

Low-SNR Capacity of MIMO Optical Intensity Channels

Anas Chaaban, Zouheir Rezki, and Mohamed-Slim Alouini

Abstract—The capacity of the multiple-input multiple-output (MIMO) optical intensity channel is studied, under both average and peak intensity constraints. We focus on low SNR, which can be modeled as the scenario where both constraints proportionally vanish, or where the peak constraint is held constant while the average constraint vanishes. A capacity upper bound is derived, and is shown to be tight at low SNR under both scenarios. The capacity achieving input distribution at low SNR is shown to be a maximally-correlated vector-binary input distribution. Consequently, the low-SNR capacity of the channel is characterized. As a byproduct, it is shown that for a channel with peak intensity constraints only, or with peak intensity constraints and individual (per aperture) average intensity constraints, a simple scheme composed of coded on-off keying, spatial repetition, and maximum-ratio combining is optimal at low SNR.

I. INTRODUCTION

Intensity modulation is a simple transmission scheme that is favorable in applications such as optical wireless communications (OWC) [1]. In this context, the transmitter modulates the intensity of a light source, and the receiver uses a photodetector to detect the intensity signal. The resulting scheme is known as intensity-modulation direct-detection (IM-DD).

This mode of operation imposes a nonnegativity constraint on the transmit signal. In addition to this, the intensity is commonly constrained by average and peak constraints due to safety and practical considerations. At the receiver side, the signal is disturbed by several sources of noise, such as ambient light and thermal noise. The resulting channel can be modeled as a Poisson channel [2] or an additive Gaussian noise channel [3], [4] depending on the regime of operation, the latter being more common in regimes with high received signal power. In this paper, we focus on the IM-DD channel with input-independent Gaussian noise [3], which is a suitable model for OWC with strong ambient light and/or thermal noise [5].

Due to the above constraints, the capacity of the Gaussian IM-DD channel is not the same as the classical Gaussian channel used for modeling radio-frequency (RF) communications. Consequently, studying the capacity of this channel is important for understanding the fundamental limits of OWC, which has witnessed increasing attention recently, see [1], [6], [7] and references therein. This has motivated researchers to study the capacity of this channel. In this context, the capacity of the single-input single-output (SISO) IM-DD channel has

been studied in [3], [8], [9]. The SISO IM-DD channel with input-dependent Gaussian noise, which models strong relative intensity noise, has been studied in [4]. The capacity of multi-user SISO IM-DD channels has been studied in [10], [11].

Similar to RF communication, one can also realize MIMO transmission in OWC using multiple transmit apertures and multiple detectors [12] or multiple colors in visible-light communication [13]. If the transmit-receiver aperture pairs do not interfere, then we obtain a system of parallel channels which was studied in [14]. Otherwise, we get a general MIMO IM-DD channel which has been studied in [12], [15]–[17] where various communication schemes were analyzed. Recently, the capacity of the Gaussian MIMO IM-DD channel has been studied in [18]–[20], where capacity bounds have been derived, and the high-SNR capacity has been characterized for the case where the number of detectors is more than the number of transmit apertures. However, these works did not consider the low-SNR capacity of the MIMO IM-DD channel. This can arise if the receiver experiences strong noise, or if the average and/or peak intensity constraints at the transmitter are small.

In this paper, we characterize the capacity of this channel at low SNR under average and peak intensity constraints. We start by deriving a capacity upper bound using the “Gaussian maximizes entropy” principle. Then, we prove that this bound is tight at low SNR by devising a capacity achieving scheme in this regime. The capacity achieving input distribution is shown to be a *vector-binary distribution with maximum correlation between its components*, whose achievable rate at low SNR is derived using results in [21]. For instance, for a transmitter with 3 transmit apertures, a peak constraint \mathcal{A} , and average constraints $\mathcal{E}_1 \geq \mathcal{E}_2 \geq \mathcal{E}_3$ for the three transmit apertures respectively, the transmit symbols are from the set $\{(0, 0, 0), (\mathcal{A}, 0, 0), (\mathcal{A}, \mathcal{A}, 0), (\mathcal{A}, \mathcal{A}, \mathcal{A})\}$. This bears resemblance to the structure of the high-SNR capacity-achieving distribution for multiple-input single-output (MISO) IM-DD channels discovered recently in [22], where one transmit aperture must be “on” on order for the remaining ones to be activated. This leads to the low-SNR capacity which we express for a channel under a total average intensity constraint with two cases: (i) proportionally vanishing average and peak constraints, and (ii) a fixed peak constraint and a vanishing average constraint. As a byproduct, we obtain the low-SNR capacity for a channel under individual (per aperture) average intensity constraints also with the two cases above.

In addition to the interesting optimal input distribution above, we conclude the following. Under individual average intensity constraints, we show that the optimal scheme at low SNR consists of *coded on-off keying (OOK)*, *spatial repetition*,

A. Chaaban and M.-S. Alouini are with the Division of Computer, Electrical, and Mathematical Sciences and Engineering (CEMSE) at King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. Email: {anas.chaaban, slim.alouini}@kaust.edu.sa.

Z. Rezki is with the Department of Electrical and Computer Engineering, University of Idaho, Moscow, ID, USA. Email: zrezki@uidaho.edu.

and maximum-ratio combining (MRC). In other words, the transmitter encodes a single binary stream, which is sent bit-by-bit by repeating each bit at all transmit apertures, i.e., for the example above, the input distribution consists of symbols in $\{(0, 0, 0), (\mathcal{A}, \mathcal{A}, \mathcal{A})\}$. The receiver performs MRC [23] to combine the received signals of the detectors into a single stream which is then decoded. This simple scheme also applies for the channel under peak constraints only.

Next, we introduce the channel model in Sec. II. Then, we state the main results of the paper in Sec. III which we support with numerical evaluations. The detailed derivation of the results is given in Sec. IV and the paper is concluded in Sec. V. Throughout the paper, we denote scalars, vectors, and matrices using normal font, boldface lower case, and boldface upper case letters, respectively. The distinction between constants and random quantities will be clear from the context. We use \mathbf{I}_M and $\mathbf{1}_M$ to denote the $M \times M$ identity matrix and the M -dimensional all-ones vector, respectively. The ℓ_2 -norm of \mathbf{x} is denoted $\|\mathbf{x}\|$, its transpose denoted \mathbf{x}^T , and the trace of \mathbf{X} is denoted $\text{Tr}(\mathbf{X})$. The different entropy and mutual information are denoted $h(\cdot)$ and $I(\cdot; \cdot)$, respectively. We write $f(x) = o(x)$ to indicate that $\lim_{x \rightarrow 0} \frac{f(x)}{x} = 0$, and we use $p(x)$ to denote the distribution of a random variable X .

II. CHANNEL MODEL

Consider an OWC system comprising M transmit and N receive apertures, employing IM-DD. Denote the light intensity of the i th transmitter by x_i , and the received signal at the j th receiver by y_j . At any time instant k , the received vector $\mathbf{y}(k) = (y_1(k), \dots, y_N(k))^T$ can be expressed in terms of the input vector $\mathbf{x}(k) = (x_1(k), \dots, x_M(k))^T$ as (Fig. 1)

$$\mathbf{y}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{z}(k), \quad k = 1, 2, \dots \quad (1)$$

where $\mathbf{z}(k) = (z_1(k), \dots, z_N(k))^T$ is $\mathcal{N}(\mathbf{0}_N, \mathbf{I}_N)$ noise independent and identically distributed (i.i.d.) over time,¹ and $\mathbf{H} \in \mathbb{R}_+^{N \times M}$ is a matrix with elements $h_{j,i} \geq 0$ denoting the channel gain from transmitter i to receiver j . We denote the i th column of \mathbf{H} by \mathbf{h}_i . This Gaussian channel model is common in the OWC context [3].

The transmitter wants to send a message by using the channel n times. The message w is chosen uniformly from the set $\mathcal{W}_n \triangleq \{1, \dots, |\mathcal{W}_n|\}$. This message is encoded into $\mathbf{x}^n = (\mathbf{x}(1), \dots, \mathbf{x}(n))$ and sent over n transmissions. The receiver uses its received signal $\mathbf{y}^n = (\mathbf{y}(1), \dots, \mathbf{y}(n))$ to decode \hat{w} . Decoding is considered successful if $\hat{w} = w$, and erroneous otherwise. The probability of error is defined as $P_n = \frac{1}{|\mathcal{W}_n|} \sum_{w \in \mathcal{W}_n} \text{Prob}(w \neq \hat{w})$. The capacity of this channel is defined as the highest achievable rate $\frac{1}{n} \log |\mathcal{W}_n|$ that can be guaranteed so that $P_n \rightarrow 0$ as $n \rightarrow \infty$.

The transmit signal $x_i(k)$ is a realization of a random variable X_i which satisfies

$$0 \leq X_i \leq \mathcal{A}, \quad (2)$$

¹Spatially correlated noise $\mathcal{N}(\mathbf{0}_N, \mathbf{\Sigma})$ can be decorrelated at the receiver by multiplying the received signal by $\mathbf{\Sigma}^{-1}$ (assumed invertible).

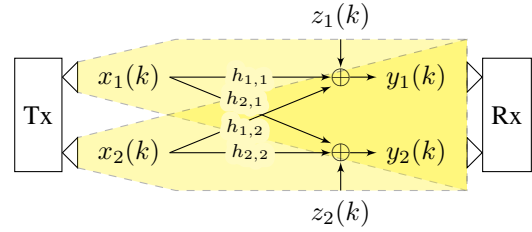


Fig. 1. A MIMO optical wireless communication system at time instant k : $x_i(k) \geq 0$ is the optical intensity, $h_{i,j} \geq 0$ is a channel gain, and $z_i(k)$ is Gaussian noise.

i.e., a peak intensity constraint.² Additionally, $\mathbf{X} = (X_1, \dots, X_N)$ is subjected to a sum average intensity constraint. Denoting $\mathbb{E}[\mathbf{X}]$ by $\mathcal{E} = (\mathcal{E}_1, \dots, \mathcal{E}_M)^T \in [0, \mathcal{A}]^M$, this constraint is given by

$$\sum_{i=1}^M \mathcal{E}_i \leq \mathcal{E} \quad (3)$$

for some $\mathcal{E} \in (0, M\mathcal{A}]$. If $\mathcal{E} > M\mathcal{A}$, then the average constraint is inactive, i.e., the channel is only peak-constrained. We denote the capacity of the channel by $C(\mathcal{A}, \mathcal{E})$.

We study the capacity of this channel at low SNR, i.e., where the signal intensity is small relative to noise power. This can come about in different ways as follows:

- 1) First, \mathcal{A} and \mathcal{E} approach zero proportionally, i.e., $\mathcal{A} \rightarrow 0$ with $\mathcal{E} = \alpha\mathcal{A}$ and $\alpha \in (0, M]$ (if $\alpha > M$, then the average constraint is redundant). In this case, we define the low-SNR capacity as $C_{\text{low},1}(\mathcal{A}, \alpha\mathcal{A})$ which satisfies

$$\lim_{\mathcal{A} \rightarrow 0} \frac{C(\mathcal{A}, \alpha\mathcal{A})}{C_{\text{low},1}(\mathcal{A}, \alpha\mathcal{A})} = 1. \quad (4)$$

- 2) Second, \mathcal{A} is held constant while $\mathcal{E} \rightarrow 0$. In this case, we define the low-SNR capacity as $C_{\text{low},2}(\mathcal{A}, \mathcal{E})$ which satisfies

$$\lim_{\mathcal{E} \rightarrow 0} \frac{C(\mathcal{A}, \mathcal{E})}{C_{\text{low},2}(\mathcal{A}, \mathcal{E})} = 1. \quad (5)$$

The third case where \mathcal{E} is held constant and $\mathcal{A} \rightarrow 0$ is equivalent to the first case with $\alpha = M$, because a constant average constraint \mathcal{E} becomes inactive as \mathcal{A} decreases. Thus, we focus on the two cases above.

Another scenario of interest is where the apertures have individual average intensity constraints, i.e., $\mathbb{E}[X_i] = \mathcal{E}_i \leq \mathcal{E}_{\text{ind}}$ with $\mathcal{E}_{\text{ind}} \in (0, \mathcal{A}]$. The low-SNR capacity of this case follows as a byproduct of the two cases above. We denote the low-SNR capacity under proportional average and peak constraints by $C_{\text{low},1}^{\text{ind}}(\mathcal{A}, \alpha_{\text{ind}}\mathcal{A})$ defined for $\mathcal{A} \rightarrow 0$ where $\mathcal{E}_{\text{ind}} = \alpha_{\text{ind}}\mathcal{A}$ and $\alpha_{\text{ind}} \in (0, 1]$. On the other hand, we denote the low-SNR capacity under a fixed peak constraint by $C_{\text{low},2}^{\text{ind}}(\mathcal{A}, \mathcal{E}_{\text{ind}})$ defined for $\mathcal{E}_{\text{ind}} \rightarrow 0$. The main results are given next.

²The case $0 \leq X_i \leq \mathcal{A}_i$ can be transformed to one with a common peak constraint \mathcal{A} by normalization.

III. MAIN RESULTS

The main results of the paper are summarized in this section. We start with a capacity upper bound which is instrumental for deriving the results.

Lemma 1: The capacity of a MIMO IM-DD channel a peak constraint \mathcal{A} and a given average intensity assignment $\mathcal{E} \in [0, \mathcal{A}]^M$, denoted $c(\mathcal{A}, \mathcal{E})$, satisfies

$$c(\mathcal{A}, \mathcal{E}) \leq \max_{\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}} \frac{1}{2} \log |\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{I}_N|, \quad (6)$$

where $\mathcal{Q}_{\mathcal{E}}$ is the set of covariance matrices \mathbf{Q} of distributions of $\mathbf{X} \in [0, \mathcal{A}]^M$ with $\mathbb{E}[\mathbf{X}] = \mathcal{E}$.

Proof: The proof is given in Sec. IV-A. ■

From Lemma 1, we obtain the following upper bound.

Theorem 1: The capacity $C(\mathcal{A}, \mathcal{E})$ satisfies

$$C(\mathcal{A}, \mathcal{E}) \leq \max_{\mathcal{E} \in \mathcal{S}} c(\mathcal{A}, \mathcal{E}), \quad (7)$$

where $\mathcal{S} = \{\mathcal{E} \in [0, \mathcal{A}]^M | \mathcal{E}_1 + \dots + \mathcal{E}_M \leq \mathcal{E}\}$.

Proof: This follows from Lemma 1 by maximizing with respect to \mathcal{E} . ■

The optimization problems in Lemma 1 is convex. To show this, note first that $\log |\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{I}_N|$ is concave with respect to positive semi-definite \mathbf{Q} . Then, note that $\mathcal{Q}_{\mathcal{E}}$ is a convex set. Namely, if $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ are two distributions of $\mathbf{X} \in [0, \mathcal{A}]^M$ with $\mathbb{E}[\mathbf{X}] = \mathcal{E}$ and covariance matrices $\mathbf{Q}_1, \mathbf{Q}_2 \in \mathcal{Q}_{\mathcal{E}}$, respectively, then $p_3(\mathbf{x}) = \mu p_1(\mathbf{x}) + (1 - \mu)p_2(\mathbf{x})$ has a covariance matrix $\mathbf{Q}_3 = \mu \mathbf{Q}_1 + (1 - \mu)\mathbf{Q}_2$, and $\mathbf{Q}_3 \in \mathcal{Q}_{\mathcal{E}}$ since $\mathbf{X} \sim p_3(\mathbf{x})$ implies $\mathbf{X} \in [0, \mathcal{A}]^M$ and $\mathbb{E}[\mathbf{X}] = \mathcal{E}$. This proves the convexity of $\mathcal{Q}_{\mathcal{E}}$, and therefore, the maximization in Lemma 1 is a convex optimization problem. The maximization in Theorem 1 can also be shown to be convex by writing it as $\max_{\mathbf{Q} \in \mathcal{Q}} \frac{1}{2} \log |\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{I}_N|$ where $\mathcal{Q} = \cup_{\mathcal{E} \in \mathcal{S}} \mathcal{Q}_{\mathcal{E}}$ and arguing similarly. Nevertheless, evaluating these bounds is difficult due to the lack of a full description of $\mathcal{Q}_{\mathcal{E}}$.

Although these bounds are difficult to evaluate, we can use the inequality $\log |\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{I}_N| \leq \text{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^T)$ (since $\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{I}_N$ is positive definite) to obtain a computable bound. Moreover, the resulting bound is tight at low SNR leading to the low-SNR capacity as given next. We start with the case of proportional average and peak constraints.

A. Proportional Average and Peak Constraints

Recall that in this case $\mathcal{E} = \alpha \mathcal{A}$, and low SNR is defined as $\mathcal{A} \rightarrow 0$ and hence also $\mathcal{E} \rightarrow 0$. To derive the low-SNR capacity for this case, we need the following lemma.

Lemma 2: The capacity of the MIMO IM-DD channel with $\mathbf{X} \in [0, \mathcal{A}]^M$ and $\sum_{i=1}^M \mathcal{E}_i \leq \mathcal{E}$ is bounded by

$$(1 + \epsilon_{\mathcal{A}}) \max_{\mathcal{E} \in \mathcal{S}} \phi(\mathcal{E}) \leq C(\mathcal{A}, \mathcal{E}) \leq \max_{\mathcal{E} \in \mathcal{S}} \phi(\mathcal{E}) \quad (8)$$

where $\lim_{\mathcal{A} \rightarrow 0} \epsilon_{\mathcal{A}} = 0$ and

$$\phi(\mathcal{E}) = \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \mathbf{h}_i^T \mathbf{h}_j \min\{\mathcal{E}_i, \mathcal{E}_j\} (\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\}). \quad (9)$$

Proof: This is proved in Sec. IV-B. ■

As a consequence of this lemma, the following theorem presents the low-SNR capacity for this case.

Theorem 2 (Proportional Constraints, $\mathcal{A} \rightarrow 0$): The low-SNR capacity of the MIMO IM-DD channel with $\mathbf{X} \in [0, \mathcal{A}]^M$ and $\sum_{i=1}^M \mathcal{E}_i \leq \mathcal{E} = \alpha \mathcal{A}$ is given by

$$C_{\text{low},1}(\mathcal{A}, \alpha \mathcal{A}) = \max_{\mathcal{E} \in \mathcal{S}} \phi(\mathcal{E}). \quad (10)$$

Proof: This follows from Lemma 2 and the definition of $C_{\text{low},1}(\mathcal{A}, \alpha \mathcal{A})$ in Sec. II. ■

The maximization in this theorem can be solved numerically.

This low-SNR capacity is achieved using coded OOK, where \mathbf{X} follows a maximally-correlated M -variate Bernoulli distribution over $\{0, \mathcal{A}\}^M$. The achievable rate is then

$$\max_{\mathcal{E} \in \mathcal{S}} I(\mathbf{X}; \mathbf{Y}) |_{\mathbf{X} \sim p^*(\mathbf{x})} \quad (11)$$

where $p^*(\mathbf{x})$ is a maximally-correlated distribution over $\{0, \mathcal{A}\}^M$. The general structure of this distribution is shown in Table I. In particular, the maximum correlation between $X_i, X_j \in \{0, \mathcal{A}\}$ given \mathcal{E}_i and \mathcal{E}_j is given by $\min\{\mathcal{E}_i, \mathcal{E}_j\} (\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\})$ as we shall see in Sec. IV-B.

If $\alpha \geq \frac{M}{2}$, then the channel is equivalent to one with only peak intensity constraints. This leads to the following corollary.

Corollary 1: If $\alpha \geq \frac{M}{2}$, then the average intensity constraint is redundant, and the low-SNR capacity is given by

$$C_{\text{low},1}(\mathcal{A}, \alpha \mathcal{A}) = C_{\text{low},1}(\mathcal{A}, M\mathcal{A}) = \frac{\mathcal{A}^2}{8} \|\mathbf{H}\mathbf{1}_M\|^2. \quad (12)$$

Proof: This follows since for $\mathcal{E} \in \mathcal{S}$, $\min\{\mathcal{E}_i, \mathcal{E}_j\} (\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\}) \leq \max\{\mathcal{E}_i, \mathcal{E}_j\} (\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\}) \leq \frac{\mathcal{A}^2}{4}$. This is achieved if $\mathcal{E}_i = \frac{\mathcal{A}}{2}$ for all $i \in \{1, \dots, M\}$, which is feasible if $\alpha \geq \frac{M}{2}$. Substituting this in Theorem 2 and noting that $\sum_{i=1}^M \sum_{j=1}^M \mathbf{h}_i^T \mathbf{h}_j = \|\mathbf{H}\mathbf{1}_M\|^2$ concludes the proof. ■

Corollary 1 highlights an interesting aspect, which can be observed by parsing the low-SNR capacity as follows. The first part of this parsing is $\frac{\mathcal{A}^2}{8}$, which is nothing but the low-SNR capacity of a peak-constrained SISO IM-DD channel [3], which is achieved with coded OOK. The second part is $\mathbf{1}_M$, which resembles a precoding of a scalar transmit signal by spatial repetition. This precoding transforms the channel into a single-input multiple-output (SIMO) one. Finally, the third part is the ℓ_2 -norm $\|\mathbf{H}\mathbf{1}_M\|$ which is the effective channel obtained after using MRC over the resulting SIMO channel. Therefore, the low-SNR capacity in Corollary 3 can be achieved using $\mathbf{X} = \mathcal{A}\mathbf{1}_M S$ where S is Bernoulli distributed with $p(s) = \frac{1}{2}$ for $s \in \{0, 1\}$, and decoding S from $a\mathbf{y}$ where $a = \frac{\|\mathbf{H}\mathbf{1}_M\|}{\|\mathbf{H}\mathbf{1}_M\|}$. The resulting achievable rate is

$$I(S; a(\mathcal{A}\mathbf{1}_M S + \mathbf{Z})). \quad (13)$$

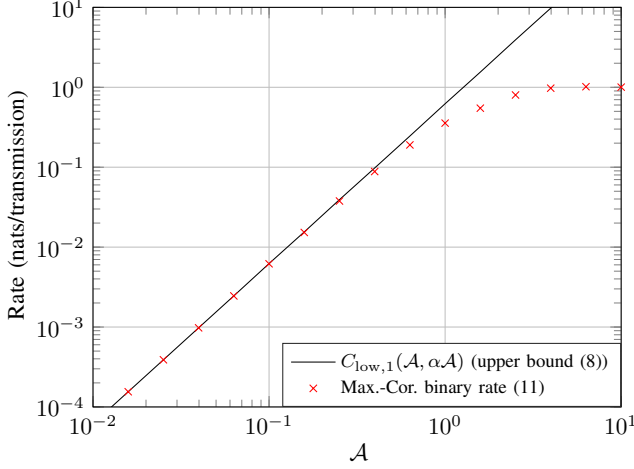
In conclusion, Corollary 1 proves the *low-SNR capacity-optimality of coded OOK, spatial repetition, and MRC* in the MIMO IM-DD channel with peak constraints only.

Fig. 2 shows the upper bound given in Lemma 2 versus \mathcal{A} for an exemplary MIMO channel with $M = N = 3$, and

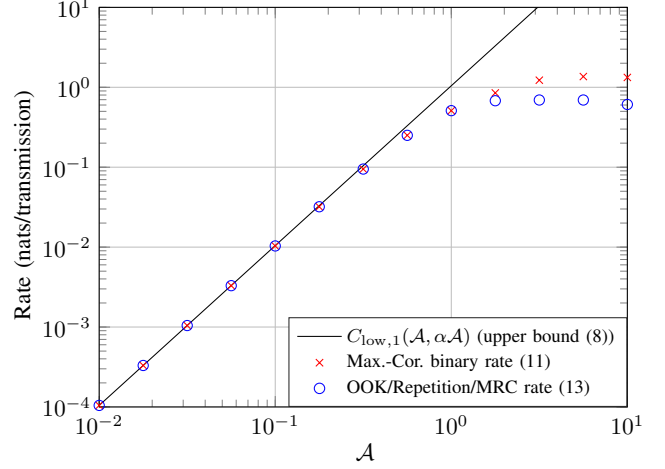
$$\mathbf{H} = \begin{bmatrix} 1 & 0.4 & 0.2 \\ 0.4 & 1 & 0.4 \\ 0.2 & 0.4 & 1 \end{bmatrix}. \quad (14)$$

TABLE I
CAPACITY ACHIEVING DISTRIBUTION AT LOW SNR FOR $\mathcal{E}_1 \geq \mathcal{E}_2 \geq \dots \geq \mathcal{E}_M$. REALIZATIONS OF \mathbf{x} WITH ZERO PROBABILITY ARE NOT SHOWN.

\mathbf{x}^\top	$(0, 0, 0, \dots, 0, 0)$	$(\mathcal{A}, 0, 0, \dots, 0, 0)$	$(\mathcal{A}, \mathcal{A}, 0, \dots, 0, 0)$	\dots	$(\mathcal{A}, \mathcal{A}, \mathcal{A}, \dots, \mathcal{A}, 0)$	$(\mathcal{A}, \mathcal{A}, \mathcal{A}, \dots, \mathcal{A}, \mathcal{A})$
$p(\mathbf{x})$	$1 - \frac{\mathcal{E}_1}{\mathcal{A}}$	$\frac{\mathcal{E}_1 - \mathcal{E}_2}{\mathcal{A}}$	$\frac{\mathcal{E}_2 - \mathcal{E}_3}{\mathcal{A}}$	\dots	$\frac{\mathcal{E}_{M-1} - \mathcal{E}_M}{\mathcal{A}}$	$\frac{\mathcal{E}_M}{\mathcal{A}}$



(a) $\alpha = 0.2M$.



(b) $\alpha = 0.5M$ (equivalently, peak constraint only).

Fig. 2. Achievable rates and upper bounds versus \mathcal{A} for a 3×3 MIMO IM-DD channel with \mathbf{H} as given in (14) and $\mathcal{E} = \alpha\mathcal{A}$.

In Fig. 2a where $\alpha = 0.2M$, this is compared with the achievable rate in (11) achieved using a maximally-correlated binary distribution according to the description in Table I. The figure clearly shows that the two coincide at low SNR verifying the statement of Theorem 2. Fig. 2b shows the same for the same channel with $\alpha = \frac{M}{2}$, which is equivalent to a channel with peak constraints only. In addition to (11), this is compared with (13), which is the achievable rate using OOK, spatial repetition, and MRC. This plot verifies Corollary 3.

Similar statements can be derived under a fixed peak constraint, although the low-SNR capacity can be simplified further as given next.

B. Fixed Peak Constraint

Here, we consider the low SNR regime where \mathcal{A} is fixed and $\mathcal{E} \rightarrow 0$. The low-SNR capacity in this case is based on the following lemma.

Lemma 3: The capacity of the MIMO IM-DD channel with $\mathbf{X} \in [0, \mathcal{A}]^M$ and $\sum_{i=1}^M \mathcal{E}_i \leq \mathcal{E}$ is bounded by

$$(1 + \epsilon_{\mathcal{E}}) \max_{\mathcal{E} \in \mathcal{S}} \phi(\mathcal{E}) \leq C(\mathcal{A}, \mathcal{E}) \leq \max_{\mathcal{E} \in \mathcal{S}} \phi(\mathcal{E}) \quad (15)$$

where $\lim_{\mathcal{E} \rightarrow 0} \epsilon_{\mathcal{E}} = 0$.

Proof: This is proved in Sec. IV-C. ■

Consequently, the following theorem presents the low-SNR capacity for this case.

Theorem 3 (Fixed Peak, $\mathcal{E} \rightarrow 0$): The low-SNR capacity of the MIMO IM-DD channel with $\mathbf{X} \in [0, \mathcal{A}]^M$ and $\sum_{i=1}^M \mathcal{E}_i \leq \mathcal{E}$ is given by

$$C_{\text{low},2}(\mathcal{A}, \mathcal{E}) = \max_{\mathcal{E} \in \mathcal{S}} \phi(\mathcal{E}). \quad (16)$$

Proof: The statement follows from Lemma 3 and the definition of $C_{\text{low},2}(\mathcal{A}, \mathcal{E})$ given in Sec. II. ■

This low-SNR capacity is also achievable using a maximally-correlated binary input distribution (see (11)). Although the low-SNR capacity expression is similar to Theorem 2 in this case, it is important to note that $C_{\text{low},2}(\mathcal{A}, \mathcal{E})$ is defined for fixed \mathcal{A} and vanishing \mathcal{E} contrary to $C_{\text{low},1}(\mathcal{A}, \alpha\mathcal{A})$ which is defined for proportionally vanishing \mathcal{A} and \mathcal{E} . Due to this, the low-SNR capacity in Theorem 3 can be simplified as follows.

Corollary 2: As $\mathcal{E} \rightarrow 0$, $C_{\text{low},2}(\mathcal{A}, \mathcal{E})$ approaches

$$\tilde{C}_{\text{low},2}(\mathcal{A}, \mathcal{E}) = \max_{\mathcal{E} \in \mathcal{S}} \tilde{\phi}(\mathcal{E}), \quad (17)$$

where

$$\tilde{\phi}(\mathcal{E}) = \frac{\mathcal{A}}{2} \sum_{i=1}^M \sum_{j=1}^M \mathbf{h}_i^\top \mathbf{h}_j \min\{\mathcal{E}_i, \mathcal{E}_j\}. \quad (18)$$

Proof: This follows since the function $\phi(\mathcal{E})$ simplifies to $\tilde{\phi}(\mathcal{E})$ as $\mathcal{E} \rightarrow 0$ since this implies $\mathcal{E}_i \rightarrow 0$ which becomes negligible with respect to \mathcal{A} . ■

The advantage of this corollary is that the involved maximization is convex. Namely, $\min\{\mathcal{E}_i, \mathcal{E}_j\}$ is concave in $(\mathcal{E}_i, \mathcal{E}_j)$ and thus $\tilde{\phi}(\mathcal{E})$ is the sum of concave functions which is concave. Thus, this maximization can be solved easily using standard solvers [24]. The resulting allocation \mathcal{E} can be used as a reliable allocation in Theorem 3.

Fig. 3 shows the upper bound given in Lemma 3 versus \mathcal{E} for an exemplary MIMO channel with $M = N = 3$, \mathbf{H} as given in (14), and $\mathcal{A} = 1$. The figure also shows the achievable rate using a maximally-correlated binary input distribution, and the low-SNR capacity expression $\tilde{C}_{\text{low},2}(\mathcal{A}, \mathcal{E})$ given in Corollary 1. All three approach each other as \mathcal{E} decreases, which demonstrates Theorem 3 and Corollary 4.

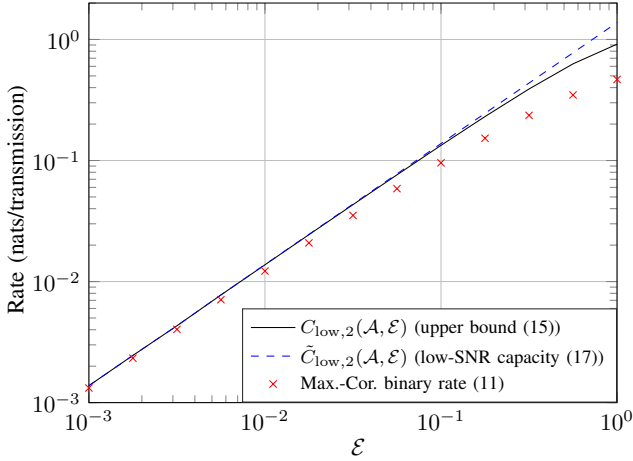


Fig. 3. Achievable rates and upper bounds versus \mathcal{E} for a 3×3 MIMO IM-DD channel with \mathbf{H} as given in (14) and fixed $\mathcal{A} = 1$.

A byproduct of these results is the low-SNR capacity of the channel under individual average intensity constraints. This is given next.

C. Individual Average Constraints

The function $\phi(\mathcal{E})$ can be interpreted as the low-SNR capacity of the channel subject to the constraints $\mathbf{X} \in [0, \mathcal{A}]^M$ and $\mathbb{E}[\mathbf{X}] = \mathcal{E}$. This can be used to obtain the following corollaries on the low-SNR capacity under individual average intensity constraints.

Corollary 3 (Proportional Constraints, $\mathcal{A} \rightarrow 0$): The low-SNR capacity $C_{\text{low},1}^{\text{ind}}(\mathcal{A}, \alpha_{\text{ind}}\mathcal{A})$ of the MIMO IM-DD channel with $\mathbf{X} \in [0, \mathcal{A}]^M$, $\mathcal{E}_i \leq \mathcal{E}_{\text{ind}} = \alpha_{\text{ind}}\mathcal{A}$ for $i \in \{1, \dots, M\}$, and $\alpha_{\text{ind}} \in (0, 1]$ is

$$C_{\text{low},1}^{\text{ind}}(\mathcal{A}, \alpha_{\text{ind}}\mathcal{A}) = \phi(\bar{\alpha}_{\text{ind}}\mathcal{A}\mathbf{1}_M) \quad (19)$$

$$= \frac{\mathcal{A}^2}{2} \bar{\alpha}_{\text{ind}}(1 - \bar{\alpha}_{\text{ind}}) \|\mathbf{H}\mathbf{1}_M\|^2, \quad (20)$$

where $\bar{\alpha}_{\text{ind}} = \min\{\alpha_{\text{ind}}, \frac{1}{2}\}$.

Proof: See Sec. IV-D1. ■

Here, if $\alpha_{\text{ind}} \geq \frac{1}{2}$, then the average constraint is redundant, and the low-SNR capacity coincides with that given in Corollary 1.

For a fixed peak constraint, we have the following.

Corollary 4 (Fixed Peak, $\mathcal{E} \rightarrow 0$): The low-SNR capacity $C_{\text{low},2}^{\text{ind}}(\mathcal{A}, \mathcal{E}_{\text{ind}})$ of the MIMO IM-DD channel with $\mathbf{X} \in [0, \mathcal{A}]^M$, $\mathcal{E}_i \leq \mathcal{E}_{\text{ind}}$ is

$$C_{\text{low},2}^{\text{ind}}(\mathcal{A}, \mathcal{E}_{\text{ind}}) = \mathcal{E}_{\text{ind}}(\mathcal{A} - \mathcal{E}_{\text{ind}}) \|\mathbf{H}\mathbf{1}_M\|^2. \quad (21)$$

Proof: See Sec. IV-D2. ■

Both these corollaries show that *coded OOK, spatial repetition, and MRC are optimal at low SNR under individual average constraints*. In particular, the low-SNR capacity is achievable by choosing $\mathbf{X} = \mathcal{A}\mathbf{1}_M\bar{S}$ where \bar{S} is Bernoulli distributed with $p(\bar{s}) = \frac{\mathcal{E}_{\text{ind}}}{\mathcal{A}}$ if $\bar{s} = 1$ and $p(\bar{s}) = 1 - \frac{\mathcal{E}_{\text{ind}}}{\mathcal{A}}$ if $\bar{s} = 0$. Then, the receiver decodes \bar{S} from $a\mathbf{y}$ where $a = \frac{\mathbf{H}\mathbf{1}_M}{\|\mathbf{H}\mathbf{1}_M\|}$, achieving

$$I(\bar{S}; a(\mathcal{A}\mathbf{1}_M\bar{S} + \mathbf{Z})). \quad (22)$$

The factors $\frac{\mathcal{A}^2}{2}\bar{\alpha}_{\text{ind}}(1 - \bar{\alpha}_{\text{ind}})$ in Corollary 3 and $\mathcal{E}_{\text{ind}}(\mathcal{A} - \mathcal{E}_{\text{ind}})$ in Corollary 4 represent the low-SNR capacity of the resulting SISO channel, and $\|\mathbf{H}\mathbf{1}_M\|^2$ is a scaling that captures the effect of spatial repetition ($\mathbf{1}_M$) and MRC (the ℓ_2 -norm).

Fig. 4 shows the low-SNR capacities in Corollaries 3 and 4 for an exemplary MIMO channel with $M = N = 3$ and \mathbf{H} as given in (14). In Fig. 4a, we fix $\alpha_{\text{ind}} = 0.2$ and we compare the low-SNR capacity with the achievable rate using coded OOK, spatial repetition, and MRC. In the same figure, we plot the achievable rate using a maximally-correlated binary distribution given by

$$\max_{\mathcal{E} \in (0, \min\{\mathcal{E}_{\text{ind}}, \frac{\mathcal{A}}{2}\})^M} I(\mathbf{X}; \mathbf{Y})|_{\mathbf{X} \sim p^*(\mathbf{x})} \quad (23)$$

where $p^*(\mathbf{x})$ has the general structure shown in Table I. Note that the three converge as \mathcal{A} decreases. Fig. 4b show the same versus \mathcal{E}_{ind} for a channel with fixed $\mathcal{A} = 1$. This verifies Corollaries 3 and 4.

Next, we give the proofs of these results.

IV. CAPACITY ANALYSIS

A. Upper Bound for a Given Intensity Allocation

The capacity of a channel with individual average constraints $\mathbb{E}[\mathbf{X}] = \mathcal{E}$ can be written as

$$c(\mathcal{A}, \mathcal{E}) = \max_{p(\mathbf{x}) \in \mathcal{P}_{\mathcal{E}}} I(\mathbf{X}; \mathbf{Y}) \quad (24)$$

$$= \max_{p(\mathbf{x}) \in \mathcal{P}_{\mathcal{E}}} h(\mathbf{H}\mathbf{X} + \mathbf{Z}) - h(\mathbf{H}\mathbf{X} + \mathbf{Z}|\mathbf{X}) \quad (25)$$

$$= \max_{p(\mathbf{x}) \in \mathcal{P}_{\mathcal{E}}} h(\mathbf{H}\mathbf{X} + \mathbf{Z}) - h(\mathbf{Z}), \quad (26)$$

where $\mathcal{P}_{\mathcal{E}}$ is the collection of all distributions of $\mathbf{X} \in [0, \mathcal{A}]^M$ satisfying $\mathbb{E}[\mathbf{X}] = \mathcal{E}$. Let us write

$$\mathcal{P}_{\mathcal{E}} = \cup_{\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}} \mathcal{P}_{\mathcal{E}, \mathbf{Q}}, \quad (27)$$

where $\mathcal{Q}_{\mathcal{E}}$ is the set of covariance matrices corresponding to distributions $p(\mathbf{x}) \in \mathcal{P}_{\mathcal{E}}$, and $\mathcal{P}_{\mathcal{E}, \mathbf{Q}}$ is the subset of $\mathcal{P}_{\mathcal{E}}$ with covariance matrix \mathbf{Q} . Then, we can write

$$c(\mathcal{A}, \mathcal{E}) = \max_{\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}} \max_{p(\mathbf{x}) \in \mathcal{P}_{\mathcal{E}, \mathbf{Q}}} h(\mathbf{H}\mathbf{X} + \mathbf{Z}) - h(\mathbf{Z}) \quad (28)$$

$$\leq \max_{\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}} \frac{1}{2} \log |\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{I}_N|, \quad (29)$$

where the last step follows since the Gaussian distribution maximizes the differential entropy under a covariance constraint [25]. This proves the statement of Lemma 1.

B. Low-SNR Capacity Under Proportional Constraints

Here, we prove Lemma 2. We start by proving the upper bound, followed by the achievability.

1) *Upper Bound:* By Theorem 1, we have

$$C(\mathcal{A}, \mathcal{E}) \leq \max_{\mathcal{E} \in \mathcal{S}} \max_{\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}} \frac{1}{2} \log |\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{I}_N|. \quad (30)$$

Since $\log |\mathcal{A}| \leq \text{Tr}(\mathbf{A} - \mathbf{I}_N)$ for a positive definite matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$, and since $\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{I}_N$ is positive definite, then we can write

$$C(\mathcal{A}, \mathcal{E}) \leq \max_{\mathcal{E} \in \mathcal{S}} \max_{\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}} \frac{1}{2} \text{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^T). \quad (31)$$

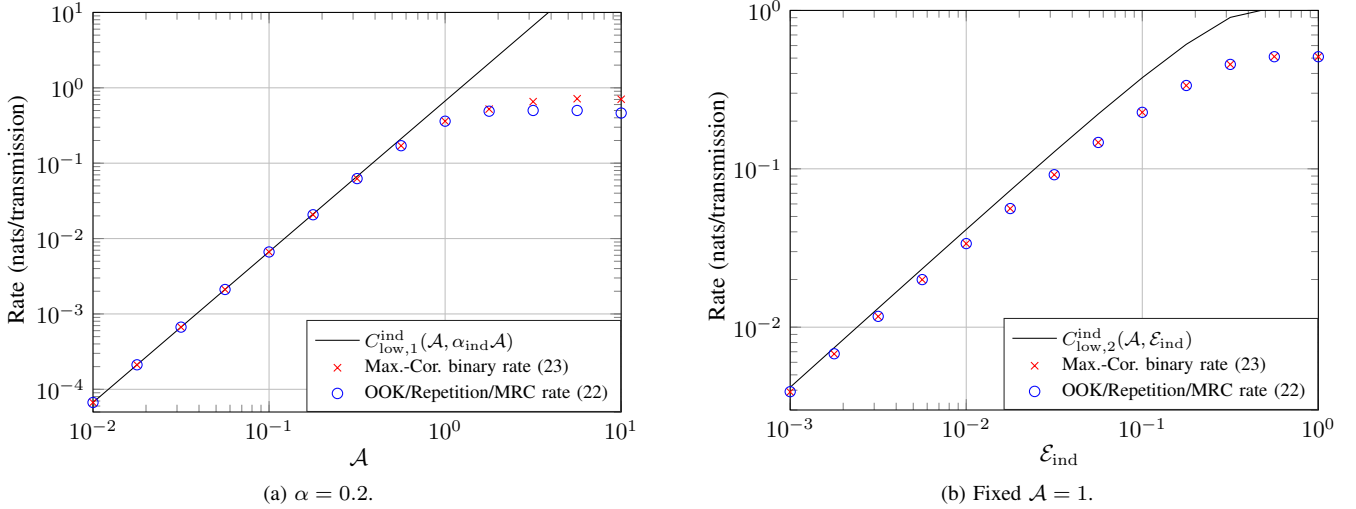


Fig. 4. Achievable rates and upper bounds versus \mathcal{A} for a 3×3 MIMO IM-DD channel with \mathbf{H} as given in (14) and $\mathcal{E} = \alpha\mathcal{A}$.

Now the proof proceeds along the following steps. First, we show that the matrix \mathbf{Q} that maximizes the above expression for a given \mathcal{E} corresponds to $\mathbf{X} \in \{0, \mathcal{A}\}^M$ with possible dependence between the components of \mathbf{X} . Then, we derive the maximum covariance between binary random variables with predefined mean. This allows us to upper bound capacity using a simple expression.

Let us consider any $p(\mathbf{x})$ with a covariance matrix $\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}$, and let us define $\mathbf{G} = \mathbf{H}^T \mathbf{H}$. Then,

$$\text{Tr}(\mathbf{H}\mathbf{Q}\mathbf{H}^T) = \text{Tr}(\mathbf{G}\mathbf{Q}) \quad (32)$$

$$= \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{j,i} \quad (33)$$

$$= \mathbb{E}_{p(\mathbf{x})}[\psi(\mathbf{x})], \quad (34)$$

where $q_{j,i}$ is the component of \mathbf{Q} in the j^{th} row and i^{th} column, $\mathbb{E}_{p(\mathbf{x})}$ is the expectation with respect to the distribution $p(\mathbf{x})$, and

$$\psi(\mathbf{x}) = \sum_{i=1}^M \sum_{j=1}^M g_{i,j} (x_i - \mathcal{E}_i)(x_j - \mathcal{E}_j). \quad (35)$$

Here, we used the definition of covariance. Due to the non-negativity of $h_{i,j}$ and hence also $g_{i,j}$, the function $\psi(\mathbf{x})$ is a convex elliptic paraboloid with minimum at $x_i = \mathcal{E}_i$ for all i . Thus, we have

$$\psi(x_1, x_2, \dots, x_M) \leq \max_{a \in \{0, \mathcal{A}\}} \psi(a, x_2, \dots, x_M). \quad (36)$$

Therefore, if we write

$$p(\mathbf{x}) = p_2(x_2, \dots, x_M) p_1(x_1 | x_2, \dots, x_M),$$

then we have

$$\mathbb{E}_{p(\mathbf{x})}[\psi(\mathbf{x})] \leq \mathbb{E}_{p^{[1]}(\mathbf{x})}[\psi(\mathbf{x})], \quad (37)$$

where $p^{[1]}(\mathbf{x}) = p_2(x_2, \dots, x_M) \tilde{p}_1(x_1 | x_2, \dots, x_M)$ and

$$\begin{aligned} & \tilde{p}_1(x_1 | x_2, \dots, x_M) \\ &= \begin{cases} 1, & x_1 = \arg \max_{a \in \{0, \mathcal{A}\}} \psi(a, x_2, \dots, x_M) \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (38)$$

This can be repeated for all x_2, \dots, x_M to conclude that

$$\mathbb{E}_{p(\mathbf{x})}[\psi(\mathbf{x})] \leq \mathbb{E}_{p^{[M]}(\mathbf{x})}[\psi(\mathbf{x})], \quad (39)$$

where $p^{[M]}(\mathbf{x})$ is a distribution of \mathbf{X} on $\{0, \mathcal{A}\}^M$.

Consequently, so far we have

$$C(\mathcal{A}, \mathcal{E}) \leq \max_{\mathcal{E} \in \mathcal{S}} \max_{\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}^b} \frac{1}{2} \text{Tr}(\mathbf{G}\mathbf{Q}), \quad (40)$$

where $\mathcal{Q}_{\mathcal{E}}^b$ is the set of covariance matrices of binary $\mathbf{X} \in \{0, \mathcal{A}\}^M$ with $\mathbb{E}[\mathbf{X}] = \mathcal{E}$.

Since $\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}^b$, then $q_{i,i} = \mathcal{E}_i(\mathcal{A} - \mathcal{E}_i)$. Next, we need the following lemma which bounds the covariance of (X_i, X_j) .

Lemma 4 (Max. Binary Covariance): The covariance of $(X_1, X_2) \in \{0, \mathcal{A}\}^2$ satisfying $\mathbb{E}[(X_1, X_2)] = (\mathcal{E}_1, \mathcal{E}_2)$ and $\mathcal{E}_1 \geq \mathcal{E}_2$ is upper bounded by

$$\text{Cov}(X_1, X_2) \leq \mathcal{E}_2(\mathcal{A} - \mathcal{E}_1),$$

with equality if $X_i = U_i \mathcal{A}$ where $U_2 \sim \text{Bern}(\frac{\mathcal{E}_2}{\mathcal{A}})$ and $U_1 = U_2 \vee V_1$ with $V_1 \sim \text{Bern}(\frac{\mathcal{E}_1 - \mathcal{E}_2}{\mathcal{A} - \mathcal{E}_2})$.

Proof: Denote the distribution of (X_1, X_2) by $p(x_1, x_2)$. We have that

$$q_{1,2} = \sum_{(x_1, x_2) \in \{0, \mathcal{A}\}^2} p(x_1, x_2) (x_1 - \mathcal{E}_1)(x_2 - \mathcal{E}_2). \quad (41)$$

TABLE II
SOLUTION OF PROBLEM (42).

	X_j	0	\mathcal{A}
X_i		$1 - \frac{\max\{\mathcal{E}_i, \mathcal{E}_j\}}{\mathcal{A}}$	$\frac{\max\{0, \mathcal{E}_j - \mathcal{E}_i\}}{\mathcal{A}}$
0		$\frac{\max\{0, \mathcal{E}_i - \mathcal{E}_j\}}{\mathcal{A}}$	$\frac{\min\{\mathcal{E}_i, \mathcal{E}_j\}}{\mathcal{A}}$
\mathcal{A}			

The maximum covariance can be obtained by solving

$$\max_{p(x_1, x_2)} q_{1,2} \quad (42a)$$

$$\text{s.t.} \quad \sum_{(x_1, x_2) \in \{0, \mathcal{A}\}^2} p(x_1, x_2) = 1, \quad (42b)$$

$$p(x_1, x_2) \geq 0, \quad \forall (x_1, x_2) \in \{0, \mathcal{A}\}^2, \quad (42c)$$

$$p(\mathcal{A}, 0) + p(\mathcal{A}, \mathcal{A}) = \frac{\mathcal{E}_1}{\mathcal{A}}, \quad (42d)$$

The last two constraints are implied by $\mathbb{E}[X_i] = \mathcal{E}_i$, $i \in \{1, 2\}$. This is a simple linear program, whose solution can be shown to be given by $p(0, 0) = 1 - \frac{\mathcal{E}_1}{\mathcal{A}}$, $p(0, \mathcal{A}) = 0$, $p(\mathcal{A}, 0) = \frac{\mathcal{E}_1 - \mathcal{E}_2}{\mathcal{A}}$, and $p(\mathcal{A}, \mathcal{A}) = \frac{\mathcal{E}_2}{\mathcal{A}}$. This leads to the desired covariance bound and its achievability. ■

Thus, in general, we can write $\text{Cov}(X_i, X_j) \leq \min\{\mathcal{E}_i, \mathcal{E}_j\}(\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\})$, which is achievable using the structure in Table II. At this point, we know that $\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}^b$ satisfies

$$q_{i,i} = \mathcal{E}_i(\mathcal{A} - \mathcal{E}_i) \quad (43a)$$

$$q_{i,j} \leq \min\{\mathcal{E}_i, \mathcal{E}_j\}(\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\}), \quad (43b)$$

and we can write

$$C(\mathcal{A}, \mathcal{E}) \leq \max_{\mathcal{E} \in \mathcal{S}} \max_{\mathbf{Q} \in \mathcal{Q}_{\mathcal{E}}^b} \frac{1}{2} \text{Tr}(\mathbf{G}\mathbf{Q}), \quad (44)$$

where $\hat{\mathcal{Q}}_{\mathcal{E}}^b$ is the set of matrices which satisfy (43). The inner maximization is a linear program (cf. (33)) whose solution is one which satisfies (43) with equality since $g_{i,j} \geq 0 \quad \forall i, j$. This leads to

$$C(\mathcal{A}, \mathcal{E}) \leq \max_{\mathcal{E} \in \mathcal{S}} \sum_{i=1}^M \sum_{j=1}^M \frac{g_{i,j}}{2} \min\{\mathcal{E}_i, \mathcal{E}_j\}(\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\}). \quad (45)$$

Since $g_{i,j} = \sum_{k=1}^N h_{k,i} h_{k,j} = \mathbf{h}_i^T \mathbf{h}_j$, this proves the upper bound in Lemma 2.

2) *Lower Bound:* The lower bound follows by using a binary input distribution which has the structure in Table II, and using the asymptotic result in [26] as described next.

We first need to show that the structure in Table II is feasible for $\mathbf{X} \in \{0, \mathcal{A}\}^M$. To show this, let us choose $\mathcal{E} \in \mathcal{S}$ with $\mathcal{A} \geq \mathcal{E}_1 \geq \mathcal{E}_2 \geq \dots \geq \mathcal{E}_M$ for simplicity. Other orderings of $\mathcal{E}_1, \dots, \mathcal{E}_M$ can be treated similarly. Then, we generate

$$V_i \sim \text{Bern}\left(\frac{\mathcal{E}_i - \mathcal{E}_{i+1}}{\mathcal{A} - \mathcal{E}_{i+1}}\right), \quad i = 1, \dots, M, \quad (46)$$

where we formally define $\mathcal{E}_{M+1} = 0$, and we set

$$U_M = V_M, \quad (47)$$

$$U_i = U_{i+1} \vee V_i, \quad i = 1, \dots, M-1. \quad (48)$$

Finally, we set $\mathbf{X} = \mathbf{U}\mathcal{A}$. It can be easily verified that this $\mathbf{X} \in \{0, \mathcal{A}\}^M$ satisfies $\mathbb{E}[\mathbf{X}] = \mathcal{E}$ and has the structure in Table II. This leads to the distribution given in Table I which has a covariance matrix \mathbf{Q}^* with

$$q_{i,j}^* = \min\{\mathcal{E}_i, \mathcal{E}_j\}(\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\}). \quad (49)$$

Now, we show that this distribution achieves a rate which coincides with the upper bound (45) at low SNR. The achievable rate is given by $I(\mathbf{X}; \mathbf{Y})$. Since the channel under consideration satisfies conditions A–F in [26], then [26, Theorem 1] applies and we have

$$I(\mathbf{X}; \mathbf{Y}) = \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \iota_{i,j} q_{i,j}^* + o(\text{Tr}(\mathbf{Q}^*)), \quad (50)$$

where $\iota_{i,j}$ is the Fisher-information-type integral

$$\iota_{i,j} = \int_{\mathbb{R}^N} \frac{1}{f(\mathbf{y}|\mathbf{x})} \frac{\partial f(\mathbf{y}|\mathbf{x})}{\partial x_i} \frac{\partial f(\mathbf{y}|\mathbf{x})}{\partial x_j} d\mathbf{y}. \quad (51)$$

This $\iota_{i,j}$ can be shown to be independent of \mathbf{x} for the channel under consideration, and is given by

$$\iota_{i,j} = \sum_{k=1}^N h_{k,i} h_{k,j} = g_{i,j}. \quad (52)$$

Therefore,

$$I(\mathbf{X}; \mathbf{Y}) = \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^* + o(\text{Tr}(\mathbf{Q}^*)) \quad (53)$$

$$= \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^* (1 + \epsilon_{\mathcal{A}}), \quad (54)$$

where $\epsilon_{\mathcal{A}} = \frac{o(\text{Tr}(\mathbf{Q}^*))}{\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^*}$. To show that this has the same form as the upper bound (45) at low SNR, it remains to show that $\lim_{\mathcal{A} \rightarrow 0} \epsilon_{\mathcal{A}} = 0$.

Now,

$$\epsilon_{\mathcal{A}} = \frac{o(\text{Tr}(\mathbf{Q}^*))}{\text{Tr}(\mathbf{Q}^*)} \frac{\text{Tr}(\mathbf{Q}^*)}{\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^*}. \quad (55)$$

Note that $\mathcal{A} \rightarrow 0$ implies that $\text{Tr}(\mathbf{Q}^*) \rightarrow 0$ and thus $\lim_{\mathcal{A} \rightarrow 0} \frac{o(\text{Tr}(\mathbf{Q}^*))}{\text{Tr}(\mathbf{Q}^*)} = 0$. Moreover,

$$\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^* \geq \frac{1}{2} \sum_{i=1}^M g_{i,i} q_{i,i}^* \quad (56)$$

$$\geq \frac{1}{2} \min_i g_{i,i} \text{Tr}(\mathbf{Q}^*). \quad (57)$$

Thus,

$$0 \leq \frac{\text{Tr}(\mathbf{Q}^*)}{\frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^*} \leq \frac{1}{\min_i g_{i,i}}, \quad (58)$$

which follows since $q_{i,j}^*$ and $g_{i,j}$ are nonnegative. Multiplying both sides by $\frac{o(\text{Tr}(\mathbf{Q}^*))}{\text{Tr}(\mathbf{Q}^*)}$ and taking the limit as $\mathcal{A} \rightarrow 0$ yields

$$0 \leq \lim_{\mathcal{A} \rightarrow 0} \epsilon_{\mathcal{A}} \leq 0, \quad (59)$$

as long as $\min_i g_{i,i} \neq 0$ which is true almost surely.

As a result, we have

$$I(\mathbf{X}; \mathbf{Y}) = \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^* (1 + \epsilon_{\mathcal{A}}), \quad (60)$$

where $\lim_{\mathcal{A} \rightarrow 0} \epsilon_{\mathcal{A}} = 0$. Maximizing with respect to $\mathcal{E} \in \mathcal{S}$ leads to the achievability of

$$R(\mathcal{A}, \mathcal{E}) = \max_{\mathcal{E} \in \mathcal{S}} \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^* (1 + \epsilon_{\mathcal{A}}). \quad (61)$$

Since $g_{i,j} = \mathbf{h}_i^T \mathbf{h}_j$, this shows the lower bound in Lemma 2 and completes its proof.

C. Low-SNR Capacity Under a Fixed Peak Constraint

The capacity upper bound given in (45) holds also in this case. Moreover, for a given intensity allocation \mathcal{E} , the expression (53) given by

$$I(\mathbf{X}; \mathbf{Y}) = \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^* + o(\text{Tr}(\mathbf{Q}^*)) \quad (62)$$

is an achievable rate in this case. To show that this achievable rate coincides with the low-SNR capacity, we need to show that it can be written as

$$I(\mathbf{X}; \mathbf{Y}) = \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^* (1 + \epsilon_{\mathcal{E}}), \quad (63)$$

where $\lim_{\mathcal{E} \rightarrow 0} \epsilon_{\mathcal{E}} = 0$. This can be shown following similar steps as (55)-(59), and noting that $\mathcal{E} \rightarrow 0$ implies that $\text{Tr}(\mathbf{Q}^*) \rightarrow 0$ and thus $\lim_{\mathcal{E} \rightarrow 0} \frac{o(\text{Tr}(\mathbf{Q}^*))}{\text{Tr}(\mathbf{Q}^*)} = 0$. Therefore, we have the achievable rate

$$R(\mathcal{A}, \mathcal{E}) = \max_{\mathcal{E} \in \mathcal{S}} \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^* (1 + \epsilon_{\mathcal{E}}), \quad (64)$$

leading to the low-SNR capacity $C_{\text{low},2} = \max_{\mathcal{E} \in \mathcal{S}} \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M g_{i,j} q_{i,j}^*$. This completes the proof of Theorem 3.

D. Low-SNR Capacity Under Individual Average Constraints

We consider proportional constraints first, where $\mathcal{E}_{\text{ind}} = \alpha_{\text{ind}} \mathcal{A}$ and $\alpha_{\text{ind}} \in (0, 1]$.

1) *Proportional Constraints*: Low SNR is defined in this case as $\mathcal{A} \rightarrow 0$. We start by writing

$$C_{\text{low},1}^{\text{ind}}(\mathcal{A}, \alpha_{\text{ind}} \mathcal{A}) = \max_{\mathcal{E} \in [0, \mathcal{E}_{\text{ind}}]^M} \phi(\mathcal{E}), \quad (65)$$

using similar analysis as in Sec. IV-B. Next, we show that the solution of this maximization is $\mathcal{E} = \min\{\alpha_{\text{ind}}, \frac{1}{2}\} \mathcal{A} \mathbf{1}_M$.

We start by upper bounding $C_{\text{low},1}^{\text{ind}}(\mathcal{A}, \alpha_{\text{ind}} \mathcal{A})$ as follows,

$$C_{\text{low},1}^{\text{ind}}(\mathcal{A}, \alpha_{\text{ind}} \mathcal{A}) \quad (66)$$

$$\begin{aligned} &= \max_{\mathcal{E} \in [0, \mathcal{E}_{\text{ind}}]^M} \sum_{i=1}^M \sum_{j=1}^M \frac{g_{i,j}}{2} \min\{\mathcal{E}_i, \mathcal{E}_j\} (\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\}) \\ &\leq \sum_{i=1}^M \sum_{j=1}^M \frac{g_{i,j}}{2} \max_{\mathcal{E}_i, \mathcal{E}_j \in [0, \mathcal{E}_{\text{ind}}]} \min\{\mathcal{E}_i, \mathcal{E}_j\} (\mathcal{A} - \max\{\mathcal{E}_i, \mathcal{E}_j\}) \\ &\leq \sum_{i=1}^M \sum_{j=1}^M \frac{g_{i,j}}{2} \max_{\mathcal{E}_{\min} \in [0, \mathcal{E}_{\text{ind}}]} \mathcal{E}_{\min} (\mathcal{A} - \mathcal{E}_{\min}) \\ &= \sum_{i=1}^M \sum_{j=1}^M \frac{g_{i,j}}{2} \mathcal{A}^2 \bar{\alpha}_{\text{ind}} (1 - \bar{\alpha}_{\text{ind}}), \end{aligned} \quad (67)$$

where $\bar{\alpha}_{\text{ind}} = \min\{\alpha_{\text{ind}}, \frac{1}{2}\}$. Since $\sum_{i=1}^M \sum_{j=1}^M g_{i,j} = \|\mathbf{H} \mathbf{1}_M\|^2$, then

$$C_{\text{low},1}^{\text{ind}}(\mathcal{A}, \alpha_{\text{ind}} \mathcal{A}) \leq \frac{\mathcal{A}^2}{2} \bar{\alpha}_{\text{ind}} (1 - \bar{\alpha}_{\text{ind}}) \|\mathbf{H} \mathbf{1}_M\|^2. \quad (69)$$

But this upper bound is achievable by setting $\mathcal{E} = \min\{\alpha_{\text{ind}}, \frac{1}{2}\} \mathcal{A} \mathbf{1}_M$. This proves the statement of Corollary 3.

2) *Fixed Peak Constraint*: Low SNR is defined in this case as $\mathcal{E}_{\text{ind}} \rightarrow 0$ with \mathcal{A} held fixed. Using similar steps as in Sec. IV-D1, we have that

$$\begin{aligned} &\max_{\mathcal{E} \in [0, \mathcal{E}_{\text{ind}}]^M} \phi(\mathcal{E}) \\ &\leq \sum_{i=1}^M \sum_{j=1}^M \frac{g_{i,j}}{2} \min\left\{\mathcal{E}_{\text{ind}}, \frac{\mathcal{A}}{2}\right\} \left(\mathcal{A} - \min\left\{\mathcal{E}_{\text{ind}}, \frac{\mathcal{A}}{2}\right\}\right), \end{aligned}$$

which is achievable by setting $\mathcal{E} = \min\{\mathcal{E}_{\text{ind}}, \frac{\mathcal{A}}{2}\} \mathbf{1}_M$.

Note that as $\mathcal{E}_{\text{ind}} \rightarrow 0$, \mathcal{E}_{ind} eventually becomes smaller than $\frac{\mathcal{A}}{2}$, and thus

$$C_{\text{low},2}^{\text{ind}}(\mathcal{A}, \mathcal{E}_{\text{ind}}) = \sum_{i=1}^M \sum_{j=1}^M \frac{g_{i,j}}{2} \mathcal{E}_{\text{ind}} (\mathcal{A} - \mathcal{E}_{\text{ind}}), \quad (70)$$

This proves the statement of Corollary 4.

V. CONCLUSION

We studied the capacity of the optical intensity MIMO channel under average and peak intensity constraint, which is an important quantity in the context of optical wireless communications. We derived a capacity upper bound and proved its achievability at low SNR. In this process, we have shown that the optimal input distribution is a maximally-correlated vector-binary distribution. We have also shown that if the transmitter is constrained by individual (per aperture) constraints or peak constraints only, then coded OOK and spatial repetition at the transmitter, and MRC at the receiver achieve the low-SNR capacity of the channel. This leads to a simple practical scheme under this scenario. The results of this paper can be useful in optical wireless communication systems with strong noise due to ambient light e.g., or with low average intensity due to light-dimming for instance.

REFERENCES

- [1] M. A. Khalighi and M. Uysal, "Survey on free space optical communications: A communication theory perspective," *Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2231–2258, 4th quarter 2014.
- [2] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel Parts I and II," *IEEE Trans. on Info. Theory*, vol. 34, no. 6, pp. 1462–1471, Nov. 1988.
- [3] A. Lapidot, S. M. Moser, and M. Wigger, "On the capacity of free-space optical intensity channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.
- [4] S. M. Moser, "Capacity results of an optical intensity channel with input-dependent Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 207–223, Jan. 2012.
- [5] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proc. of the IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.
- [6] H. Burchardt, N. Serafimovski, D. Tsonev, S. Videv, and H. Haas, "VLC: Beyond point-to-point communication," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 98–105, July 2014.
- [7] P. Pathak, X. Fen, P. Hu, and P. Mohapatra, "Visible light communication, networking, and sensing: A survey, potential and challenges," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2047–2077, 4th quarter 2015.
- [8] A. A. Farid and S. Hranilovic, "Capacity bounds for wireless optical intensity channels with Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6066–6077, Dec. 2010.
- [9] A. Chaaban, J.-M. Morvan, and M.-S. Alouini, "Free-space optical communications: Capacity bounds, approximations, and a new sphere-packing perspective," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1176–1191, Mar. 2016.
- [10] A. Chaaban, Z. Rezeki, and M.-S. Alouini, "On the capacity of the intensity-modulation direct-detection optical broadcast channel," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3114–3130, May 2016.
- [11] A. Chaaban, O. M. S. Al-Ebraheemy, T. Y. Al-Naffouri, and M. S. Alouini, "Capacity bounds for the Gaussian IM-DD optical multiple-access channel," vol. 16, no. 5, pp. 3328–3340, May 2017.
- [12] T. Fath and H. Haas, "Performance comparison of MIMO techniques for optical wireless communications in indoor environments," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 733–742, Feb. 2013.
- [13] E. Monteiro and S. Hranilovic, "Design and implementation of color-shift keying for visible light communications," *J. Lightw. Technol.*, vol. 32, no. 10, pp. 2053–2060, May 2014.
- [14] A. Chaaban, Z. Rezeki, and M.-S. Alouini, "Fundamental limits of parallel optical wireless channels: Capacity results and outage formulation," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 296–311, Jan. 2017.
- [15] M. Safari, *MIMO Free-Space Optical Communication*. Springer International Publishing, 2016.
- [16] K. Xu, H. Y. Yu, Y. J. Zhu, and H. B. Cai, "Channel-adaptive space-collaborative constellation design for MIMO VLC with fast maximum likelihood detection," *IEEE Access*, vol. 5, pp. 842–852, Mar. 2017.
- [17] L. Zeng, D. O'Brien, H. Minh, G. Faulkner, K. Lee, D. Jung, Y. Oh, and E. T. Won, "High data rate multiple input multiple output (MIMO) optical wireless communications using white LED lighting," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 9, pp. 1654–1662, Dec. 2009.
- [18] A. Chaaban, Z. Rezeki, and M.-S. Alouini, "MIMO intensity-modulation channels: Capacity bounds and high SNR characterization," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017.
- [19] —, "Capacity Bounds and High-SNR Capacity of MIMO Intensity-Modulation Optical Channels," *KAUST Technical Report* <http://hdl.handle.net/10754/624895>, June 2017.
- [20] S. M. Moser, M. Mylonakis, L. Wang, and M. Wigger, "Asymptotic capacity results for MIMO wireless optical communication," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017.
- [21] V. V. Prelov and E. C. van der Meulen, "An asymptotic expression for the information and capacity of a multidimensional channel with weak input signals," *IEEE Trans. on Info. Theory*, vol. 39, no. 5, pp. 1728–1735, Sep. 1993.
- [22] S. M. Moser, L. Wang, and M. Wigger, "Asymptotic high-SNR capacity of MISO optical intensity channels," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 6–10.
- [23] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.
- [24] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [25] T. Cover and J. Thomas, *Elements of Information Theory (Second Edition)*. John Wiley and Sons, Inc., 2006.
- [26] V. V. Prelov and E. C. van der Meulen, "An asymptotic expression for the information and capacity of a multidimensional channel with weak input signals," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1728–1735, Sep. 1993.