IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 18, 2020, accepted June 7, 2020, date of publication June 22, 2020, date of current version July 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3004135

Diversity-Multiplexing Tradeoff for MIMO-FSO System Under Different Transmission Scenarios With Limited Quantized Feedback

PRANAV SHARDA, (Student Member, IEEE), AND MANAV R. BHATNAGAR, (Senior Member, IEEE)
Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi 110016, India

Corresponding author: Pranav Sharda (shardapranav73@gmail.com)

This work was supported in part by the Media Lab Asia (Sir Visvesvaraya Young Faculty Research Fellowship) through the Ministry of Electronics and Information Technology (MeitY), Government of India, and in part by the Brigadier Bhopinder Singh Chair, Indian Institute of Technology Delhi, New Delhi, India.

ABSTRACT The diversity-multiplexing tradeoff (DMT) in a multiple-input multiple-output (MIMO) free-space optical (FSO) communication with limited channel state information at the transmitter (CSIT) is investigated. Using the limited CSIT based power and rate control strategy, we optimally allocate the power among the good and bad channels in such a fashion that the DMT performance of the system enhances significantly unlike the no-CSIT based MIMO-FSO DMT. In this way, a new limited CSIT based technique/model with optimal power and rate control strategy is proposed to enhance the DMT performance. The optimal DMT is studied for two different transmission scenarios: single-rate and adaptive-rate transmission. It is shown that how the optimal DMT is influenced when the concept of minimum guaranteed multiplexing gain in the forward link is taken into account. It is illustrated that power control based on the feedback plays a vital role in attaining the optimal DMT, and rate adaptation is significant in obtaining a high diversity gain, especially at high rates. Moreover, the analysis of upper and lower bounds on the optimal DMT is done by giving useful insights. Furthermore, a novel study based on the optimal tradeoff between the degrees of freedom and the number of transmit apertures in a coherent MIMO-FSO channel is also done. To validate the results of the proposed model, we compare the derived results with the no-CSIT based MIMO-FSO DMT. It is observed that the proposed technique/model outperforms the no-CSIT based MIMO-FSO DMT.

INDEX TERMS Adaptive-rate, diversity-multiplexing tradeoff (DMT), free-space optical (FSO) communication, gamma-gamma channel, log-normal channel, multiple-input multiple-output (MIMO), negative exponential channel, single-rate.

I. INTRODUCTION

Communication over wireless multiple-input multiple-output (MIMO) channels has fascinated more and more researchers over the last decade. As compared to conventional single-input single-output (SISO) systems, MIMO systems offer better reliability in terms of the availability of many independent propagation paths characterized by the term 'diversity gain'. Moreover, MIMO systems make use of parallel spatial modes characterized by the term 'spatial multiplexing gain' or simply 'multiplexing gain'. However, it is proven in [1], that there exists a fundamental tradeoff between diversity and multiplexing gains, known as diversity-multiplexing tradeoff (DMT). DMT is characterized as a significant performance metric for comparing different MIMO techniques. Additionally, DMT inspires to design new MIMO schemes that are DMT optimal [2]–[4]. Moreover, the performance of a MIMO system is highly dependent on the assumption of channel state information (CSI) at both sides of the communication link. Obtaining perfect CSI at the transmitter (CSIT) is a challenge in practical scenarios, while partial CSIT is commonly available in practice in the form of a few feedback bits.

In order to improve various performance metrics of multiple-antenna systems, different useful methods employing partial CSIT have been explored in [5]–[9]. Under the

The associate editor coordinating the review of this manuscript and approving it for publication was Congduan Li.

assumption of partial CSIT, characterization of the DMT over a multiple-antenna channel are discussed in [10], [11] where the feedback information consists of the scalar quantized singular values of the channel matrix. However, this assumption is restrictive and the results in [10], [11] only demonstrate the asymptotic outage corresponding to a particular class of feedback schemes. In [12], an automatic retransmission request (ARQ) scheme for MIMO systems and its performance tradeoffs are studied. Combining ARQ with transmit power control [13]–[15] offers a superior tradeoff compared to the no-CSIT case [12]. In [16], it is shown that the tradeoff between the diversity and multiplexing gains can be simultaneously achieved over a slowly fading multiple-antenna channel with partial CSIT. Partial power control is shown to be instrumental in achieving the optimal DMT. Moreover, the results in [16] indicate that the diversity gain can be increased significantly with limited quantized CSI, especially at low multiplexing gains. Further, direct partial CSIT via resolution constrained feedback scheme is more useful than the indirect channel knowledge provided by an ARQ scheme.

A. MOTIVATION

In the last two decades, there has been a tremendous interest in the research of free-space optical (FSO) communication. Hence, it has attracted a significant attention in the research community. This is because contrary to radio-frequency (RF) communication, where spectrum usage is a big constraint, FSO communication is completely independent of spectrum licensing. Due to this, FSO links possess a very high optical bandwidth, low implementation cost and inherent security [17]. Despite the major advantages of FSO technology, its performance is highly dependent on the atmospheric turbulence (AT) [18]. AT is a random phenomenon, which occurs due to changes in the refractive index of the atmosphere with time. In FSO communication, AT is the primary source of random fluctuations in the received optical signal. There are many channel models for FSO systems to characterize the irradiance fluctuation. The best-suited channel models under weak, moderate-to-strong, and saturation AT regime are log-normal, gamma-gamma, and negative exponential, respectively. So far, there is a very limited study of DMT in MIMO-FSO systems. In [19], DMT for coherent SIMO-FSO systems is analysed for the first time. In [20], DMT for indoor RF environments and terrestrial MIMO-FSO are studied under the consideration of log-normal channel. However, the channel model considered in [20] is valid only for weak AT regime of FSO systems. To overcome this problem, a study of optimal DMT for intensity modulation direct detection (IM/DD) MIMO-FSO system is done in [21] covering all the ranges of AT. However, in the open literature, no CSIT-dependent power controller strategy using limited quantized feedback has been adopted to derive the optimal DMT of MIMO-FSO system. Therefore, considering all the above-mentioned limitations, the major contributions of this work are discussed in the subsequent subsection.

TABLE 1. List of mathematical notation used in this paper.

τ	Average electrical SNR at each receiver aperture
$\mathcal{I}$	Feedback Index
K	Feedback resolution or Number of feedback levels
$P_{\mathcal{I}=i}$	Average total transmit power corresponding to the received feedback index $\mathcal{I}=i$ by the transmitter
$f_Z(.)$	Probability density function (pdf) of random variable Z
$\mathbf{A}^T$	Transpose of matrix $\mathbf{A}$
$\mathbf{A}^\dagger$	Hermitian of matrix $\mathbf{A}$
$\Gamma(.)$	Gamma function
$G_{p,q}^{m,n}(.)$	Meiger-G function
$\mathrm{Tr}(\mathbf{A})$	Trace of matrix $\mathbf{A}$
$\lim f(z)$	Limit of a function z
$\doteq$	Exponential equality
$\mathbb{E}(.)$	Expectation operator
$P_{out}(.)$	Outage error probability
inf	Infimum of a set
Υ_i	Back-off multiplexing gain

B. NOVEL CONTRIBUTIONS

Motivated by the aforementioned background, this paper takes the basic study of DMT for MIMO-FSO system to a higher level by introducing the concept of CSIT-dependent power controller with resolution constrained quantized feedback. Moreover, we consider a practical scenario of CSIT, i.e., partial CSIT. Further, this paper aims at investigating the optimum DMT for a MIMO-FSO system over all AT regimes. The novel contributions of this work are as follows:

1. We consider two different scenarios of transmission: single-rate transmission and adaptive-rate transmission. For single-rate transmission, the information rate is kept constant, a good model for constant bit-rate services. Further, this transmission scenario has no dependence on the CSIT. For adaptive rate systems, motivated by the concept of minimum rate [22], [23], we introduce the concept of minimum multiplexing gain that makes ''reliability '' more meaningful in the limit of high signal-to-noise ratios (SNRs).
2. By performing joint power and rate control, we derive the optimal DMT of MIMO-FSO system in terms of the outage upper bound in both cases (single-rate and adaptive-rate) in a recursive fashion.
3. Our results give useful insights into the design of adaptive-rate based MIMO-FSO system with sufficiently long codewords at very high SNR. The optimal DMT corresponding to single-rate based MIMO-FSO system can be attained by a single codebook and a CSIT-dependent power controller. On the other side, an adaptive-rate based MIMO-FSO system may attain performance close to the optimal with only two different codebooks, even if the resolution of the feedback link is higher.
4. Furthermore, a novel insight into the optimal tradeoff between the degrees of freedom (d.o.f) and the number of transmit apertures/lasers is also given.

In other words, this paper for the first time, develops the resolution-constrained quantized feedback based DMT of MIMO-FSO system under all ranges of AT ranging from weak-to-saturation regime. In this way, this paper gives the generalised insights into the study of quantized feedback based DMT for IM/DD MIMO-FSO system under two different transmission scenarios.

The remainder of this paper is organized as follows: In Section II, we describe the system and channel models under consideration. Moreover, we introduce some basic definitions that will be used throughout the paper. The optimal DMT for a single-rate and adaptive-rate based MIMO-FSO system in terms of the outage upper bound is studied in Section III. Section IV discusses the lower bounds (back-off bounds and expurgated bounds) on the optimal DMT. In Section V, optimal tradeoff between the d.o.f and the number of transmit apertures is studied. Section VI presents the derived numerical results. Section VII draws concluding remarks. The paper contains three appendices.

II. PRELIMINARIES

In this section, we present the system model, different channel models and provide some basic definitions, which will be used in this paper.

A. SYSTEM MODEL

Consider a MIMO-FSO system with N_t transmit lasers/apertures and N_r photodetectors. The channel is constant during a fading block consisting of T channel uses, but changes independently from one block to the next. During a fading block l, the channel is represented by an $N_r \times N_t$ random matrix $\mathbf{H}$. In this work, we consider IM/DD technique for MIMO-FSO system employing line-of-sight (LOS) links. Therefore, all the coefficients of $\mathbf{H}$ are assumed to be real and positive. The element of $\mathbf{H}$, given by $h_{yz} \geq 0$, represents the channel coefficient between z^{th} $(z = 1, 2, \cdots, N_t)$ laser and y^{th} $(y = 1, 2, \cdots, N_r)$ photodetector. Let $\mathbf{X} \in \Re^{N_t \times T}$ and $\mathbf{Y} \in \Re^{N_r \times T}$ represent the transmitted and received codewords, respectively, then the received signal can be written in the matrix form as:

$$\mathbf{Y} = \sqrt{\frac{\tau}{N_t^2}}\mathbf{H}\mathbf{X} + \mathbf{E}, \tag{1}$$

where $\mathbf{E} \in \Re^{N_r \times T}$ denotes the additive white Gaussian noise (AWGN) matrix. Further, each element of $\mathbf{E}$ is Gaussian distributed with zero mean and unity variance, $\tau = (\rho P_{avg})^2$ represents the average electrical SNR at each receiver aperture, where ρ is the responsivity of the photodetector, and P_{avg} denotes the average received optical power.

It is assumed that the receiver knows the channel matrix perfectly. Let us denote the random variable representing the feedback index as $\mathcal{I}$. Further, $\mathcal{I}$ takes values on the set $\{1, 2, \cdots, K\}$ with K being a positive integer, termed as the feedback resolution. For a given channel condition, the receiver feedbacks the index $\mathcal{I}(\mathbf{H})$ through a noiseless, zero-delay feedback link to the transmitter. It should be noted that $\mathcal{I}(\mathbf{H})$ is a deterministic mapping from a channel matrix to an integer index. In simple words, the index mapping $\mathcal{I}(\mathbf{H})$ corresponds to the partitioning of the set of all possible channel matrices into K number of quantized feedback regions. Furthermore, prior to the transmission, the channel matrix corresponding to a particular quantized region is exactly known to the transmitter. Depending upon the received feedback index $\mathcal{I} = i$ by the transmitter, the codeword $\mathbf{X}$ is transmitted from a codebook $C_i = \{\mathbf{X}_i(1), \mathbf{X}_i(2), \cdots, \mathbf{X}_i(M_i)\}$ of rate R_i, where all codewords are uniformly drawn from the codebook. The $\mathbf{X}_i(k)'s$ are matrices of size $N_t \times T$. Let us define

$$P_{\mathcal{I}=i} \triangleq \frac{1}{TM_i}\sum_{k=1}^{M_i}\|\mathbf{X}_i(k)\|_F^2, \tag{2}$$

where $\|\mathbf{X}\|_F$ denotes the Frobenius norm of matrix $\mathbf{X}$. Note that $P_{\mathcal{I}=i}$ is the average total transmit power defined over the event that the feedback index $\mathcal{I} = i$ is received by the transmitter. Moreover, we impose a power constraint [13], [14] over infinitely many fading blocks which is given as:

$$\lim_{L\to\infty}\frac{1}{L}\sum_{l=1}^{L}\frac{1}{T}\|\mathbf{X}\|_F^2 = \mathbb{E}_H\left[P_{\mathcal{I}(\mathbf{H})}\right] \leq \tau, \tag{3}$$

where the first equality holds with probability one. Moreover, $\mathcal{I}(\mathbf{H})$ and the codebooks C_i's are SNR dependent. Thus, P_is and R_is are also SNR dependent. Throughout this paper, we focus on the average rate over infinitely many fading blocks

$$R \triangleq \lim_{L\to\infty}\frac{1}{L}\sum_{l=1}^{L}R_{\mathcal{I}(H)} = \sum_{i=1}^{K}\Pr(\mathcal{I}=i)R_i \tag{4}$$

Let r denotes the multiplexing gain, then

$$r = \lim_{\tau\to\infty}\frac{R(\tau)}{\frac{1}{2}\log\tau}, \tag{5}$$

where $R(\tau) = \frac{r}{2}\log(\tau)$ is the target data rate for a given τ. However, the system is said to have a multiplexing gain of r, in an average sense, if

$$\lim_{\tau\to\infty}\frac{2R}{\log\tau} = \sum_{i=1}^{K} r_i \lim_{\tau\to\infty}\Pr(\mathcal{I}=i) = r. \tag{6}$$

Moreover, we consider two different scenarios of transmission as discussed below:

- Single-rate transmission: For single-rate transmission, corresponding to each value of SNR, the transmission rate is independent of the feedback index. In simple words, it is constrained that

$$r_1 = \cdots = r_K = r, \tag{7}$$

where r_is, $i = 1, 2, \cdots, K$ are referred to as the 'individual multiplexing gains'. This model of transmission is specifically designed to support constant-rate services, such as voice or video transmission [13], [14].

• Adaptive-rate transmission: The values of $r_1, \cdots, r_K$ can be optimized subject to (6), i.e., a variable-rate MIMO-FSO system is considered. Additionally, a constraint on the individual multiplexing gains is imposed

$$r_i \geq r_{\min}, \quad \forall i \in \{1, \cdots, K\}, \tag{8}$$

where $0 \leq r_{\min} \leq \min(N_r, N_t)$ is a constant. We term it as the 'minimum multiplexing gain'.

Remark 1: From (8), it can be inferred that an acceptable quality of service is only attained at a certain minimum rate. Furthermore, without considering (8), it is not possible to predict the outage, i.e., the event that a particular LOS link of a MIMO-FSO system cannot support the target data rate. However, (8) can also be imposed for the single-rate transmission, but in that case $r_{\min}$ has no influence on the optimal DMT. On the other side, this does not hold for the adaptive-rate case, where $r_{\min}$ significantly influences the optimal DMT.

Remark 2: Since we assume perfect CSIT in this work, therefore the proposed model is valid for even long distances (5-8 Km). As far as the strong turbulence effect is concerned, we consider negative exponential channel model (specific for strong turbulence conditions) and gamma-gamma channel model (with strong turbulence regime) in our results to ensure the validity of the proposed model for strong turbulence regime as well.

B. CHANNEL MODELS

In this subsection, we provide different FSO channel models covering all the AT regimes. For simplifying the presentation, we drop indexes of the channel gains. In this work, the different FSO channel models under consideration are as follows:

1. Log-Normal Channel: This channel model is valid for weak AT regime. Under weak AT, the fading of optical signal is modelled as $h = \exp(g)$, where g is a normal random variable with mean and variance of μ_g and σ_g^2, respectively [18]. Therefore, h follows the log-normal probability distribution function (pdf) which is given as:

$$f_h(h) = \frac{1}{h\sqrt{2\pi\sigma_g^2}} \exp\left(-\frac{(\ln(h) - \mu_g^2)}{2\sigma_g^2}\right), \tag{9}$$

Since the transmitted optical power is neither attenuated nor amplified by the turbulence, $\mathbb{E}(h) = 1$ which requires $\mu_g = -\sigma_g^2/2$ [18, Eq.(3.115)]. Note that $\mathbb{E}(.)$ denotes the expectation operator.

2. Gamma-Gamma Channel: Under moderate-to-strong AT regime, h is distributed as:

$$f_h(h) = \frac{(\alpha\beta)^{\frac{\alpha+\beta}{2}}}{\Gamma(\alpha)\Gamma(\beta)} h^{\frac{\alpha+\beta}{2}-1} G_{0,2}^{2,0}\left(\alpha\beta h \middle| \begin{matrix} \cdot, \cdot \\ \frac{\alpha-\beta}{2}, \frac{\beta-\alpha}{2} \end{matrix}\right), \tag{10}$$

where α and β are the turbulence parameters meant for characterizing the irradiance fluctuations, which depend upon the log irradiance variance (σ_g^2). Further, $\Gamma(.)$ is the gama function, and $G_{p,q}^{m,n}(.|)$ is the Meiger-G function [24, Subsection 2.24]. In this work, the parameter (α, β) is taken as (4.2,1.4) to model the strong AT regime.

3. Negative Exponential Channel: The negative exponential distribution is given by [18]

$$f_h(h) = \exp(-h). \tag{11}$$

This channel model is valid only for modeling of saturation AT regime.

C. BASIC DEFINITIONS

Before aheading towards the derivations of the optimal DMT, we wish to introduce some important definitions and notations that will be used throughout this paper.

1. For any function of $\tau, f(\tau)$, the following equality:

$$\lim_{\tau\to\infty} \frac{\log f(\tau)}{\log \tau} = b, \tag{12}$$

is denoted as $f(\tau) \doteq \tau^b$, where $\doteq$ represents exponential equality. Symbols $\dot{\leq}$ and $\dot{\geq}$ are similarly defined.

2. Let d_{out} denote the outage diversity gain, then

$$d_{out} = -\lim_{\tau\to\infty} \frac{\log P_{out}(R, P_{\mathcal{I}})}{\log \tau}, \tag{13}$$

where $P_{out}(R, P_{\mathcal{I}})$ is the outage error probability. The supremum of the diversity gain achieved over all the feedback schemes for a particular r is denoted as $d^*(r)$.

3. An outage event has the interpretation that the channel does not support a target data rate, or in other words, when the instantaneous channel capacity, $C(\mathbf{H}, P_{\mathcal{I}})$, is less than the target data rate, $R(\tau)$. Mathematically,

$$P_{out}(R, P_{\mathcal{I}}) \triangleq \Pr(C(\mathbf{H}, P_{\mathcal{I}}) < R(\tau)), \tag{14}$$

where the probability is defined over the random channel matrix $\mathbf{H}$. In this work, we use Telatar's capacity formula, which is commonly used to find capacity of any wireless communication system and is given by [20], [25]. However, the capacity expression given in [20], [25] is valid for an uninformed transmitter, whereas the expression given in (15) corresponds to the capacity of an informed transmitter.

$$C(\mathbf{H}, P_{\mathcal{I}}) \triangleq \frac{1}{2}\log_2\left(\det\left(I_{N_r} + \mathbf{H}\mathbf{L}_{\mathcal{I}(\mathbf{H})}\mathbf{H}^T\right)\right), \tag{15}$$

where det(.) denotes the determinant, and the identity matrix of size N_r is denoted as I_{N_r}, and $\mathbf{H}^T$ denotes the transpose of matrix $\mathbf{H}$. Since the channel gains are real in IM/DD FSO system, hence (15) contains $\mathbf{H}^T$ instead of $\mathbf{H}^\dagger$, where $\mathbf{H}^\dagger$ denotes hermitian of the matrix $\mathbf{H}$. Note that the maximization of the capacity is done over the total transmit power constraint, which depends upon the peak power of the input signal, and it is positive for an IM/DD FSO system. Moreover, $\{\mathbf{L}_i\}_{i=1}^K$ are the positive semidefinite matrices which correspond to the covariance matrices of the input to the channel depending upon the received feedback index ($\mathcal{I} = i$) by the transmitter.

4. Moreover, we define

$$\begin{aligned} m &\triangleq \max(N_r, N_t), \\ n &\triangleq \min(N_r, N_t), \\ d^*_{\max} &\triangleq d^*(0), \\ r^*_{\max} &\triangleq \sup\left\{r : d^*(r) > 0\right\}, \end{aligned}$$

where $d^*_{\max}$ and $r^*_{\max}$ are the maximal diversity gain and maximal spatial multiplexing gain respectively in the channel.

5. The dependence of the target data rates on the SNR is explicitly given by

$$R_i = \frac{r_i}{2}\log\tau, \quad i = 1, \cdots, K, \tag{16}$$

where the r_is are some real values in $(0, \min(N_r, N_t))$, independent of SNR.

III. OPTIMAL DMT FOR OUTAGE BASED UPPER BOUND

An outage event can be considered as a scenario when the channel is so bad that no scheme can offer reliable communication at a certain fixed data rate. Thus, it motivates us to study the DMT for such a practical scenario. For an outage event and a given feedback scheme, the outage probability, which is also SNR dependent, is thus defined as:

$$P_{out,K} \triangleq \Pr\left(\frac{1}{2}\log_2\left(\det\left(I_{N_r} + \mathbf{H}\mathbf{L}_{\mathcal{I}(\mathbf{H})}\mathbf{H}^T\right)\right) < R\right), \tag{17}$$

where Pr(.) denotes the probability. As far as DMT is concerned, outage probability is quite related to our discussion because of the upper bound yielded by an application of Fano's inequality, similar to that in[1, Lemma 5].

$$P_e \dot{\geq} P_{out,K} \doteq \tau^{-d_{out,K}(r)} \geq \tau^{-d^*_{out,K}(r)}, \tag{18}$$

where P_e denotes the average probability of error corresponding to a transmitted codeword being incorrectly decoded at the receiver. Further, (18) is valid for any feedback scheme and any code of finite length T. Moreover, the SNR exponent of the minimum outage probability P^*_{out} is denoted by $-d^*_{out,K}(r)$ (the optimal outage diversity order). Further, this SNR exponent is valid for all feedback schemes with feedback resolution K. In order to determine $d^*_{out,K}(r)$, for each given SNR, we characterize the resolution-constrained feedback schemes that minimize the outage probability, and then study the asymptotic behavior of the solution corresponding to the characterized feedback scheme for each SNR.

A. SINGLE-RATE TRANSMISSION

Single-rate transmission is taken into account when the constraint (7) is imposed. In order to determine $d^*_{out,K}(r)$, a joint optimization over $\mathcal{I}(\mathbf{H})$,$\{P_i\}_{i=1}^K$, and $\{\mathbf{L}_i\}_{i=1}^K$ is required. Moreover, $\mathrm{Tr}(\mathbf{L}_i) \leq P_i$, where $\mathrm{Tr}(\cdot)$ denotes the trace of $\mathbf{L}_i$. With $\tau \to \infty$, $\mathbf{L}_i$s are considered to be scaled identity matrices dependent on the partial CSIT and the SNR. For a given $\mathcal{I}(\mathbf{H})$ and $\{P_i\}_{i=1}^K$, by choosing $\mathbf{L}_i = \frac{P_i}{N_t^2} I_{N_t}, \forall i \in \{1, \cdots, K\}$. Note that I_{N_t} is a $N_t \times N_t$ identity matrix. Therefore, the different bounds on the outage probability are defined as:

$$\begin{aligned} &\lim_{\tau\to\infty} \frac{\log\Pr\left(\frac{1}{2}\log_2\left(\det\left(I_{N_r} + P_{\mathcal{I}}\mathbf{H}\mathbf{H}^T\right)\right) < R\right)}{\log\tau}, \\ &\leq \lim_{\tau\to\infty} \frac{\log P_{out,K}}{\log\tau}, \\ &\leq \lim_{\tau\to\infty} \frac{\log\Pr\left(\frac{1}{2}\log_2\left(\det\left(I_{N_r} + \frac{P_{\mathcal{I}}}{N_t^2}\mathbf{H}\mathbf{H}^T\right)\right) < R\right)}{\log\tau}, \end{aligned} \tag{19}$$

where $R = \frac{r}{2}\log\tau$.

Implication 1: If power allocation for all the feedback schemes is done in such a way that $P_i \doteq \tau^{\ q_i}$ where $0 < q_i < \infty, \forall i$, it then follows from (19) that we can restrict the analysis to the case $\mathbf{L}_i = \frac{P_i}{N_t^2} I_{N_t}, \forall i$. Moreover, in an asymptotic scenario, the SNR exponent of the outage probability is not affected when allocating the transmit power over all spatial directions. Further, the right singular vectors of the channel matrix [14], [15], provides an SNR gain that is significant in some range of SNR. However, in this work, we are only interested in characterizing the diversity gain in the framework of the DMT, i.e., as $\tau \to \infty$. Thus, we can exclusively focus on the asymptotic behavior of systems utilizing partial CSIT in order to control only the transmit power such that the outage probability is minimized. Such a system is efficiently determined by an optimal index mapping and requires an outage-minimizing power codebook $\{P_i\}_{i=1}^K$ for a given feedback scheme with feedback resolution K. This motivates us to introduce the following lemma.

Lemma 1 (Outage Minimization): The outage-minimizing power codebook $\left\{P_i^*\right\}_{i=1}^K$ corresponding to a given SNR and rate R solves the optimization problem which is defined as:

$$\begin{aligned} &\max P_K \\ &s.t.\ \left[P_{out}(R, P_K) + 1 - P_{out}(R, P_1)\right] P_1 \\ &\quad + \sum_{i=2}^{K} \left[P_{out}(R, P_{i-1}) + 1 - P_{out}(R, P_i)\right] P_i \leq \tau, \\ &\quad 0 \leq P_1 < \cdots < P_K \end{aligned} \tag{20}$$

where $P_{out}(R, P_{\mathcal{I}})$ is defined as in (14). The optimal deterministic index mapping is given by

$$\mathcal{I}^*(\mathbf{H}) = \begin{cases} 1, & \text{if} C(\mathbf{H}, P_K^*) < R \\ \min\left\{i : i \in \{1, \cdots, K\},\ C(\mathbf{H}, P_i^*) \geq R\right\}, & \\ \quad \text{otherwise}, & \end{cases} \tag{21}$$

where $C(\mathbf{H}, P_{\mathcal{I}})$ is defined by (15). The minimum outage probability is $P_{out}(R, P_K^*)$.

Proof: Refer to Appendix A for the proof.

From (21), we observe that the optimal index mapping follows a " circular" structure wherein the "best" and the " worst" channel conditions share a common index. Further, the region corresponding to $\mathcal{I}^*(\mathbf{H})$=1 is the only region where

an outage event is likely to occur. However, unlike in the perfect-CSIT case [13], [14], the optimal transmitter in the partial-CSIT or limited feedback case does not essentially switch off transmission, i.e., P_1^* is generally nonzero. The reason behind is that switching off transmission requires a zero power level in the power codebook, which is instead costly when the resolution of the feedback link is finite.

Now we are left with determining the asymptotic SNR exponent of P_K^* that solves (20). At high SNR, most channels do not undergo outage, thus leading to a scenario wherein some quantization regions may have a power level with an SNR exponent strictly larger than one can be employed without violating (3). This inspires us to introduce the following lemma, which determines the SNR exponent of the outage probability when the asymptotic SNR exponent of the transmit power is given.

Lemma 2 (Outage Diversity Order): Let Λ be a function of SNR such that $\Lambda \doteq \mathrm{SNR}^q$ where q is a finite constant and $q \geq 1$. For any $r \in (0, n)$, with $(x)^+ = \max(x, 0)$, we have

$$P_{out}(R, \Lambda) \doteq \mathrm{SNR}^{-J_i(r,q)},$$

where $P_{out}(R, \Lambda)$ is defined as in (14) and $J_i(r, q)$ is the outage diversity order, given by

$$\begin{aligned} J_1(r, q) &\triangleq \left\{ \inf_{\alpha \in \mathcal{A}} \sum_{i=1}^{n} \frac{\left(|m-n|+2i-1\right)}{2}\alpha_i \right. \\ &\qquad \left. - \frac{3mn\alpha_n}{4} + \frac{mn\alpha_n^2}{8\sigma_g^2}\log\tau \right\}, \\ J_2(r, q) &\triangleq \left\{ \inf_{\alpha \in \mathcal{A}} \sum_{i=1}^{n} \frac{\left(|m-n|+2i-1\right)}{2}\alpha_i \right. \\ &\qquad \left. + \frac{mn\alpha_n}{2}\left(\beta - 1\right) \right\}, \\ J_3(r, q) &\triangleq \left\{ \inf_{\alpha \in \mathcal{A}} \sum_{i=1}^{n} \frac{\left(|m-n|+2i-1\right)}{2}\alpha_i \right\}, \end{aligned} \tag{22}$$

where $J_i(r, q), i = 1, \cdots, 3$, is the outage diversity order corresponding to log-normal channel, gamma-gamma channel and negative exponential channel.

and

$$\mathcal{A} \triangleq \left\{ \alpha : \alpha_1 \geq \cdots \geq \alpha_n \geq 0, \sum_{i=1}^{n} (q - \alpha_i)^+ \leq r \right\}.$$

Here, α_is indicate the level of singularity of channel matrix H. The larger the α_is are, the more singular H is. Further, the set $\mathcal{A}$ describes the outage event in terms of singularity levels.

Proof: $J_1(r, q)$, $J_2(r, q)$ and $J_3(r, q)$ are derived in [21]. Further, on substituting (15) into (14), and applying eigenvalue decomposition of $\mathbf{H}\mathbf{H}^T$, we get

$$P_{out}(R, \Lambda) \triangleq \Pr\left(\sum_{k=1}^{n} \log\left(1 + \frac{\Lambda}{N_t}\lambda_k\right) < R \right),$$

where $\lambda_1 \leq \cdots \leq \lambda_n$ are the real and non-negative eigen values of $\mathbf{H}\mathbf{H}^T$. By applying change of variables $\lambda_k = \mathrm{SNR}^{-\alpha_k}$, $k = 1, \cdots, n$, we get

$$\begin{aligned} P_{out}(R, q) &\doteq \Pr\left(\prod_{k=1}^{n} \mathrm{SNR}^{(q-\alpha_k)^+} < \mathrm{SNR}^r \right) \\ &= \Pr\left(\sum_{k=1}^{n} (q - \alpha_k)^+ < r \right). \end{aligned}$$

Moreover, the region $\mathcal{A}$ only contains $\alpha_k \geq 0, \forall k$, since outside this region, the outage probability decays exponentially as $\mathrm{SNR} \to \infty$.

For a single-rate MIMO-FSO system, we define the following proposition which recursively defines the SNR exponent of the minimum outage probability (optimal outage diversity order) using quantized feedback.

Proposition 1: The optimal DMT of a single-rate MIMO-FSO system with K quantization regions in the feedback link is upper-bounded by the outage bound

$$d_{out,K}^*(r) = J(r, 1 + d_{out,K-1}^*(r)), \tag{23}$$

where $d_{out,0}^*(r) \triangleq 0, \forall r$.

Proof: Refer to Appendix B for the proof.

$J(r, q)$ is defined as in (22) according to the channel model. When asymptotic scenario is taken into account, mostly the channel is not in outage, and hence the transmit power $P_1 = \tau$ is used. So, these "good" channel conditions are mapped to $\mathcal{I} = 1$ (first quantization region). On the other side, the set of "bad" channel conditions in outage if P_1 is applied, results in a probability measure in the order of $\tau^{-J(r,1)}$ are mapped to $\mathcal{I} = 2$ (second quantization region). Thus, for the bad channel conditions, we may use a power P_2 in the order of $\tau^{1+J(r,1)}$ without violating (3). However, this leads to an outage probability in the order of $\tau^{-J(r,1+J(r,1))}$. We provide a formal proof given in Appendix B, that computes a lower bound and an upper bound on $d_{out,K}^*(r)$. The lower bound is obtained by choosing

$$P_1 = \frac{\tau}{K}, \quad P_2 = \frac{\tau}{KP_{out}(R, P_1)}, \cdots, \quad P_K = \frac{\tau}{KP_{out}(R, P_{K-1})}. \tag{24}$$

It follows from (24) that we can employ the following index mapping together with the power codebook $\{P_i\}_{i=1}^K$ in order to attain the outage bound:

$$\mathcal{I}^*(\mathbf{H}) = \begin{cases} K, & \text{if} C(\mathbf{H}, P_K) < R \\ \min\{i : i \in \{1, \cdots, K\},\ C(\mathbf{H}, P_i) \geq R\}, \\ \qquad \text{otherwise.} \end{cases}$$

Corollary 1: For a single-rate based MIMO-FSO system, we have

(a) Log-Normal Channel:

$$\lim_{r \to 0} d_{out,K}^*(r) = \sum_{k=1}^{K} \frac{N_t N_r}{2} - \frac{3}{4}N_t N_r + \frac{N_t N_r}{8\sigma_g^2}\log(\tau)^k$$

As $\tau \to \infty$, $\lim_{r\to 0} d_{out,K}^*(r) = \infty, \forall K$, and

$$\lim_{r \to n} d_{out,K}^*(r) = \infty, \quad \forall K. \tag{25}$$

(b) Gamma-Gamma Channel:

$$\lim_{r\to 0} d^*_{out,K}(r) = \sum_{k=1}^{K} \left(\frac{N_t N_r \min(\alpha,\beta)}{2}\right)^k$$

and

$$\lim_{r\to n} d^*_{out,K}(r) = 0, \forall K.$$

(c) Negative Exponential Channel:

$$\lim_{r\to 0} d^*_{out,K}(r) = \sum_{k=1}^{K} \left(\frac{N_t N_r}{2}\right)^k$$

and

$$\lim_{r\to n} d^*_{out,K}(r) = 0, \forall K.$$

Proof: We consider the negative exponential channel for the illustration purpose and on similar lines, the proof for the other channel models can be established. For r sufficiently close to zero and any $q \geq 1$, (22) is minimized by $\alpha_i^* = q$, i=1, $\cdots$, $n-1$, and $\alpha_n^* = q - r$ leading to $\lim_{r\to 0} J(r,q) = \left(\frac{N_t N_r q}{2}\right)$. Applying Proposition 1, we obtain

$$\lim_{r\to 0} d^*_{out,K}(r) = \sum_{k=1}^{K} \left(\frac{N_t N_r}{2}\right)^k, \quad \forall K.$$

Moreover, it can also be verified that $\lim_{r\to n} J(r,1) = 0$, Using Proposition 1 gives

$$\lim_{r\to n} d^*_{out,K}(r) = 0, \quad \forall K.$$

Remark 3: A significant conclusion that can be drawn from corollary 1 is the impact of partial CSIT in the MIMO-FSO system, which is much more as compared to the SISO-FSO system. This is because, over a SISO channel, the optimal diversity gain has a linear proportionality to the number of feedback levels K. This is contradictory to multiple apertures (at either transmitter or receiver, or both), wherein the optimal diversity gain has an exponential proportionality to the feedback levels.

B. ADAPTIVE-RATE TRANSMISSION

Under the well-behaved channel conditions, rate adaptation is significant for transmitting more information. For an adaptive-rate based MIMO-FSO system, one may try to perform the joint optimization (optimal power and rate allocation) corresponding to each SNR. This is however a cumbersome task, even in the SISO case with perfect CSIT [22]. Thus, we provide an alternative approach and find an upper bound on the outage bound itself, and which allows us to reuse the results of the single-rate transmission.

Proposition 2: The optimal DMT of an adaptive-rate MIMO-FSO system with $K \geq 2$ quantization regions in the feedback link and a minimum multiplexing gain $r_{\min}$, where $r_{\min} \in (0, n)$, and $r \in [r_{\min}, n)$ is upper-bounded by the outage bound

$$d^*_{out,K}(r, r_{\min}) = J(r_{\min}, 1 + d^*_{out,K-1}(r, r_{\min})), \quad (26)$$

where $d^*_{out,1}(r, r_{\min}) \triangleq J(r,1), \forall r \geq r_{\min}$.

Proof: Refer to Appendix C for the proof.

Efficient rate adaptation relies on the proper tracking and prediction of the channel at the transmitter. By performing the joint optimization, the outage bound is obtained by considering a two-rate system. As far as the rate corresponding to a particular quantization region is concerned, the rate of one quantization region dominates the average multiplexing gain of the system in asymptotic conditions. Moreover, power control is employed over the rest $K-1$ quantized feedback regions. However, minimum rate is employed for these $K-1$ quantized feedback regions. The alternative approach introduced for the adaptive-rate problem employs an optimal combination of channel inversion and water-filling to reduce outage and improve throughput, respectively [22]. At very high SNR, this alternative approach reduces to channel inversion alone, resulting in a similar structure to the single-rate transmission case. Further, as $\tau \to \infty$ the effect of water-filling has almost no effect.

Corollary 2: For an adaptive-rate based MIMO-FSO system, in the extreme conditions, when $r_{\min} \to 0$ and $K \geq 2$, a simple outage bound can be obtained

(a) Log-Normal Channel:

$$\lim_{r_{\min}\to 0} d^*_{out,K}(r, r_{\min}) = \left(\frac{N_t N_r}{2} - \frac{3}{4} N_t N_r + \frac{N_t N_r}{8\sigma_g^2}\log(\tau)\right)^{K-1} \times J(r,1) + \sum_{k=1}^{K-1} \left(\frac{N_t N_r}{2} - \frac{3}{4} N_t N_r + \frac{N_t N_r}{8\sigma_g^2}\log(\tau)^k\right).$$

(b) Gamma-Gamma Channel:

$$\lim_{r_{\min}\to 0} d^*_{out,K}(r, r_{\min}) = \left(\frac{N_t N_r \min(\alpha,\beta)}{2}\right)^{K-1} \times J(r,1) + \sum_{k=1}^{K-1} \left(\frac{N_t N_r \min(\alpha,\beta)}{2}\right)^k.$$

(c) Negative Exponential Channel:

$$\lim_{r_{\min}\to 0} d^*_{out,K}(r, r_{\min}) = \left(\frac{N_t N_r}{2}\right)^{K-1} + \sum_{k=1}^{K-1} \left(\frac{N_t N_r}{2}\right)^k.$$

The proof is almost identical to that of corollary 1 and therefore omitted for brevity.

Remark 4: We have provided an alternative approach to an adaptive-rate problem where joint optimization (optimal power and rate allocation) is a challenging task. To counter this problem, a simple outage bound for the different channel models covering all the AT regimes is given.

IV. GENERALIZED LOWER BOUNDS ON THE OPTIMAL DMT

In this section, we develop two lower bounds on the optimal DMT covering all the AT regimes of a MIMO-FSO system.

Due to the presence of quantized CSIT, and under the consideration of IM/DD FSO system, the derivations of these lower bounds are completely different. Moreover, the derived lower bounds quickly approach the outage upper bound even when the codeword length is moderate. Furthermore, our results give a novel insight into the approximate universality of codes drawn from random ensembles, as well as into the DMT performance of such codes in a scenario of limited feedback.

A. BACK-OFF BOUNDS

The motivation behind employing an idea of back-off bounds is to feed back when the channel is efficient enough to support a rate larger than the transmission rate. This happens when the MIMO-FSO link is having the least AT effect. Further, in order to gain in terms of error exponent, we exploit the gap between the code rate and the instantaneous mutual information of the channel. As far as the feedback thresholds are concerned, their optimization can be done for any given codeword length T and multiplexing gain r. For simplifying the presentation, we first discuss the case for a single rate MIMO-FSO system with $K = 2$, then generalize the results to $K > 2$ and adaptive rate MIMO-FSO system.

Consider the following sequence of index mappings and the two power codebooks for $K = 2$:

$$\mathcal{I}(\mathbf{H}) = \begin{cases} 1, & \text{if} C(\mathbf{H}, P_2) < \left(\frac{r}{2} + \Upsilon_2\right) \log \tau \\ \min\left\{i : i \in \{1, \cdots, K\}, C(\mathbf{H}, P_i) \geq \left(\frac{r}{2} + \Upsilon_i\right) \log \tau\right\}, \\ \text{else.} \end{cases}$$

and

$$P_1 = \frac{\tau}{2}, \quad P_2 = \frac{\tau}{2P_{out}\left(\left(\frac{r}{2} + \Upsilon_1\right) \log \tau, P_1\right)}. \tag{27a}$$

The Υ_i's are non-negative back-off multiplexing gains that are functions of r and T. Moreover, Υ_i's are independent of SNR. By construction $P_1 \doteq \tau \equiv \tau^{q_1}$ and $P_2 \doteq \tau^{1+J\left(\frac{r}{2}+\Upsilon_1, 1\right)} \equiv \tau^{q_2}$. Depending upon the received feedback index $\mathcal{I} = i$ by the transmitter, the transmit signals are constructed as

$$S_i(k) = \sqrt{\frac{P_i}{N_t}} X_i(k), \tag{28}$$

where $X_i(k)$ are codewords belonging to a random codebook with independent and identically distributed (i.i.d.) components $\sim \mathcal{N}(0,1)$. The error probabilities, averaged over the codebook, the channel, and the code ensembles are

$$\Pr(\text{error}, \mathcal{I}(\mathbf{H}) = 1, \quad C(\mathbf{H}, P_1) \geq \left(\frac{r}{2} + \Upsilon_1\right) \log \tau \doteq \tau^{-d_{B,1}(r)} \tag{29a}$$

and

$$\Pr(\text{error}, \mathcal{I}(\mathbf{H}) = 2) \dot{\leq} \tau^{-d_{B,2}(r)} \tag{29b}$$

where

(a) The back-off bound $d_{B,i}(r)$ for Log-normal channel is given as:

$$d_{B,i}(r) \triangleq \left\{ \inf_{\alpha \in \mathcal{A}} \sum_{i=1}^{n} \frac{(|m-n|+2i-1)}{2} \alpha_i - \frac{3mn\alpha_n}{4} + \frac{mn\alpha_n^2}{8\sigma_g^2} \log \tau + \frac{T}{2}\left(\sum_{i=1}^{n} (q_i - \alpha_i)^+ - \frac{r}{2}\right)\right\},$$

(b) The back-off bound $d_{B,i}(r)$ for Gamma-Gamma channel is given as:

$$d_{B,i}(r) \triangleq \left\{ \inf_{\alpha \in \mathcal{A}} \sum_{i=1}^{n} \frac{(|m-n|+2i-1)}{2} \alpha_i + \frac{mn\alpha_n}{2}(\beta - 1) + \frac{T}{2}\left(\sum_{i=1}^{n} (q_i - \alpha_i)^+ - \frac{r}{2}\right)\right\},$$

(c) The back-off bound $d_{B,i}(r)$ for Negative Exponential channel is given as:

$$d_{B,i}(r) \triangleq \left\{ \inf_{\alpha \in \mathcal{A}} \sum_{i=1}^{n} \frac{(|m-n|+2i-1)}{2} \alpha_i + \frac{T}{2}\left(\sum_{i=1}^{n} (q_i - \alpha_i)^+ - \frac{r}{2}\right)\right\},$$

and

$$\mathcal{A} \triangleq \left\{\alpha : \alpha_1 \geq \cdots \geq \alpha_n \geq 0, \sum_{i=1}^{n} (q_i - \alpha_i) \geq \frac{r}{2} + \Upsilon_i\right\}.$$

Proof: The derivation of $d_{B,i}(r)$ for the above channel models is based on the Theorem 3 given in [21].

For $T \geq N_t + N_r - 1$, the optimum α_i's always satisfy $\sum_{i=1}^{n} \left(q_i - \alpha_i^*\right)^+ = \frac{r}{2} + \Upsilon_i$ and thus we have

$$d_{B,1}(r) = J\left(\frac{r}{2} + \Upsilon_1, q_1\right) + \frac{T}{2}\Upsilon_1 = J\left(\frac{r}{2} + \Upsilon_1, 1\right) + \frac{T}{2}\Upsilon_1$$

$$d_{B,2}(r) = J\left(\frac{r}{2} + \Upsilon_2, q_2\right) + \frac{T}{2}\Upsilon_2 = J\left(\frac{r}{2} + \Upsilon_2, 1 + J\left(\frac{r}{2} + \Upsilon_1, 1\right)\right) + \frac{T}{2}\Upsilon_2.$$

Furthermore, by construction

$$\Pr(\text{outage}) \doteq \tau^{J\left(\frac{r}{2}+\Upsilon_2, q_2\right)} = \tau^{J\left(\frac{r}{2}+\Upsilon_2, 1+J\left(\frac{r}{2}+\Upsilon_1, 1\right)\right)} = \tau^{-d_{out}(r)}.$$

Moreover, corresponding to each r, the optimization of the back-off multiplexing gains is done as follows:

$\sup_{\Upsilon_1^2 \in [0,n-r)^2} \min(d_{B,1}(r), d_{B,2}(r), d_{out}(r))$. It is to be noticed that $d_{B,1}(r)$ does not depend on Υ_2. Also

$$d_{B,2}(r) = d_{out}(r) + \frac{T}{2}\Upsilon_2 \geq d_{out}(r),$$

and

$$\begin{aligned} d_{out}(r) &= J\Big(\frac{r}{2} + \Upsilon_2, 1 + J\Big(\frac{r}{2} + \Upsilon_1, 1\Big)\Big) \\ &\leq J\Big(r, 1 + J\Big(\frac{r}{2} + \Upsilon_1, 1\Big)\Big), \end{aligned}$$

where both inequalities become equalities if $\Upsilon_2 = 0$. Moreover, we conclude that that $\Upsilon_2^* = 0$ is optimal[1] and the corresponding optimization can be rewritten as

$$\sup_{\Upsilon_1 \in [0,n-r)} \min\Big\{J\Big(\frac{r}{2} + \Upsilon_1, 1\Big) + \frac{T}{2}\Upsilon_1, J\Big(r, 1 + J\Big(\frac{r}{2} + \Upsilon_1, 1\Big)\Big)\Big\}. \quad (30)$$

Taking (30) into account, for $T \geq N_t + N_r - 1$, and any $r \in (0, n)$, $d_{B,1}(r)$ is an increasing function of Υ_1 while $d_{B,2}(r) = d_{out}(r)$ is a decreasing function of Υ_1. In order to balance the SNR exponents ($d_{B,1}(r)$ and $d_{B,2}(r)$) corresponding to the outage events, optimization over Υ_1 is required. However, if we choose Υ_1 to be very small, it leads to $d_{B,1}(r) < d_{out}(r)$, and most errors occur when $C(H, P_1) \geq \frac{r}{2} + \Upsilon_1$, i.e., channel is not in outage. On the other side, if we choose $\Upsilon_1 > \Upsilon_1^*$, it results in an enlarged outage region.

B. GENERALISED BACK-OFF BOUNDS

Now, we generalize our discussion to $K > 2$ with the help of the following two cases:

Case 1: Single-rate MIMO-FSO system: For T $\geq N_t + N_r - 1$, $K \geq 2$, and $k = 1, \cdots, K$, recursively define

$$d_{B,k}(r) = J\Big(\frac{r}{2} + \Upsilon_k, 1 + d_{B,k-1}(r) - \frac{T}{2}\Upsilon_{k-1}\Big) + \frac{T}{2}\Upsilon_k,$$

where $d_{B,0}(r) \triangleq 0$ and $\Upsilon_0 \triangleq 0$, $\Upsilon_K \triangleq 0$. Thus, the lower bound on the optimal DMT of a single-rate MIMO-FSO system with feedback resolution K is given by

$$\begin{aligned} d_{B,K}^*(r) \triangleq \sup_{\Upsilon_1^{K-1}} &\min\big\{d_{B,1}(r), \cdots, d_{B,K}(r)\big\} \\ &s.t. \Upsilon_1^{K-1} \in [0, n-r)^{K-1} \end{aligned}$$

Case 2: Adaptive-rate MIMO-FSO system: For T $\geq N_t + N_r - 1$, $K \geq 2$, and $k = 2, \cdots, K$, recursively define

$$\begin{aligned} d_{B,k}(r, r_{\min}) = J\Big(\frac{r_{\min}}{2} + \Upsilon_k, 1 + d_{B,k-1}(r, r_{\min}) \\ - \frac{T}{2}\Upsilon_{K-1}\Big) + \frac{T}{2}\Upsilon_K, \end{aligned}$$

where $d_{B,1}(r, r_{\min}) \triangleq J(r, 1) + \frac{T}{2}\Upsilon_1$ and $\Upsilon_0 \triangleq 0$, $\Upsilon_K \triangleq 0$. Thus, the lower bound on the optimal DMT of an adaptive-rate MIMO-FSO system with feedback resolution K is given by

$$\begin{aligned} d_{B,K}^*(r, r_{\min}) \triangleq \sup_{\Upsilon_1^{K-1}} &\min \\ &\big\{d_{B,1}(r, r_{\min}), \cdots, d_{B,K}(r, r_{\min})\big\} \\ &s.t. \Upsilon_1^{K-1} \in [0, n-r) \times [0, n-r_{\min})^{K-2}. \end{aligned}$$

Furthermore, the back-off bound is asymptotically tight as $T \to \infty$. This can be illustrated by taking the case of $K = 2$. For this particular case, the back-off bound is given by (30). We rewrite (30) for completeness as:

$$\sup_{\Upsilon_1 \in [0,n-r)} \min\Big\{J\Big(\frac{r}{2} + \Upsilon_1, 1\Big) + \frac{T}{2}\Upsilon_1, J\Big(r, 1 + J\Big(\frac{r}{2} + \Upsilon_1, 1\Big)\Big)\Big\}. \quad (31)$$

By choosing a particular value of Υ_1, i.e.,

$$\tilde{\Upsilon}_1 = \frac{d_{out,2}^*(r) - J(r,1)}{T/2} = \frac{J(r, 1 + J(r,1)) - J(r,1)}{T/2} \quad (32)$$

Moreover, the value of Υ_1 should be chosen in accordance with a large value of T so that $\tilde{\Upsilon}_1 < n - r$. Thus, we get

$$d_{B,2}^*(r) \geq \min\Big\{J\Big(\frac{r}{2} + \tilde{\Upsilon}_1, 1\Big) + d_{out,2}^*(r) - J(r,1), J\Big(r, 1 + J\Big(\frac{r}{2} + \tilde{\Upsilon}_1, 1\Big)\Big)\Big\}. \quad (33)$$

On the other side, $\lim_{T\to\infty} \tilde{\Upsilon}_1 = 0$, therefore

$$\begin{aligned} \lim_{T\to\infty} d_{B,2}^*(r) &\geq \min\Big\{d_{out,2}^*(r), J(r, 1 + J(r,1))\Big\} \\ &= d_{out,2}^*(r). \end{aligned} \quad (34)$$

However, the lower bound $d_{B,2}^*(r)$ cannot exceed the outage bound $d_{out,2}^*(r)$ for any T, hence

$$\lim_{T\to\infty} d_{B,2}^*(r) = d_{out,2}^*(r), \quad \forall r \in (0, n). \quad (35)$$

Remark 5: The back-off multiplexing gains represent the small gaps between the code rate and the instantaneous mutual information. Now the point of interest is the minimization of these gaps (appproximately equal to zero). This is done by computing the union bound, averaged over a random ensemble, as done for the no-CSIT case [1]. However, to our surprise, such a methodology concludes that for any $K \geq 1$, and $T \geq N_t + N_r - 1$, we have $\Pr(\text{error}) \dot{\leq} \tau^{-J(r,1)}$ [16]. However, optimizing these gaps allows us to tighten the bounds gradually as T increases.

C. EXPURGATED BOUNDS

In the previous subsection, we gave an insight into the back-off multiplexing gains wherein we feed back when the channel is under a well-behaved condition. In this scenario, the channel is efficient enough to support a rate larger than the transmission rate. Now, rather backing off and balancing the SNR exponents of the error probability in each quantization region, we pursue another

[1]This can be observed as a generalization of the no-CSIT case, where a back-off multiplexing again $\Upsilon_1^* = 0$ makes the outage and outage-free regions have the same SNR exponent, solving $\sup_{\Upsilon_1 \in [0,n-r)} \min(d_{B,1}(r), d_{out}(r))$.

approach. We introduce this approach by deriving the motivation from the approximately universal condition and expurgation techniques [1], [26], [27]. A given codeword W is said to be an expurgated codeword, if there exists at least one other codeword in the codebook so that $\sum_{j=1}^{n}(\gamma_j)^+ > r$. Further, γ_js correspond to the SNR exponents of the squared singular values of an arbitrary codeword difference matrix ΔW. To start with, we consider a scenario where the codes are drawn from a random ensemble and then expurgate bad codewords that do not satisfy the approximately universal condition given by (36)

$$\left(\min_{C}\prod_{j=1}^{n}\tau^{-(\gamma_j)^+}\right)\dot{\geq}\tau^{-r}, \tag{36}$$

where C is a sequence of codes with rate $r\log\tau$ (bits per channel use). For an arbitrary pair of codewords, let $\mu_1 \leq \cdots \leq \mu_n$ be the n smallest squared singular values of the codeword difference matrix such that $\mu_j = \tau^{-\gamma_j}$. Recall that $(\gamma_j)^+ = \max(\gamma_j, 0)$. Moreover, the average energy of a component in a codeword matrix is normalized such that

$$\frac{1}{TMN_t}\sum_{k=1}^{M}\|W(k)\|_F^2 \leq 1, \tag{37}$$

where $k = 1, 2, \cdots, M$ with $M \triangleq \lfloor \tau^{rT} \rfloor$, are matrices of size $N_t \times T$, with $T \geq N_t$. Since the codeword matrix size is fixed ($N_t \times T$) whereas the number of codewords grow without any restriction, codes drawn from a random ensemble will lead to the violation of the power constraint almost on every codeword as $\tau \to \infty$. Hence, it is not correct to discuss the approximate universality of such a random code outside our considered power constraint, i.e, (36)-(37). This technique of expurgating all the bad codewords lead to lower rate codewords with high probability. So, expurgated codes cannot be considered as approximately universal. However, due to the presence of quantized CSIT, employing a combination of the expurgated codes with rate back-off results in a much improved bound. This can be observed especially at low multiplexing gains.

Let us consider a scenario where the codes are drawn from a random ensemble and each codeword is a matrix of size $N_t \times T$ with i.i.d. gaussian components $\mathcal{N}(0,1)$. The number of codewords is $\tau^{T\hat{r}}$. Hence, the rate is $\hat{r}\log\tau$ bits per channel use. Furthermore, let $t \doteq \Pr\left(\sum_{j=1}^{n}(\gamma_j)^+ \geq r\right)$ where the probability is over the code ensemble. Since ΔW is a matrix of size $(N_t \times T)$ with i.i.d. zero-mean real Gaussian elements with variance 2, we have

$$t \doteq \tau^{-\inf_{\gamma}\sum_{j=1}^{N_t}(2j-1+T-N_t)\gamma_j} \tag{38}$$

where

$$\gamma \triangleq \left\{\gamma_1^{N_t} : \gamma_1 \geq \cdots \geq \gamma_{N_t} \geq 0, \sum_{j=1}^{n}(\gamma_j)^+ \geq r\right\}. \tag{39}$$

This results in $t \doteq \tau^{-(T-N_t+1)r}$ with optimum minimizers

$$\gamma_1^* = r, \gamma_2^* = \cdots \gamma_{N_t}^* = 0. \tag{40}$$

The probability that a codeword is expurgated can be union-bounded by

$$\Pr(W\text{expurgated}) \leq \tau^{T\hat{r}}t. \tag{41}$$

If $T\hat{r} = (T - N_t + 1)\,r - \delta$ for any arbitrarily small $\delta > 0$, then

$$\Pr(W\text{expurgated}) \dot{\leq}\tau^{-\delta} \tag{42}$$

This implies that we obtain a code with multiplexing gain arbitrarily close to

$$\hat{r} = \left(1 - \frac{N_t - 1}{T}\right)r, \tag{43}$$

such that $\left(\sum_{j=1}^{n}(\gamma_j)^+ \leq r\right)$ for all pairs of codewords. The result give us two useful insights which are as below:

1. For $N_t = 1$, there exists at least one expurgated code of length $T \geq 1$ that satisfies (36).
2. On the other side, as $T \to \infty$, the expurgated codes become closer to universal. Moreover, for any finite T and $N_t \geq 2$, a code with a smaller rate, i.e., $\hat{r} < r$ is obtained. However, such a code will have no significance without CSIT.

To overcome this problem, we combine the expurgated codes with CSIT. For this, consider the expurgated codes $\mathcal{I}_E(H)$ for the following feedback scheme:

$$\begin{aligned}\mathcal{I}_E(H) = \Bigg\{ & 1, \text{if} C(H, P_K) < \frac{r}{1-\frac{N_t-1}{T}}\log\tau \\ & \min\Big\{i : i \in \{1, \cdots, K\}, \\ & C(H, P_i) \geq \frac{r}{1-\frac{N_t-1}{T}}\log\tau\Big\}\Bigg\}, \text{otherwise.}\end{aligned} \tag{44}$$

Moreover, the constraint $\frac{r}{1-\frac{N_t-1}{T}} < n$, has to be considered. Because we cannot expect expurgate codes to offer very high rates. Therefore,

$$\begin{aligned}P_1 &= \frac{\tau}{K}, \\ P_2 &= \frac{\tau}{KP_{out}\left(\frac{r}{1-\frac{N_t-1}{T}}\log\tau, P_1\right)}, \cdots \\ P_K &= \frac{\tau}{KP_{out}\left(\frac{r}{1-\frac{N_t-1}{T}}\log\tau, P_{K-1}\right)}.\end{aligned} \tag{45}$$

Herein, the transmit signals are constructed as

$$S_i(k) = \sqrt{\frac{P_i}{N_t}}W_E(k), \tag{46}$$

where $W_E(k)$'s correspond to the codewords of the expurgated code. Furthermore, by construction

$$\Pr\left(\text{error}, \mathcal{I}=i, C(H, P_i) \geq \frac{r}{1-\frac{N_t-1}{T}})\log\tau\right) \doteq \tau^{-\infty},\ \forall i. \tag{47}$$

Hence,

$$P_e \dot{\leq} \tau^{-d_{out}(r)} \equiv \tau^{-d_{E,K}(r)}. \quad (48)$$

We define recursively

$$d_{E,k}(r) \triangleq J\left(\frac{r}{1-\frac{N_t-1}{T}}, 1+d_{E,k-1}(r)\right), \quad (49)$$

where $d_{E,0}(r)=0$. Moreover, from the discussion of back-off bounds, we know that there is existence of codes of length $T \geq N_t + N_r - 1$ such that $\Pr(\text{error}, \mathcal{I}=K) \doteq \tau^{-d_{out}(r)}, \forall r$. Therefore, there is no need to use expurgated codes corresponding to $\mathcal{I}=K$.

Now for completeness, we summarize the results of the derivation with the help of following cases:

Case 1: Single-rate MIMO-FSO system: For $k = 1,\cdots,K-1$, define recursively

$$d_{E,k}(r) \triangleq J\left(\frac{r}{1-\frac{N_t-1}{T}}, 1+d_{E,k-1}(r)\right),$$

where $d_{E,0}(r)=0$. For $r < \left(1-\frac{N_t-1}{T}\right)n$, the optimal DMT of a single-rate MIMO-FSO system with feedback resolution K is lower-bounded by

$$J\left(\frac{r}{1-\frac{N_t-1}{T}}, 1+d_{E,K-1}(r)\right), \quad \text{for } T \geq N_t \quad (50a)$$

and

$$J(r, 1+d_{E,K-1}(r)), \quad \text{for } T \geq N_t + N_r - 1. \quad (50b)$$

The same approach can be applied in the case of adaptive-rate MIMO-FSO system. For brevity, we omit the derivation and summarize the results as follows:

Case 2: Adaptive-rate MIMO-FSO system: For $k = 2,\cdots,K-1$, define recursively

$$d_{E,k}(r, r_{\min}) = J\left(\frac{r_{\min}}{1-\frac{N_t-1}{T}}, 1+d_{E,k-1}(r, r_{\min})\right),$$

where

$$d_{E,1}(r, r_{\min}) = J\left(\frac{r}{1-\frac{N_t-1}{T}}, 1\right), \quad \forall r_{\min}.$$

For $r \in [r_{\min}, \left(1-\frac{N_t-1}{T}\right)n)$, the optimal DMT of an adaptive-rate MIMO-FSO system with feedback resolution K is lower-bounded by

$$J\left(\frac{r_{\min}}{1-\frac{N_t-1}{T}}, 1+d_{E,K-1}(r, r_{\min})\right), \text{ for } T \geq N_t \quad (51a)$$

and

$$J\left(r_{\min}, 1+d_{E,K-1}(r, r_{\min})\right), \text{ for } T \geq N_t + N_r - 1. \quad (51b)$$

Remark 6: Since the derived results are valid for lower bounds, these codes drawn from a random ensemble do not appear to be sufficient to complete the DMT analysis unlike the no-CSIT case [1]. Though, such random codes even with moderate codeword lengths and/or careful expurgation technique may quickly approach the outage bounds, when combined with a well-designed feedback scheme.

V. DEGREES OF FREEDOM: THE OPTIMAL TRADEOFF PERSPECTIVE

In this section, we provide a novel insight into the optimal tradeoff between the number of transmit apertures and the corresponding d.o.f for a coherent MIMO-FSO channel. Note that d.o.f can be viewed as the finite number of parallel spatial channels to communicate. More importantly, there should be minimization of the cost related to the installation of the number of transmit apertures. Further, along with the cost minimization, we must not compromise with the reliability (in terms of efficient transmission and reception). This motivates us to provide an optimal tradeoff between d.o.f and N_t.

For simplifying the presentation, we consider the channel model given by (11) in order to derive the coherent capacity and the numerical formula for the d.o.f. Likewise, the coherent capacity can be derived for other channel models also. Further, the channel model is going to affect only the numerical computation of the coherent capacity that depends upon the channel's distribution function and not the d.o.f, which is independent of the channel's distribution function. The coherent capacity of the channel model (11) is computed using (15), and the final result is summarized in the following lemma.

Lemma 3: Coherent Capacity and d.o.f.

Assume the fading coefficient matrix $\mathbf{H}$ is known to the receiver (known as "coherent assumption"), the channel capacity (bps/Hz) of a system with N_t transmit and N_r receive apertures is given by

$$C_{coherent}(P_{\mathcal{I}}) = \frac{1}{2}\mathbb{E}\left[\log_2\left(det\left(I_{N_r} + \frac{P_{\mathcal{I}}}{N_t^2}\mathbf{H}\mathbf{H}^T\right)\right)\right]. \quad (52)$$

Defining $U \triangleq \min(N_t, N_r)$. Then the coherent capacity can be lower-bounded as:

$$C_{coherent}(P_{\mathcal{I}}) \geq \frac{1}{2}\left[U\log_2\frac{P_{\mathcal{I}}}{N_t^2} + \mathbb{E}\left[\log_2 f_G(t)\right]\right], \quad (53)$$

where $f_G(t)$ is the distribution function corresponding to $\mathbf{H}\mathbf{H}^T$, $\mathbb{E}[.]$ denotes the expectation operator, and the d.o.f are given by

$$\text{d.o.f} = \frac{U}{2} = \frac{\min(N_t, N_r)}{2}. \quad (54)$$

Since the optimal tradeoff between the degrees of freedom and the number of transmit apertures is obtained at $N_t = N_r$, we can write the coherent capacity for the case $N_t = N_r$ as

$$C_{coherent}(P_{\mathcal{I}}) \geq \frac{1}{2}\left[N_t\log_2\frac{P_{\mathcal{I}}}{N_t^2} + \mathbb{E}\left[\log_2 f_G(t)\right]\right]. \quad (55)$$

$$\therefore \text{d.o.f} = \frac{N_t}{2} = \frac{N_r}{2}. \quad (56)$$

Proof : For completeness, we rewrite (15) as follows:

$$C(H,P_{\mathcal{I}}) = \frac{1}{2}\left[\log_2\left(det\left(I_{N_r}+\frac{P_{\mathcal{I}}}{N_t^2}\mathbf{H}\mathbf{H}^T\right)\right)\right],$$
$$\therefore C_{coherent}(P_{\mathcal{I}}) = \frac{1}{2}\mathbb{E}\left[\log_2\left(det\left(I_{N_r}+\frac{P_{\mathcal{I}}}{N_t^2}\mathbf{H}\mathbf{H}^T\right)\right)\right]. \quad (57)$$

In order to determine the distribution of $\mathbf{H}\mathbf{H}^T$, let us define a random variable G s.t.

$$G \triangleq h_1^2+h_2^2+\cdots+h_U^2, = Y_1+Y_2+\cdots+Y_U, \quad (58)$$

where $h_1^2, h_2^2, \cdots, h_U^2$ are i.i.d. Further, Y_1, Y_2 and Y_U are the random variables corresponding to the distribution of h_1^2, h_2^2 and h_U^2, respectively. Let

$$Y_1 = h_1^2, \quad \text{s.t.} f_{h_1}(h_1) = e^{-h_1}. \quad (59)$$

By applying the transformation of random variables, the distribution of Y_1 is obtained as follows:

$$f_{Y_1}(y_1) = \frac{e^{-\sqrt{y_1}}-e^{\sqrt{y_1}}}{2\sqrt{y_1}}. \quad (60)$$

Therefore,

$$f_G(g) = f_{Y_1}(y_1) \circledast f_{Y_2}(y_2) \circledast \cdots f_{Y_U}(y_U), \quad (61)$$

where $\circledast$ denotes the convolution operation. In Laplace domain, $f_G(g)$ can be written as:

$$f_G(s) = f_{Y_1}(s)\times f_{Y_2}(s)\times\cdots\times f_{Y_U}(s), = \prod_{i=1}^{U} f_{Y_i}(s), = \prod_{i=1}^{U} -\left[\frac{e^{\frac{1}{4s}}\sqrt{\pi}\,\text{Erf}\left(\frac{1}{2\sqrt{s}}\right)}{\sqrt{s}}\right], = -\left[\frac{e^{\frac{1}{4s}}\sqrt{\pi}\,\text{Erf}\left(\frac{1}{2\sqrt{s}}\right)}{\sqrt{s}}\right]^U. \quad (62)$$

Further,

$$f_G(s) = (-\sqrt{\pi})^U\left[(f_1(s))^U\times(f_2(s))^U\times(f_3(s))^U\right], \quad (63)$$

where $f_1(s) = e^{\frac{1}{4s}}, f_2(s) = \text{Erf}\left(\frac{1}{2\sqrt{s}}\right)$, and $f_3(s) = \frac{1}{\sqrt{s}}$. Thus, $f_G(t) = (-\sqrt{\pi})^U\left[f_1(t)\circledast f_2(t)\circledast f_3(t)\right]$. Note that $f_1(t)$, $f_2(t)$ and $f_3(t)$ can be obtained by taking the inverse laplace transform of $(f_1(s))^U$, $(f_2(s))^U$ and $(f_3(s))^U$, respectively.

$$f_1(t) = L^{-1}\left[(f_1(s))^U\right] = \delta(t), \quad (64)$$

Because, $\lim_{s\to\infty}(e^{\frac{1}{4s}})^U = 1 = \delta(t)$. In order to determine $f_2(t) = L^{-1}\left[(f_2(s))^U\right]$, we use the following Maclaurin series expansion of Erf (z):

$$\text{Erf}\,(z) = \frac{2}{\sqrt{\pi}}\left(\sum_{n=0}^{\infty} a_n z^{2n+1}\right), \quad (65)$$

where $a_n = \frac{(-1)^n}{n!(2n+1)}$, and

$$\left(\text{Erf}\left(\frac{1}{2\sqrt{s}}\right)\right)^U = \left(\frac{2}{\sqrt{\pi}}\right)^U\left(\sum_{n=0}^{\infty}\frac{a_n}{2^{2n+1}}s^{-\left(n+\frac{1}{2}\right)}\right)^U, = \left(\frac{2}{\sqrt{\pi}}\right)^U\left\{\sum_{n=0}^{\infty}\frac{a_n}{2^{2n+1}}s^{-\left(n+\frac{1}{2}\right)} \times\sum_{n=0}^{\infty}\frac{a_n}{2^{2n+1}}s^{-\left(n+\frac{1}{2}\right)}\times\cdots\cdots \times\sum_{n=0}^{\infty}\frac{a_n}{2^{2n+1}}s^{-\left(n+\frac{1}{2}\right)}\right\}. \quad (66)$$

$$f_2(t) = L^{-1}\left[\left(\text{Erf}\left(\frac{1}{2\sqrt{s}}\right)\right)^U\right] = \left(\frac{2}{\sqrt{\pi}}\right)^U \times\left\{\sum_{n=0}^{\infty}\tilde{u}t^{-\left(n+\frac{1}{2}\right)} \circledast\sum_{n=0}^{\infty}\tilde{u}t^{-\left(n+\frac{1}{2}\right)}\circledast\cdots\cdots\circledast\sum_{n=0}^{\infty}\tilde{u}t^{-\left(n+\frac{1}{2}\right)}\right\}, = \left(\frac{2}{\sqrt{\pi}}\right)^U \times\sum_{n=0}^{\infty}\frac{(\tilde{u})^U 4^{-n}\sqrt{\pi}\Gamma(n+\frac{1}{2})^{U-1}\Gamma(2n+1)\,t^{Un+\frac{U-2}{2}}}{\Gamma(n+1)\Gamma\left(Un+(\frac{U-2}{2})+1\right)}, \quad (67)$$

where $\tilde{u} = \frac{a_n}{2^{2n+1}\Gamma(n+\frac{1}{2})}$, and $f_3(t) = L^{-1}\left[(f_3(s))^U\right] = \frac{t^{-1+\frac{U}{2}}}{\Gamma\left(\frac{U}{2}\right)}$.

$$\therefore f_G(t) = -(\sqrt{\pi})^U\left\{\left(f_1(t)\circledast f_3(t)\right)\circledast f_2(t)\right\}, = -(\sqrt{\pi})^U\left\{\left(\int_0^t\delta(t).\frac{(t-v)^{-1+\frac{U}{2}}}{\Gamma\left(\frac{U}{2}\right)}dv\right)\circledast f_2(t)\right\}. = \left\{\left(\frac{\left(2t^{\frac{U}{2}}\delta(t)\right)}{U\Gamma\left(\frac{U}{2}\right)}\right)\circledast f_2(t)\right\}. \quad (68)$$

Thus, a lower bound on the coherent capacity can be written as follows:

$$C_{coherent}(P_{\mathcal{I}}) \geq \frac{1}{2}\left[U\log_2\frac{P_{\mathcal{I}}}{N_t^2}+\mathbb{E}\left[\log_2 f_G(t)\right]\right], \quad (69)$$

and the d.o.f are given as:

$$\text{d.o.f} = \frac{U}{2} = \frac{\min(N_t,N_r)}{2}. \quad (70)$$

For the case $N_t = N_r$, at high SNR,

$$C_{coherent}(P_{\mathcal{I}}) \geq \frac{1}{2}\left[N_t\log_2\frac{P_{\mathcal{I}}}{N_t^2}+\mathbb{E}\left[\log_2 f_G(t)\right]\right], \quad (71)$$

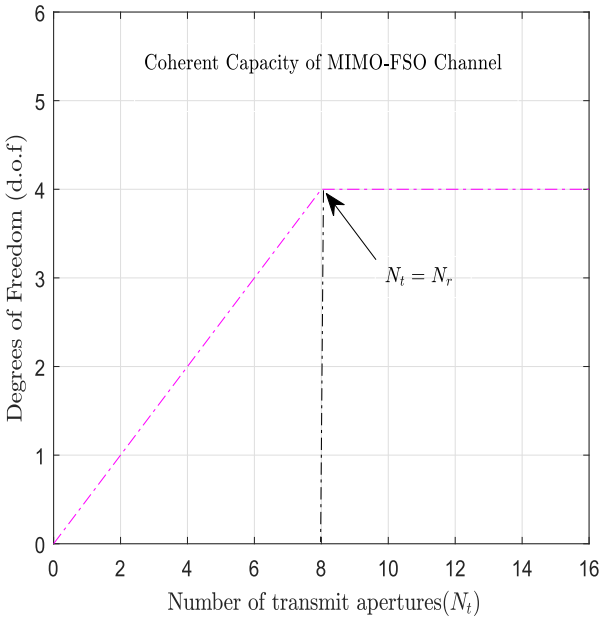

FIGURE 1. The optimal tradeoff between the degrees of freedom (d.o.f) and the number of transmit apertures (N_t).

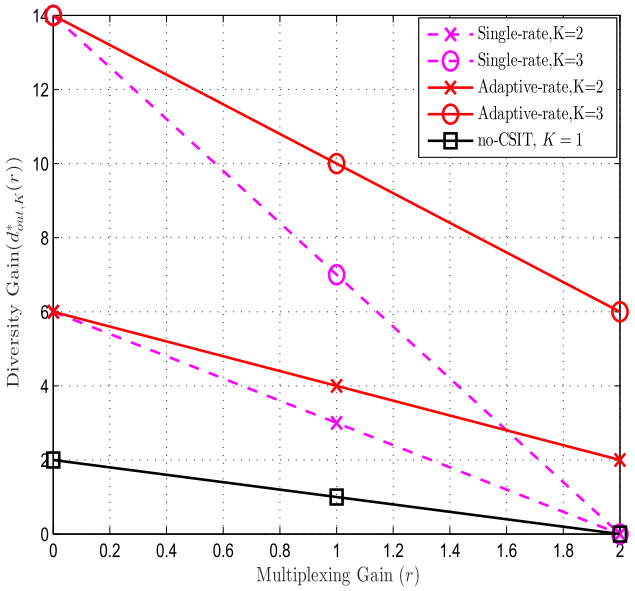

FIGURE 2. Optimal DMT of single-rate and adaptive-rate ($r_{\min} = 0.001$) transmission over a 2 × 2 negative exponential MIMO-FSO channel with different feedback resolution K.

and the d.o.f corresponding to this case is given as:

$$\text{d.o.f} = \frac{N_t}{2} = \frac{N_r}{2}. \tag{72}$$

Remark 7: The result suggests that any MIMO-FSO channel can be viewed as $\frac{U}{2}$ parallel spatial channels; hence the number $\frac{U}{2} = \frac{\min(N_t,N_r)}{2}$ is the total number of d.o.f to communicate. Now, it is good enough to transmit independent information symbols in parallel through these spatial channels. This idea is also known as 'spatial multiplexing'.

Numerical Example: In this particular numerical example, we fix the number of receive apertures (N_r) and vary the number of transmit apertures (N_t). It can be clearly seen from Fig. 1 that the number of d.o.f increases linearly with N_t till $N_t = N_r$ (peak value), and then it saturates. The optimal tradeoff is observed at $N_t = N_r$ and it is given as (N_t, d.o.f) = (8, 4). It employs that by installing eight number of transmit and receive apertures, the MIMO-FSO channel would offer four d.o.f. These four d.o.f further implies that we can transmit four independent information symbols in parallel through the spatial channels. Another useful insight that can be drawn is that in a coherent channel, by adding more transmit apertures, although the d.o.f is not increased proportionally all the time, the capacity increases by a constant.

VI. RESULTS AND DISCUSSION

In this section, we present and discuss the numerical results derived in the previous sections. It is discussed in the previous sections that the log-normal channel is used for modelling of low AT regime. The typical value of σ_l taken for a FSO system is less than one. Further, to model the strong AT regime in case of gamma-gamma channel, the parameter (α, β) is taken as (4.2,1.4), as given in [18].To validate the results of our proposed model, we compare our results with the no-CSIT based MIMO-FSO DMT system model [21].

In Fig. 2, the optimal DMT curves for single-rate and adaptive-rate based 2×2 MIMO-FSO system with a near-zero minimum multiplexing gain ($r_{\min} = 0.001$) are compared. We first compare the optimal DMT curves for a single-rate MIMO-FSO system employing negative exponential channel model with different values of feedback resolution (K). For $K = 2$, it can be seen from Fig. 2 that the maximum optimal diversity gain that can be attained is given as $d^*_{out,K}(0) = 6$. The minimum diversity gain is given as $d^*_{out,K}(2) = 0$, at a maximum multiplexing gain given by $r^*_{\max} = \min(N_t, N_r) = 2$, for a 2 × 2 MIMO-FSO system. This minimum diversity gain is due to the optimal tradeoff between the diversity gain and the multiplexing gain. For $K = 3$, the maximum optimal diversity gain that can be attained is given as $d^*_{out,K}(0) = 14$ (a drastic increase). Now, we compare the DMT curves for an adaptive-rate based 2 × 2 MIMO-FSO system with different values of feedback resolution. It can be clearly seen from the figure that for an adaptive-rate based 2 × 2 MIMO-FSO system, the maximum optimal diversity gain for $K = 2$ and $K = 3$ is same as that of the single-rate based 2 × 2 MIMO-FSO system, i.e., $d^*_{out,K}(0)$ is same for $K = 2$ and $K = 3$, respectively. However, for $K = 2$, the minimum optimal diversity gain is given as $d^*_{out,K}(2) = 2$ (nonzero). Further, for $K = 3$, the minimum optimal diversity gain is given as $d^*_{out,K}(2) = 6$. So, we observe that in both the cases of $K = 2$ and $K = 3$, the minimum diversity gain is nonzero for an adaptive-rate based MIMO-FSO system. Additionally, there is a significant improvement in the achievement of the minimum optimal diversity gain due to the adaptive-rate based transmission.

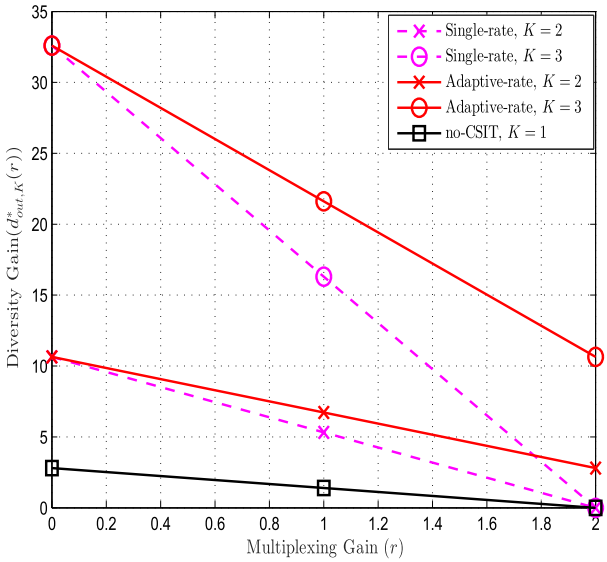

FIGURE 3. Optimal DMT of single-rate and adaptive-rate ($r_{\min} = 0.001$) transmission over a 2×2 gamma-gamma MIMO-FSO channel with different feedback resolution K.

We also consider the case of $K = 1$ (no-CSIT). Even with this considered scenario of no-CSIT [21], we attain a finite optimal diversity gain given as $d^*_{out,K}(0) = 2$.

Figure 3 illustrates the comparison of optimal DMT curves for single-rate and adaptive-rate based 2×2 MIMO-FSO system employing gamma-gamma channel model with a near-zero minimum multiplexing gain ($r_{\min} = 0.001$). Further, we consider the strong AT scenario. We first consider the case of single-rate based 2×2 MIMO-FSO system to compare the DMT curves with different values of feedback resolution (K). For $K = 2$, the maximum optimal diversity gain that can be achieved is given as $d^*_{out,K}(0) \approx 11$. The minimum diversity gain is given as $d^*_{out,K}(2) = 0$, at a maximum multiplexing gain given by $r^*_{\max} = \min(N_t, N_r) = 2$, for a 2×2 MIMO-FSO system. For $K = 3$, the maximum optimal diversity gain that can be achieved is given as $d^*_{out,K}(0) \approx 33$ (a drastic increase). Now, we compare the DMT curves for an adaptive-rate based 2×2 MIMO-FSO system for different values of K. It is quite clear from the figure that for an adaptive-rate transmission, the maximum optimal diversity gain for $K = 2$ and $K = 3$ is same as that of the single-rate based transmission. On the other side, for $K = 2$, the minimum optimal diversity gain is given as $d^*_{out,K}(2) = 3$ (more as compared to the negative exponential channel model). Further, for $K = 3$, the minimum optimal diversity gain is given as $d^*_{out,K}(2) \approx 11$ (same as that of the maximum optimal diversity gain of single-rate and adaptive-rate 2×2 system with $K = 2$. Additionally, we consider the case of no-CSIT ($K = 1$). This scenario of CSI attains a finite optimal diversity gain of $d^*_{out,K}(0) = 3$ (more as compared to the negative exponential channel model).

Remark 8: In Figs. 2 and 3, it can be observed that the DMT performance corresponding to the proposed technique/model in this work is significantly better as compared to the no-CSIT case [21]. This is because [21] does not employ any CSIT dependent power and rate control technique. Moreover, the improved DMT performance is valid for both single-rate as well as adaptive-rate transmission. The main differences between Fig.2 and Fig.3 are: 1) Fig.2 shows the DMT performance of a 2×2 negative exponential MIMO-FSO channel whereas Fig.3 shows the DMT performance of a 2×2 gamma-gamma MIMO-FSO channel. 2) The DMT performance of gamma-gamma MIMO-FSO channel is significantly better as compared to the negative exponential MIMO-FSO channel for both single-rate as well as adaptive-rate transmission scenarios.

Implication 2: Furthermore, we give some useful insights into the above discussion. A few bits of feedback information can significantly increase the diversity gain of a MIMO-FSO channel. A significant increase in the diversity gain can be observed even with coarsely quantized feedback. This implies that from a diversity gain perspective, increasing the feedback resolution is more efficient than adding apertures, provided that transmit power control is possible. The justification behind this " power-control diversity" is that we can segregate certain " bad" channels into regions with polynomially small (in SNR) probability measures and employ polynomially large powers over those regions without violating the power constraint (cf. also [12]). It can be observed from Figs. 2 and 3 that the penalty of keeping the rate independent of the feedback index is relatively large, especially at high multiplexing gain. Importantly, it is possible to achieve nonzero diversity gain with rate adaptation, even at the "maximal" multiplexing gain. Thus, we conclude that rate adaptation is essential to achieve a high throughput together with a optimal nonzero diversity gain. Furthermore, we compare the optimal DMT curves for the channel models considered in Figs. 2 and 3. It is observed that irrespective of the feedback resolution, optimal DMT curves of gamma-gamma channel model outperforms the optimal DMT curves of negative exponential channel model for both single-rate and adaptive-rate based 2×2 MIMO-FSO system.

Figure 4 illustrates the comparison of optimal DMT curves for single-rate and adaptive-rate based 2×2 MIMO-FSO system employing log-normal channel model with a near-zero minimum multiplexing gain ($r_{\min} = 0.001$). It can be clearly seen from the figure that the maximum optimal diversity gain obtained is infinite, $\forall r$ for single-rate as well as adaptive-rate system. However, this infinite diversity gain is valid only for weak turbulence regime, since the log-normal channel model is valid for weak turbulence regime. This achievement does not hold for strong and moderate turbulence regimes. Hence, this is the limitation of the log-normal channel model. Furthermore, for certain values of α and β, the gamma-gamma channel model reduces to log-normal channel model. So, it is almost redundant to discuss the optimal tradeoff for the log-normal channel in the subsequent discussion. Therefore, from now onwards, the main focus would be on the negative exponential and gamma-gamma channel models which covers the entire turbulence regime (weak to strong).

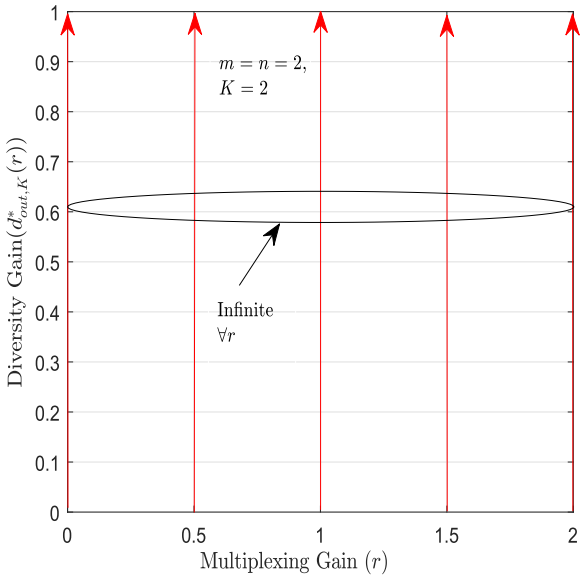

FIGURE 4. Optimal DMT of single-rate and adaptive-rate ($r_{\min} = 0.001$) transmission over a 2 × 2 log-normal MIMO-FSO channel with different feedback resolution K.

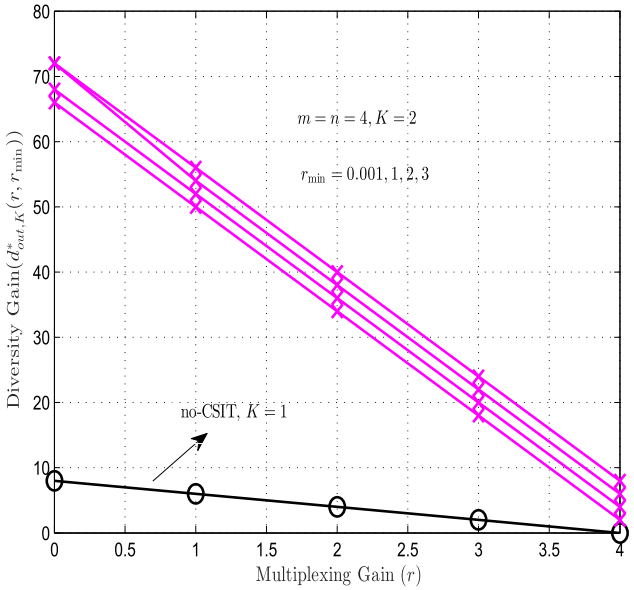

FIGURE 5. Optimal DMT for adaptive-rate transmission over a 4 × 4 negative exponential MIMO-FSO channel with minimum multiplexing gain $r_{\min} = 0.001, 1, 2, 3$.

Moreover, the detailed description related to the limitation of the log-normal channel in terms of accuracy can be found in [18].

In Fig. 5, the optimal DMT curves for an adaptive-rate based 4×4 MIMO-FSO system employing negative exponential channel model with different values of $r_{\min}$ are compared. The different values of $r_{\min}$ are taken as $r_{\min} = 0.001, 1, 2, 3$. Moreover, we consider $K = 2$. For $r_{\min} = 0.001$, the maximum optimal diversity gain that can be attained by an adaptive-rate based 4×4 MIMO-FSO system is given as $d^*_{out,K}(0, 0.001) = 72$. The minimum optimal diversity gain that can be attained is given as

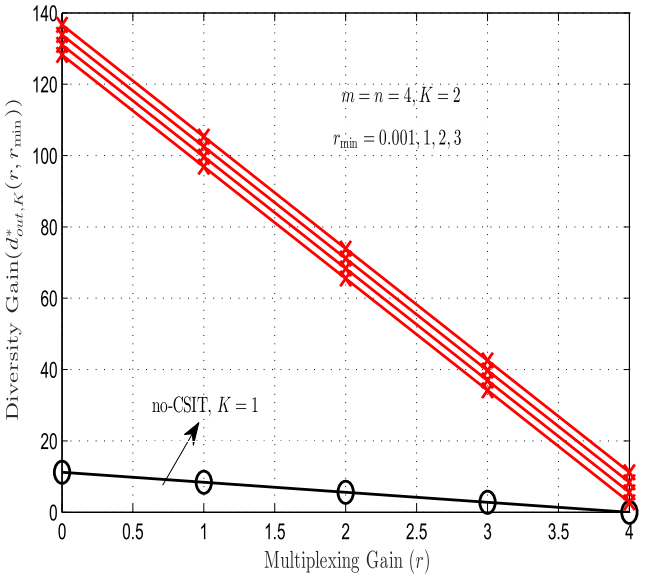

FIGURE 6. Optimal DMT for adaptive-rate transmission over a 4 × 4 gamma-gamma MIMO-FSO channel with minimum multiplexing gain $r_{\min} = 0.001, 1, 2, 3$.

$d^*_{out,K}(4, 0.001) = 8$, at a maximum multiplexing gain given by $r^*_{\max} = \min(N_t, N_r) = 4$, for a 4×4 MIMO-FSO system. For $r_{\min} = 1$, the maximum optimal diversity gain that can be attained is given as $d^*_{out,K}(0, 1) = 72$ (same as that attained with $r_{\min} = 0.001$). However, the minimum optimal diversity gain is given as $d^*_{out,K}(4, 1) = 6$. For $r_{\min} = 2$, the maximum optimal diversity gain achieved by the considered system is given as $d^*_{out,K}(0, 2) = 68$, whereas the minimum optimal diversity gain is given as $d^*_{out,K}(4, 2) = 4$. Further, with $r_{\min} = 3$, the maximum attainable optimal diversity gain is given as $d^*_{out,K}(0, 3) = 66$, and the minimum optimal diversity gain that can be attained is given as $d^*_{out,K}(4, 3) = 2$. Moreover, we also compare the adaptive-rate based optimal DMT scenario with the optimal DMT curve corresponding to the no-CSIT [21] scenario. It is observed that when no-CSIT scenario is considered, the 4×4 MIMO-FSO system employing negative exponential channel model is efficient enough to provide an optimum maximum diversity gain of $d^*_{out,1}(0) = 8$.

Figure 6 illustrates the comparison of optimal DMT curves for an adaptive-rate based 4×4 MIMO-FSO system employing gamma-gamma channel model with different values of $r_{\min}$, and $K = 2$. For $r_{\min} = 0.001$, the maximum optimal diversity gain that can be attained by an adaptive-rate based 4×4 MIMO-FSO system is given as $d^*_{out,K}(0, 0.001) \approx 137$. The minimum optimal diversity gain that can be attained is given as $d^*_{out,K}(4, 0.001) \approx 12$, at a maximum multiplexing gain given by $r^*_{\max} = \min(N_t, N_r) = 4$ for a 4×4 MIMO-FSO system. For $r_{\min} = 1$, the maximum optimal diversity gain that can be attained is given as $d^*_{out,2}(0, 1) \approx 134$. However, the minimum optimal diversity gain is given as $d^*_{out,K}(4, 1) \approx 9$. For $r_{\min} = 2$, the maximum optimal diversity gain achieved by the considered system is given as $d^*_{out,K}(0, 2) \approx 129$, whereas the minimum optimal

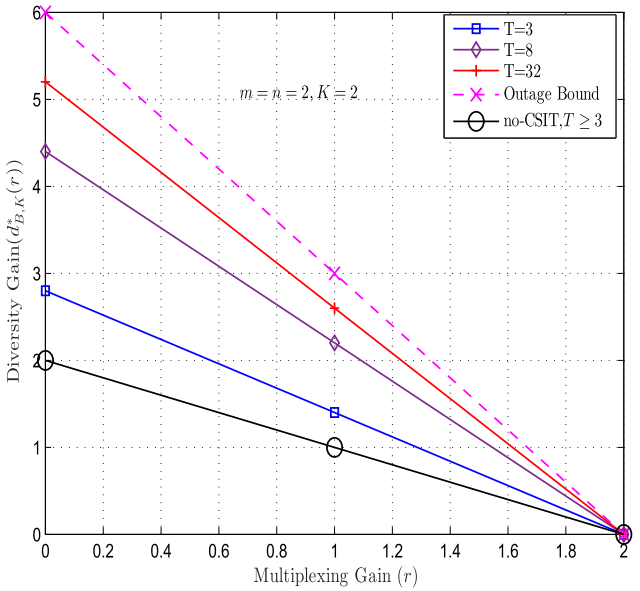

FIGURE 7. Optimal DMT for back-off bounds over a 2 × 2 single-rate based negative exponential MIMO-FSO channel with feedback resolution $K = 2$.

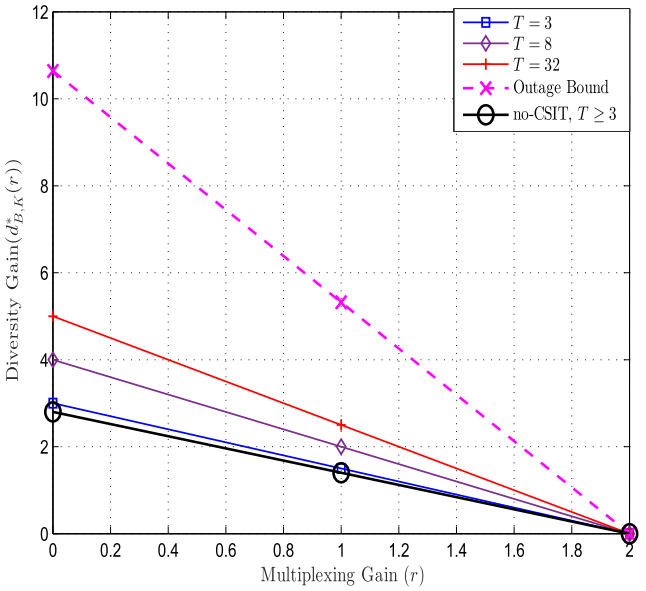

FIGURE 8. Optimal DMT for back-off bounds over a 2 × 2 single-rate based gamma-gamma MIMO-FSO channel with feedback resolution $K = 2$.

diversity gain is given as $d^*_{out,K}(4, 2) \approx 6$. Further, with $r_{\min} = 3$, the maximum attainable optimal diversity gain is given as $d^*_{out,K}(0, 3) \approx 128$, and the minimum optimal diversity gain that can be attained is given as $d^*_{out,K}(4, 3) = 3$. Furthermore, when no-CSIT [21] scenario is considered, the 4 × 4 MIMO-FSO system employing gamma-gamma channel model is efficient enough to provide an optimal maximum diversity gain of $d^*_{out,1}(0) \approx 12$.

Implication 3: The discussion related to Figs. 5 and 6 reflects some useful insights. First, increasing the minimum threshold on the individual rates leads to a degradation in reliability. However, it should not be taken as granted that a small $r_{\min}$ is preferable. It can be clearly seen from the figures that the optimal DMT with very coarse feedback ($K = 2$) is far better than the no-CSIT case [21]. Second, the optimal DMT curves of the gamma-gamma channel model outperforms the optimal DMT curves of the negative exponential channel model by providing a drastic increase in the diversity gain for all different values of $r_{\min}$. Third, when the no-CSIT scenario is considered, even then the optimal DMT of the gamma-gamma channel model outperforms the optimal DMT of negative exponential channel model.

In Fig. 7, we plot the optimal DMT curves for the back-off bounds over a 2 × 2 single-rate based MIMO-FSO system employing negative exponential channel model. These optimal DMT curves are plotted for different codeword lengths ($T = 3, 8, 32$). For $T = 3$, the maximum optimal diversity gain that can be attained is given as $d^*_{B,K}(0) = 2.8$. For other values of T, the maximum optimal diversity gain that can be achieved is given as $d^*_{B,K}(0) = 4.4$ and $d^*_{B,K}(0) = 5.6$, respectively. However, for all the considered values of T, the minimum diversity gain offered by the system is $d^*_{B,K}(2) = 0$, at a maximum multiplexing gain given by $r^*_{\max} = \min(N_t, N_r) = 2$ for a 2 × 2 MIMO-FSO system. Furthermore, a deep inspection of Fig. 7 reveals that even for moderate values of T, the back-off bounds are very tight at high rates, and quickly approach the single-rate outage bound. Figure 8 illustrates the optimal DMT curves for the back-off bounds over a 2 × 2 single-rate based MIMO-FSO system employing gamma-gamma channel model. The optimal DMT curves are plotted for ($T = 3, 8, 32$). For $T = 3$, the maximum optimal diversity gain is given as $d^*_{B,2}(0) = 2.8$ (same as that obtained for negative exponential channel model). For other values of T, the maximum optimal diversity gain that can be achieved is given as $d^*_{B,K}(0) = 4$ and $d^*_{B,K}(0) = 5$, respectively. Moreover, we observe that the optimal diversity gain offered by single-rate based 2 × 2 gamma-gamma channel for $T = 8$ and $T = 32$ is comparatively less as compared to the negative exponential channel (Fig.7). So, one can conclude that for the larger codeword lengths, the negative exponential channel model based back-off bounds provide better diversity gain as compared to the gamma-gamma channel model. Further, a deep inspection of Fig. 8 reveals that gamma-gamma channel model is able to approach the outage bound for very large values of T. On the other side, negative exponential channel model quickly approaches the outage bound for even moderate values of T.

We plot the optimal DMT curves for the back-off bounds over a 2 × 2 adaptive-rate based MIMO-FSO system employing negative exponential channel in Fig. 9. The optimal DMT curves are plotted for $T = 3, 8, 32$. For $T = 3$, the maximum optimal diversity gain that can be attained is given as $d^*_{B,K}(0.5, 0.5) \approx 4$. For other values of T, it is given as $d^*_{B,K}(0.5, 0.5) \approx 9$, and $d^*_{B,K}(0.5, 0.5) \approx 32$, respectively. This indicates the extreme optimality offered by adaptive-rate

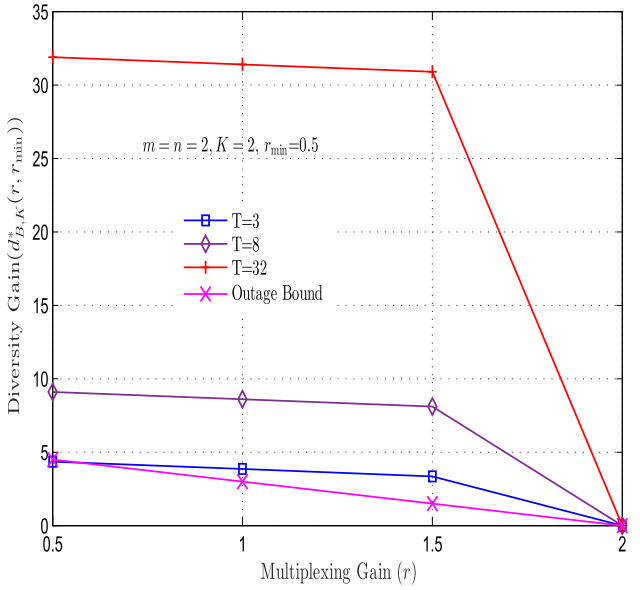

FIGURE 9. Optimal DMT for back-off bounds over a 2 × 2 adaptive-rate based negative exponential MIMO-FSO channel with feedback resolution $K = 2$ and $r_{\min} = 0.5$.

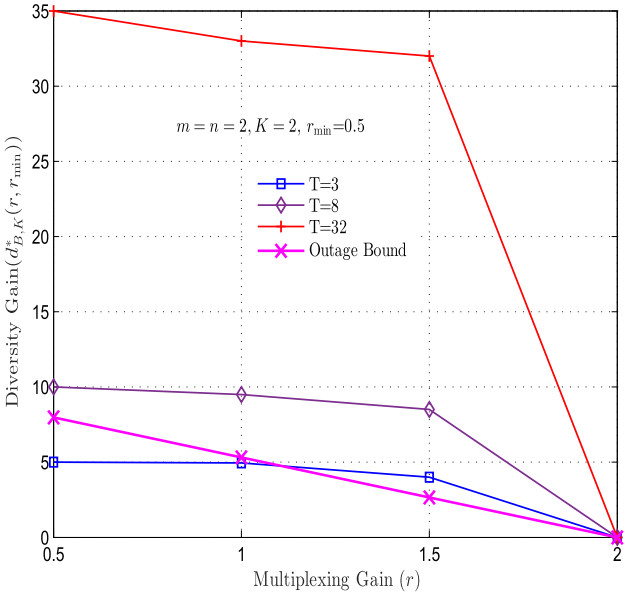

FIGURE 10. Optimal DMT for back-off bounds over a 2 × 2 adaptive-rate based gamma-gamma MIMO-FSO channel with feedback resolution $K = 2$ and $r_{\min} = 0.5$.

transmission as compared to single-rate transmission in context of back-off bounds. Moreover, we observe that even for a small value of T ($T = 3$), the optimal diversity gain corresponding to back-off bound quickly approaches the optimal diversity gain corresponding to outage bound. This is due to the joint optimality offered by the adaptive-rate transmission and backing-off strategy. Even a moderate value of T, i.e., $T = 8$ surpasses the outage bound quickly. Another useful insight that can be drawn from Fig. 9 is the piecewise nature of the optimal DMT curves. For $T = 3$, the slope corresponding to the first segment $0.5 \leq r < 1$ is ≈ -1 and the second segment $1 \leq r < 1.5$ is $= -1$. Further, the third segment $1.5 \leq r \leq 2$ is $= -6.7$. One can conclude that at low rates, adaptive-rate transmission based back-off bounds employing negative exponential channel provide almost a constant optimal diversity gain for a particular value of T. However, this does not hold for high rates which show an abrupt decrement in the diversity gain.

Figure 10 illustrates the optimal DMT curves for the back-off bounds over a 2×2 adaptive-rate based MIMO-FSO system employing gamma-gamma channel model. For $T = 3$, the maximum optimal diversity gain that can be attained is given as $d^*_{B,2}(0.5, 0.5) = 5$. For other values of T, the optimal diversity gain is given as $d^*_{B,K}(0.5, 0.5) = 10$, and $d^*_{B,K}(0.5, 0.5) = 35$, respectively. For all the considered values of T, the optimal diversity gains are more as compared to the negative exponential channel model (Fig. 9). Moreover, a useful insight that can be drawn from Fig. 10 is that even for a small value of $T = 8$, the optimal diversity gain corresponding to back-off bound surpasses the optimal diversity gain corresponding to the outage bound. Furthermore, unlike the negative exponential channel model (Fig. 9), the slope corresponding to different segments of the optimal tradeoff curves is not same at low rates. So, for certain applications where a constant diversity gain is required at low rates, one can prefer to use backing-off strategy accompanied by adaptive-rate transmission based MIMO-FSO system employing a negative exponential channel model.

Remark 9: Figs. 9 and 10, the significance of including the outage bound curve is to compare the DMT performance of back-off bounds with the outage bound. It can be observed form the figures that by applying the back-off strategy, the DMT performance of back-off bounds is significantly better as compared to the outage bound.

In Fig. 11, we plot in the optimal DMT curves for the expurgated bounds over a 2×2 single-rate transmission based MIMO-FSO system employing negative exponential channel model. The expurgated bounds are plotted for ($T = 3$, 8, and 32). It can be easily seen from the figure that the optimal diversity gain $d^*_{E,K}(0) = 6$ for all the considered values of T. For the purpose of completeness, we give the description related to Fig. 12 along with Fig. 11. Fig. 12 illustrates the optimal DMT curves for the expurgated bounds over a 2×2 single-rate transmission based MIMO-FSO system employing gamma-gamma channel model. The optimal diversity gain that can be attained over a gamma-gamma channel model is given as $d^*_{E,K}(0) = 10.64$ irrespective of the value of T. However, the minimum optimal diversity gain is nonzero for all the different values of T but not attained over an optimal multiplexing gain, i.e., $r^*_{\max} = \min(N_t, N_r) = 2$ for a 2×2 system. The implication is that the single-rate based expurgated bounds are very tight at low multiplexing gains (regardless of the codeword length T) as compared to the back-off bounds. This is because the optimization is done over both the codes and the feedback link using expurgation techniques. Furthermore, on comparing the performance of

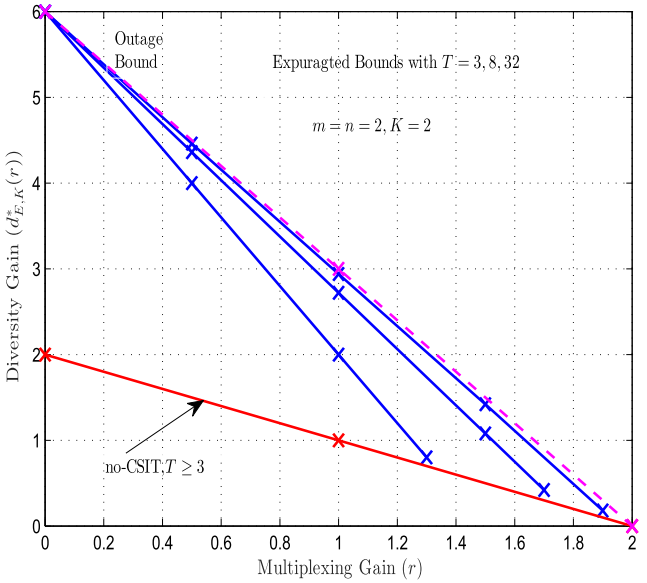

FIGURE 11. Optimal DMT for expurgated bounds over a 2 × 2 single-rate based negative exponential MIMO-FSO channel with feedback resolution $K = 2$.

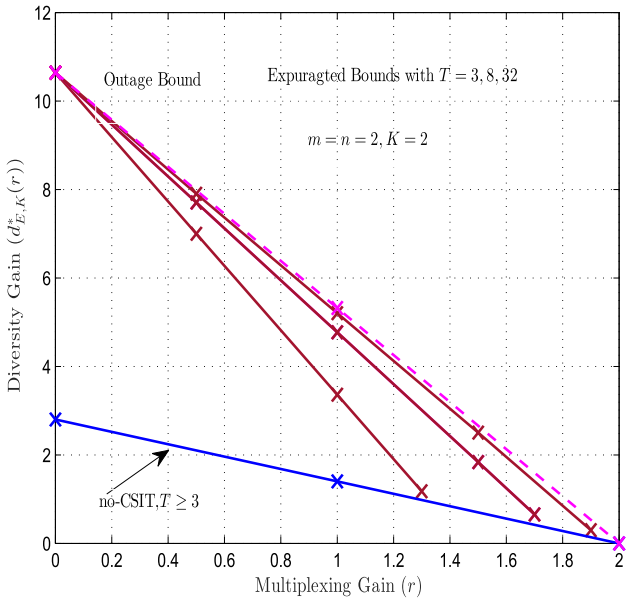

FIGURE 12. Optimal DMT for expurgated bounds over a 2 × 2 single-rate based gamma-gamma MIMO-FSO channel with feedback resolution $K = 2$.

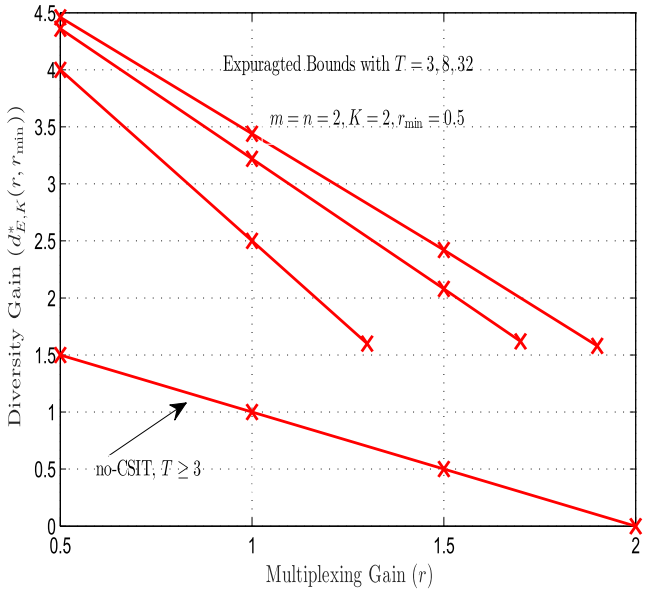

FIGURE 13. Optimal DMT for expurgated bounds over a 2 × 2 adaptive-rate based negative exponential MIMO-FSO channel with feedback resolution $K = 2$ and $r_{\min} = 0.5$.

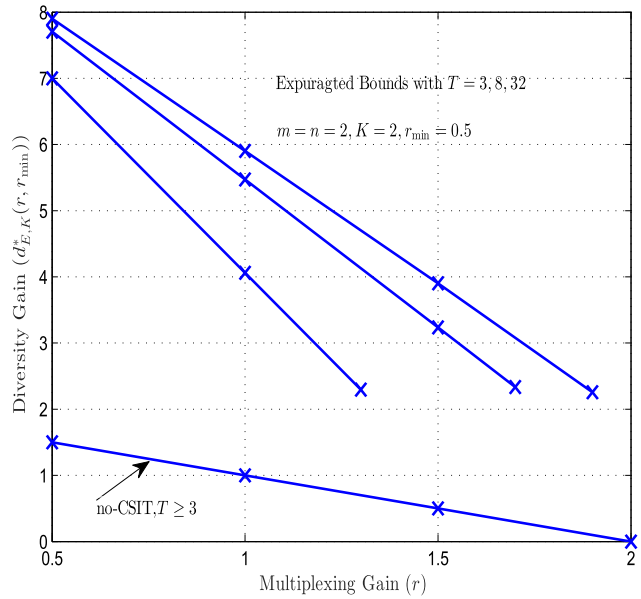

FIGURE 14. Optimal DMT for expurgated bounds over a 2 × 2 adaptive-rate based gamma-gamma MIMO-FSO channel with feedback resolution $K = 2$ and $r_{\min} = 0.5$.

single-rate based back-off bounds with expurgated bounds, we observe that back-off bounds are defined over the entire $(0, n)$ whereas the expurgated bounds only exist for sufficiently small multiplexing gains.

Remark 10: The main differences between Fig.11 and Fig.12 are: 1) Fig.11 shows the optimal DMT for expurgated bounds corresponding to a 2 × 2 single-rate based negative exponential MIMO-FSO channel whereas Fig.12 shows the optimal DMT for expurgated bounds corresponding to a 2×2 single-rate based gamma-gamma MIMO-FSO channel. 2) The optimal DMT performance for expurgated bounds corresponding to gamma-gamma MIMO-FSO channel is significantly better as compared to the negative exponential MIMO-FSO channel.

In Fig. 13, we plot the optimal DMT curves for the expurgated bounds over a 2 × 2 adaptive-rate transmission based MIMO-FSO system employing negative exponential channel model. It can be clearly seen from the figure that the maximum optimal diversity gain for $T = 3$ is given as $d^*_{E,K}(0.5, 0.5) = 4$. However, for $T = 8$ and $T = 32$,

the maximum optimal diversity gain is approximately same, and it is given as $d^*_{E,K}(0.5, 0.5) = 4.5$. One useful insight that can be drawn from Fig. 13 is the same minimum optimal diversity gain offered at high rates for all values of T. This is quite unique as compared to single-rate based expurgated bounds corresponding to Fig. 11. Figure 14 illustrates the optimal DMT for the expurgated bounds over a 2×2 adaptive-rate transmission based MIMO-FSO system employing gamma-gamma channel model. For $T = 3$, the maximum optimal diversity gain is given as $d^*_{E,K}(0) = 7$ (more as compared to Fig. 13). However, for $T = 8$ and $T = 32$, the maximum optimal diversity gain is approximately same, and it is given as $d^*_{E,K}(0.5, 0.5) \approx 8$. One can conclude that the gamma-gamma channel model is also efficient enough to provide the same minimum optimal diversity gain at high rates for all the considered values of T. Furthermore, this minimum optimal diversity gain is comparatively more as compared to Fig. 13.

VII. CONCLUSION AND FUTURE SCOPE

A novel investigation of the optimal DMT for the MIMO-FSO system for different channel models under different transmission scenarios (single-rate and adaptive-rate transmission) has been done. The concept of CSIT-dependent power controller using limited quantized feedback has been adopted to derive the optimal DMT. Moreover, we have analyzed the optimal DMT for the outage based upper bound. Herein, we observed that the gamma-gamma channel model outperforms the negative exponential channel model in terms of the optimal DMT. It is worth mentioning that the results in [10], [11] only demonstrate the asymptotic outage corresponding to a particular class of feedback schemes. However, our proposed technique is valid for any feedback scheme. In addition to this, to validate our results, we compare our results with the No-CSIT based MIMO-FSO DMT [21]. It has been observed that the proposed technique offers much superior performance as compared to the No-CSIT based MIMO-FSO DMT [21] for both single-rate and adaptive-rate transmission scenarios. Further, we have analyzed the lower bounds (back-off bounds and expurgated bounds) on the optimal DMT by giving useful insights. Furthermore, a novel study based on the optimal tradeoff between the degrees of freedom (d.o.f) and the number of transmit apertures (N_t) is also done. Finally, our results provide a comprehensive understanding of the adaptive-rate based MIMO-FSO systems in the asymptotically high-SNR regime. As far as the future research direction is concerned, the erroneous feedback based transmission scenario and its impact on the DMT performance would be an interesting research problem.

APPENDIX A
PROOF OF LEMMA 1

Let $\{P_i\}_{i=1}^K$ be an arbitrary power codebook and $\mathcal{I}(H)$ be a deterministic index mapping form a channel matrix to an integer feedback index required for the transfer of feedback information. Moreover, we consider two constraints such that $0 \leq P_1 < \cdots < P_K$ and $\sum_{i=1}^K \Pr(\mathcal{I}(H) = i)\, P_i \leq \text{SNR}$. It should be noticed that for an optimal feedback scheme, the above mentioned power constraints must be considered. This is because the joint pdf of the singular values of the channel matrix are continuous and takes on positive values. Moreover, we need to show that the optimal index mapping must have the form (21).

Consider another feedback scheme using the same power codebook and the following index mapping:

$$\mathcal{I}^*(H) = \begin{cases} 1, & \text{if} C(H, P_{\mathcal{I}(H)}) < R \\ \min\{i : i \in \{1, \cdots, \mathcal{I}(H)\},\ C(H, P_i) \geq R\}, & \\ \quad \text{otherwise.} & \end{cases} \tag{73}$$

This results in the same outage probability as for $\mathcal{I}(H)$. However, the average transmit power of the newly constructed feedback scheme is

$$\sum_{i=1}^{K} \Pr\left(\mathcal{I}^*(H) = i\right) P_i \leq \sum_{i=1}^{K} \Pr(\mathcal{I}(H) = i)\, P_i \tag{74}$$

Taking this average transmit power constraint into consideration, we must deal with two probabilistic events. First, all the channels that are in outage are mapped to $\mathcal{I} = 1$. Second, an event that the channel is not in outage is mapped to the smallest power in the power codebook that has the ability to '' invert'' H. It means that the mutual information is greater than R when this power level is applied at the transmitter.

We still need to show that no set of channel conditions with positive probability measure that can be inverted by some P_i, $i \geq 2$, is mapped to $\mathcal{I} = 1$. On the other side, let us consider a set $\mathcal{S}$ and an index $j \geq 2$ such that $\Pr(H \in \mathcal{S}) = \tau_S > 0$ and $C(H, P_j) \geq R$, $\mathcal{I}(H) = 1$, $\forall H \in \mathcal{S}$. With this, there exists a $P_j^* \in (P_{j-1}, P_j)$such that $\Pr\left(P^m(H; R_j) \in (P_j^*, P_j)\right) = \tau_S$ where $P^m(H; R)$ satisfies $C\left(H, P^m(H; R)\right) = R$, i.e., the minimum power required to invert H. With all choices of $\mathcal{S}$ subject to $\Pr(H \in \mathcal{S}) = \tau_S$ are same in terms of both average power and outage probability, we can consider

$$\mathcal{S} = \left\{H : P^m(H; R) \in (P_j^*, P_j)\right\}. \tag{75}$$

Let us consider another feedback scheme with optimal index mapping $\mathcal{I}^*(H)$ and the power codebook

$$\left\{P_1, \cdots, P_{j-1}, P_j^*, P_{j+1}, \cdots, P_K\right\}. \tag{76}$$

By construction, this results in giving the same outage probability as that obtained by the codebook $\{P_i\}_{i=1}^K$ together with $\mathcal{I}^*(H)$. On the other side, the newly constructed feedback scheme uses less average power, because $P_j^* < P_j$, i.e., the power constraint is inactive. Hence, the optimal index mapping must have the form (21), which results in outage probability of $P_{out}(R, P_K^*)$.

Furthermore, since $P_1^* < \cdots < P_K^*$, the event $C(P_{i-1}^*, H) > R$ also implies $C(P_i^*, H) > R$. Hence,

$$\begin{aligned}&\Pr\Big(C(P_i^*, H) > R, C(P_{i-1}^*, H) < R\Big)\\ &= \Pr\Big(C(P_{i-1}^*, H) < R\Big) - \Pr\Big(C(P_i^*, H) < R\Big)\\ &= P_{out}\Big(R, P_{i-1}^*\Big) - P_{out}\Big(R, P_i^*\Big). \end{aligned} \tag{77}$$

Eventually, the average power of the system is given by

$$\begin{aligned}&\left[P_{out}\left(R, P_K^*\right) + 1 - P_{out}(R, P_1^*)\right] P_1^*\\ &\quad + \sum_{i=2}^{K}\left[P_{out}\left(R, P_{i-1}^*\right) + 1 - P_{out}(R, P_i^*)\right] P_i^*. \end{aligned} \tag{78}$$

Since $P_{out}(R, P_K)$ is a monotonically decreasing function of P_K for any given $R > 0$, the optimal power codebook is the solution to (20).

APPENDIX B
PROOF OF PROPOSITION 1

Consider an optimal power codebook with largest power level P_K^*. Corresponding to P_K^*, we first derive an upper bound on the SNR exponent. If SNR is denoted as τ and $\{\bar{P}_i\}$ represents the solution to the optimization problem of (20). Then, a more relaxed version of (20) can be written as:

$$\begin{aligned}&\max P_K\\ &s.t.\ \left[P_{out}\left(R, P_K\right) + 1 - P_{out}(R, P_1)\right] P_1 \le \tau\\ &\quad \left[P_{out}\left(R, P_{i-1}\right) - P_{out}(R, P_i)\right] P_i \le \tau, i \ge 2\\ &\quad 0 \le P_1 < \cdots < P_K. \end{aligned} \tag{79}$$

Moreover, it is intuitive that $\bar{P}_K \ge P_K^*$. Further, the constraints of (79) imply $\sum_{i=1}^{K} \frac{\tau}{\bar{P}_i} \ge 1$. This leads to $\bar{P}_1 \le K\tau$; otherwise, $\sum_{i=1}^{K} \frac{\tau}{\bar{P}_i} < K\frac{1}{K} = 1$. Since K is a finite constant, we have $\bar{P}_1 \le \tau$. The implication of Lemma 2 results in

$$P_{out}(R, \bar{P}_1) \dot{\ge} \tau^{-J(r,1)} = \tau^{-d_{out,1}^*(r)}. \tag{80}$$

Furthermore, the constraints of (79) require $\frac{\tau}{\bar{P}_2} + P_{out}(R, \bar{P}_2) \ge P_{out}(R, \bar{P}_1)$. This results in

$$\frac{\tau}{\bar{P}_2} + P_{out}(, \bar{P}_2) \dot{\ge} \tau^{-d_{out,1}^*(r)}. \tag{81}$$

For any $\delta > 0$, if $\bar{P}_2 \doteq \tau^{-J(r,1+d_{out,1}^*(r)+\delta)}$, then

$$\frac{\tau}{\bar{P}_2} + P_{out}(\bar{P}_2) \doteq \tau^{-d_{out,1}^*(r)-\delta} + \tau^{-J(r,1+d_{out,1}^*(r)+\delta)}$$

which contradicts (79). This is because

$$J(r, 1 + d_{out,1}^*(r) + \delta) > J(r, 1) = d_{out,1}^*(r). \tag{82}$$

Hence, $\bar{P}_2 \dot{\le} \tau^{1+d_{out,1}^*(r)}$, and thus

$$P_{out}(R, \bar{P}_2) \dot{\ge} \tau^{-J(r,1+d_{out,1}^*(r))} = \tau^{-d_{out,2}^*(r)}. \tag{83}$$

Finally, by induction, we get $\bar{P_K} \dot{\le} \tau^{1+d_{out,K-1}^*(r)}$ and

$$\begin{aligned}P_{out}(R, P_K^*) &\ge P_{out}(R, \bar{P_K}) \dot{\ge} \tau^{-J(r,1+d_{out,K-1}^*(r))}\\ &= \tau^{-d_{out,K}^*(r)}. \end{aligned} \tag{84}$$

Eventually, a lower bound on P_K^* is obtained according to the following selection:

$$P_1 = \frac{\tau}{K}, \quad P_2 = \frac{\tau}{KP_{out}(P_1)}, \cdots, P_K = \frac{\tau}{KP_{out}(P_{K-1})}. \tag{85}$$

As the above P_i's satisfy the constraints of (20), we have $P_K^* \ge P_K$. By construction, $P_1 \doteq \tau$, $P_2 \doteq \tau^{1+d_{out,1}^*(r)}, \cdots, P_K \doteq \tau^{1+d_{out,K-1}^*(r)}$. Therefore,

$$P_{out}(R, P_K^*) \le P_{out}(R, P_K) \doteq \tau^{-d_{out,K}^*(r)}. \tag{86}$$

Hence, proved.

APPENDIX C
PROOF OF PROPOSITION 2

Consider a arbitrary class of deterministic feedback schemes $\mathcal{F}$ providing a multiplexing gain of r. For a certain class of feedback scheme, the outage probability can be written as $P_{out,\mathcal{F}} \doteq \tau^{-d_{out,\mathcal{F}}(r)}$. We first derive an upper bound on $d_{out,\mathcal{F}}(r)$. Let $p_i, i = 1, \cdots, K$, such that $P_i \doteq \tau^{p_i}$, where $p_i > 0$. Using the long-term power constraint (3), we have

$$\sum_{i=1}^{K} \Pr(\mathcal{I} = i)\tau^{p_i} \dot{\le} \tau. \tag{87}$$

Without violating the generality, assume that $0 < p_1 \le \cdots \le p_K$. Let us consider $\bar{K}$ be any integer $\in \{1, \cdots, K\}$ such that $p_{\bar{K}} \le 1$ and $p_{\bar{K}+1} > 1$. In order to have the existence of such a $\bar{K}$, we use the convention $p_{K+1} = \infty$, else (81) will be violated. Further, (81) indicates that $\Pr(\mathcal{I} = i)\,\tau^{p_i} \le \tau, \forall i$. Hence, $\Pr(\mathcal{I} = i) \le \tau^{1-pi}$, and

$$\begin{aligned}\lim_{T\to\infty} \frac{\Pr(\mathcal{I} = i)R_i}{\log \tau} &= r_i \lim_{T\to\infty} \Pr(\mathcal{I} = i)\\ &= 0, \quad \forall i \ge \bar{K} + 1. \end{aligned} \tag{88}$$

Herein, $\bar{K}$ corresponds to the number of regions having a significant contribution to the overall throughput. On the other side, $K - \bar{K}$ can be seen as the number of regions that correspond to improving the overall reliability.

Let $r_{\max} = \max(r_1, \cdots, r_{\bar{K}})$. By (5), $r \le r_{\max}$. Let us consider the first $\bar{K}$ regions of $\mathcal{F}$. Moreover, the power levels associated with these regions have dominant SNR exponent less than or equal to 1. For the purpose of completeness, we consider another class of feedback schemes $\hat{\mathcal{F}}$ such that $\hat{r_1} = r < r_{\max}$, $\hat{r}_i = r_{\min} \le r_i$, $\forall i \ge 2$, and employing $\hat{P}_1 \doteq \tau$. With this, we can write $P_{out,\hat{\mathcal{F}}} \doteq \tau^{-d_{\hat{\mathcal{F}}}(r)}$. Maximizing $d_{\hat{\mathcal{F}}}(r)$ will result an upper bound on $d_{out,\mathcal{F}}(r)$.

Our aim is to minimize the SNR exponent of outage. For this, we consider $\hat{\mathcal{F}}$ such that

$$\hat{P}_i > 1, \quad \forall i \ge 2 \tag{89}$$

where $\hat{P}_i \doteq \tau^{p_i}$. On the other side, assume the contrary, i.e., $\hat{P}_2 \doteq \tau$. Thus, $d_{\hat{\mathcal{F}}}(r)$ can be upper bounded by $d_{\bar{K}-1}^*(r_{\min})$. It can be easily shown that such an upper bound lies strictly below $d_{\hat{\mathcal{F}}}(r)$ when (87) is assumed. Further, assuming (87) implies, $\bar{K} = 1$.

REFERENCES

[1] 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.
[2] H. ElGamal, G. Caire, and M. O. Damen, "Lattice coding and decoding achieve the optimal diversity–multiplexing tradeoff of MIMO channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 968–985, Jun. 2004.
[3] S. Sfar, L. Dai, and K. B. Letaief, "Optimal diversity-multiplexing tradeoff with group detection for MIMO systems," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1178–1190, Jul. 2005.
[4] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.
[5] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.
[6] G. Jongren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.
[7] G. Jongren and M. Skoglund, "Quantized feedback information in orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2473–2486, Oct. 2004.
[8] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.
[9] K. Kiran Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.
[10] A. Khoshnevis and A. Sabharwal, "On diversity and multiplexing gain of multiple antenna systems with transmitter channel information," in *Proc. 42nd Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Sep. 2004.
[11] A. Khoshnevis and A. Sabharwal, "Achievable diversity and multiplexing in multiple antenna systems with quantized power control," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Seoul, South Korea, May 2005, pp. 463–467.
[12] H. El Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ channel: Diversity–multiplexing–delay tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3601–3621, Aug. 2006.
[13] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.
[14] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1273–1289, May 2001.
[15] A. Khoshnevis and A. Sabharwal, "Performance of quantized power control in multiple antenna systems," in *Proc. IEEE Int. Conf. Commun.*, Paris, France, Jun. 2004, pp. 803–807.
[16] T. T. Kim and M. Skoglund, "Diversity-multiplexing tradeoff of MIMO systems with partial power control," in *Proc. Int. Zurich Seminar Commun.*, Zurich, Switzerland, Feb. 2006, pp. 106–109.
[17] L. Andrews and R. L. Philips, *Laser Beam Propagation through Random Media*. Bellingham, WA, USA: SPIE, 2005.
[18] Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, *Optical Wireless Communications System and Channel Modelling with MATLAB*. Boca Raton, FL, USA: CRC Press, 2013.
[19] S. Aghajanzadeh and M. Uysal, "Diversity–multiplexing trade-off in coherent free-space optical systems with multiple receivers," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 2, no. 12, pp. 1087–1094, Dec. 2010.
[20] 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, Jul. 2016.
[21] 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, Feb. 2019.
[22] J. Luo, L. Lin, R. Yates, and P. Spasojevic, "Service outage based power and rate allocation," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 323–330, Jan. 2003.
[23] N. Jindal and A. Goldsmith, "Capacity and optimal power allocation for fading broadcast channels with minimum rates," *IEEE Trans. Inf. Theory*, vol. 49, no. 11, pp. 2895–2909, Nov. 2003.
[24] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series*, vol. 3. New York, NY, USA: Gordon and Breach, 1990.
[25] J. Zhang, L. Dai, Y. Han, Y. Zhang, and Z. Wang, "On the ergodic capacity of MIMO free-space optical systems over turbulence channels," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1925–1934, Sep. 2015.
[26] R. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inf. Theory*, vol. 11, no. 1, pp. 3–18, Jan. 1965.
[27] S. Tavildar and P. Viswanath, "Approximately universal codes over slow-fading channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233–3258, Jul. 2006.

PRANAV SHARDA (Student Member, IEEE) received the B.Tech. degree in electronics and communication engineering from Lovely Professional University, India, in 2015, and the M.Tech. degree in electronics and communication engineering (communication systems) from Guru Nanak Dev University (GNDU), India, in 2017. He is currently pursuing the Ph.D. degree in optical wireless communications from the Department of Electrical Engineering, Indian Institute of Technology (IIT) Delhi, New Delhi, India. He was amongst the top six percent UGC-NET qualified candidates in electronic science subject, India. His research interests include free-space optical communications, visible light communications, and multiple-input multiple-output systems for wireless communications. He received the Gold Medal for getting the First Rank from the M.Tech. degree, GNDU.

MANAV R. BHATNAGAR (Senior Member, IEEE) received the M.Tech. degree in communications engineering from the Indian Institute of Technology Delhi, New Delhi, India, in 2005, and the Ph.D. degree in wireless communications from the Department of Informatics, University of Oslo, Oslo, Norway, in 2008. From 2008 to 2009, he was a Postdoctoral Research Fellow with the University Graduate Center (UNIK), University of Oslo. He held visiting appointments with the Wireless Research Group, Indian Institute of Technology Delhi, the Signal Processing in Networking and Communications (SPiNCOM) Group, University of Minnesota Twin Cities, Minneapolis, MN, USA, the Alcatel-Lucent Chair, SUPELEC, France, the Department of Electrical Computer Engineering, Indian Institute of Science, Bengaluru, India, UNIK, University of Oslo, the Department of Communications and Networking, Aalto University, Espoo, Finland, and the INRIA/IRISA Laboratory, University of Rennes, Lannion, France. He is currently a Professor with the Department of Electrical Engineering, Indian Institute of Technology Delhi, where he is also a Brigadier Bhopinder Singh Chair Professor. His research interests include signal processing for multiple-input-multiple-output systems, cooperative communications, non-coherent communication systems, distributed signal processing for cooperative networks, multiuser communications, ultrawideband-based communications, free-space optical communication, cognitive radio, software defined radio, power line communications, molecular communications, and satellite communications. He is a Fellow of the Institution of Engineering and Technology (IET), U.K., the Indian National Academy of Engineering (INAE), the National Academy of Sciences, India (NASI), the Institution of Electronics and Telecommunication Engineers (IETE), India, and the Optical Society of India (OSI). He received the NASI-Scopus Young Scientist Award in engineering category, in 2016, the Shri Om Prakash Bhasin Award in electronics and information technology, in 2016, and the Hari Om Ashram Prerit Dr. Vikram Sarabhai Research Award, in 2017. He serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, from 2011 to 2014. He was selected as an Exemplary Reviewer of the IEEE COMMUNICATIONS LETTERS, from 2010 to 2012. He was also an Exemplary Reviewer of the IEEE TRANSACTIONS ON COMMUNICATIONS, in 2015.