Average Worst-Case PEP Optimality of Repetition Coding among Rate-1 DC-Offset STBCs for MIMO Optical Intensity Channels

Anas Chaaban, *Senior Member, IEEE*, Yerzhan Sapenov, Zouheir Rezki, *Senior Member, IEEE*, and Mohamed-Slim Alouini, *Fellow, IEEE*

Abstract—An optical wireless intensity-modulation direct-detection multiple-input multiple-output communication system is considered. The performance of M -PAM rate-1 direct current offset space-time block codes is studied in terms of average worst-case pairwise error probability (WC-PEP) in quasi-static channels. It is shown that within this code class, the average WC-PEP is minimized by repetition coding (RC) under both electrical and optical individual power constraints, irrespective of channel statistics. This agrees with previously published results related to on-off keying RC. This is further extended to sum power constraints, where it is shown that spatial beamforming minimizes the average WC-PEP within this code class, which simplifies to RC if the channel matrix has independent and identically distributed columns and a sum electrical power constraint. Under a sum optical power constraint, this also holds true at high signal-to-noise ratio (SNR), but not at low SNR. Generally, the time dimension of this code class is redundant from an average WC-PEP perspective. Numerical results are provided to support the theoretical findings and to show that the average WC-PEP leads to a good approximation of the actual error probability at high SNR.

Index Terms—Intensity-modulation direct-detection, optical wireless communications, repetition coding, space-time block codes, maximum likelihood detection.

I. INTRODUCTION

The continuous increase in the demand for higher data rates stresses current radio frequency (RF) wireless communication systems and requires new solutions for coping with this increase. One such solution is optical wireless communications (OWC) in which light is used to convey information between communicating nodes. OWC enjoys many attractive features including free broadband spectrum, enhanced security, and reduced electromagnetic interference, and therefore, it has witnessed increasing research attention from academic and

industrial worlds (see [2], [3] and references therein). Furthermore, OWC operating in the visible-light spectrum, known as visible-light communications (VLC) [4], [5], allows simultaneous communication and illumination leading to increased energy efficiency. These features make OWC a promising solution for future wireless communication systems, especially in its simple and practical intensity-modulated direct-detection (IM/DD) form.

Recent studies on IM/DD OWC focus on various performance criteria including achievable rates [6]–[8], outage probability [9], and error rates [10]–[13], for instance. In addition to these criteria, diversity order and pair-wise error probability (PEP) are important wireless communication performance criteria [14], which can be optimized by exploiting channel variation in space, time, frequency, or combinations thereof. Exploiting space and time diversity by proper design of space-time block codes (STBC) [15] is known to provide excellent performance in terms of PEP and diversity order, and that orthogonal STBC (OSTBC) are optimal in RF communications in terms of diversity gain [16]. With this in mind, the following question arises: Does this also hold true in IM/DD OWC?

The answer to this question can be obtained by analyzing the error probability of STBCs in IM/DD OWC. This question has been discussed in [11], [17]. In [11], an on-off keying (OOK) multiple input single output (MISO) IM/DD OWC system with individual (per-aperture) optical power constraints was studied, and it was demonstrated that simple spatial repetition coding (RC)¹ interestingly outperforms OSTBC. This has been demonstrated by a numerical simulation under independent and identically distributed (i.i.d.) log-normally distributed channels. This has also been demonstrated by an analytic comparison of the error probability which shows that OOK-RC is superior to OOK-OSTBC for the given system under any channel statistics. In [17], an OOK system with 2 transmit and N receive apertures, individual optical power constraints, and i.i.d. Gamma-Gamma distributed channels was studied, and it was shown that RC is quasi-optimal at high signal-to-noise ratio (SNR) in terms of average PEP among all OOK-STBCs and OOK trellis codes.

By restricting attention to space codes, the authors of [12], [13] prove the diversity-optimality of RC among all *space codes* for a MIMO IM/DD OWC system, with a channel

A. Chaaban is with the School of Engineering, University of British Columbia, Kelowna, BC Canada V1Y 1V7. Email: anas.chaaban@ubc.ca. His work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Y. Sapenov is with Samruk-Energy, Astana, Kazakhstan. Email: e.sapenov@samruk-energy.kz. He contributed to this work while he was a Masters student with the Computer, Electrical, and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia.

Z. Rezki is with the Department of Electrical and Computer Engineering, University of Idaho, Moscow, ID 83844-1023, USA. Email: zrezki@uidaho.edu.

M.-S. Alouini is with the CEMSE Division, KAUST, Thuwal 23955-6900, Saudi Arabia. Email: slim.alouini@kaust.edu.sa.

This paper is an extension of the work [1] presented at ISIT 2017.

¹Throughout the paper, we refer to spatial RC simply as RC.

matrix with independent log-normally distributed components and identically distributed columns, and under sum electrical or optical power constraints. Contrary to [11], [17], this result is not restricted to OOK.

Note that RC can be considered as a special case of STBC, where the temporal dimension is ignored and the transmit symbol is repeated spatially. Thus, results in [11], [17] indicate that the time dimension of an STBC is redundant in IM/DD OWC under the restrictions considered in these works (OOK, individual constraints, etc.). This is supported by intuition due to the nonnegativity of the channel gains and transmit symbols in IM/DD OWC. This nonnegativity rules out the possibility of destructive interference which is the main factor that promotes STBCs in RF communications. Does the same hold true under more general assumptions?

In this paper, we aim to address this problem analytically. Namely, we consider a multiple input multiple output (MIMO) IM/DD OWC system with N_t and N_r transmit and receiver apertures, respectively, under individual electrical, sum electrical, individual optical, and sum optical power constraints. From the code perspective, we consider the class of M -ary pulse-amplitude modulation (M -PAM) rate-1 direct current (DC) offset STBC,² where the transmit signal is constructed linearly as $\mathbf{s} = \mathbf{G}\mathbf{x} + \mathbf{d}$ where $\mathbf{s}$ is the LN_t -dimensional transmit vector spanning L channel uses, $\mathbf{x}$ is an L -dimensional vector of PAM symbols, $\mathbf{G}$ is a coding matrix, and $\mathbf{d}$ is a DC offset. This is referred to henceforth as an $(M, 1)$ DC-STBC. Moreover, we allow the channel state to follow general distributions of IM/DD OWC, with the restriction that the temporal variation is quasi-static as commonly considered in STBC literature [15]–[17]. This quasi-static assumption applies for VLC systems with limited mobility or free-space optical (FSO) systems with turbulence and pointing errors, where the channel varies very slowly in comparison to the symbol duration [2]. We assume the availability of channel-state information at the receiver and the lack thereof at the transmitter.

Note that this setup is more general than [11], [17], since OOK-RC and OOK-STBCs are special cases of the considered $(M, 1)$ DC-STBC, and we do not impose restrictions on the number of apertures and channel statistics. Moreover, apart from the signal alphabet restriction to M -PAM herein, this setup is also more general than [12], [13] because we consider space-time codes and arbitrary channel distributions, whereas [12], [13] consider space-codes and log-normally distributed channels. As a performance criterion, we use the *average worst-case PEP (WC-PEP)*, where the average is with respect to channel statistics, and the ‘worst-case’ refers to the minimum distance between codeword pairs. We focus on the average WC-PEP for several reasons. First, it dictates the diversity order. Second, it can be used to bound and also approximate the actual error probability [18], [19]. Finally, it is a tractable upper bound on the worst-case average PEP (considered in [17] e.g.) and thus it provides a performance guarantee thereon.

The main results of the paper are summarized in Table I and explained next. We prove that, subject to *individual power constraints* and for a given channel state, the WC-PEP is minimized by RC among all $(M, 1)$ DC-STBC. The proof is obtained by deriving an upper bound on the minimum distance of an arbitrary $(M, 1)$ DC-STBC satisfying the constraints, and then showing that this upper bound is indeed achievable by RC. Using this result, we then show that for a system with individual power constraints and a quasi-static channel with arbitrary distribution, RC minimizes the *average WC-PEP*. This agrees with [17] which shows the same for independent and identically distributed (i.i.d.) Gamma-Gamma distributed channels in terms of the worst-case average PEP. Then, we take this result one step further by proving that the time dimension of $(M, 1)$ DC-STBC is redundant from an average WC-PEP point of view for a system with an arbitrary quasi-static channel and with a *sum power constraint*. In this case, spatial beamforming minimizes the average WC-PEP,³ and the optimal beamformer can be obtained by the transmitter if it has access to channel statistics. All these results hold under electrical and optical power constraints. Furthermore, subject to a *sum electrical power constraint*, we show that if the channel matrix has i.i.d. columns (channels from a transmit aperture to all receiver apertures), then the optimal beamformer coincides with RC. This conclusion applies generally to channel matrices with i.i.d. columns and not necessarily i.i.d. components. This also applies to channels with a *sum optical power constraint at high SNR*, but interestingly is not true at low SNR where sending from a single aperture is optimal.

We verify the results numerically by showing that the solution obtained by analysis coincides with the one obtained numerically, and by showing that the minimum WC-PEP provides a good approximation at high SNR of the actual error probability obtained by Monte Carlo simulation.

These results are interesting because under these considerations, the $(M, 1)$ DC-STBC reduces to a simple space code which leads to lower transmitter complexity, and reduces the vector detection of $(M, 1)$ DC-STBCs to scalar detection leading to lower detection complexity. As side remarks, we state the solutions for quasi-static channels with channel-state information at the transmitter (CSIT). Under individual constraints, RC remains optimal from an average WC-PEP point of view. Under sum constraints, beamforming is optimal where the direction varies between electrical and optical power constrained systems.

Finally, it is worth to note that although the time dimension is redundant given the $(M, 1)$ DC-STBC construction, this may not be the case in a more general space-time code construction. For instance, Zhang *et al.* show in [20] that choosing $\mathbf{s}$ generally as a vector of nonnegative integers provides performance gains over time-disjoint design.

The rest of the paper is organized as follows. In Sec. II we introduce the system model and formulate the problem. The summary of main results is provided in Sec. III. The results are proved in Sec. IV and V. Then, in Sec. VI, we

²Throughout the paper, rate-1 codes are codes, whose rate is 1 symbol per channel use on average.

³Note that spatial beamforming can be interpreted as ‘weighted’ RC, since the same symbol is transmitted from multiple apertures possibly with different scaling.

TABLE I
SUMMARY OF RESULTS. BEAMFORMING HERE REFERS TO A GENERAL BEAMFORMING DIRECTION, WHICH INCLUDES SELECTION AND REPETITION AS SPECIAL CASES.

	Sum Constraint			Individual Constraint
	Electrical	Optical		
		High SNR	Low SNR	
No CSIT, i.i.d. columns	Repetition	Repetition	Selection	Repetition
No CSIT, arbitrary channels	Beamforming	Beamforming		
With CSIT		Selection		

provide numerical evaluations and simulations for different OWC scenarios verifying the main results of the paper. Finally, we conclude in Sec. VII, where we also highlight some interesting open problems.

Notations: Throughout the paper, we use bold lower-case and upper-case letters to denote vectors and matrices, respectively. We use $\mathbb{R}^{M \times N}$ to denote the set of $M \times N$ real matrices, and $\mathbb{R}_+^{M \times N}$ to denote that subset of $\mathbb{R}^{M \times N}$ with nonnegative elements. We write $\mathbf{x}^T$ and $\|\mathbf{x}\|_p$ to denote the transpose and the ℓ_p -norm of $\mathbf{x}$. We also use $\|\mathbf{X}\|_F$ to denote the Frobenius norm of $\mathbf{X}$. We denote the $L \times L$ identity matrix and the M -dimensional all-ones vector by $\mathbf{I}_L$ and $\mathbf{1}_M$, respectively, and we use $\otimes$ to denote the Kronecker product.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. Channel Model

Consider an IM/DD OWC system consisting of N_t transmit apertures and N_r receive aperture as shown in Fig. 1. The transmission can be represented as a discrete-time channel (after sampling at the receiver) where the received signal at time t can be written as

$$\mathbf{y}(t) = \mathbf{H}(t)\mathbf{s}(t) + \mathbf{n}(t), \quad t = 1, 2, \dots \quad (1)$$

where $\mathbf{s}(t) = [s_1(t), \dots, s_{N_t}(t)]^T \in \mathbb{R}_+^{N_t \times 1}$ is the vector of transmit signals, $\mathbf{H}(t) \in \mathbb{R}_+^{N_r \times N_t}$ is the matrix of the channel coefficients,⁴ and $\mathbf{n}(t)$ is an N_r -dimensional vector of independent Gaussian noises with zero mean and variance σ^2 . Noise is independent through time, and combines thermal noise and background radiation. Here, the nonnegativity of $\mathbf{s}(t)$ and $\mathbf{H}(t)$ follows due to the IM/DD operation [21]. Note that $s_i(t)$ can be interpreted as optical intensity or electric current due to their linear relation (assuming unity electrical-to-optical and optical-to-electrical conversion efficiency without loss of generality).

The channel coefficients combine several physical effects pertaining to OWC, such as path-loss (due to transmitter beam divergence, receiver field of view, and transmitter–receiver distance and relative position) and/or turbulence induced fading (due to variations in the transmission medium properties such as temperature). All these effects are abstracted in $\mathbf{H}$. We consider quasi-static channels, where $\mathbf{H}(t)$ remains fixed for a block of L transmissions and changes independently to another state in the next block. Thus, $\mathbf{H}(mL + 1) = \dots = \mathbf{H}((m + 1)L) = \mathbf{H}_m$ for $m \in \mathbb{N}$. This assumption is motivated by VLC with limited mobility and outdoor FSO

Fig. 1. A 3×2 MIMO OWC system at time t : $s_j(t) \geq 0$ is the optical intensity of aperture j , $h_{ij}(t) \geq 0$ is the channel from aperture j to the detector i , and $n_i(t)$ is Gaussian noise.

channels with turbulence and/or pointing errors, in both of which the channels vary very slowly in comparison to the symbol duration and hence can be assumed constant over a block [2]. It is assumed that the channel state information (CSI) is known at the receiver side but not at the transmitter side.

The nonnegativity of $\mathbf{s}(t)$ can be ensured using a DC offset. This signal is constrained by a power constraint. Letting S_i denote the random variable representing $s_i(t)$, common power constraints considered in this context can be expressed as follows [22]:

- i) Individual electrical power constraint: $\mathbb{E}[S_i^2] \leq P_e, \forall i \in \{1, \dots, N_t\}$;
- ii) Sum electrical power constraint: $\sum_{i=1}^{N_t} \mathbb{E}[S_i^2] \leq N_t P_e$;
- iii) Individual optical power constraint: $\mathbb{E}[S_i] \leq P_o, \forall i \in \{1, \dots, N_t\}$;
- iv) Sum optical power constraint: $\sum_{i=1}^{N_t} \mathbb{E}[S_i] \leq N_t P_o$.

All these constraints are practical in OWC communications,⁵ and will be considered in our analysis. Since we are interested in the performance of M -PAM rate-1 DC-STBC, we define it next.

B. M -PAM rate-1 DC-offset STBC

1) *Transmission:* Let $\mathbf{x} \in \mathbb{R}^{L \times 1}$ denote a vector of symbols to be sent to the receiver in L transmissions $t \in \{1, \dots, L\}$ (block $m = 0$). This is chosen uniformly and independently from the alphabet $\mathcal{X}_M$, which is an L -fold Cartesian product of an M -PAM constellation $\mathcal{P}_M = \{2i - (M - 1) | i = 0, \dots, M - 1\}$ where M is even, i.e.,

$$\mathcal{X}_M = \mathcal{P}_M^L. \quad (2)$$

⁵Note that the dissipated power at the transmitter is proportional to the square of the current. This is why an electrical power constraint depends on $\mathbb{E}[S_i^2]$. On the other hand, the optical power (intensity) varies linearly with the current, leading to an optical power constraint which depends on $\mathbb{E}[S_i]$.

Then, $\mathbf{x}$ is encoded in the transmit signals $\mathbf{s}(1), \dots, \mathbf{s}(L)$. This is a rate-1 code where one constellation symbol is sent per transmission on average. If we stack the vectors $\mathbf{s}(1), \dots, \mathbf{s}(L)$ into one vector $\bar{\mathbf{s}} = [\mathbf{s}(1)^T, \dots, \mathbf{s}(L)^T]^T$, then, a DC-offset STBC is constructed as follows:

$$\bar{\mathbf{s}} = \mathbf{G}\mathbf{x} + \mathbf{d}, \quad (3)$$

where $\mathbf{G} = [\mathbf{g}_1, \dots, \mathbf{g}_{N_t L}]^T \in \mathbb{R}^{N_t L \times L}$ is the coding matrix, and $\mathbf{d} = [d_1, \dots, d_{N_t L}]^T \in \mathbb{R}_+^{N_t L}$ is a DC-offset. It is assumed that $\mathbf{G}$ has rank L (rate-1 code), and $\mathbf{G}$ and $\mathbf{d}$ are known at the receiver and transmitter. Under this construction, the nonnegativity constraint can be stated as follows

$$\mathbf{g}_i^T \mathbf{x} + d_i \geq 0, \quad \forall i \in \{1, \dots, N_t L\}, \quad \forall \mathbf{x} \in \mathcal{X}_M, \quad (4)$$

Due to this, the minimum DC-offset which satisfies the non-negativity constraint (4) is

$$d_i = (M - 1) \|\mathbf{g}_i\|_1. \quad (5)$$

Furthermore, the power constraints can be written as:

- i) Individual electrical power constraint: $\frac{M^2-1}{3} \|\mathbf{g}_i\|_2^2 + d_i^2 \leq P_e, \forall i \in \{1, \dots, N_t L\};$
- ii) Sum electrical power constraint: $\sum_{i=(k-1)N_t+1}^{kN_t} \left(\frac{M^2-1}{3} \|\mathbf{g}_i\|_2^2 + d_i^2 \right) \leq N_t P_e, \forall k \in \{1, \dots, L\};$
- iii) Individual optical power constraint: $d_i \leq P_o, \forall i \in \{1, \dots, N_t L\};$
- iv) Sum optical power constraint: $\sum_{i=(k-1)N_t+1}^{kN_t} d_i \leq N_t P_o, \forall k \in \{1, \dots, L\}.$

The transmitter sends $\bar{\mathbf{s}}$ through L transmissions, by sending $\mathbf{s}(t)$ in time t .

2) *Reception*: The received signal in L transmissions can be written as

$$\bar{\mathbf{y}} = [\mathbf{y}(1)^T, \dots, \mathbf{y}(L)^T]^T = \bar{\mathbf{H}}\bar{\mathbf{s}} + \bar{\mathbf{n}}, \quad (6)$$

where $\bar{\mathbf{n}} = [\mathbf{n}(1)^T, \dots, \mathbf{n}(L)^T]^T$, and $\bar{\mathbf{H}} = \mathbf{I}_L \otimes \mathbf{H}_m$ is the extended-channel matrix for block m . Substituting (3) in (6) yields

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}\mathbf{G}\mathbf{x} + \bar{\mathbf{H}}\mathbf{d} + \bar{\mathbf{n}}, \quad (7)$$

as shown in Fig. 2. Since $\bar{\mathbf{H}}$ and $\mathbf{d}$ are known at the receiver, $\bar{\mathbf{H}}\mathbf{d}$ can be subtracted leading to

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{x} + \bar{\mathbf{n}}, \quad (8)$$

where $\mathbf{A} = \bar{\mathbf{H}}\mathbf{G}$. The signal is then detected using a maximum likelihood (ML), i.e.

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}_0 \in \mathcal{X}_M} \|\dot{\mathbf{y}} - \mathbf{A}\mathbf{x}_0\|_2. \quad (9)$$

The goal is to design $\mathbf{G}$ so that the system performance is optimized in terms of a criterion of interest defined as follows.

C. The Worst-Case PEP Criterion

We choose the WC-PEP as the main performance criterion in our analysis, i.e. the PEP corresponding to the two codewords of the DC-STBC at the minimum distance. For a given transmission block m with $\mathbf{H}_m = \mathbf{H}$, the PEP between

symbols $\mathbf{x}_a$ and $\mathbf{x}_b \neq \mathbf{x}_a$ in $\mathcal{X}_M$ is obtained as for the additive white Gaussian noise channel [18], [23], i.e.,

$$\mathbb{P}\{\mathbf{x}_a \leftrightarrow \mathbf{x}_b | \mathbf{H}\} = Q\left(\frac{\|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2}{2\sigma}\right), \quad (10)$$

where $Q(\cdot)$ is the Q-function. Accordingly, the WC-PEP is given by

$$\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H}) = \max_{\substack{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M \\ \mathbf{x}_a \neq \mathbf{x}_b}} Q\left(\frac{\|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2}{2\sigma}\right). \quad (11)$$

This criterion is interesting for the following reasons. First, the word-error probability of the system given by

$$\mathbb{P}^{\text{E}}(\mathbf{G}, \mathbf{H}) = \frac{1}{|\mathcal{X}_M|} \sum_{\substack{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M \\ \mathbf{x}_a \neq \mathbf{x}_b}} \mathbb{P}\{(\mathbf{x}, \hat{\mathbf{x}}) = (\mathbf{x}_a, \mathbf{x}_b) | \mathbf{H}\}, \quad (12)$$

can be approximated in terms of the WC-PEP using union bound estimate as follows [19]

$$\mathbb{P}^{\text{E}}(\mathbf{G}, \mathbf{H}) \approx k_{\min} \mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H}), \quad (13)$$

where $k_{\min}$ is the average number of neighbors in the constellation at the minimum distance. This approximation is very accurate at high SNR, and is mainly dictated by $\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})$ since $k_{\min}$ has a ‘‘much milder impact’’ [19]. Second, the WC-PEP can be used to bound $\mathbb{P}^{\text{E}}(\mathbf{G}, \mathbf{H})$ by

$$\frac{\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})}{|\mathcal{X}_M|} \leq \mathbb{P}^{\text{E}}(\mathbf{G}, \mathbf{H}) \leq (|\mathcal{X}_M| - 1) \mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H}), \quad (14)$$

where the upper bound is given in [18] and the lower bound can be easily obtained from (12).

This formulation is for a given channel state. Next, we extend this criterion to the quasi-static channel by defining the average WC-PEP.

D. Problem Formulation

Since the channel is quasi-static where $\mathbf{H}$ changes in each block, a practical performance criterion is the average error rate $\mathbb{E}_{\mathbf{H}}[\mathbb{P}^{\text{E}}(\mathbf{G}, \mathbf{H})]$. Thus, according to the discussion above, $\mathbb{E}_{\mathbf{H}}[\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})]$ can serve as a surrogate performance indicator. Therefore, our objective is to study the following problem

$$\bar{\mathbb{P}}_{\min}^{\text{PE}} = \min_{\mathbf{G}} \mathbb{E}_{\mathbf{H}}[\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})], \quad (15)$$

i.e., the *minimum average WC-PEP*.⁶

According to the discussion above, minimizing this average WC-PEP also minimizes the bounds of the error probability in (14), which consequently dictates the diversity order. This also minimizes the approximation (13). Note that the average WC-PEP is an upper bound on the worst-case average PEP studied in [17], and hence provides a performance guarantee on this latter criterion. For these reasons, we choose the WC-PEP as our performance criterion. The resulting problem is formulated in the next subsection.

⁶Other practical performance criteria are the worst-case average PEP studied in [17] and the average PEP studied in [12] for MIMO OWC systems with some conditions on the channel distribution.

Fig. 2. Signal processing of $N_t \times N_t$ DC-STBC.

As an intermediate step towards this objective, we define the *minimum WC-PEP* for a given channel as

$$\mathbb{P}_{\min}^{\text{PE}}(\mathbf{H}) = \min_{\mathbf{G}} \mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H}) \quad (16)$$

$$= \min_{\mathbf{G}} \max_{\substack{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M \\ \mathbf{x}_a \neq \mathbf{x}_b}} Q\left(\frac{\|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2}{2\sigma}\right). \quad (17)$$

Since $Q(x)$ is a decreasing function of x , then $\mathbb{P}_{\min}^{\text{PE}}(\mathbf{H})$ can be derived by solving

$$\max_{\mathbf{G}} \min_{\substack{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M \\ \mathbf{x}_a \neq \mathbf{x}_b}} \|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2^2, \quad (18)$$

where $\mathbf{G}$ is subject to one of the four power constraints. Note that we maximize only with respect to $\mathbf{G}$ since $\mathbf{d}$ is a function of $\mathbf{G}$ as specified in (5). This provides a formulation which applies for static channels.

In the rest of the paper, we study the optimal M -PAM rate-1 DC-STBC that solves these optimization problems.

Definition 1. The M -PAM rate-1 DC-offset STBC will be referred to simply as an $(M, 1)$ DC-STBC henceforth.

Definition 2. To avoid repetition, we say an $(M, 1)$ DC-STBC is ‘optimal’ if it is optimal in terms of the above criteria, i.e., WC-PEP for a given $\mathbf{H}$, and average WC-PEP for quasi-static $\mathbf{H}$.

We summarize the main results of the paper in the following section.

III. SUMMARY OF MAIN RESULTS

The main result of the paper can be summarized as follows. The optimal $(M, 1)$ DC-STBC is in fact a simple *DC-offset space code* in the form of RC or beamforming. This is true for all 4 constraints given in Sec. II-A. This owes to the fact that the channels are positive-valued in IM/DD OWC, which prevents any ‘destructive interference’ of symbols at the receiver side. This is contrary to RF where this effect can occur, and thus, one should capitalize on multiple transmission to harness benefit from multiple channel realizations.

We first introduce the results under electrical power constraints, followed by optical power constraints.

A. Electrical Power Constraints

1) *Individual Power Constraints:* Under an individual electrical power constraint and for a given channel state $\mathbf{H}$, we show that the optimal $(M, 1)$ DC-STBC coincides with spatial RC. This general statement is summarized as follows.

Lemma 1. Given a channel $\mathbf{H}$, the optimal $\mathbf{G}$ under an individual electrical power constraint P_e is $\mathbf{G}_e = \gamma \mathbf{I}_L \otimes \mathbf{1}_{N_t}$ (RC), and achieves a minimum WC-PEP of

$$\mathbb{P}_{\min, e}^{\text{PE}}(\mathbf{H}) = Q\left(\frac{\gamma}{\sigma} \|\mathbf{H} \mathbf{1}_{N_t}\|_2\right), \quad (19)$$

where $\gamma = \sqrt{\frac{3P_e}{4M^2 - 6M + 2}}$.

Proof: The proof is based on upper bounding $\min_{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M} \|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2^2$ in (18) and showing that the obtained upper bound is achievable using RC. Details are given in Sec. IV-A. ■

This lemma which focuses on a given channel state $\mathbf{H}$ forms the basis for the following results. In a quasi-static channel, spatial RC is optimal for a system with individual power constraints under any channel distribution as described next.

Theorem 1. For a quasi-static channel with an individual electrical power constraint P_e , the optimal $(M, 1)$ DC-STBC corresponds to $\mathbf{G}_e = \gamma \mathbf{I}_L \otimes \mathbf{1}_{N_t}$ (RC) where γ is defined in Lemma 1, and achieves a minimum average WC-PEP of

$$\bar{\mathbb{P}}_{\min, e}^{\text{PE}} = \mathbb{E}_{\mathbf{H}} \left[Q\left(\frac{\gamma}{\sigma} \|\mathbf{H} \mathbf{1}_{N_t}\|_2\right) \right]. \quad (20)$$

Proof: We have $\bar{\mathbb{P}}_{\min, e}^{\text{PE}} = \min_{\mathbf{G}} \mathbb{E}_{\mathbf{H}}[\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})] \geq \mathbb{E}_{\mathbf{H}}[\min_{\mathbf{G}} \mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})] = \mathbb{E}_{\mathbf{H}}[\mathbb{P}_{\min, e}^{\text{PE}}(\mathbf{H})]$, where the minimization is with respect to $\mathbf{G}$ that satisfy the individual power constraint, and $\mathbb{P}_{\min, e}^{\text{PE}}(\mathbf{H})$ is defined in Lemma 1. This lower bound is achievable by setting $\mathbf{G} = \gamma \mathbf{I}_L \otimes \mathbf{1}_{N_t}$ which concludes the proof. ■

2) *Sum Power Constraint:* On the other hand, under a sum power constraint, spatial beamforming is optimal as stated next.

Theorem 2. For a quasi-static channel with a sum electrical power constraint $N_t P_e$, the optimal $(M, 1)$ DC-STBC corresponds to $\mathbf{G}_{\text{sum}}(\mathbf{r}^*) = \gamma \sqrt{N_t} \mathbf{I}_L \otimes \mathbf{r}^*$, and achieves a minimum average WC-PEP of

$$\bar{\mathbb{P}}_{\min, \text{sum}}^{\text{PE}} = \min_{\substack{\mathbf{r} \in \mathbb{R}_+^{N_t \times 1} \\ \|\mathbf{r}\|_2 = 1}} \mathbb{E}_{\mathbf{H}} \left[Q\left(\frac{\gamma \sqrt{N_t}}{\sigma} \|\mathbf{H} \mathbf{r}\|_2\right) \right] \quad (21)$$

where $\mathbf{r}^*$ is the optimal solution of the above minimization and γ is defined in Lemma 1.

Proof: The proof is obtained by developing an achievable lower bound for (15) under the given constraints. Details are given in Sec. IV-B. ■

The optimal $\mathbf{r}^*$ in Theorem 2 can be found numerically in general for a given distribution of $\mathbf{H}$. However, this can be specified explicitly if $\mathbf{H}$ has *i.i.d.* columns, where it is optimal to use spatial RC as given next.

Corollary 1. *If $\mathbf{H}$ has i.i.d. columns, then the optimal $\mathbf{r}$ in Theorem 2 is $\mathbf{r}^* = \frac{1}{\sqrt{N_t}} \mathbf{1}_{N_t}$ (RC), and achieves a minimum average WC-PEP of*

$$\bar{\mathbb{P}}_{\min, \text{e sum}}^{\text{PE}} = \mathbb{E}_{\mathbf{H}} \left[Q \left(\frac{\gamma}{\sigma} \|\mathbf{H} \mathbf{1}_{N_t}\|_2 \right) \right], \quad (22)$$

where γ is defined in Lemma 1.

Proof: The proof is obtained using majorization techniques [24] and is detailed in Sec. IV-C. ■

Similar statements as above can be made under optical power constraints as stated next.

B. Optical Power Constraints

1) *Individual Power Constraints:* We start by stating a general lemma similar to Lemma 1.

Lemma 2. *Given a channel $\mathbf{H}$, the optimal $\mathbf{G}$ under an individual optical power constraint P_o is $\mathbf{G} = \frac{P_o}{M-1} \mathbf{I}_L \otimes \mathbf{1}_{N_t}$ (RC), and achieves a minimum WC-PEP of*

$$\bar{\mathbb{P}}_{\min, o}^{\text{PE}}(\mathbf{H}) = Q \left(\frac{P_o}{(M-1)\sigma} \|\mathbf{H} \mathbf{1}_{N_t}\|_2 \right). \quad (23)$$

Proof: The proof is similar to that of Lemma 1. Details are given in Sec. V-A. ■

We apply this to quasi-static channels with individual constraints next, where spatial RC is optimal.

Theorem 3. *For a quasi-static channel with an individual optical power constraint P_o , the optimal $(M, 1)$ DC-STBC corresponds to $\mathbf{G} = \frac{P_o}{M-1} \mathbf{I}_L \otimes \mathbf{1}_{N_t}$ (RC), and achieves a minimum average WC-PEP of*

$$\bar{\mathbb{P}}_{\min, o}^{\text{PE}} = \mathbb{E}_{\mathbf{H}} \left[Q \left(\frac{P_o}{(M-1)\sigma} \|\mathbf{H} \mathbf{1}_{N_t}\|_2 \right) \right]. \quad (24)$$

Proof: The proof is similar to that of Theorem 1 and is hence omitted. ■

2) *Sum Power Constraint:* Under a sum power constraint, we can show that spatial beamforming is optimal as stated in the following theorem.

Theorem 4. *For a quasi-static channel with a sum optical power constraint $N_t P_o$, the optimal $(M, 1)$ DC-STBC corresponds to $\mathbf{G} = \frac{P_o}{M-1} \mathbf{I}_L \otimes \mathbf{r}^*$ and achieves a minimum average WC-PEP of*

$$\bar{\mathbb{P}}_{\min, o \text{ sum}}^{\text{PE}} = \min_{\substack{\mathbf{r} \in \mathbb{R}_+^{N_t \times 1} \\ \|\mathbf{r}\|_1 = 1}} \mathbb{E}_{\mathbf{H}} \left[Q \left(\frac{N_t P_o}{(M-1)\sigma} \|\mathbf{H} \mathbf{r}\|_2 \right) \right] \quad (25)$$

where $\mathbf{r}^*$ is the optimal $\mathbf{r}$ for the above minimization.

Proof: The proof is similar to that of Theorem 2 and is given in Sec. V-B. ■

Interestingly, specializing Theorem 4 to the case of $\mathbf{H}$ with i.i.d. columns does not lead to the same result as in Corollary 1. Recall that Corollary 1 asserts that the optimal $(M, 1)$ DC-STBC under a sum electrical power constraint when $\mathbf{H}$ is quasi-static with i.i.d. columns is RC. While this holds universally for any sum electrical power constraint, the same can not be said about the case with a sum optical power constraint. In fact, in the latter case, if the sum optical power

constraint is small then it is optimal to transmit from only one aperture. This holds true not only for channels with i.i.d. columns, but for any distribution of $\mathbf{H} \in \mathbb{R}_+^{N_r \times N_t}$. Generally, in this case, the optimal $(M, 1)$ DC-STBC at low SNR transmits from the aperture which has the largest average channel magnitude to the receiver. If $\mathbf{H}$ has i.i.d. columns, then all apertures are equally strong and it is optimal to transmit from any aperture. This prevents proving a general result like that in Corollary 1. This result is stated next.

Corollary 2. *For any distribution of $\mathbf{H}$ we have that*

$$\lim_{\frac{P_o}{\sigma} \rightarrow 0} \left(\bar{\mathbb{P}}_{\min, o \text{ sum}}^{\text{PE}} - \left(\frac{1}{2} - \frac{N_t P_o \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_q\|_2]}{\sqrt{2\pi}(M-1)\sigma} \right) \right) = 0, \quad (26)$$

where $q = \arg \max_i \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_i\|_2]$ (ties are resolved randomly) and $\tilde{\mathbf{h}}_i$ is the i^{th} column of $\mathbf{H}$. This asymptotic average WC-PEP is achieved for $\frac{P_o}{\sigma} \rightarrow 0$ by setting $\mathbf{r}$ equal to the q^{th} column of $\mathbf{I}_{N_t}$.

Proof: The proof is based on using the expansion $Q(x) = \frac{1}{2} - \frac{1}{\sqrt{2\pi}}x + O(x^3)$ as $x \rightarrow 0$ [26], as given in Sec. V-C. ■

Although the low SNR regime is not a desired regime of operation due to the high PEP, this corollary serves as a counter example which demonstrates the RC is not universally optimal under a sum optical power constraint. Nevertheless, in the high SNR regime $\frac{P_o}{\sigma} \rightarrow \infty$, RC is optimal if $\mathbf{H}$ has i.i.d. columns as stated next.

Corollary 3. *If $\mathbf{H}$ has i.i.d. columns, then we have*

$$\lim_{\frac{P_o}{\sigma} \rightarrow \infty} \frac{\bar{\mathbb{P}}_{\min, o \text{ sum}}^{\text{PE}}}{\Omega \left(\frac{P_o}{\sigma} \right)} = 1, \quad (27)$$

where $\Omega \left(\frac{P_o}{\sigma} \right) = \mathbb{E}_{\mathbf{H}} \left[\frac{e^{-\frac{P_o^2}{2(M-1)^2 \sigma^2} \|\mathbf{H} \mathbf{1}_{N_t}\|_2^2}}{\sqrt{2\pi} \frac{P_o}{(M-1)\sigma} \|\mathbf{H} \mathbf{1}_{N_t}\|_2} \right]$. This asymptotic average WC-PEP is achieved for $\frac{P_o}{\sigma} \rightarrow \infty$ using $\mathbf{r} = \frac{1}{N_t} \mathbf{1}_{N_t}$ (RC).

Proof: The proof is given in Sec. V-D. ■

C. Discussion

The structure of the optimal $(M, 1)$ DC-STBC given in all statements above implies that the time dimension is redundant under the considered construction. With this structure, each symbol of $\mathbf{x}$ is sent (possibly with some scaling) from all transmit apertures in one and only one channel use. The $(M, 1)$ DC-STBC thus reduces to a simple DC-offset space code. This significantly reduces the encoding and decoding complexity since the transmitter and receiver need to process vectors of length N_t and N_r , respectively, instead of LN_t and LN_r in the case of a general $(M, 1)$ DC-STBC. Consequently, vector ML detection simplifies to scalar ML detection. For instance, under individual electrical power constraints, instead of detecting $\hat{\mathbf{x}}$ as in (9), we can detect component-wise as

$$\hat{x}_i = \arg \max_{x_i \in \mathcal{P}_M} \|\mathbf{y}(t) - (x_i + (M-1)\gamma) \mathbf{H} \mathbf{1}_{N_t}\|_2,$$

for $i = 1, \dots, L$, where x_i and $\hat{x}_i$ are the i^{th} components of $\mathbf{x}$ and $\hat{\mathbf{x}}$. Similarly for the other power constraints.

The statements above apply to quasi-static channels with no CSIT. While this is the classical scenario of application of STBC, it is worth to extend the statements to a system with CSIT. First, nothing changes under individual constraints with CSIT since the optimal $\mathbf{G}$ for a given $\mathbf{H}$ has the RC structure independent of $\mathbf{H}$ (cf. Lemmas 1 and 2). Thus, Theorems 1 and 3 continue to apply in this case. Under a sum power constraint with CSIT, we have the following remarks.

Remark 1. Similar to Lemma 1, one can show that given a channel $\mathbf{H}$, the optimal $\mathbf{G}$ under a sum electrical power constraint $N_t P_e$ is $\mathbf{G}(\mathbf{H}) = \gamma \sqrt{N_t} \mathbf{I}_L \otimes \mathbf{w}_{\max}(\mathbf{H})$ (spatial beamforming) achieving a minimum WC-PEP of $Q\left(\frac{\gamma}{\sigma} \sqrt{N_t \lambda_{\max}(\mathbf{H})}\right)$, where $\gamma = \sqrt{\frac{3P_e}{4M^2 - 6M + 2}}$, and $\lambda_{\max}(\cdot)$ and $\mathbf{w}_{\max}(\cdot)$ are the largest eigenvalue and the associated eigenvector of $\mathbf{H}^T \mathbf{H}$, respectively. In a quasi-static channel with CSIT, this leads to an average WC-PEP of $\mathbb{E}_{\mathbf{H}} \left[Q\left(\frac{\gamma}{\sigma} \sqrt{N_t \lambda_{\max}(\mathbf{H})}\right) \right]$.

Remark 2. Similar to Lemma 2, one can show that given a channel $\mathbf{H}$, the optimal $\mathbf{G}$ under a sum optical power constraint $N_t P_o$ is $\mathbf{G}(\mathbf{H}) = \frac{N_t P_o}{M-1} \mathbf{I}_L \otimes \mathbf{e}_{q(\mathbf{H})}$ (aperture selection) achieving a minimum WC-PEP of $Q\left(\frac{N_t P_o}{(M-1)\sigma} \max_i \|\tilde{\mathbf{h}}_i\|_2\right)$, where $q(\mathbf{H}) = \arg \max_i \|\tilde{\mathbf{h}}_i\|_2$, and $\tilde{\mathbf{h}}_i$ and $\mathbf{e}_i$ are the i^{th} columns of $\mathbf{H}$ and $\mathbf{I}_{N_t}$, respectively. In a quasi-static channel with CSIT, this leads to an average WC-PEP of $\mathbb{E}_{\mathbf{H}} \left[Q\left(\frac{N_t P_o}{(M-1)\sigma} \max_i \|\tilde{\mathbf{h}}_i\|_2\right) \right]$.

These statements can be generally represented as shown in Table I.

Corollary 1 agrees with results in [11] which demonstrate that RC outperforms orthogonal STBC under i.i.d. log-normal channels from bit-error rate point-of-view when using OOK. This corollary also agrees with the result of [12] when restricted to DC-offset M -PAM constellations,⁷ which demonstrates the optimality of RC in terms of diversity order among all *space codes* when $\mathbf{H}$ has independent log-normal components, and identically distributed columns. Corollary 1 thus extends this statement to general distributions of channel matrices $\mathbf{H}$ with i.i.d. columns, and proves the optimality of RC in terms of average WC-PEP (and hence diversity order) among $(M, 1)$ DC-STBC (and hence DC-offset space codes). This makes the result more general. The MISO case considered in [11] can be obtained as a special case when $N_r = 1$, i.e., $\mathbf{H}$ is replaced with a row vector $\mathbf{h}^T$.

On the other hand, Corollaries 2 and 3 highlight an interesting aspect. Although RC has good performance in terms of bit-error rate and diversity order [11], [12], it is not universally optimal among $(M, 1)$ DC-STBCs. These corollaries indicate that the optimal $(M, 1)$ DC-STBC depends on SNR. However, for practical SNR values, RC is a good choice which becomes optimal as SNR increases.

The following sections prove these statements.

IV. MIMO CHANNEL WITH ELECTRICAL CONSTRAINTS

Here, we prove the results for a MIMO IM/DD system with electrical power constraints. We start by proving Lemma 1 which forms the basis for proving the other results.

A. Individual Constraints given $\mathbf{H}$ (Proof of Lemma 1)

Problem (18) with an individual electrical power constraint can be written as follows

$$\Theta_e \triangleq \max_{\mathbf{G} \in \mathcal{G}_{e,M}} \min_{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M} \|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2^2. \quad (28)$$

where

$$\begin{aligned} \mathcal{G}_{e,M} & \\ &= \left\{ \mathbf{G} \in \mathbb{R}^{N_t L \times L} \mid \frac{M^2 - 1}{3} \|\mathbf{g}_i\|_2^2 + (M - 1)^2 \|\mathbf{g}_i\|_1^2 \leq P_e, \right. \\ & \quad \left. \forall i \in \{1, \dots, N_t L\} \right\}. \end{aligned} \quad (29)$$

Our approach towards solving this max-min problem starts by upper bounding it, and then showing that the upper bound is achievable. We start by upper bounding the inner minimization.

Note that

$$\min_{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M} \|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2^2 \leq \min_{\mathbf{x}_b \in \mathcal{X}_M} \|\mathbf{A}(\mathbf{1}_L - \mathbf{x}_b)\|_2^2 \quad (30)$$

$$\leq \min_{\mathbf{x}_b \in \mathcal{X}'_M} \|\mathbf{A}(\mathbf{1}_L - \mathbf{x}_b)\|_2^2 \quad (31)$$

where $\mathcal{X}'_M = \{\mathbf{1}_L - 2\mathbf{e}_i \mid i = 1, \dots, L\}$ with $\mathbf{e}_i$ being the i^{th} column of $\mathbf{I}_L$. The first inequality follows by fixing $\mathbf{x}_a = \mathbf{1}_L$, and the second follows by restricting $\mathbf{x}_b$ to a specific structure, both of which can only increase the minimum value. This implies that

$$\min_{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M} \|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2^2 \leq \min_{i=1, \dots, L} 4\|\mathbf{a}_i\|_2^2 \quad (32)$$

where $\mathbf{a}_i$ is the i^{th} column of $\mathbf{A}$. Since the minimum is less than or equal the average, then

$$\min_{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M} \|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2^2 \leq \frac{1}{L} \sum_{i=1}^L 4\|\mathbf{a}_i\|_2^2 = \frac{4}{L} \|\mathbf{A}\|_F^2. \quad (33)$$

Thus, Θ_e in (28) is upper bounded by

$$\Theta_e \leq \max_{\mathbf{G} \in \mathcal{G}_{e,M}} \frac{4}{L} \|\mathbf{A}\|_F^2. \quad (34)$$

Now we consider the maximization with respect to $\mathbf{G}$. Let us write $\mathbf{G}$ and $\mathbf{A}$ as

$$\mathbf{G} = [\mathbf{F}(1)^T, \dots, \mathbf{F}(L)^T]^T, \quad (35)$$

$$\mathbf{A} = [\mathbf{A}(1)^T, \dots, \mathbf{A}(L)^T]^T, \quad (36)$$

⁷The result of [12] applies to more general positive constellations.

where $\mathbf{F}(k) \in \mathbb{R}^{N_t \times L}$ consists of the k^{th} block of N_t rows of $\mathbf{G}$ and $\mathbf{A}(k) = \mathbf{H}\mathbf{F}(k) \in \mathbb{R}^{N_r \times L}$ consists of the k^{th} block of N_r rows of $\mathbf{A}$. Note that $\mathbf{F}(k) \in \mathcal{F}_{e,M}$ where

$$\mathcal{F}_{e,M} = \left\{ \mathbf{F} \in \mathbb{R}^{N_t \times L} \left| \frac{M^2 - 1}{3} \sum_{j=1}^L f_{ij}^2 + (M - 1)^2 \left(\sum_{j=1}^L |f_{ij}| \right)^2 \leq P_e, \right. \right. \\ \left. \left. \forall i \in \{1, \dots, N_t\} \right\}, \quad (37)$$

with f_{ij} being the i^{th} component in $\mathbf{f}_j$, the j^{th} column of $\mathbf{F}$. Thus, the upper bound in (34) can be written as

$$\Theta_e \leq \max_{\mathbf{G} \in \mathcal{G}_{e,M}} \frac{4}{L} \sum_{k=1}^L \|\mathbf{A}(k)\|_F^2 \quad (38)$$

$$= \frac{4}{L} \sum_{k=1}^L \max_{\mathbf{F}(k) \in \mathcal{F}_{e,M}} \|\mathbf{A}(k)\|_F^2 \quad (39)$$

$$= \frac{4}{L} \sum_{k=1}^L \max_{\mathbf{F}(k) \in \mathcal{F}_{e,M}} \sum_{j=1}^{N_r} \|\mathbf{h}_j^T \mathbf{F}(k)\|_2^2 \quad (40)$$

$$\leq \frac{4}{L} \sum_{k=1}^L \sum_{j=1}^{N_r} \max_{\mathbf{F}(k) \in \mathcal{F}_{e,M}} \sum_{i=1}^L (\mathbf{h}_j^T \mathbf{f}_i(k))^2, \quad (41)$$

where $\mathbf{h}_j^T$ is the j^{th} row of $\mathbf{H}$. Since $\mathbf{h}_j$ is positive component-wise, then a necessary condition for the optimal $\mathbf{f}_i(k)$ is that it has components of similar sign, assumed positive henceforth without loss of generality. Since $\sum_{j=1}^L f_{ij}^2 \leq \left(\sum_{j=1}^L |f_{ij}| \right)^2$, then we have that

$$\Theta_e \leq \frac{4}{L} \sum_{k=1}^L \sum_{j=1}^{N_r} \max_{\mathbf{F}(k) \in \tilde{\mathcal{F}}_{e,M}} \sum_{i=1}^L (\mathbf{h}_j^T \mathbf{f}_i(k))^2, \quad (42)$$

where

$$\tilde{\mathcal{F}}_{e,M} = \left\{ \mathbf{F} \in \mathbb{R}_+^{N_t \times L} \left| \sum_{j=1}^L f_{ij}^2 \leq \gamma^2, \forall i \in \{1, \dots, N_t\} \right. \right\} \\ \supseteq \mathcal{F}_{e,M}, \quad (43)$$

and $\gamma = \sqrt{\frac{3P_e}{4M^2 - 6M + 2}}$. Now consider feasible $\mathbf{F}(k), \mathbf{F}'(k) \in \tilde{\mathcal{F}}_{e,M}$ with $\mathbf{f}'_j(k) = \mathbf{f}_j(k)$ for $j = 3, \dots, L$, $\mathbf{f}'_1(k) = \mathbf{0}_{N_t \times 1}$, and $f'_{i2}(k) = \sqrt{f_{i1}^2(k) + f_{i2}^2(k)}$ $\forall i \in \{1, \dots, N_t\}$. Both $\mathbf{F}(k)$ and $\mathbf{F}'(k)$ have rows of equal ℓ_2 -norm. One can easily verify that $\sum_{i=1}^L (\mathbf{h}_j^T \mathbf{f}_i(k))^2 \leq \sum_{i=1}^L (\mathbf{h}_j^T \mathbf{f}'_i(k))^2$. This can be repeated to show that the optimal $\mathbf{F}(k)$ for (42) has one nonzero column. This yields

$$\Theta_e \leq \frac{4}{L} \sum_{k=1}^L \sum_{j=1}^{N_r} \max_{\substack{\mathbf{f}_1(k) \in \mathbb{R}_+^{N_t \times 1} \\ f_{i1}(k)^2 \leq \gamma^2 \forall i}} (\mathbf{h}_j^T \mathbf{f}_1(k))^2. \quad (44)$$

Since the objective is increasing in $f_{i1}(k)$, then the maximum is achieved when the constraint is met with equality. Hence,

we can write the optimal $\mathbf{f}_1(k)$ as $\gamma \mathbf{1}_{N_t}$ independent of k (the maximization in (45) is equivalent for all k). This yields

$$\Theta_e \leq \sum_{k=1}^L \sum_{j=1}^{N_r} \frac{4\gamma^2}{L} (\mathbf{h}_j^T \mathbf{1}_{N_t})^2 = 4\gamma^2 \|\mathbf{H} \mathbf{1}_{N_t}\|_2^2. \quad (45)$$

Going back to (28) and setting $\mathbf{G} = \gamma \mathbf{I}_L \otimes \mathbf{1}_{N_t}$ leads to $\min_{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M} \|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2^2 = 4\gamma^2 \|\mathbf{H} \mathbf{1}_{N_t}\|_2^2$ which coincides with the upper bound. This proves the statement of Lemma 1.

B. Quasi-Static Channels with a Sum Constraint (Proof of Theorem 2)

In this case, we need to solve

$$\bar{\mathbb{P}}_{\min}^{\text{PE}} = \min_{\mathbf{G} \in \mathcal{G}_{e,\text{sum},M}} \mathbb{E}_{\mathbf{H}}[\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})], \quad (46)$$

where $\mathbf{G}$ is independent of $\mathbf{H}$ and

$$\mathcal{G}_{e,\text{sum},M} = \left\{ \mathbf{G} \in \mathbb{R}^{N_t L \times L} \left| \sum_{i=(k-1)N_t+1}^{kN_t} \left(\frac{M^2 - 1}{3} \|\mathbf{g}_i\|_2^2 + (M - 1)^2 \|\mathbf{g}_i\|_1^2 \right) \leq N_t P_e, \right. \right. \\ \left. \left. \forall k \in \{1, \dots, L\} \right\}.$$

The constraint of $\mathcal{G}_{e,\text{sum},M}$ can be written as $\|\mathbf{r}_k\|_2 \leq 1, \forall k \in \{1, \dots, L\}$, where we define $\mathbf{r}_k = [\rho_{(k-1)N_t+1}, \dots, \rho_{kN_t}]^T$ and

$$\rho_i = \sqrt{\frac{M^2 - 1}{3N_t P_e} \|\mathbf{g}_i\|_2^2 + \frac{(M - 1)^2}{N_t P_e} \|\mathbf{g}_i\|_1^2}, \quad (47)$$

which can be restricted to be nonnegative. Since $\mathbf{H}$ preserves its value throughout a block, then by symmetry, the optimal solution satisfies $\mathbf{r}_k = \mathbf{r} \forall k$ for some $\mathbf{r}$. Let us write $\mathbf{G} = \bar{\mathbf{R}} \tilde{\mathbf{G}}$ where $\bar{\mathbf{R}} = \mathbf{I}_L \otimes \mathbf{R}$, and $\mathbf{R}$ is a diagonal matrix with $\mathbf{r}$ as its diagonal. The resulting $(M, 1)$ DC-STBC can be expressed as $\bar{\mathbf{R}}(\tilde{\mathbf{G}}\mathbf{x} + \tilde{\mathbf{d}})$, where $\tilde{\mathbf{G}} = [\tilde{\mathbf{g}}_1, \dots, \tilde{\mathbf{g}}_{N_t L}]^T$, $\tilde{\mathbf{d}} = [\tilde{d}_1, \dots, \tilde{d}_{N_t L}]^T$,

$$\tilde{\mathbf{g}}_i = \frac{\mathbf{g}_i}{\rho_i}, \quad \text{and} \quad \tilde{d}_i = \frac{(M - 1) \|\mathbf{g}_i\|_1}{\rho_i}. \quad (48)$$

Note that using (48) and (47), we have

$$\frac{M^2 - 1}{3} \|\tilde{\mathbf{g}}_i\|_2^2 + (M - 1)^2 \|\tilde{\mathbf{g}}_i\|_1^2 = N_t P_e, \quad \forall i. \quad (49)$$

Thus, $\tilde{\mathbf{G}}$ is individual power constrained (cf. Sec. IV-A).

Now we can write

$$\bar{\mathbb{P}}_{\min}^{\text{PE}} = \min_{\mathbf{G} \in \mathcal{G}_{e,\text{sum},M}} \mathbb{E}_{\mathbf{H}}[\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})] \quad (50)$$

$$= \min_{\mathbf{r}, \tilde{\mathbf{G}}} \mathbb{E}_{\mathbf{H}}[\mathbb{P}^{\text{PE}}(\bar{\mathbf{R}} \tilde{\mathbf{G}}, \mathbf{H})] \quad (51)$$

$$\geq \min_{\mathbf{r}} \mathbb{E}_{\mathbf{H}}[\min_{\tilde{\mathbf{G}}} \mathbb{P}^{\text{PE}}(\bar{\mathbf{R}} \tilde{\mathbf{G}}, \mathbf{H})], \quad (52)$$

where $\mathbf{r} \in \mathbb{R}_+^{N_t \times 1}$ satisfies $\|\mathbf{r}\|_2 \leq 1$ and $\tilde{\mathbf{G}}$ satisfies (49). Using the definition of $\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})$ in (11), we can write this lower bound as

$$\bar{\mathbb{P}}_{\min}^{\text{PE}} \geq \min_{\mathbf{r}} \mathbb{E}_{\mathbf{H}}[\min_{\tilde{\mathbf{G}}} \mathbb{P}^{\text{PE}}(\tilde{\mathbf{G}}, \mathbf{H} \mathbf{R})]. \quad (53)$$

The inner problem corresponds to optimizing an $(M, 1)$ DC-STBC given a channel matrix $\mathbf{H}\mathbf{R}$ and an individual power constraint $N_t P_e$ (cf. Sec. IV-A). Thus, the optimal $\tilde{\mathbf{G}}$ corresponds to RC, i.e., $\tilde{\mathbf{G}} = \gamma\sqrt{N_t}\mathbf{I}_L \otimes \mathbf{1}_{N_t}$ as given in Lemma 1. The corresponding minimum WC-PEP given $\mathbf{H}\mathbf{R}$ is $\mathbb{P}_{\min, e}^{\text{PE}}(\mathbf{H}\mathbf{R}) = Q\left(\frac{\gamma\sqrt{N_t}}{\sigma}\|\mathbf{H}\mathbf{r}\|_2\right)$. Since this quantity is decreasing in $\|\mathbf{r}\|_2$, then,

$$\mathbb{P}_{\min}^{\text{PE}} \geq \min_{\mathbf{r} \in \mathbb{R}_+^{N_t \times 1} : \|\mathbf{r}\|_2=1} \mathbb{E}_{\mathbf{H}} \left[Q\left(\frac{\gamma\sqrt{N_t}}{\sigma}\|\mathbf{H}\mathbf{r}\|_2\right) \right]. \quad (54)$$

This lower bound is achievable by using $\mathbf{G} = \gamma\sqrt{N_t}\mathbf{I}_L \otimes \mathbf{r}^*$, where $\mathbf{r}^*$ is the optimal solution of the minimization in (54). This proves Theorem 2.

C. Proof of Corollary 1

We need to find the optimal solution of the optimization in (54) when $\mathbf{H}$ has i.i.d. columns. First we rewrite this optimization as

$$\min_{\substack{\mathbf{u} \in \mathbb{R}_+^{N_t \times 1} \\ \|\mathbf{u}\|_1=1}} \mathbb{E}_{\mathbf{H}} \left[Q\left(\frac{\gamma\sqrt{N_t}}{\sigma} \sqrt{\sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_t} h_{ji}\sqrt{u_i}\right)^2}\right) \right], \quad (55)$$

where $\mathbf{u} = [u_1, \dots, u_{N_t}] = [r_1^2, \dots, r_{N_t}^2]$. Since $\sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_t} h_{ji}\sqrt{u_i}\right)^2$ is concave in $\mathbf{u}$ (recall that $h_{ji} \geq 0$), $\sqrt{x}$ is concave and nondecreasing in $x \geq 0$, and $Q(x)$ is convex and nonincreasing in $x \geq 0$, then the objective of (55) is convex in $\mathbf{u}$ by the composition rules [25]. Moreover, this objective function is symmetric in u_i when $\mathbf{H}$ has i.i.d. columns. Therefore, the objective is Schur-convex [24, Proposition 2.8], and is minimized by $\mathbf{u} = \frac{1}{N_t}\mathbf{1}_{N_t}$ since $\frac{1}{N_t}\mathbf{1}_{N_t}$ is majorized by any $\mathbf{u}$ satisfying $\|\mathbf{u}\|_1 = 1$ (see [24, Theorem 2.21]). This solution corresponds to $\mathbf{r} = \frac{1}{\sqrt{N_t}}\mathbf{1}_{N_t}$ which proves Corollary 1.

V. MIMO CHANNEL WITH OPTICAL CONSTRAINTS

In this section, we prove the results related to optical power constraints. We start with Lemma 2.

A. Individual Constraints given $\mathbf{H}$ (Proof of Lemma 2)

In the case of a sum optical power constraint, we need to solve the following problem

$$\Theta_o \triangleq \max_{\mathbf{G} \in \mathcal{G}_{o,M}} \min_{\mathbf{x}_a, \mathbf{x}_b \in \mathcal{X}_M} \|\mathbf{A}(\mathbf{x}_a - \mathbf{x}_b)\|_2^2, \quad (56)$$

where

$$\mathcal{G}_{o,M} = \left\{ \mathbf{G} \in \mathbb{R}^{N_t \times L} \mid (M-1)\|\mathbf{g}_i\|_1 \leq P_o, \forall i \in \{1, \dots, N_t L\} \right\}. \quad (57)$$

After applying similar steps as (30)–(34), Θ_o can be upper bounded by

$$\Theta_o \leq \max_{\mathbf{G} \in \mathcal{G}_{o,M}} \frac{4}{L} \|\mathbf{A}\|_F^2 \quad (58)$$

$$\leq \frac{4}{L} \sum_{k=1}^L \sum_{j=1}^{N_r} \max_{\mathbf{F}(k) \in \mathcal{F}_{o,M}} \|\mathbf{h}_j^T \mathbf{F}(k)\|_2^2, \quad (59)$$

where $[\mathbf{F}(1)^T, \dots, \mathbf{F}(L)^T]^T = \mathbf{G}$, $\mathbf{h}_j^T$ is the j^{th} row of $\mathbf{H}$,

$$\mathcal{F}_{o,M} = \left\{ \mathbf{F} \in \mathbb{R}^{N_t \times L} \mid (M-1) \sum_{j=1}^L f_{ij} \leq P_o \right\}. \quad (60)$$

Similar to Sec. IV-A, the maximum of this upper bound is achieved when only one column of $\mathbf{F}(k)$ is nonzero. This can be restricted to be positive component-wise, leading to

$$\Theta_o \leq \frac{4}{L} \sum_{k=1}^L \sum_{j=1}^{N_r} \max_{\substack{\mathbf{f}_1(k) \in \mathbb{R}_+^{N_t \times 1} \\ (M-1)f_{i1}(k) \leq P_o \quad \forall i}} (\mathbf{h}_j^T \mathbf{f}_1(k))^2. \quad (61)$$

The objective is maximized when the constraints are met with equality, because it is increasing in $f_{i1}(k)$. Thus, we can write $\mathbf{f}_1(k)$ as $\frac{P_o}{(M-1)}\mathbf{1}_{N_t}$. Hence,

$$\Theta_o \leq \frac{4P_o^2}{(M-1)^2 L} \sum_{k=1}^L \sum_{j=1}^{N_r} (\mathbf{h}_j^T \mathbf{1}_{N_t})^2 \quad (62)$$

$$= \frac{4P_o^2}{(M-1)^2} \|\mathbf{H}\mathbf{1}_{N_t}\|_2^2. \quad (63)$$

This upper bound is achievable by setting $\mathbf{G} = \frac{P_o}{M-1}\mathbf{I}_L \otimes \mathbf{1}_{N_t}$ which proves Lemma 2.

B. Quasi-Static Channels with a Sum Constraint (Proof of Theorem 4)

We need to solve

$$\mathbb{P}_{\min}^{\text{PE}} = \min_{\mathbf{G} \in \mathcal{G}_{\text{osum},M}} \mathbb{E}_{\mathbf{H}}[\mathbb{P}^{\text{PE}}(\mathbf{G}, \mathbf{H})], \quad (64)$$

where

$$\mathcal{G}_{\text{osum},M} = \left\{ \mathbf{G} \in \mathbb{R}^{N_t \times L} \mid (M-1) \sum_{i=(k-1)N_t+1}^{kN_t} \|\mathbf{g}_i\|_1 \leq N_t P_o, \right. \\ \left. \forall k \in \{1, \dots, L\} \right\}. \quad (65)$$

Defining $\mathbf{r}_k = [\rho_{(k-1)N_t+1}, \dots, \rho_{kN_t}]^T$, $\rho_i = \frac{M-1}{N_t P_o} \|\mathbf{g}_i\|_1$, and $\tilde{\mathbf{G}} = [\tilde{\mathbf{g}}_1, \dots, \tilde{\mathbf{g}}_{N_t L}]^T$ where $\tilde{\mathbf{g}}_i = \frac{\mathbf{g}_i}{\rho_i}$, and proceeding similar to Sec. IV-B, we can derive the lower bound

$$\mathbb{P}_{\min}^{\text{PE}} \geq \min_{\mathbf{r} : \|\mathbf{r}\|_1=1} \mathbb{E}_{\mathbf{H}} \left[Q\left(\frac{N_t P_o}{(M-1)\sigma} \|\mathbf{H}\mathbf{r}\|_2\right) \right]. \quad (66)$$

This lower bound is achievable by using $\mathbf{G} = \frac{N_t P_o}{M-1}\mathbf{I}_L \otimes \mathbf{r}^*$, where $\mathbf{r}^*$ is the optimal solution of the minimization in (66). This proves Theorem 4.

C. Proof of Corollary 2

We shall show that, as $\frac{P_o}{\sigma} \rightarrow 0$, the minimization in (66) is achieved by $\mathbf{r} = \mathbf{e}_q$ where $q = \arg \max_i \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_i\|_2]$,⁸ $\mathbf{e}_i$ is the i^{th} column of $\mathbf{I}_{N_t}$, and $\tilde{\mathbf{h}}_i$ is the i^{th} column of $\mathbf{H}$.

Let $a = \frac{N_t P_o}{(M-1)\sigma}$, and note that we have the expansion $Q(x) = \frac{1}{2} - \frac{1}{\sqrt{2\pi}}x + O(x^3)$ (cf. [26, p. 297]), where $O(x^3)$

⁸Ties are resolved randomly, e.g. using a coin flip.

stands for a function $e(x)$ satisfying $|e(x)| < M|x^3|$ for $0 < |x| < \delta$, for some positive δ and M . Then,

$$\mathbb{E}_{\mathbf{H}} [Q(a\|\mathbf{H}\mathbf{r}\|_2)] = \frac{1}{2} - \frac{a}{\sqrt{2\pi}} \mathbb{E}_{\mathbf{H}} [\|\mathbf{H}\mathbf{r}\|_2] + O(a^3). \quad (67)$$

Note that $\mathbb{E}_{\mathbf{H}} [\|\mathbf{H}\mathbf{r}\|_2]$ can be written as $\mathbb{E}_{\mathbf{H}} \left[\left\| \mathbf{H} \sum_{i=1}^{N_t} r_i \mathbf{e}_i \right\|_2 \right]$. Since the ℓ_2 norm is convex, then we have

$$\mathbb{E}_{\mathbf{H}} \left[\left\| \mathbf{H} \sum_{i=1}^{N_t} r_i \mathbf{e}_i \right\|_2 \right] \leq \sum_{i=1}^{N_t} r_i \mathbb{E}_{\mathbf{H}} [\|\mathbf{H}\mathbf{e}_i\|_2] \quad (68)$$

$$= \sum_{i=1}^{N_t} r_i \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_i\|_2] \quad (69)$$

$$\leq \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_q\|_2], \quad (70)$$

where q is as defined above and where we used $\sum_{i=1}^{N_t} r_i = 1$. Since this applies for any $\mathbf{r}$, then

$$\min_{\mathbf{r}: \|\mathbf{r}\|_1=1} \mathbb{E}_{\mathbf{H}} [Q(a\|\mathbf{H}\mathbf{r}\|_2)] \geq \frac{1}{2} - \frac{a}{\sqrt{2\pi}} \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_q\|_2] + O(a^3).$$

and we have that

$$\lim_{a \rightarrow 0} \left(\min_{\mathbf{r}: \|\mathbf{r}\|_1=1} \mathbb{E}_{\mathbf{H}} [Q(a\|\mathbf{H}\mathbf{r}\|_2)] - \frac{1}{2} + \frac{a}{\sqrt{2\pi}} \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_q\|_2] \right) \geq 0. \quad (71)$$

On the other hand, choosing $\mathbf{r} = \mathbf{e}_q$ leads to

$$\begin{aligned} \min_{\mathbf{r}: \|\mathbf{r}\|_1=1} \mathbb{E}_{\mathbf{H}} [Q(a\|\mathbf{H}\mathbf{r}\|_2)] &\leq \mathbb{E}_{\mathbf{H}} [Q(a\|\tilde{\mathbf{h}}_q\|_2)] \\ &= \frac{1}{2} - \frac{a}{\sqrt{2\pi}} \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_q\|_2] + O(a^3). \end{aligned} \quad (72)$$

This leads to

$$\lim_{a \rightarrow 0} \left(\min_{\mathbf{r}: \|\mathbf{r}\|_1=1} \mathbb{E}_{\mathbf{H}} [Q(a\|\mathbf{H}\mathbf{r}\|_2)] - \frac{1}{2} + \frac{a}{\sqrt{2\pi}} \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_q\|_2] \right) \leq 0 \quad (73)$$

Therefore,

$$\lim_{a \rightarrow 0} \left(\min_{\mathbf{r}: \|\mathbf{r}\|_1=1} \mathbb{E}_{\mathbf{H}} [Q(a\|\mathbf{H}\mathbf{r}\|_2)] - \frac{1}{2} + \frac{a}{\sqrt{2\pi}} \mathbb{E}_{\mathbf{H}} [\|\tilde{\mathbf{h}}_q\|_2] \right) = 0. \quad (74)$$

By noting that $a \rightarrow 0$ is equivalent to $\frac{P_o}{\sigma} \rightarrow 0$, the result is proved.

D. Proof of Corollary 3

We show that at high SNR, the objective of the optimization in (66) is minimized by $\mathbf{r} = \frac{1}{N_t} \mathbf{1}_{N_t}$ when $\mathbf{H}$ has i.i.d. columns. To this end, we show that

$$\lim_{a \rightarrow \infty} \frac{\min_{\mathbf{r}} \mathbb{E}[Q(a\|\mathbf{H}\mathbf{r}\|_2)]}{\mathbb{E} \left[\omega_{a,\mathbf{H}} \left(\frac{1}{N_t} \mathbf{1}_{N_t} \right) \right]} = 1. \quad (75)$$

where $a = \frac{N_t P_o}{(M-1)\sigma}$ and where $\omega_{a,\mathbf{H}}(\mathbf{r}) = \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{a^2 \|\mathbf{H}\mathbf{r}\|_2^2}{2}}}{a\|\mathbf{H}\mathbf{r}\|_2}$.

To this end, we start by using the lower bound $Q(x) \geq \frac{x}{1+x^2} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ for $x > 0$. Thus,

$$Q(a\|\mathbf{H}\mathbf{r}\|_2) \geq \omega_{a,\mathbf{H}}(\mathbf{r})(1+o(1)), \quad (76)$$

where $o(1)$ represents a function $e(a)$ that satisfies $\lim_{a \rightarrow \infty} e(a) = 0$. Next, we show that $\omega_{a,\mathbf{H}}(\mathbf{r})$ is convex in $\mathbf{r}$ for large a , by showing that its Hessian matrix is positive semidefinite in the limit as $a \rightarrow \infty$. Let $\eta_i = \sum_{k=1}^{N_r} \sum_{\ell=1}^{N_t} h_{ki} h_{k\ell} r_{\ell}$. Using simple analysis, one can show that

$$\begin{aligned} \frac{\partial^2 \omega_{a,\mathbf{H}}(\mathbf{r})}{\partial r_i^2} &= \omega_{a,\mathbf{H}}(\mathbf{r}) \left(\left(a^4 + \frac{2a^2}{\|\mathbf{H}\mathbf{r}\|_2^2} + \frac{3}{\|\mathbf{H}\mathbf{r}\|_2^4} \right) \eta_i^2 \right. \\ &\quad \left. - \left(\frac{1}{\|\mathbf{H}\mathbf{r}\|_2^2} + a^2 \right) \sum_{k=1}^{N_r} h_{ki}^2 \right) \\ &= a^3 \omega_{a,\mathbf{H}}(\mathbf{r}) \eta_i^2 (1+o(1)), \end{aligned} \quad (77)$$

and

$$\begin{aligned} \frac{\partial^2 \omega_{a,\mathbf{H}}(\mathbf{r})}{\partial r_i \partial r_j} &= \omega_{a,\mathbf{H}}(\mathbf{r}) \left(\left(a^4 + \frac{2a^2}{\|\mathbf{H}\mathbf{r}\|_2^2} + \frac{3}{\|\mathbf{H}\mathbf{r}\|_2^4} \right) \eta_i \eta_j \right. \\ &\quad \left. - \left(\frac{1}{\|\mathbf{H}\mathbf{r}\|_2^2} + a^2 \right) \sum_{k=1}^{N_r} h_{ki} h_{kj} \right) \\ &= a^3 \omega_{a,\mathbf{H}}(\mathbf{r}) \eta_i \eta_j (1+o(1)). \end{aligned} \quad (78)$$

This leads to

$$\mathcal{H}(\omega_{a,\mathbf{H}}(\mathbf{r})) = a^3 \omega_{a,\mathbf{H}}(\mathbf{r}) \boldsymbol{\eta}^T \boldsymbol{\eta} (1+o(1)), \quad (79)$$

where $\mathcal{H}(\cdot)$ denotes the Hessian matrix, and $\boldsymbol{\eta} = (\eta_1, \dots, \eta_{N_t})^T$. Since $\boldsymbol{\eta}^T \boldsymbol{\eta}$ is positive semidefinite and $a^3 \omega_{a,\mathbf{H}}(\mathbf{r})$ is positive, then $\mathcal{H}(\omega_{a,\mathbf{H}}(\mathbf{r}))$ is positive semidefinite as $a \rightarrow \infty$. Hence $\omega_{a,\mathbf{H}}(\mathbf{r})$ is convex in $\mathbf{r}$ for a large enough.

The convexity of $\omega_{a,\mathbf{H}}(\mathbf{r})$ at large a , combined with the symmetry of $\mathbb{E}[\omega_{a,\mathbf{H}}(\mathbf{r})]$ in $\mathbf{r}$ when $\mathbf{H}$ has i.i.d. columns, implies that $\mathbb{E}[\omega_{a,\mathbf{H}}(\mathbf{r})]$ is Schur-convex for a large enough. Hence, $\mathbb{E}[\omega_{a,\mathbf{H}}(\mathbf{r})]$ is minimized by $\mathbf{r} = \frac{1}{N_t} \mathbf{1}_{N_t}$ in this regime since $\frac{1}{N_t} \mathbf{1}_{N_t}$ is majorized by any $\mathbf{r}$ satisfying $\|\mathbf{r}\|_1 = 1$ [24, Theorem 2.21]. Consequently, using (76), we can write

$$\min_{\mathbf{r}} \mathbb{E}[Q(a\|\mathbf{H}\mathbf{r}\|_2)] \geq \mathbb{E} \left[\omega_{a,\mathbf{H}} \left(\frac{1}{N_t} \mathbf{1}_{N_t} \right) \right] (1+o(1)), \quad (80)$$

which implies that

$$\lim_{a \rightarrow \infty} \frac{\min_{\mathbf{r}} \mathbb{E}[Q(a\|\mathbf{H}\mathbf{r}\|_2)]}{\mathbb{E} \left[\omega_{a,\mathbf{H}} \left(\frac{1}{N_t} \mathbf{1}_{N_t} \right) \right]} \geq 1. \quad (81)$$

On the other hand, using RC, i.e., $\mathbf{r} = \frac{1}{N_t} \mathbf{1}_{N_t}$, and since $Q(x) \leq \frac{1}{x} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ for $x > 0$, we achieve an average WC-PEP $\mathbb{E} \left[Q \left(a \left\| \mathbf{H} \frac{1}{N_t} \mathbf{1}_{N_t} \right\|_2 \right) \right]$ which satisfies

$$\min_{\mathbf{r}} \mathbb{E} [Q(a\|\mathbf{H}\mathbf{r}\|_2)] \leq \mathbb{E} \left[Q \left(a \left\| \mathbf{H} \frac{1}{N_t} \mathbf{1}_{N_t} \right\|_2 \right) \right] \quad (82)$$

$$\leq \mathbb{E} \left[\omega_{a,\mathbf{H}} \left(\frac{1}{N_t} \mathbf{1}_{N_t} \right) \right]. \quad (83)$$

(b) $\mathbf{H} = [\mathbf{h}_a^T \ \mathbf{h}_b^T]^T$, i.e., $(N_t, N_r) = (2, 2)$.

(b) $(N_t, N_r) = (2, 2)$.

(a) $(N_t, N_r) = (2, 1)$.

(b) $(N_t, N_r) = (2, 2)$.

Fig. 5. Minimum average WC-PEP and average error probability of beamforming and RC for a system with a sum electrical power constraint $N_t P_e$, $L = 2$, and independent but not identically distributed log-normal channels.

(a) $(N_t, N_r) = (2, 1)$.

(b) $(N_t, N_r) = (2, 2)$.

Fig. 6. Minimum average WC-PEP and average error probability of beamforming and RC for a system with a sum electrical power constraint $N_t P_e$, $L = 2$, and i.i.d. log-normally distributed channels.

set $h_{ij} = h_{ij}^{[d]} h_{ij}^{[r]}$ where $h_{ij}^{[d]}$ is a deterministic quantity (path loss e.g.) and $h_{ij}^{[r]}$ is a random variable (turbulence induced fading e.g.). We normalize the system so that $h_{ij}^{[d]} = 1 \forall i, j$, motivated by a scenario where the transmitter-receiver distance is much larger than the transmit and receiver aperture separation. We choose $h_{ij}^{[r]}$, $i = 1, \dots, N_r$, $j = 1, \dots, N_t$, to be i.i.d. log-normally distributed [27], [28]

$$f(h_{ij}^{[r]}) = \frac{1}{h_{ij}^{[r]} \sqrt{2\pi\sigma_R^2}} \exp\left(-\frac{(\ln(h_{ij}^{[r]}) + \frac{\sigma_R^2}{2})^2}{2\sigma_R^2}\right), \quad (85)$$

where σ_R^2 is the Rytov variance, assumed here equal to $\frac{1}{4}$.⁹ We can observe that the minimum average WC-PEP obtained from our analysis (Theorem 1) coincides with the numerically optimized average WC-PEP where $\mathbf{G}$ is optimized using grid

⁹The independence of $h_{ij}^{[r]}$ holds if the transmit and receive apertures are sufficiently spaced [29].

search. Moreover, at high SNR, the error probability of RC matches the minimum average WC-PEP for $M = 2$, and is $\frac{3}{2}$ times the minimum average WC-PEP for $M = 4$, which coincides with the approximated minimum average error probability. This demonstrates the importance of the minimum average WC-PEP defined in (15) as an approximation of the average error probability.

Fig. 5 shows a similar comparison under a sum electrical power constraint $N_t P_e$, and an independent but not identically distributed log-normal channel. In particular, we choose $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2]$ where $\mathbf{h}_2 = \mathbf{h}_1/5$ and the components of $\mathbf{h}_1$ and $\mathbf{h}_2$ are i.i.d. distributed according to (85) with $\sigma_R^2 = \frac{1}{4}$. The figure shows the analytical minimum average WC-PEP (Theorem 2), the numerical one (grid search), the approximated error probability, and the Monte Carlo simulated error probability of spatial beamforming (BF) where $\mathbf{G}$ is chosen as in Theorem 2 with $\mathbf{r}^*$ obtained numerically for the given channel distribution. The same observations we saw in

Fig. 7. Average WC-PEP versus r_1 for a system with a sum optical power constraint, $N_t = N_r = 2$, and i.i.d. log-normally distributed channels.

Fig. 8. Minimum average WC-PEP and average error probability for a system with a sum optical power constraint $N_t P_o$, $L = 2$, and i.i.d. log-normally distributed channels.

Fig. 3 and Fig. 4 can be observed here. We have also marked two points with their respective optimal $\mathbf{r}^*$, which indicate, as intuition suggests, that an M -PAM symbol is sent at higher power from the first aperture than the second since the second channel is weaker (by a factor of $1/5$). In the same figure, we plot the Monte Carlo simulated error probability of RC, where $\mathbf{G}$ is chosen as in Theorem 1, which is clearly worse than the optimal beamforming especially at high SNR.

The gap between RC and beamforming in Fig. 5 shrinks as the columns of $\mathbf{H}$ become more ‘similar’ in terms of distribution. In the extreme, where the columns are i.i.d., the two coincide. This is shown in Fig. 6 which confirms Corollary 1.

These numerical evaluations confirm the results for the channel subject to electrical power constraints. Similar comparisons can be performed to confirm Lemma 2, Theorem 3, and Theorem 4, for the channel subject to optical power constraints. To support Corollaries 2 and 3 numerically, we plot the average WC-PEP $\mathbb{E}_{\mathbf{H}} \left[Q \left(\frac{N_t P_o}{(M-1)\sigma} \|\mathbf{H}\mathbf{r}\|_2 \right) \right]$ for i.i.d. log-normal channels as described above with $N_t = N_r = 2$. Since the vector $\mathbf{r}$ can be written as $[r_1, 1 - r_1]^T$ ($\|\mathbf{r}\|_1 = 1$), we plot the average WC-PEP versus r_1 in order to reflect the optimal $\mathbf{r}$. Note that at low SNR, the optimal r_1 is 0 or 1, i.e., the transmitter sends only from one aperture. This confirms Corollary 2. As SNR increases, the optimal r_1 becomes $\frac{1}{2}$,

i.e., $\mathbf{r} = [\frac{1}{2}, \frac{1}{2}]^T$ (RC) is optimal which confirms Corollary 3. Although the difference between choosing $r_1 \in \{0, 1\}$ or $r_1 = \frac{1}{2}$ (RC) is not major ($< 0.6\%$) at low SNR, and although this low SNR is not a practical regime of operation due to the high error rate, this subtlety prevents proving the universal optimality of RC under a sum optical power constraint. However, it is important to note that RC still provides good performance overall, and is optimal at high SNR for channels with i.i.d. columns (Corollary 3) as demonstrated in Fig. 8. The inset is included to demonstrate that RC is not optimal at low SNR.

VII. CONCLUSION

In this paper, we studied MIMO IM/DD OWC employing DC-offset STBCs with M -PAM and with unit rate in a quasi-static environment. The average worst-case pairwise error probability was considered as the main criterion for performance evaluation. We analyzed the problem of minimizing this criterion under electrical and optical power constraints. We showed that spatial repetition coding is optimal from this perspective, among all DC-offset STBCs, for a system with individual power constraints. Then, we showed that under sum power constraints, spatial beamforming is optimal in terms of this criterion in general. If the channel has independent and identically distributed columns, then this beamforming reduces

to spatial repetition under a sum electrical power constraint for any SNR, and under a sum optical power constraints at high SNR.

This result implies that the temporal dimension is redundant in DC-offset STBCs under these considerations, and hence, it is enough to code spatially. This result is important since space coding is much simpler in practice than more general STBCs, thus reducing the system's computational complexity.

While the average worst-case PEP is analytically more tractable than the *worst-case average PEP*, deriving the optimal DC-STBC structure from a worst-case average PEP point-of-view remains an interesting open problem for future investigation. Another interesting question is whether the same can be proved for the family of trellis codes instead of DC-STBCs.

REFERENCES

- [1] Y. Sopenov, A. Chaaban, Z. Rezki, and M.-S. Alouini, "Optical MISO IM/DD channels: Optimality of spatial repetition codes among DC-offset STBCs," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 1355–1359.
- [2] M. A. Khalighi and M. Uysal, "Survey on free space optical communications: A communication theory perspective," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 2231–2258, 4th quarter 2014.
- [3] D. Karunatilaka, F. Zafar, V. Kalavally, and R. Parthiban, "LED based indoor visible light communications: State of the art," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1649–1678, 3rd quarter 2015.
- [4] H. Haas, L. Yin, Y. Wang, and C. Chen, "What is LiFi?" *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1533–1544, Mar. 2016.
- [5] H. Elgala, R. Mesleh, and H. Haas, "Indoor optical wireless communication: Potential and state-of-the-art," *IEEE Comm. Mag.*, vol. 49, no. 9, pp. 56–62, Sep. 2011.
- [6] R. Jiang, Z. Wang, Q. Wang, and L. Dai, "A tight upper bound on channel capacity for visible light communications," *IEEE Commun. Lett.*, vol. 20, no. 1, pp. 97–100, Jan. 2016.
- [7] A. Chaaban, Z. Rezki, 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.
- [8] —, "Fundamental limits of parallel optical wireless channels: Capacity results and outage formulation," *IEEE Trans. Commun.*, vol. 65, no. 1, pp. 296–311, Jan. 2017.
- [9] N. Letzepis and A. G. i Fabregas, "Outage probability of the Free-Space Optical channel with doubly stochastic scintillation," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 2899–2902, Oct. 2009.
- [10] T. Tsiftsis, H. G. Sandalidis, G. K. Karagiannidis, and M. Uysal, "Optical wireless links with spatial diversity over strong atmospheric turbulence channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 951–957, Feb. 2009.
- [11] M. Safari and M. Uysal, "Do we really need OSTBCs for free-space optical communication with direct detection?" *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4445–4448, Nov. 2008.
- [12] Y. Y. Zhang, H. Y. Yu, J. K. Zhang, Y. J. Zhu, J. L. Wang, and X. S. Ji, "On the optimality of spatial repetition coding for MIMO optical wireless communications," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 846–849, May. 2016.
- [13] Y. Y. Zhang, H. Y. Yu, J. K. Zhang, Y. J. Zhu, J. L. Wang, and T. Wang, "Full large-scale diversity space codes for MIMO optical wireless communications," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, June 2015, pp. 1671–1675.
- [14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.
- [15] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.
- [16] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.
- [17] E. Bayaki and R. Schober, "On space-time coding for free-space optical systems," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 58–62, Jan. 2010.
- [18] M. Salehi and J. Proakis, *Digital Communications*. McGraw-Hill Education, 2007.
- [19] J. R. Barry, D. G. Messerschmitt, and E. A. Lee, *Digital Communication: Third Edition*. Norwell, MA, USA: Kluwer Academic Publishers, 2003.
- [20] Y.-Y. Zhang, H.-Y. Yu, J.-K. Zhang, Y.-J. Zhu, and T. Wang, "Energy-efficient space-time modulation for indoor miso visible light communications," *Opt. Lett.*, vol. 41, no. 2, pp. 329–332, Jan. 2016.
- [21] S. Hranilovic and F. R. Kschischang, "Optical intensity-modulated direct detection channels: Signal space and lattice codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 6, pp. 1385–1399, June 2003.
- [22] J. Armstrong and B. J. C. Schmidt, "Comparison of asymmetrically clipped optical OFDM and DC-Biased optical OFDM in AWGN," *IEEE Commun. Lett.*, vol. 12, no. 5, pp. 343–345, May 2008.
- [23] G. L. Stüber, *Principles of Mobile Communication*, 3rd ed. Springer Publishing Company, Incorporated, 2011.
- [24] E. Jorswieck and H. Boche, "Majorization and Matrix-Monotone Functions in Wireless Communications," *Found. Trend Commun. Inf. Theory*, vol. 3, no. 6, pp. 553–701, 2007.
- [25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [26] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. Dover Publications, 1964.
- [27] A. A. Farid and S. Hranilovic, "Outage capacity optimization for free-space optical links with pointing errors," *J. Lightw. Technol.*, vol. 25, no. 7, pp. 1702–1710, July 2007.
- [28] L. Yang, J. Cheng, and J. F. Holzman, "Optical communications over lognormal fading channels using OOK," in *Proc. 2nd Int. Workshop on Opt. Wireless Commun. (IWOW)*, Oct. 2013, pp. 99–103.
- [29] S. Arnon, J. Barry, G. Karagiannidis, R. Schober, and M. Uysal, Eds., *Advanced Optical Wireless Communication Systems*. Cambridge University Press, 2012.

Anas Chaaban (S'09 - M'14 - SM'17) received the *Maîtrise ès Sciences* degree in electronics from Lebanese University, Lebanon, in 2006, the M.Sc. degree in communications technology and the Dr. Ing. (Ph.D.) degree in electrical engineering and information technology from the University of Ulm and the Ruhr-University of Bochum, Germany, in 2009 and 2013, respectively. From 2008 to 2009, he was with the Daimler AG Research Group On Machine Vision, Ulm, Germany. He was a Research Assistant with the Emmy-Noether Research Group on Wireless Networks, University of Ulm, Germany, from 2009 to 2011, which relocated to the Ruhr-University of Bochum in 2011. He was a PostDoctoral Researcher with the Ruhr-University of Bochum from 2013 to 2014, and with King Abdullah University of Science and Technology from 2015 to 2017. He joined the School of Engineering at the University of British Columbia as an Assistant Professor in 2018. His research interests are in the areas of information theory and wireless communications.

Yerzhan Sopenov Yerzhan was born in Kazakhstan. He received his bachelor's degree in Electrical and Electronic Engineering from Nazarbayev University, Astana, Kazakhstan, in 2015 and the M.Sc. degree in Electrical Engineering from the King Abdullah University of Science and Technology, Saudi Arabia, in 2017. He is currently a Modelling Specialist with Samruk-Energy, Astana, Kazakhstan.

Department.

Zouheir Rezki (S'01-M'08-SM'13) was born in Casablanca, Morocco. He received the Diplôme d'Ingénieur degree from the École Nationale de l'Industrie Minérale (ENIM), Rabat, Morocco, in 1994, the M.Eng. degree from École de Technologie Supérieure, Montreal, Québec, Canada, in 2003, and the Ph.D. degree in electrical engineering from École Polytechnique, Montreal, Québec, in 2008. After a few years of experience as a postdoctoral fellow and a research scientist at KAUST, he joined University of Idaho as an Assistant Professor in the ECE

Mohamed-Slim Alouini (S94-M98-SM03-F09) was born in Tunis, Tunisia. He received the Ph.D. degree in Electrical Engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1998. He served as a faculty member in the University of Minnesota, Minneapolis, MN, USA, then in the Texas A& M University at Qatar, Education City, Doha, Qatar before joining King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia as a Professor of Electrical Engineering in 2009.

His current research interests include the modelling, design, and performance analysis of wireless communication systems.