\star}$ is always 0, which indicates the extra term $n_{\text{R}}n_{\text{T}}(\rho-1)a_{n_{\text{R}}}/2$ plays no role on $d_{\text{out}}^{\text{gg}}(r,a)$ and $d_{\text{te}}^{\text{gg}}(r,a)$. Hence when $n_{\text{R}}>1$, we always have $d_{\text{out}}^{\text{gg}}(r)=d_{\text{out}}^{\text{exp}}(r)$ and $d_{\text{te}}^{\text{gg}}(r)=d_{\text{te}}^{\text{exp}}(r)$.

VII. 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 gain and the optimal DMT are characterized in the following.

A. Outage Diversity Gain Characterization

Substituting (5) and (8) into (41), we have

$$f(\mathrm{H}) = \prod_{i=1}^{n_{\text{R}}}\prod_{j=1}^{n_{\text{T}}}\left\{\frac{1}{h_{ij}\sqrt{2\pi\sigma_l^2}}\exp\left(-\frac{\left(\log(\exp(\nu d_{ij})h_{ij})-\mu_l\right)^2}{2\sigma_l^2}\right)\right\}. \tag{90}$$

Following similar arguments as in Section V-A, we simplify $f(\mathrm{H})$ to

$$\begin{aligned} &f(\mathrm{H}) \\ &\doteq (\mathsf{OSNR})^{\frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}}{2}} \\ &\quad\times\prod_{i=1}^{n_{\text{R}}}\prod_{j=1}^{n_{\text{T}}}\left\{\exp\left(-\frac{\left(\log\left(\exp(\nu d_{ij})(\mathsf{OSNR})^{-\frac{a_{n_{\text{R}}}}{2}}\right)-\mu_l\right)^2}{2\sigma_l^2}\right)\right\} \\ &\doteq (\mathsf{OSNR})^{\frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}}{2}} \\ &\quad\times\prod_{i=1}^{n_{\text{R}}}\prod_{j=1}^{n_{\text{T}}}\left\{\exp\left(-\frac{\left(\log\left(\mathsf{OSNR}\right)^{-\frac{a_{n_{\text{R}}}}{2}}-\left(\mu_l-\nu d_{ij}\right)\right)^2}{2\sigma_l^2}\right)\right\} \\ &= (\mathsf{OSNR})^{\frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}}{2}}\times\exp\Bigg\{-\log(\mathsf{OSNR}) \\ &\quad\times\frac{n_{\text{R}}n_{\text{T}}}{2\sigma_l^2}\left(\frac{a_{n_{\text{R}}}^2}{4}\log(\mathsf{OSNR})+a_{n_{\text{R}}}\beta_1+\frac{\beta_2}{\log(\mathsf{OSNR})}\right)\Bigg\} \\ &= (\mathsf{OSNR})^{\frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}}{2}-\frac{n_{\text{R}}n_{\text{T}}}{2\sigma_l^2}\left(\frac{a_{n_{\text{R}}}^2}{4}\log(\mathsf{OSNR})+a_{n_{\text{R}}}\beta_1+\frac{\beta_2}{\log(\mathsf{OSNR})}\right)} \qquad (91) \\ &\doteq (\mathsf{OSNR})^{-\frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}^2}{8\sigma_l^2}\log(\mathsf{OSNR})}, \qquad (92) \end{aligned}$$

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

$$\beta_1=\frac{\sum_{i=1}^{n_{\text{R}}}\sum_{j=1}^{n_{\text{T}}}(\mu_l-\nu d_{ij})}{n_{\text{R}}n_{\text{T}}}, \tag{93}$$

$$\beta_2=\frac{\sum_{i=1}^{n_{\text{R}}}\sum_{j=1}^{n_{\text{T}}}(\mu_l-\nu d_{ij})^2}{n_{\text{R}}n_{\text{T}}}, \tag{94}$$

and (91) follows from

$$\begin{aligned} (\mathsf{OSNR})^{\log_{\mathsf{OSNR}}e^{b\times\log(\mathsf{OSNR})}} &= (\mathsf{OSNR})^{b\times(\log_{\mathsf{OSNR}}e^{\log(\mathsf{OSNR})})} \qquad (95) \\ &= (\mathsf{OSNR})^{b}. \qquad (96) \end{aligned}$$

Substituting (92) into (26) and (27), we bound the outage diversity gain as

$$\inf_{a\in\mathcal{A}} d_{\text{out}}^{\text{ln}}(r,a) \le d_{\text{out}}^{\text{ln}}(r) \le \inf_{a\in\mathcal{B}} d_{\text{out}}^{\text{ln}}(r,a), \tag{97}$$

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

$$d_{\text{out}}^{\text{ln}}(r,a)=\sum_{i=1}^{n_{\text{R}}}\frac{(n_{\text{T}}-n_{\text{R}}+2i-1)}{2}a_i+\frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}^2}{8\sigma_l^2}\log(\mathsf{OSNR}), \tag{98}$$

and sets $\mathcal{A}$ and $\mathcal{B}$ are defined as in (29) and (30), respectively.

B. Optimal DMT Characterization

We present the main results in the following theorem.

Theorem 5 (DMT of Log-Normal Channel)**.** Given a channel with distribution in (90), and denote $\theta\triangleq\log(\mathsf{OSNR})/2\sigma_l^2+1$. When $n_{\text{R}}=1$, if $l\geq l_{\text{ln}}=\theta n_{\text{T}}$, the optimal diversity gain is

$$d_{\text{ln}}^{\star}(r)=\theta n_{\text{T}}(1-r), \tag{99}$$

otherwise,

$$l(1-r)\leq d_{\text{ln}}^{\star}(r)\leq\theta n_{\text{T}}(1-r). \tag{100}$$

When $n_{\text{R}}>1$, if $l\geq l_{\text{ln}}=n_{\text{T}}-n_{\text{R}}+1$, then

$$d_{\text{ln}}^{\star}(r)=(n_{\text{T}}-n_{\text{R}}+1)(n_{\text{R}}-r); \tag{101}$$

otherwise,

$$l(n_{\text{R}}-r)\leq d_{\text{ln}}^{\star}(r)\leq(n_{\text{T}}-n_{\text{R}}+1)(n_{\text{R}}-r). \tag{102}$$

Proof: We first calculate $d_{\text{out}}^{\text{ln}}(r)$. In (98), $d_{\text{out}}^{\text{ln}}(r,a)$ is a nonlinear function to vector a due to the extra term

$$u(a_{n_{\text{R}}})\triangleq\frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}^2}{8\sigma_l^2}\log(\mathsf{OSNR}). \tag{103}$$

Notice that at high OSNR, $u(a_{n_{\text{R}}})$ tends to be infinite, and dominates the sum of rest terms. For this reason, we prioritize small $a_{n_{\text{R}}}^2$ to optimize $d_{\text{out}}^{\text{ln}}(r,a)$ over $\mathcal{A}$ and $\mathcal{B}$ in (97). Since this extra term $u(a_{n_{\text{R}}})$ only relates to $a_{n_{\text{R}}}$, we still separately consider $n_{\text{R}}=1$ and $n_{\text{R}}>1$ to distinguish $a_1=a_{n_{\text{R}}}$ and $a_1\neq a_{n_{\text{R}}}$.

When $n_{\text{R}}=1$, the integral domains $\mathcal{A}$ and $\mathcal{B}$ in (27) coincide as $\{a\geq 2(n_{\text{R}}-r)\}$. We can obtain the optimal point $a_{\text{out}}^{\star}=2(n_{\text{R}}-r)$, and we have

$$d_{\text{out}}^{\text{ln}}(r)=\left(n_{\text{T}}-n_{\text{R}}+1+\frac{n_{\text{R}}n_{\text{T}}}{2\sigma_l^2}\log(\mathsf{OSNR})\right)(n_{\text{R}}-r)=\theta n_{\text{T}}(1-r). \tag{104}$$

When $n_{\text{R}}>1$, a necessary condition to have finite infimums in (97) is the n_{R}th entry of a to be zero, i.e., $a_{n_{\text{R}}}=0$. Then the rest of entries in optimal point $a_{\text{out}}^{\star}$ can be computed by linear optimization, and we obtain $a_{\text{out}}^{\star}=[2(n_{\text{R}}-r),0,\cdots,0]$. Substituting $a_{\text{out}}^{\star}$ into (97), the outage diversity gain is given by

$$d_{\text{out}}^{\text{ln}}(r)=(n_{\text{T}}-n_{\text{R}}+1)(n_{\text{R}}-r). \tag{105}$$

Now we calculate $d_{\text{te}}^{\text{ln}}(r)$. Modifying the integral domain in (35) as $\mathcal{A}^c$ or $\mathcal{B}^c$, we have

$$\inf_{\mathcal{B}^c}d_{\text{te}}^{\text{ln}}(r,a)\leq d_{\text{te}}^{\text{ln}}(r)\leq\inf_{\mathcal{A}^c}d_{\text{te}}^{\text{ln}}(r,a), \tag{106}$$

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

$$d_{\text{te}}^{\text{ln}}(r,a)=\sum_{i=1}^{n_{\text{R}}}\frac{n_{\text{T}}-n_{\text{R}}+2i-1}{2}a_i+\frac{n_{\text{R}}n_{\text{T}}a_{n_{\text{R}}}^2}{8\sigma_l^2}\log(\mathsf{OSNR})+\frac{l}{2}\left(\sum_{i=1}^{n_{\text{R}}}(2-a_i)^+-2r\right). \tag{107}$$

It is direct to see that $d_{\text{te}}^{\text{ln}}(r,a)$ is also a nonlinear function due to the extra term $u(a_{n_{\text{R}}})$. Using similar method argued above, we prioritize small $a_{n_{\text{R}}}^2$ to optimize $d_{\text{te}}^{\text{ln}}(r,a)$ over $\mathcal{A}^c$ and $\mathcal{B}^c$. We show that when $n_{\text{R}}=1$, if $l<\theta n_{\text{T}}$, both infimums of $d_{\text{te}}^{\text{ln}}(r,a)$ over $\mathcal{A}^c$ and $\mathcal{B}^c$ are achieved at $a_{\text{te}}^{\star}=0$, which yields

$$d_{\text{te}}^{\text{ln}}(r)=l(n_{\text{R}}-r); \tag{108}$$

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

$$d_{\text{te}}^{\text{ln}}(r)=\left(n_{\text{T}}-n_{\text{R}}+1+\frac{n_{\text{R}}n_{\text{T}}}{2\sigma_l^2}\log(\mathsf{OSNR})\right)(n_{\text{R}}-r)=\theta n_{\text{T}}(1-r). \tag{109}$$

When $n_{\text{R}}>1$, following similar arguments as in the derivation of $d_{\text{out}}^{\text{ln}}(r)$, we obtain the optimal point $a_{\text{te}}^{\star}$ has the same expressions as in (66), (67), and (68), and $d_{\text{te}}^{\text{ln}}(r)$ has the same expressions as in (69) and (70). Then the proof is concluded. ■

Remark 5. The term $u(a_{n_{\text{R}}})$ causes the non-linearity of $d_{\text{out}}^{\text{ln}}(r,a)$ in (98). When $n_{\text{R}}>1$, a necessary condition to have a finite $d_{\text{out}}^{\text{ln}}(r)$ is $a_{n_{\text{R}}}=0$, and hence $u(a_{n_{\text{R}}})=0$. Eliminating the influence of $a_{n_{\text{R}}}$, the optimization problems in (97) are essentially equivalent to the linear problems in (51).

One issue in the above results is that when $n_{\text{R}}=1$ and $r=0$, the outage diversity gain $d_{\text{out}}^{\text{ln}}(r)$ or optimal diversity gain $d_{\text{ln}}^{\star}(r)$ tends to infinity at high OSNR, which reveals that the conventional definitions in (15) and (16) are not applicable for the log-normal channel. In fact, similar phenomena are also observed in the existing literature [12]–[14], [16]. This issue is due to the fact that $d_{\text{out}}^{\text{ln}}(r)$ and $d_{\text{ln}}^{\star}(r)$ contain the OSNR-related term, while they essentially measure the system performance at high OSNR. To deal with this problem, we modify the conventional definitions in (15) and (16), and adopt the asymptotically relative outage diversity gain and asymptotically relative diversity gain, which are proposed in [12]. The asymptotically relative outage diversity gain is defined as:

$$\hat{d}_{\text{out}}^{\text{ln}}(r)=\lim_{\mathsf{OSNR}\to\infty}\frac{d_{\text{out}}^{\text{ln}}(r)}{d_{\text{BM}}(r)}, \tag{110}$$

and the asymptotically relative diversity gain is defined as:

$$\hat{d}_{\text{ln}}^{\star}(r)=\lim_{\mathsf{OSNR}\to\infty}\frac{d_{\text{ln}}^{\star}(r)}{d_{\text{BM}}(r)}, \tag{111}$$

where $d_{\text{BM}}(r)$ denotes the benchmark. Referring to [13], [16], we instantiate the benchmark to be the outage diversity gain at $r=0$ and $n_{\text{T}}=n_{\text{R}}=1$: $d_{\text{ln,out}}^{\text{SISO}}(0)$, i.e.,

$$\hat{d}_{\text{out}}^{\text{ln}}(r)=\lim_{\mathsf{OSNR}\to\infty}\frac{d_{\text{out}}^{\text{ln}}(r)}{d_{\text{ln,out}}^{\text{SISO}}(0)}, \tag{112}$$

and

$$\hat{d}^{\star}_{\text{ln}}(r) = \lim_{\text{OSNR}\to\infty} \frac{d^{\star}_{\text{ln}}(r)}{d^{\text{SISO}}_{\text{ln,out}}(0)}. \tag{113}$$

Substituting $d^{\text{ln}}_{\text{out}}(r)$ in (104) into (112), and $d^{\star}_{\text{ln}}(r)$ in (99) and (100) into (113), we have the following theorem.

Theorem 6. Given a MISO channel with distribution in (90), under a large block length we have

$$\hat{d}^{\star}_{\text{ln}}(r) = \hat{d}^{\text{ln}}_{\text{out}}(r) = n_{\text{T}}(1-r), \tag{114}$$

under a small block length we have

$$0 \le \hat{d}^{\star}_{\text{ln}}(r) \le n_{\text{T}}(1-r). \tag{115}$$

VIII. Numerical Results and Discussions

A. Optimal DMT Analysis

In this section, we present numerical examples of the derived DMT results. Figures 1-3 depict the curve of (asymptotically relative) outage diversity gain under different fadings for MISO and MIMO channels. Among them, Figure 2 corresponds to the asymptotically relative outage diversity gain $\hat{d}_{\text{out}}(r)$ for the log-normal fading channel. It should be noted that these curves in Figures 1-3 can also represent the optimal (asymptotically relative) diversity gain when the channel block length satisfies $l \ge l_{\text{th}}$, where l_{th} represents the block length threshold of corresponding channels, such as l_{exp}, l_{gg}, l_{ln}.

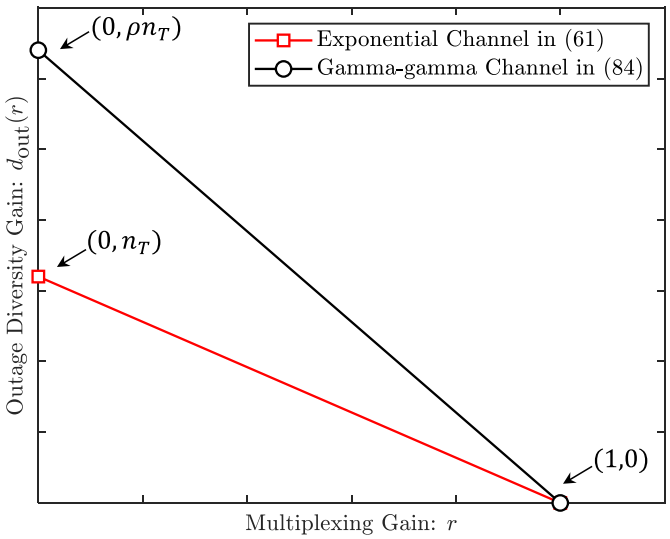

Fig. 1: Outage diversity gain for $n_{\text{T}} \times 1$ MISO channel.

From Figure 1, we observe that the outage diversity gain of the exponential channel is less than that of gamma-gamma channel at the same multiplexing gain, which is due to the fact that gamma-gamma and exponential distributions correspond to moderate and strong atmospheric turbulence intensities, respectively. Hence, we can conclude that increasing turbulence intensity may hamper the diversity gain. We also observe that the maximum (asymptotically relative) outage and multiplexing gains are always achieved at $r = 0$ and $r = n_{\text{R}}$, respectively. Moreover, when $n_{\text{R}} > 1$, the outage

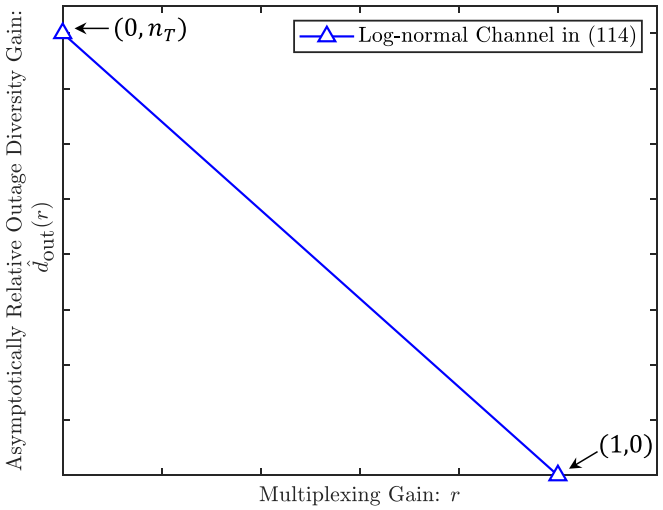

Fig. 2: Asymptotically relative outage diversity gain for $n_{\text{T}} \times 1$ MISO channel.

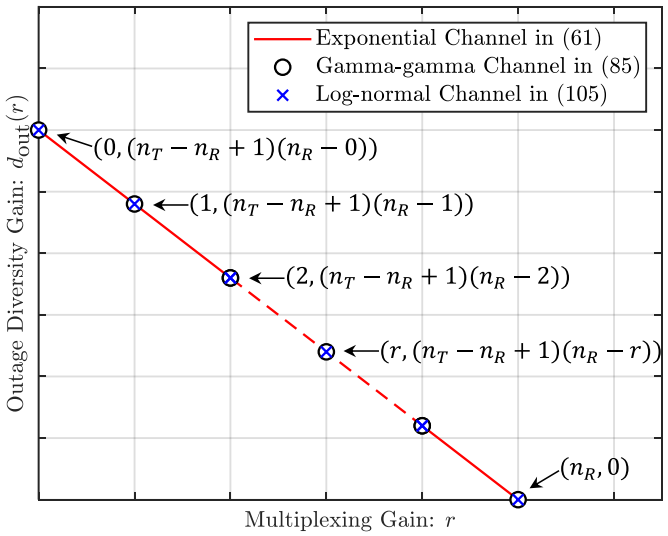

Fig. 3: Outage diversity gain for $n_{\text{T}} \times n_{\text{R}}$ MIMO channel ($n_{\text{R}} > 1$).

diversity gains are equal under these three different fading MIMO channels.

Although different distributions may lead to different outage probabilities, Figure 3 reveals that the outage diversity gains are identical regardless of the fading properties when $n_{\text{R}} > 1$. As a result, similar phenomenon has been appeared in [11, Thm. 21]. Specifically, at high OSNR, as long as the exponents of OSNR corresponding to the outage probabilities are identical, the derived outage diversity gains for different channels will be the same.

When $l < l_{\text{th}}$, Figure 4 depicts the bounds on optimal $d^{\star}(r)$ for a 16×4 MIMO exponential channel when $l = 7$ and $l = 10$. Here, $l_{\text{th}} = 13$. From Figure 4, it is direct to see $d^{\star}(r)$ is upper-bounded by $d_{\text{out}}(r)$ and lower-bounded by $d_{\text{te}}(r)$. For any $l < l_{\text{th}}$ and $r < n_{\text{R}}$, the lower bound on diversity gain is strictly below the upper bound, and the difference between the

This article has been accepted for publication in IEEE Transactions on Information Theory. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIT.2023.3244567

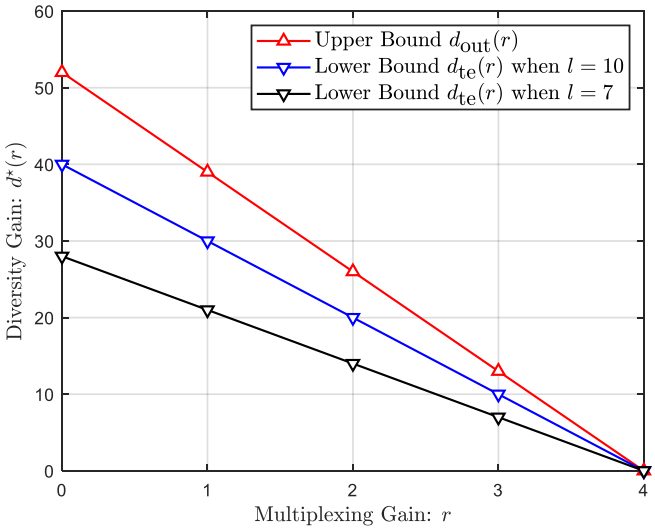

Fig. 4: Bounds on optimal diversity gain for 16×4 MIMO channel with small block length.

upper and lower bounds decreases as the block length or the multiplexing gain increases. Combined with (60), the maximal upper limit of the block length and multiplexing gain when the difference between $d_{\text{out}}(r)$ and $d_{\text{te}}(r)$ is minimal are given by $l = n_{\text{T}} - n_{\text{R}} + 1$ and $r = n_{\text{R}}$, respectively.

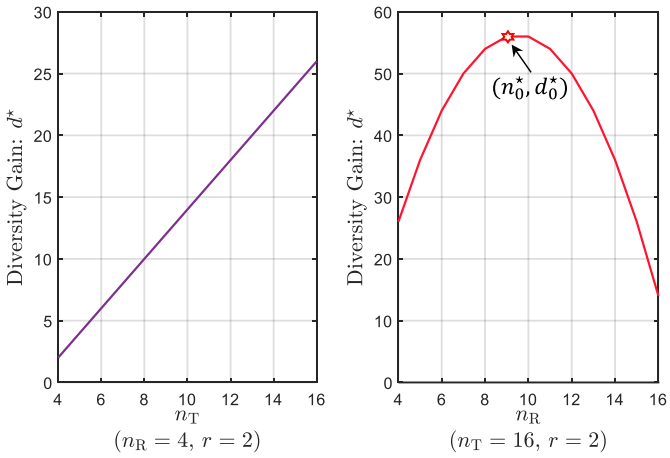

Fig. 5: Optimal diversity gain at fixed n_{R} or n_{T}.

We also depict the optimal $d^\star(r)$ as a function of the number of transmit antenna n_{T} or receive antenna n_{R} in Figure 5. The left curve of Figure 5 depicts the optimal diversity gain $d^\star(r)$ as a function of n_{T} with fixed parameters $n_{\text{R}} = 4$ and $r = 2$, while the right curve depicts $d^\star(r)$ as a function of n_{R} with fixed parameters $n_{\text{T}} = 16$ and $r = 2$. We observe that $d^\star(r)$ monotonically increases as n_{T} increases, while increasing n_{R} does not always guarantee larger diversity gain. In fact, our derived $d^\star(r)$ is a linear function with n_{T}, but a second-order function of n_{R}, which is fundamentally different from the RF DMT result established in [9]. Since the optimal diversity gain $d^\star(r)$ is a second-order function of n_{R}, hence, installing $n_0^\star = \lfloor \frac{n_{\text{T}}+1+r}{2} \rfloor$ antennas at the receiving end may achieve the best performance.

B. DMT Comparison with RF Results

In this section, we present numerical comparisons between the OWC exponential and RF Rayleigh channels [9]. Figures 6 and 7 depict their DMT curves under a large channel block length. For any multiplexing gain r, we can observe that the optimal diversity gain $d^\star(r)$ of the Rayleigh channel is always larger than that of the exponential channel. This implies that the optical signals usually suffer more severe fading than their traditional RF counterparts.

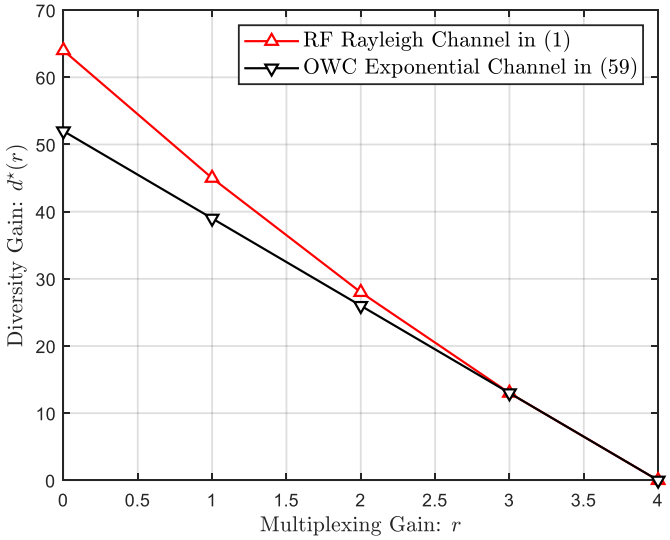

Fig. 6: Comparisons of the optimal DMT curve for 16×4 MIMO channel with large block length.

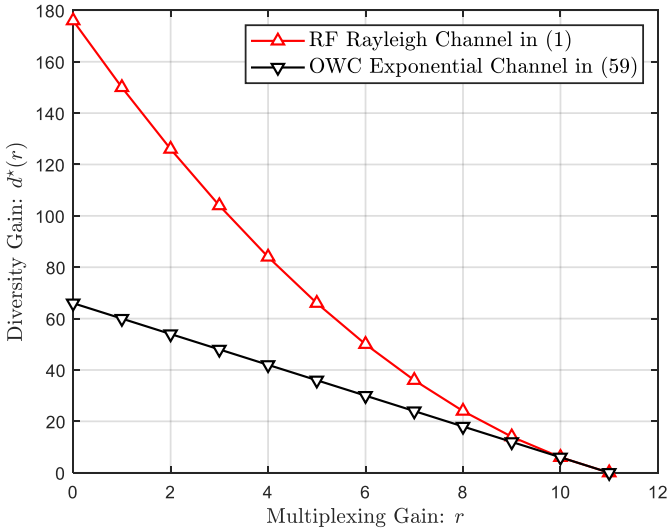

Fig. 7: Comparisons of the optimal DMT curve for 16×11 MIMO channel with large block length.

Figures 8 and 9 depict the optimal DMT bounds under a small channel block length. For a 16×4 MIMO channel, the thresholds of block length for the Rayleigh and exponential fading are $l_{\text{ra}} = 19$ and $l_{\text{exp}} = 13$, respectively. Hence, we can choose the channel block length (1) $l = 15$ and (2) $l = 10$ as small block lengths. From Figures 8 and 9, we find that the

This article has been accepted for publication in IEEE Transactions on Information Theory. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIT.2023.3244567

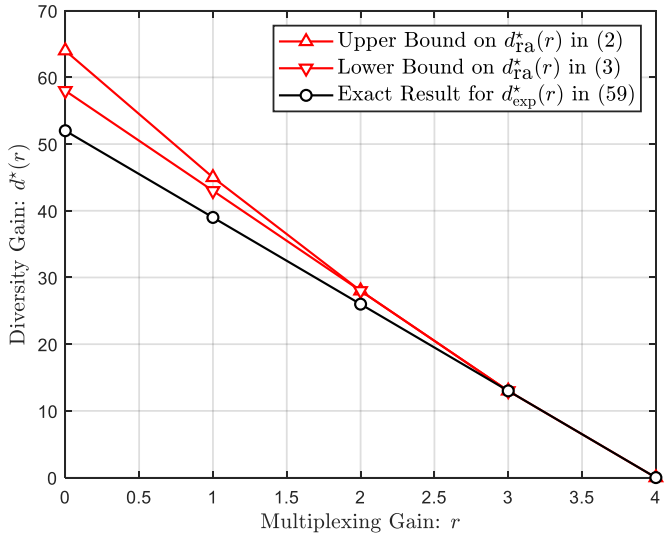

Fig. 8: Comparisons of the optimal DMT curve for 16×4 MIMO channel when $l = 15$.

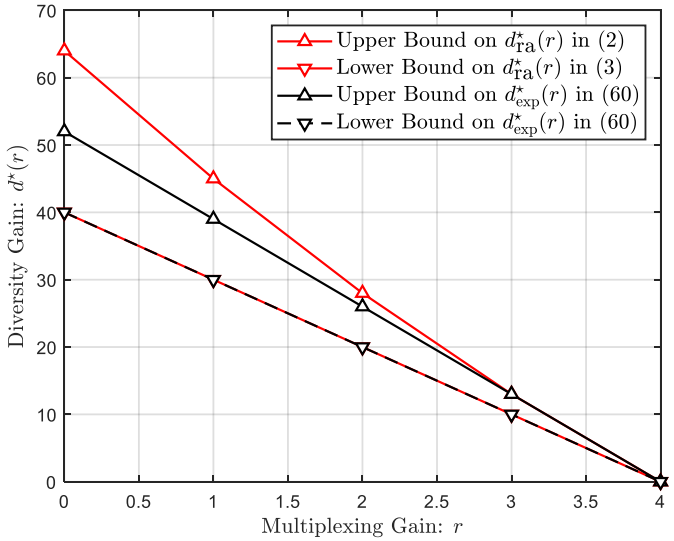

Fig. 9: Comparisons of the optimal DMT curve for 16×4 MIMO channel when $l = 10$.

maximum difference between the upper and lower bounds on $d^{\star}_{\rm ra}(r)$ for the Rayleigh channel is larger than that on $d^{\star}_{\rm exp}(r)$ for the exponential channel. Moreover, the lower bounds for the Rayleigh and exponential channels coincide in Figure 9.

To explain the above phenomena, we analyze the cases for $n_{\rm T} - n_{\rm R} + 1 \le l < n_{\rm T} + n_{\rm R} - 1$ and $l < n_{\rm T} - n_{\rm R} + 1$ corresponding to Figures 8 and 9, respectively. Under a small block length, we use the following denotations to represent the bounds on $d^{\star}_{\rm ra}(r)$:

$$d^{\rm ra}_{\rm ub} \triangleq (n_{\rm T} - r)(n_{\rm R} - r), \tag{116}$$

$$d^{\rm ra}_{\rm lb} \triangleq l(n_{\rm ra} - r), \tag{117}$$

and the bounds on $d^{\star}_{\rm exp}(r)$:

$$d^{\rm exp}_{\rm ub} \triangleq (n_{\rm T} - n_{\rm R} + 1)(n_{\rm R} - r), \tag{118}$$

$$d^{\rm exp}_{\rm lb} \triangleq l(n_{\rm R} - r), \tag{119}$$

where $n_{\rm ra} = \frac{r_{\rm ra}^2}{l} + \frac{l - n_{\rm T} - n_{\rm R}}{l} r_{\rm ra} + \frac{n_{\rm T} n_{\rm R}}{l}$ and $r_{\rm ra} = \min\{n_{\rm R}, n_{\rm R} - \lceil \frac{l - (n_{\rm T} - n_{\rm R}) - 1}{2} \rceil\}$.

When $n_{\rm T} - n_{\rm R} + 1 \le l < n_{\rm T} + n_{\rm R} - 1$, $d^{\star}_{\rm exp}(r)$ can be uniquely characterized by (59), while $d^{\star}_{\rm ra}(r)$ can only be upper-bounded by $d^{\rm ra}_{\rm ub}$ in (116) and lower-bounded by $d^{\rm ra}_{\rm lb}$ in (117). Obviously, the difference $d^{\rm ra}_{\rm ub} - d^{\rm ra}_{\rm lb} > 0$ in this case. When $l < n_{\rm T} - n_{\rm R} + 1$, we first compute $r_{\rm ra} = n_{\rm R}$ and $n_{\rm ra} = n_{\rm R}$. Substituting them into (117), we can obtain $d^{\rm ra}_{\rm lb} = d^{\rm exp}_{\rm lb}$. Furthermore, combined with (116) and (118), we can directly obtain $d^{\rm ra}_{\rm ub} > d^{\rm exp}_{\rm ub}$. These discussions provide a reasonable explanation for the phenomena presented in Figures 8 and 9, and validate the truncated exponential input achieves efficient estimation on the average error probability.

IX. 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 counterparts in traditional RF channels. These differences may be due to the unique input constraints in the optical wireless channels. As an extension of this work, it would be interesting to consider the channel fading caused by pointing errors and study the impact of pointing errors on the analysis of DMT for the OWC channels, which could refer to [25], [26].

Appendix A
Derivation of Instantaneous Capacity Bounds

A. Proof of Proposition 1

Recall $\mathbf{X}_j = [\mathbf{X}_{1j}, \mathbf{X}_{2j}, \ldots, \mathbf{X}_{n_{\rm T}j}]^{\rm T}$, it is straightforward to obtain

$$\mathbb{E}[\mathbf{X}_{ij}] = {\rm A}\left(\frac{1}{\mu} - \frac{e^{-\mu}}{1 - e^{-\mu}}\right), \ \forall i \in \{1, 2, \ldots, n_{\rm T}\}, \tag{120}$$

and

$$\mathrm{h}(\mathbf{X}_{ij}) = \log {\rm A} + 1 - \frac{\mu e^{-\mu}}{1 - e^{-\mu}} - \log \frac{\mu}{1 - e^{-\mu}}. \tag{121}$$

Next, we define a new random vector $\widetilde{\mathbf{X}}_j = [\widetilde{\mathbf{X}}_{1j}, \widetilde{\mathbf{X}}_{2j}, \ldots, \widetilde{\mathbf{X}}_{n_{\rm T}j}]^{\rm T}$, whose entries are i.i.d. according to Gaussian distribution with expectation 0, and variance

$$\mathbb{V}[\widetilde{\mathbf{X}}_{ij}] = \frac{1}{2\pi e} e^{2\,\mathrm{h}(\mathbf{X}_{ij})}. \tag{122}$$

By the Gaussian differential entropy formula, we can immediately show

$$\mathrm{h}(\widetilde{\mathbf{X}}_{ij}) = \mathrm{h}(\mathbf{X}_{ij}), \ \forall i \in \{1, \ldots, n_{\rm T}\}. \tag{123}$$

For a given channel realization $\mathbf{H} = \mathrm{H}$, apply the generalized EPI in [23], and we have

$$\mathsf{h}(\mathrm{H}\mathbf{X}_j) \geq \mathsf{h}(\mathrm{H}\widetilde{\mathbf{X}}_j) \tag{124}$$

$$= \frac{n_{\mathrm{R}}}{2}\log\left(2\pi e\left|\mathrm{H}\mathrm{K}_{\widetilde{\mathbf{X}}_j\widetilde{\mathbf{X}}_j}\mathrm{H}^{\mathsf{T}}\right|^{\frac{1}{n_{\mathrm{R}}}}\right), \tag{125}$$

where $\mathrm{K}_{\widetilde{\mathbf{X}}_j\widetilde{\mathbf{X}}_j}$ denotes the covariance matrix of $\widetilde{\mathbf{X}}_j$ which is diagonal with its ith entry

$$\left[\mathrm{K}_{\widetilde{\mathbf{X}}_j\widetilde{\mathbf{X}}_j}\right]_{ii} = \mathbb{V}\left[\widetilde{\mathbf{X}}_{ij}\right], \quad \forall i \in \{1,2,\ldots,n_{\mathrm{T}}\}. \tag{126}$$

Substituting (121) into (122), (126) and (125), we have

$$\mathsf{h}(\mathrm{H}\mathbf{X}_j) \geq \frac{n_{\mathrm{R}}}{2}\log\left(2\pi e\left|\frac{e\sigma^2}{2\pi\alpha^2}\mathrm{T}^2(\alpha,n_{\mathrm{T}})(\mathsf{OSNR})^2\mathrm{H}\mathrm{H}^{\mathsf{T}}\right|^{\frac{1}{n_{\mathrm{R}}}}\right), \tag{127}$$

where

$$\mathrm{T}(\alpha,n_{\mathrm{T}}) = \frac{1-e^{-\mu}}{\mu}e^{-\frac{\mu e^{-\mu}}{1-e^{-\mu}}}, \tag{128}$$

with μ satisfying (18).

Now we lower-bound the instantaneous capacity as

$$\mathsf{C}(\mathbf{X}_j;\mathbf{Y}_j|\mathbf{H}) \geq \mathsf{I}(\mathbf{X}_j;\mathbf{Y}_j|\mathbf{H}) \tag{129}$$

$$= \mathsf{h}(\mathbf{H}\mathbf{X}_j+\mathbf{Z}_j|\mathbf{H}) - \mathsf{h}(\mathbf{Z}_j) \tag{130}$$

$$\geq \frac{1}{2}\log\left(e^{2\,\mathsf{h}(\mathbf{H}\mathbf{X}_j)}+e^{2\,\mathsf{h}(\mathbf{Z}_j)}\right) - \mathsf{h}(\mathbf{Z}_j) \tag{131}$$

$$= \frac{1}{2}\log\left(1+\frac{e^{2\,\mathsf{h}(\mathbf{H}\mathbf{X}_j)}}{(2\pi e\sigma^2)^{n_{\mathrm{R}}}}\right) \tag{132}$$

$$\geq \frac{1}{2}\log\big(1+L_l(\mathsf{OSNR})^{2n_{\mathrm{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|\big), \tag{133}$$

where (129) follows from the fact that the instantaneous capacity is always greater than the instantaneous mutual information achieved by any feasible input, (130) follows from the independence between $\mathbf{H}$ and $\mathbf{Z}_j$, (131) by applying the EPI [36], and (133) by substituting (127) into (132). The proof is concluded.

B. Proof of Proposition 2

At high OSNR, by the asymptotic upper bound in [24, Thm. 21], we have

$$\mathsf{C}(\mathbf{X}_j;\mathbf{Y}_j|\mathbf{H}) \leq n_{\mathrm{R}}\log\mathsf{A} + \frac{1}{2}\log\left(\frac{\mathsf{V}_{\mathbf{H}}^2}{(2\pi e\sigma^2)^{n_{\mathrm{R}}}}\right), \tag{134}$$

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

We further upper bound the RHS of (134) as

$$n_{\mathrm{R}}\log\mathsf{A} + \frac{1}{2}\log\left(\frac{\mathsf{V}_{\mathbf{H}}^2}{(2\pi e\sigma)^{n_{\mathrm{R}}}}\right)$$

$$\leq n_{\mathrm{R}}\log\mathsf{A} + \frac{1}{2}\log\left(\frac{\binom{n_{\mathrm{T}}}{n_{\mathrm{R}}}\left(\sum_{\mathcal{U}}|\mathbf{H}_{\mathcal{U}}|^2\right)}{(2\pi e\sigma^2)^{n_{\mathrm{R}}}}\right) \tag{135}$$

$$= n_{\mathrm{R}}\log\mathsf{A} + \frac{1}{2}\log\left(\frac{\binom{n_{\mathrm{T}}}{n_{\mathrm{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|}{(2\pi e\sigma^2)^{n_{\mathrm{R}}}}\right), \tag{136}$$

where (135) holds by Cauchy–Schwarz inequality, and (136) by Cauchy–Binet formula: $\sum_{\mathcal{U}}|\mathbf{H}_{\mathcal{U}}|^2 = |\mathbf{H}\mathbf{H}^{\mathsf{T}}|$. The proof is concluded.

APPENDIX B
DERIVATION OF OUTAGE PROBABILITY BOUNDS

A. Preliminary

1) Decomposition on Channel Matrix: Given the channel matrix $\mathrm{H}\in\mathbb{R}_+^{n_{\mathrm{R}}\times n_{\mathrm{T}}}$, we first decompose it by LQ factorization

$$\mathrm{H} = \mathrm{L}\mathrm{Q}, \tag{137}$$

where $\mathrm{L}\in\mathbb{R}^{n_{\mathrm{R}}\times n_{\mathrm{R}}}$ is lower-triangular and $\mathrm{Q}\in\mathbb{R}^{n_{\mathrm{R}}\times n_{\mathrm{T}}}$ is orthogonal satisfying $\mathrm{Q}\mathrm{Q}^{\mathsf{T}} = \mathbf{I}_{n_{\mathrm{R}}\times n_{\mathrm{R}}}$. Combined with (137), we can rewrite $\mathrm{H}\mathrm{H}^{\mathsf{T}}$ as

$$\mathrm{H}\mathrm{H}^{\mathsf{T}} = \mathrm{L}\mathrm{L}^{\mathsf{T}}. \tag{138}$$

Then perform eigenvalue decomposition on $\mathrm{H}\mathrm{H}^{\mathsf{T}}$, we obtain

$$\mathrm{H}\mathrm{H}^{\mathsf{T}} = \mathrm{U}\Lambda\mathrm{U}^{\mathsf{T}}, \tag{139}$$

where $\mathrm{U}\in\mathbb{R}^{n_{\mathrm{R}}\times n_{\mathrm{R}}}$ is orthogonal satisfying $\mathrm{U}\mathrm{U}^{\mathsf{T}} = \mathbf{I}_{n_{\mathrm{R}}\times n_{\mathrm{R}}}$ and $\Lambda\in\mathbb{R}^{n_{\mathrm{R}}\times n_{\mathrm{R}}}$ is diagonal with $[\Lambda]_{ii} = \lambda_i$, $\forall i = \{1,\cdots,n_{\mathrm{R}}\}$. Combining (138) with (139), we can decompose $\mathrm{L}\mathrm{L}^{\mathsf{T}}$ as

$$\mathrm{L}\mathrm{L}^{\mathsf{T}} = \left(\mathrm{U}\Lambda^{\frac{1}{2}}\right)\left(\mathrm{U}\Lambda^{\frac{1}{2}}\right)^{\mathsf{T}}. \tag{140}$$

Then we have $\mathrm{L} = \mathrm{U}\Lambda^{\frac{1}{2}}$. Further substitute L into (137), and we have

$$\mathrm{H} = \mathrm{U}\Lambda^{\frac{1}{2}}\mathrm{Q}. \tag{141}$$

Denote $\mathrm{D}\triangleq\Lambda^{\frac{1}{2}}$, and rewrite λ_i in terms of a_i, $\forall i = \{1,\cdots,n_{\mathrm{R}}\}$, then we obtain

$$\mathrm{D} = \mathrm{diag}[(\mathsf{OSNR})^{-\frac{a_1}{2}},\ldots,(\mathsf{OSNR})^{-\frac{a_{n_{\mathrm{R}}}}{2}}]. \tag{142}$$

2) Derivation of Ep. (26) *:* By [16], [34], we have

$$f(a) = \xi\ [\log(\mathsf{OSNR})]^{n_{\mathrm{R}}}\times\left(\prod_{i=1}^{n_{\mathrm{R}}}(\mathsf{OSNR})^{-\frac{(n_{\mathrm{T}}-n_{\mathrm{R}}+1)}{2}a_i}\right)$$
$$\times\prod_{i<j}\left|(\mathsf{OSNR})^{-a_i}-(\mathsf{OSNR})^{-a_j}\right|$$
$$\times\int_{\mathrm{V}_{n_{\mathrm{R}},n_{\mathrm{R}}}}\int_{\mathrm{V}_{n_{\mathrm{R}},n_{\mathrm{T}}}} f(\mathrm{U}\mathrm{D}\mathrm{Q})d\mathrm{Q}d\mathrm{U}, \tag{143}$$

where ξ is a normalization constant.

Form the definition in (15), we observe that the outage diversity gain $d_{\text{out}}(r)$ is fully determined by the corresponding OSNR exponent of outage probability. Note that

$$\lim_{\mathsf{OSNR}\to\infty}\frac{\log(\xi[\log(\mathsf{OSNR})]^{n_{\mathrm{R}}})}{\log(\mathsf{OSNR})} = 0. \tag{144}$$

Thus, at high OSNR, we have

$$\xi[\log(\mathsf{OSNR})]^{n_{\mathrm{R}}} \doteq \mathsf{OSNR}^0. \tag{145}$$

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

$$\prod_{i<j}\left|(\mathsf{OSNR})^{-a_i}-(\mathsf{OSNR})^{-a_j}\right| \doteq \prod_{i=1}^{n_{\mathrm{R}}}(\mathsf{OSNR})^{-(i-1)a_i}. \tag{146}$$

Substituting (144) and (146) into (143), the proof is concluded.

B. Proof of Theorem 1

1) Upper Bound: To prove the new bounds on outage probability in Theorem 1, we employ the bounds obtained in Proposition 1. We first prove the upper bound on outage probability, and then prove the lower bound.

Denote $R = r\log\mathsf{OSNR}$. Substitute (19) into the RHS of (13), and we obtain

$$\begin{aligned}&\mathrm{P}_{\text{out}}(\mathsf{OSNR})\\ &\le \mathrm{P}\left[\frac{1}{2}\log\big(1+L_l(\mathsf{OSNR})^{2n_{\mathrm{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|\big)\le r\log(\mathsf{OSNR})\right] &(147)\\ &= \mathrm{P}\left[1+L_l(\mathsf{OSNR})^{2n_{\mathrm{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|\le \mathsf{OSNR}^{2r}\right]. &(148)\end{aligned}$$

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

$$|\mathbf{H}\mathbf{H}^{\mathsf{T}}| = \prod_{i=1}^{n_{\mathrm{R}}}\lambda_i = \mathsf{OSNR}^{-\sum_{i=1}^{n_{\mathrm{R}}}a_i}. \tag{149}$$

Substituting (149) into (148), we obtain

$$\begin{aligned}&\mathrm{P}_{\text{out}}(\mathsf{OSNR})\\ &\le \mathrm{P}\left[1+L_l(\mathsf{OSNR})^{2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_i}\le(\mathsf{OSNR})^{2r}\right] &(150)\\ &\doteq \mathrm{P}\left[(\mathsf{OSNR})^{\left(2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_i\right)^+}\le(\mathsf{OSNR})^{2r}\right] &(151)\\ &= \mathrm{P}\left[(2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_i)^+\le 2r\right] &(152)\\ &= \int_{\mathcal{A}} f(a)da, &(153)\end{aligned}$$

where (151) follows from the fact that $1+c\cdot\mathsf{OSNR}^b \doteq \mathsf{OSNR}^{b^+}$, and (153) by the definition in (29). The proof is concluded.

2) Lower Bound: Substituting (21) into the RHS of (13), and following the similar arguments as in Section B-B1, we have

$$\begin{aligned}&\mathrm{P}_{\text{out}}(\mathsf{OSNR})\\ &\ge \mathrm{P}\left[\frac{1}{2}\log\big(L_u(\mathsf{OSNR})^{2n_{\mathrm{R}}}|\mathbf{H}\mathbf{H}^{\mathsf{T}}|\big)\le r\log(\mathsf{OSNR})\right] &(154)\\ &= \mathrm{P}\left[L_u(\mathsf{OSNR})^{2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_i}\le(\mathsf{OSNR})^{2r}\right] &(155)\\ &\doteq \mathrm{P}\left[(\mathsf{OSNR})^{\left(2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_i\right)}\le(\mathsf{OSNR})^{2r}\right] &(156)\\ &\doteq \mathrm{P}\left[2n_{\mathrm{R}}-\sum_{i=1}^{n_{\mathrm{R}}}a_i\le 2r\right] &(157)\\ &= \int_{\mathcal{B}} f(a)da. &(158)\end{aligned}$$

The proof is concluded.

Appendix C
Proof of Eq. (32)

Denote the codebook as $\{\mathrm{X}(0),\mathrm{X}(1),\dots,\mathrm{X}(\mathsf{OSNR}^{lr}-1)\}$. Suppose the sent codeword is $\mathrm{X}(0)$, and we consider the event when the ML decoder decides erroneously in favor of $\mathrm{X}(1)$. This event occurs only if the projection distance of $\mathbf{Y(0)}-\mathrm{HX}(0)$ on the direction of $\mathrm{HX}(1)-\mathrm{HX}(0)$ is larger than $d_{01}/2$, where $d_{01} = \parallel \mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}$ is the distance between $\mathrm{HX}(1)$ and $\mathrm{HX}(0)$. The probability on the occurrence of this event is[3]

$$\begin{aligned}&\mathrm{P}(\mathrm{X}(0)\to\mathrm{X}(1)|\mathbf{H}=\mathrm{H})\\ &\quad=\mathrm{P}\Bigg(\frac{\langle\mathbf{Y(0)}-\mathrm{HX}(0),\mathrm{HX}(1)-\mathrm{HX}(0)\rangle}{\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}}\\ &\qquad\qquad\ge\frac{1}{2}\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}\Bigg). &(159)\end{aligned}$$

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

$$\begin{aligned}&\mathrm{P}(\mathrm{X}(0)\to\mathrm{X}(1)|\mathbf{H}=\mathrm{H})\\ &\quad=\mathrm{P}\Bigg(\frac{\boldsymbol{z}}{\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}}\ge\frac{1}{2}\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}\Bigg) &(160)\\ &\quad=Q\left(\sqrt{\frac{\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}^2}{4\sigma^2}}\right) &(161)\\ &\quad\le\exp\left(-\frac{\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}^2}{8\sigma^2}\right), &(162)\end{aligned}$$

where (162) holds because $Q(t)\le 1/2\exp(-t^2/2)$.

Note that

$$\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}^2=\sum_{k=1}^{n_{\mathrm{R}}}\sum_{j=1}^{l}|\sum_{i=1}^{n_{\mathrm{T}}}[\mathrm{X}_{ij}(1)-\mathrm{X}_{ij}(0)]h_{ki}|^2, \tag{163}$$

where $\mathrm{X}_{ij}(\cdot)=[\mathrm{X}(\cdot)]_{ij}$, and $h_{ki}=[\mathrm{H}]_{ki}$. Setting $\Omega_j = [\mathrm{X}_{1j}(1)-\mathrm{X}_{1j}(0),\cdots,\mathrm{X}_{n_{\mathrm{T}}j}(1)-\mathrm{X}_{n_{\mathrm{T}}j}(0)]^{\mathsf{T}}$, we rewrite (163) as

$$\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathsf{F}}^2=\sum_{j=1}^{l}\Omega_j^{\mathsf{T}}\mathrm{H}^{\mathsf{T}}\mathrm{H}\Omega_j. \tag{164}$$

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 (164) as

$$\parallel\mathrm{HX}(1)-\mathrm{HX}(0)\parallel_{\mathrm{F}}^2=\sum_{j=1}^{l}\sum_{i=1}^{n_{\mathrm{T}}}\lambda_i\beta_{ji}^2. \tag{165}$$

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

Substituting (165) into (162), we have

$$\mathrm{P}(\mathrm{X}(0)\to\mathrm{X}(1)|\mathbf{H}=\mathrm{H})\le\exp\left(-\frac{1}{8\sigma^2}\sum_{j=1}^{l}\sum_{i=1}^{n_\mathrm{T}}\lambda_i\beta_{ji}^2\right). \tag{166}$$

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

$$\begin{aligned}&\mathrm{P}(\mathbf{X(0)}\to\mathbf{X(1)}|\mathbf{H}=\mathrm{H})\\&\le\mathbb{E}_{\boldsymbol{\beta}_{ji}}\left[\exp\left(-\frac{1}{8\sigma^2}\sum_{j=1}^{l}\sum_{i=1}^{n_\mathrm{T}}\lambda_i\boldsymbol{\beta}_{ji}^2\right)\right] &(167)\\&=\prod_{j=1}^{l}\mathbb{E}_{\boldsymbol{\beta}_j}\left[\exp(-\frac{1}{8\sigma^2}\sum_{i=1}^{n_\mathrm{T}}\lambda_i\boldsymbol{\beta}_{ji}^2)\right] &(168)\\&=\det\left(\mathbf{I}+\frac{g(\alpha,n_\mathrm{T})}{2}(\mathsf{OSNR})^2\mathrm{H}\mathrm{H}^\intercal\right)^{-\frac{l}{2}}, &(169)\end{aligned}$$

where (168) follows from the independence between vectors $\boldsymbol{\beta}_p$ and $\boldsymbol{\beta}_q$, $\forall p\neq q$, and (169) by the derivation in the following Appendix C-1.

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

$$\begin{aligned}&\mathrm{P}(\mathrm{error}|\mathbf{H}=\mathrm{H})\\&\le(\mathsf{OSNR})^{lr}\det\left(\mathbf{I}+\frac{g(\alpha,n_\mathrm{T})}{2}(\mathsf{OSNR})^2\mathrm{H}\mathrm{H}^\intercal\right)^{-\frac{l}{2}} &(170)\\&=(\mathsf{OSNR})^{lr}\prod_{i=1}^{n_\mathrm{R}}\left(1+\frac{g(\alpha,n_\mathrm{T})}{2}(\mathsf{OSNR})^2\lambda_i\right)^{-\frac{l}{2}}. &(171)\end{aligned}$$

1) Proof of Eq. (169)*:* Denote the matrix $\mathrm{V}=[\mathrm{V}_1,\cdots,\mathrm{V}_{n_\mathrm{T}}]$, where $\mathrm{V}_p=[v_{1p},\cdots,v_{n_\mathrm{T}p}]^\intercal$ is the pth column of V. Recall that V is an orthogonal matrix, i.e.,

$$\sum_{i=1}^{n_\mathrm{T}}v_{ip}v_{ip}=1,\quad\forall p\in\{1,\cdots,n_\mathrm{T}\}, \tag{172}$$

$$\sum_{i=1}^{n_\mathrm{T}}v_{ip}v_{iq}=0,\quad\forall p\neq q. \tag{173}$$

Note that

$$\boldsymbol{\beta}_{jp}=\Delta\mathbf{X}_{1j}v_{1p}+\cdots+\Delta\mathbf{X}_{n_\mathrm{T}j}v_{n_\mathrm{T}p} \tag{174}$$

$$\boldsymbol{\beta}_{jq}=\Delta\mathbf{X}_{1j}v_{1q}+\cdots+\Delta\mathbf{X}_{n_\mathrm{T}j}v_{n_\mathrm{T}q}, \tag{175}$$

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

Recall that $\mathbf{X}_{ij}$, $\forall i\in\{1,\cdots,n_\mathrm{T}\}$ and $j\in\{1,\cdots,l\}$, follows i.i.d. truncated exponential distribution in (17), and hence it is directly to verify

$$\mathbb{E}[\boldsymbol{\beta}_{ji}]=0, \tag{176}$$

$$\begin{aligned}\mathbb{V}[\boldsymbol{\beta}_{ji}]&=2\mathrm{A}^2\left(\frac{1}{\mu^2}-\frac{e^{-\mu}}{(1-e^{-\mu})^2}\right) &(177)\\&\triangleq 2\mathrm{A}^2g(\alpha,n_\mathrm{T}). &(178)\end{aligned}$$

By the central limit theorem, we can approximate $\boldsymbol{\beta}_{ji}$ as a Gaussian random variable with expectation (176) and variance (178), and hence by averaging over the distribution of $\boldsymbol{\beta}_{ji}$ in (168), we have

$$\begin{aligned}&\prod_{j=1}^{l}\mathbb{E}_{\boldsymbol{\beta}_j}\left[\exp(-\frac{1}{8\sigma^2}\sum_{i=1}^{n_\mathrm{T}}\lambda_i\boldsymbol{\beta}_{ji}^2)\right]\\&\qquad=\det\left(\mathbf{I}+\frac{g(\alpha,n_\mathrm{T})}{2}(\mathsf{OSNR})^2\mathrm{H}\mathrm{H}^\intercal\right)^{-\frac{l}{2}}. &(179)\end{aligned}$$

The proof is concluded.

REFERENCES

[1] 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. Areas Commun.*, vol. 34, no. 4, pp. 883-893, Apr. 2016.

[2] H. Kaushal and G. Kaddoum, "Optical communication in space: Challenges and mitigation techniques," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 1, pp. 57-96, Firstquarter 2017.

[3] V. W. S. Chan, "Free-space optical communications," *J. Light. Technol.*, vol. 24, no. 12, pp. 4750-4762, Dec. 2006.

[4] M. A. Khalighi and M. Uysal, "Survey on free space optical communication: a communication theory perspective," *IEEE Commun. Surv. Tutor.*, vol. 16, no. 4, pp. 2231-2258, Fourthquarter 2014.

[5] 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.

[6] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation through Random Media*. 2nd ed. SPIE Press: Bellingham, WA, USA, 2005.

[7] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Area. Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] M. Safari and M. Uysal, "Cooperative diversity over log-normal fading channels: Performance analysis and optimization," *IEEE Trans. Wirel. Commun.*, vol. 7, no. 5, pp. 1963–1972, May 2008.

[13] 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.

[14] A. W. Shaban and O. Damen, "The diversity-multiplexing tradeoff of log-normal channels as a function of the dB spread," in *Proc. Wirel. Commun. Netw. Conf. (WCNC)*, Marrakesh, Morocco, Apr. 2019, pp. 1–7.

[15] P. Sharda and M. R. Bhatnagar, "Diversity-multiplexing tradeoff for indoor visible light communication," in *Int. Conf. Wirel. Mobile Comput., Netw. Commun. (WiMob)*, Thessaloniki, Greece, Oct. 2020, pp. 1–6.

[16] 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.

[17] P. Sharda and M. R. Bhatnagar, "Diversity-multiplexing tradeoff for MIMO-FSO system under different transmission scenarios with limited quantized feedback," *IEEE Access*, vol. 8, pp. 114266-114286, Jun. 2020.

[18] P. Sharda, A. Jaiswal, M. R. Bhatnagar, and A. Garg, "Two laser selection scheme for FSO communication," in *Proc. IEEE Int. Conf. Adv. Netw. Telecommun. Syst. (ANTS)*, New Delhi, India, Dec. 2020, pp. 1–6.

[19] P. Sharda, A. Jaiswal, M. R. Bhatnagar, A. Garg, and L. Song, "Clustering based diversity improving transmit laser selection schemes using quantized feedback for FSO system," *IEEE Trans. Veh. Technol.*, vol.70, no. 7, pp. 6855-6868, Jul. 2021.

[20] 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. Light. Technol.*, vol. 36, no. 19, pp. 4354–4360, Apr. 2018.
[21] 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. 13 927–13 931, Aug. 2020.
[22] 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.
[23] 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.
[24] 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.
[25] 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.
[26] A. A. Farid and S. Hranilovic, "Diversity gain and outage probability for MIMO free-space optical links with misalignment," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 479-487, Feb. 2012.
[27] E. W. Weisstein, "Modified Bessel function of the second kind," 2002. [Online]. Available: https://mathworld.wolfram.com/
[28] 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.
[29] 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.
[30] M. Riedl, *Optical Design Fundamentals for Infrared Systems*. 2nd ed. SPIE Press: Bellingham, WA, USA, 2001.
[31] 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.
[32] A. Lapidoth, S. M. Moser and M. A. Wigger, "On the capacity of free-space optical intensity channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4449-4461, Oct. 2009.
[33] D. Dowson and A. Wragg, "Maximum-entropy distributions having prescribed first and second moments (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 19, no. 5, pp. 689-693, Sep. 1973.
[34] A. Edelman, "Eigenvalues and condition numbers of random matrices," *SIAM J. Matrix Anal. Appl.*, vol. 9, no. 4, p. 543–560, Dec. 1988.
[35] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.
[36] 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.

Sufang Yang (Graduate Student Member, IEEE) received the B.E. degree from the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China, in 2020. She is currently pursuing the Ph.D. degree with the Department of Electronic Engineering, Tsinghua University, Beijing, China. Her current research interests include wireless communications and information theory.

Longguang Li (Member, IEEE) received the Ph.D. degree in communication engineering from Telecom ParisTech in 2019. He is currently an Associate Professor with East China Normal University. His research interests include information theory and wireless communications.

Haoyue Tang (Student Member, IEEE) received the B.E. and Ph.D. degrees from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2017 and 2022, respectively. She was a Visiting Student at Technische Universitat München from September 2015 to February 2016 and Télécom Paris from January 2019 to March 2019. She is currently a Post-Doctoral Research Associate with Yale University. Her research interests include age of information, stochastic network optimization, and statistical learning theory.

Jintao Wang (Senior Member, IEEE) received the B.E. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2001 and 2006, respectively. From 2006 to 2009, he was an Assistant Professor at the Department of Electronic Engineering, Tsinghua University. Since 2009, he has been an Associate Professor and Ph.D. Supervisor. He is the Standard Committee Member of the Chinese National Digital Terrestrial Television Broadcasting Standard. He has authored or coauthored more than 100 journal and conference papers and holds more than 40 national invention patents. His research interests include space-time coding, MIMO, and OFDM systems.