Signaling for MISO Channels Under First- and Second-Moment Constraints

Shuai Ma*, Stefan M. Moser†‡, Ligong Wang§, and Michèle Wigger*
mashuai001@cumt.edu.cn, moser@isi.ee.ethz.ch, ligong.wang@ensea.fr, michele.wigger@telecom-paris.fr
*LTCI, Telecom Paris, IP Paris, 91120 Palaiseau, France
†Signal and Information Processing Lab, ETH Zurich, Switzerland
‡Institute of Communications Engineering at National Yang Ming Chiao Tung University, Hsinchu, Taiwan
§ETIS—CY Cergy Paris Univeristy, ENSEA, CNRS, Cergy-Pontoise, France

***Abstract*—Consider a multiple-input single-output system, where the nonnegative, peak-limited inputs $X_1, \ldots, X_{n_\mathrm{T}} \in [0, \mathsf{A}]$ are subject to first- and second-moment sum-constraints on all antennas. The paper characterizes all probability distributions that can be induced for the "channel image," which is given by the inner product of the input vector with a given channel vector. Key to this result is the description of input vectors that achieve a given deterministic channel image with the smallest energy, where "energy" of an input vector refers to a weighted sum of its one- and two-norms. Minimum-energy input vectors have an interesting structure: depending on the desired channel image, some of the weakest antennas are silenced, and the remaining antennas are chosen according to a shifted and amplitude-constrained beamforming rule.**

I. Introduction

Consider a multiple-input single-output (MISO) antenna system of the form:

$$\overline{X} = \mathbf{h}^\mathsf{T}\mathbf{X}, \tag{1}$$

where $\mathbf{h} = (h_1, \ldots, h_{n_\mathrm{T}})^\mathsf{T}$ is a constant channel state vector and $\mathbf{X}$ a random n_T-dimensional channel input vector $\mathbf{X} = (X_1, \ldots, X_{n_\mathrm{T}})^\mathsf{T}$.

The channel image $\overline{X}$ in (1) is useful in describing the input-output relation of MISO (optical or radio-frequency) wireless channels, where the channel output is modeled as $Y = \overline{X} + Z$, with Z being additive noise [1]–[4]. In radio-frequency communication, the inputs $X_1, \ldots, X_{n_\mathrm{T}}$ correspond to the modulated electromagnetic field, and as such both the inputs and the channel coefficients can take values over the entire real line $\mathbb{R}$. In this context, battery and power limitations impose a second-moment constraint $\mathsf{E}\left[\|\mathbf{X}\|_2^2\right] \le \mathsf{P}$, and modulation schemes can be restricted without loss of optimality to beamforming input-vectors [4], [5]

$$\mathbf{x}_{\text{beamformer}}(\bar{x}) = \frac{\bar{x}}{\|\mathbf{h}\|_2^2} \cdot \mathbf{h}, \quad \bar{x} \in \mathbb{R}, \tag{2}$$

since $\mathbf{x}_{\text{beamformer}}(\bar{x})$ has the smallest two-norm $\|\mathbf{x}\|_2^2$ among all input vectors $\mathbf{x}$ inducing the same channel image $\bar{x} = \mathbf{h}^\mathsf{T}\mathbf{x}$. In particular, because $\overline{X}$ forms a sufficient statistic for $\mathbf{X}$ with respect to the output $Y = \overline{X} + Z$, the capacity-achieving input distribution for such a channel can be restricted to taking value only in the set of beamforming vectors like (2).

In intensity-modulated direct-detection (IM/DD) optical communication systems, the inputs are directly proportional to the intensity of the emitted light. They therefore cannot be negative: $X_1, \ldots, X_{n_\mathrm{T}} \ge 0$ with probability one. The channel coefficients are also nonnegative in this case. To avoid degeneracy, we henceforth assume that no two coefficients are equal, and, without loss of generality, that they are ordered as

$$h_1 > h_2 > \cdots > h_{n_\mathrm{T}} > 0. \tag{3}$$

Furthermore, optical-power limitations are captured by the first-moment constraints $\mathsf{E}[\|\mathbf{X}\|_1] \le \mathsf{E}$. The input vector with the smallest one-norm inducing a target channel image $\bar{x}$ is

$$\mathbf{x}_{\text{strongest}}(\bar{x}) = \left(\frac{\bar{x}}{h_1}, 0, \ldots, 0\right)^\mathsf{T}, \quad \bar{x} \in [0, \infty). \tag{4}$$

Therefore, modulation systems for first-moment constrained MISO additive noise channels as well as the capacity-achieving input distributions of these channels can be restricted in such a way that only the first input antenna is used.

Safety considerations and technical limitations often impose strict peak constraints $\mathsf{A} > 0$ on the inputs:

$$X_1, \ldots, X_{n_\mathrm{T}} \in [0, \mathsf{A}] \quad \text{with probability one}, \tag{5}$$

in which case the random channel image $\overline{X}$ takes value in the interval $\overline{\mathcal{X}} \triangleq [0, h_{\text{sum}}\mathsf{A}]$, for $h_{\text{sum}} \triangleq \sum_{k=1}^{n_\mathrm{T}} h_k$. Under such peak constraints, the input vector $\mathbf{x}$ that induces a target channel image $\bar{x}$ with the smallest one-norm is [6], [7]

$$\mathbf{x}_{\text{peak}}(\bar{x}) = \left(\mathsf{A}, \ldots, \mathsf{A}, \frac{\bar{x} - \sum_{k=1}^{i-1} \mathsf{A}h_k}{h_i}, 0, \ldots, 0\right)^\mathsf{T}, \tag{6}$$

where the single input that is neither 0 nor A is at position $i \in \{1, \ldots, n_\mathrm{T}\}$ if $\bar{x} \in \left(\mathsf{A}\sum_{k=1}^{i-1} h_k, \mathsf{A}\sum_{k=1}^{i} h_k\right)$.

In this paper, we consider the peak constraint (5) together with both first- and second-moment input constraints:

$$\mathsf{E}[\|\mathbf{X}\|_1] \le \alpha_1 \mathsf{A}, \tag{7a}$$

$$\mathsf{E}\left[\|\mathbf{X}\|_2^2\right] \le \alpha_2 \mathsf{A}^2. \tag{7b}$$

Our main contribution is a characterization of the set of all random channel images $\overline{X}$ that can be induced under constraints (5) and (7). A main step in our analysis is determining the "minimum-energy" input vectors $\mathbf{x}_{\min,\lambda}(\bar{x})$ that, among

all input vectors inducing a (deterministic) channel image $\bar{x}$, minimize the weighted sum-norm

$$\lambda \frac{\|\mathbf{x}\|_1}{\mathsf{A}} + (1-\lambda)\frac{\|\mathbf{x}\|_2^2}{\mathsf{A}^2} \tag{8}$$

for some $\lambda \in [0,1]$. For $\lambda = 1$, $\mathbf{x}_{\min,1}(\bar{x}) = \mathbf{x}_{\text{peak}}(\bar{x})$, as determined in [6]. A key intermediate result of the present paper is the minimizer to (8) for $\lambda \in [0,1)$: it sets a number of strongest antennas to the maximum value A, switches off a number of the weakest antennas, and applies "shifted beamforming" on the remaining antennas. The exact solution is given in Lemma 9 ahead.

The motivation for this study is that many IM/DD systems have limitations on both the optical power (first moment) and the power consumed by the electronic circuit (second moment) [8]–[12]. Our characterization of all possible channel images $\overline{X}$ finds application in capacity calculations of additive-noise MISO channels with first- and second-moment constraints: it reduces the capacity problem to a simpler optimization problem over all possible[1] $\overline{X}$. To illustrate this application, we look at a specific example of a MISO channel (Example 13 in Section V), and determine the set of constraint parameters α_1, α_2 for which the first- and and second-moment constraints (7) do not affect the capacity in the high signal-to-noise ratio (SNR) regime. We note that the capacity of additive Gaussian noise MISO channels with peak or first-moment constraint, or both (but without a second-moment constraint), has also been studied using different techniques; see [7], [11], [13]–[24] and references therein.

Our minimum-energy solutions are also useful in the design of energy-efficient modulation and communication systems. To illustrate this, we study another specific channel (Example 11 in Section III), and present the most energy-efficient way to induce an 8-ary pulse-amplitude modulation (PAM) on $\overline{X}$.

The remainder of this paper is arranged as follows. In Section II, we recapitulate our problem formulation and our notation. In Section III, we show that the problem of characterizing all achievable $\overline{X}$ can be simplified to a class of optimization problems for deterministic channel images $\bar{x} \in \overline{\mathcal{X}}$ and some parameter $\lambda \in [0,1]$. In Section IV, we then present the solution to these optimization problems. Section V combines these findings to present an explicit characterization of the set of channel images $\overline{X}$ that are achievable under constraints (5) and (7).

II. Notation and Problem Formulation

A random variable is denoted by a capital Roman letter, e.g., X, while its realization is denoted by the corresponding small Roman letter, e.g., x. Vectors are boldfaced, e.g., $\mathbf{X}$ denotes a random vector and $\mathbf{x}$ its realization. Sets are denoted by calligraphic letters like $\mathcal{I}$. We further use small Greek letters like α and λ, as well as capital Roman letters of a special font like A, to denote certain constants.

[1] A similar technique was already applied in [6], [7] to channels without a second-moment constraint.

We summarize our problem formulation, which has already been introduced in Section I. We consider the MISO system (1) under the peak constraint (5) and the first- and second-moment constraints (7). Given any desired distribution for $\overline{X}$ as in (1), our goal is to find out whether it is achievable under these constraints or not.

III. Pareto-Optimal Inputs

Since we have two power constraints, it is convenient to think of the first and second moments of a random vector as two utility functions, and introduce the notion of Pareto optimality.

Definition 1: A random vector $\mathbf{X}$ is said to be *Pareto optimal* if there exists no other random vector $\mathbf{X}'$ such that $\mathbf{h}^{\mathsf{T}}\mathbf{X}'$ has the same distribution as $\mathbf{h}^{\mathsf{T}}\mathbf{X}$, while

$$\mathsf{E}[\|\mathbf{X}'\|_1] \le \mathsf{E}[\|\mathbf{X}\|_1] \tag{9a}$$

$$\mathsf{E}\big[\|\mathbf{X}'\|_2^2\big] \le \mathsf{E}\big[\|\mathbf{X}\|_2^2\big] \tag{9b}$$

with at least one of the two inequalities in (9) being strict.

Lemma 2 below says that we can restrict our attention to Pareto-optimal input distributions. Lemma 3 characterizes these Pareto-optimal distributions as solutions to a class of optimization problems. The proofs of these two lemmas are elementary and omitted.

Lemma 2: Consider a random channel image $\overline{X}$ with a desired distribution. If there exists an input vectors $\mathbf{X}$ satisfying (7) and inducing $\overline{X}$, then there must exist a Pareto-optimal input vector $\mathbf{X}^*$ satisfying (7) and inducing $\overline{X}$.

Lemma 3: If a random vector $\mathbf{X}^*$ is Pareto optimal, then there exists a $\lambda \in [0,1]$ such that $\mathbf{X}^*$ minimizes

$$\lambda \cdot \frac{\mathsf{E}[\|\mathbf{X}\|_1]}{\mathsf{A}} + (1-\lambda)\cdot \frac{\mathsf{E}\big[\|\mathbf{X}\|_2^2\big]}{\mathsf{A}^2} \tag{10}$$

among all $\mathbf{X}$ for which $\mathbf{h}^{\mathsf{T}}\mathbf{X}$ has the same distribution as $\mathbf{h}^{\mathsf{T}}\mathbf{X}^*$.

The following lemma characterizes the minimum of (10).

Lemma 4: Fix any $\lambda \in [0,1]$. For any $\bar{x} \in \overline{\mathcal{X}}$, denote

$$g(\bar{x},\lambda) \triangleq \min_{\mathbf{x}:\ \mathbf{h}^{\mathsf{T}}\mathbf{x}=\bar{x}} \left\{ \lambda \cdot \frac{\|\mathbf{x}\|_1}{\mathsf{A}} + (1-\lambda)\cdot \frac{\|\mathbf{x}\|_2^2}{\mathsf{A}^2} \right\}. \tag{11}$$

For any target distribution for $\overline{X}$,

$$\min_{\mathbf{X}:\ \mathbf{h}^{\mathsf{T}}\mathbf{X}\sim\overline{X}} \left\{ \lambda \cdot \frac{\mathsf{E}[\|\mathbf{X}\|_1]}{\mathsf{A}} + (1-\lambda)\cdot \frac{\mathsf{E}\big[\|\mathbf{X}\|_2^2\big]}{\mathsf{A}^2} \right\} = \mathsf{E}\big[g(\overline{X},\lambda)\big]. \tag{12}$$

Proof: For any $\mathbf{X}$ such that $\mathbf{h}^{\mathsf{T}}\mathbf{X} \sim \overline{X}$, we have

$$\lambda \cdot \frac{\mathsf{E}[\|\mathbf{X}\|_1]}{\mathsf{A}} + (1-\lambda)\cdot \frac{\mathsf{E}\big[\|\mathbf{X}\|_2^2\big]}{\mathsf{A}^2} = \mathsf{E}\left[\lambda \cdot \frac{\|\mathbf{X}\|_1}{\mathsf{A}} + (1-\lambda)\cdot \frac{\|\mathbf{X}\|_2^2}{\mathsf{A}^2}\right] \tag{13}$$

$$= \mathsf{E}\left[\mathsf{E}\left[\lambda \cdot \frac{\|\mathbf{X}\|_1}{\mathsf{A}} + (1-\lambda)\cdot \frac{\|\mathbf{X}\|_2^2}{\mathsf{A}^2}\,\middle|\,\mathbf{h}^{\mathsf{T}}\mathbf{X} = \overline{X}\right]\right] \tag{14}$$

$$\geq \mathsf{E}\big[g(\overline{X},\lambda)\big]. \tag{15}$$

Equality in the above is achieved by choosing, for every $\bar{x} \in \overline{\mathcal{X}}$, a corresponding input vector $\mathbf{x}$ that achieves $g(\bar{x},\lambda)$. ■

Recall that our goal is to find out whether a desired distribution for $\overline{X}$ is achievable under the constraints (7) or not. Using Lemmas 2, 3, and 4, we see that this task can be simplified as follows. We first determine, for every $\lambda \in [0,1]$ and $\bar{x} \in \overline{\mathcal{X}}$, the input vectors that achieve the minimum in (11). (This can be done without knowing the target distribution for $\overline{X}$.) We refer to such input vectors as *minimum-energy signaling* (with respect to parameter λ). If a target distribution for $\overline{X}$ is achievable under (7), then there must exist some $\lambda \in [0,1]$ such that an $\mathbf{X}$ taking values only on minimum-energy signaling input vectors with respect to parameter λ induces $\overline{X}$ and satisfies (7). In fact, as we shall later see, such $\mathbf{X}$ is uniquely determined by $\overline{X}$ and λ. Hence, instead of considering distributions over $[0,\mathsf{A}]^{n_\mathrm{T}}$, we only need to examine these specific distributions for each λ.

IV. Characterization of Minimum-Energy Signaling

In this section we characterize, for every $\lambda \in [0,1)$ and $\bar{x} \in \overline{\mathcal{X}}$, the solution to the following minimization problem:[2]

$$\min_{\mathbf{x}\in[0,\mathsf{A}]^{n_\mathrm{T}}\colon\, \mathbf{h}^\mathsf{T}\mathbf{x}=\bar{x}} \left\{\lambda\cdot\frac{\|\mathbf{x}\|_1}{\mathsf{A}} + (1-\lambda)\cdot\frac{\|\mathbf{x}\|_2^2}{\mathsf{A}^2}\right\}. \tag{16}$$

Recall that for $\lambda = 1$, the optimization problem in (16) was solved in [6].

Fix a $\lambda \in [0,1)$ and denote

$$\nu \triangleq \frac{\lambda}{1-\lambda}. \tag{17}$$

We introduce some indices and threshold values.

Definition 5: Set $\kappa_0 \triangleq 0$. For each index $\ell \in \{1,\dots,n_\mathrm{T}\}$, define the integer

$$\kappa_\ell \triangleq \max\left\{j \in \{\ell,\dots,n_\mathrm{T}\}\colon \frac{\nu}{2}\left(\frac{h_\ell}{h_j}-1\right) < 1\right\}; \tag{18}$$

define the point

$$t_\ell \triangleq \mathsf{A}\sum_{i=1}^{\ell} h_i + \mathsf{A}\sum_{i=\ell+1}^{\kappa_\ell}\left(\frac{h_i^2}{h_\ell} + \frac{\nu}{2}\left(\frac{h_i^2}{h_\ell}-h_i\right)\right); \tag{19}$$

and finally, for every $k \in \{\kappa_{\ell-1}+1,\dots,\kappa_\ell\}$, define

$$s_k \triangleq \mathsf{A}\sum_{i=1}^{\ell-1} h_i + \mathsf{A}\sum_{i=\ell}^{k-1}\frac{\nu}{2}\left(\frac{h_i^2}{h_k}-h_i\right). \tag{20}$$

Remark 6: Notice that κ_k, t_k, and s_k are all nondecreasing in $k \in \{1,\dots,n_\mathrm{T}\}$. Moreover, if $\kappa_\ell < \kappa_{\ell+1}$, then

$$s_{\kappa_\ell} \leq t_\ell \leq s_{\kappa_\ell+1}, \tag{21}$$

[2]It is easily verified that A simply acts as a multiplicative factor for the optimal input. That means, if for fixed λ the input $\mathbf{x}_{\mathsf{A}_1}$ is a minimizer when $\mathsf{A}=\mathsf{A}_1$, then $\mathbf{x}_{\mathsf{A}_2} = \frac{\mathsf{A}_2}{\mathsf{A}_1}\mathbf{x}_{\mathsf{A}_1}$ is a minimizer for $\mathsf{A}=\mathsf{A}_2$.

because

$$\left(\frac{h_i^2}{h_\ell}+\frac{\nu}{2}\left(\frac{h_i^2}{h_\ell}-h_i\right)\right) - \frac{\nu}{2}\left(\frac{h_i^2}{h_k}-h_i\right) = \frac{h_i^2}{h_\ell}\underbrace{\left(1-\frac{\nu}{2}\left(\frac{h_\ell}{h_k}-1\right)\right)}_{\triangleq\,\Gamma}, \tag{22}$$

where Γ is positive for any $k \leq \kappa_\ell$ and $i \in \{1,\dots,k\}$ and is negative for $k = \kappa_\ell+1$ and $i \in \{1,\dots,k\}$. Thus we have an ordering of the form:

$$\begin{aligned} 0 = s_1 &\leq \cdots \leq s_{\kappa_1} \leq t_1 \\ &\leq s_{\kappa_1+1} \leq \cdots \leq s_{\kappa_2} \leq t_2 \\ &\leq s_{\kappa_2+1} \leq \cdots \leq s_{\kappa_3} \leq t_3 \\ &\leq \cdots \leq s_{n_\mathrm{T}} \leq \cdots \leq t_{n_\mathrm{T}} = h_{\mathrm{sum}}\mathsf{A}. \end{aligned} \tag{23}$$

As will be shown later, for each $k \in \{1,\dots,n_\mathrm{T}\}$, s_k is the threshold on $\bar{x}$ at which the kth antenna should be switched on (i.e., the optimal choice should have $x_k = 0$ for $\bar{x} \leq s_k$ and $x_k > 0$ for $\bar{x} > s_k$); and t_k is the threshold on $\bar{x}$ at which the kth antenna should be set to the maximum value A (i.e., the optimal choice should have $x_k = \mathsf{A}$ for $\bar{x} \geq t_k$). The integer κ_k indicates the number of antennas that should be switched on before the kth antenna is fixed to its maximum value A.

Definition 7: For any $\ell \in \{1,\dots,n_\mathrm{T}\}$ for which $\kappa_\ell > \kappa_{\ell-1}$, and for any $k \in \{\kappa_{\ell-1},\dots,\kappa_\ell\}$ (and $k > 0$), define the subintervals

$$\mathcal{I}_{\ell,k} \triangleq \begin{cases} [t_{\ell-1}, s_{k+1}] & \text{if } k = \kappa_{\ell-1}, \\ [s_k, s_{k+1}] & \text{if } \kappa_{\ell-1} < k < \kappa_\ell, \\ [s_k, t_\ell] & \text{if } k = \kappa_\ell. \end{cases} \tag{24}$$

For any $\ell \in \{2,\dots,n_\mathrm{T}\}$ for which $\kappa_\ell = \kappa_{\ell-1}$, define the subinterval

$$\mathcal{I}_{\ell,\emptyset} \triangleq [t_{\ell-1}, t_\ell]. \tag{25}$$

The next lemma follows immediately from Remark 6.

Lemma 8: The set of intervals

$$\begin{aligned} \mathcal{P} \triangleq\ & \{\mathcal{I}_{\ell,\kappa_{\ell-1}+1},\dots,\mathcal{I}_{\ell,\kappa_\ell}\}_{\{\ell\colon\,\kappa_{\ell-1}<\kappa_\ell\}} \\ & \cup \{\mathcal{I}_{\ell,\emptyset}\}_{\{\ell\colon\,\kappa_{\ell-1}=\kappa_\ell\}} \end{aligned} \tag{26}$$

overlap on a set of Lebesgue measure zero, and their union is $\overline{\mathcal{X}}$.

We can now present the solution to (16).

Lemma 9: Fix $\lambda \in [0,1)$. For any $\bar{x} \in \overline{\mathcal{X}}$, the optimization problem (16) has the following unique[3] solution $\mathbf{x}_{\min,\lambda}(\bar{x}) = (x_{\min,\lambda,1},\dots,x_{\min,\lambda,n_\mathrm{T}})^\mathsf{T}$. If $\bar{x}$ lies in an interval $\mathcal{I}_{\ell,k}$ as defined in (24), then the solution is given by

$$x_{\min,\lambda,i}(\bar{x}) = \mathsf{A}, \qquad i \leq \ell-1, \tag{27a}$$

$$\begin{aligned} x_{\min,\lambda,i}(\bar{x}) = &\left(\bar{x} - \mathsf{A}\sum_{j=1}^{\ell-1} h_j\right)\cdot\frac{h_i}{\sum_{j=\ell}^{k} h_j^2} \\ &+ \frac{\nu\mathsf{A}}{2}\left(\frac{\sum_{j=\ell}^{k} h_j}{\sum_{j=\ell}^{k}\frac{h_j^2}{h_i}} - 1\right), \quad i = \ell,\dots,k, \end{aligned} \tag{27b}$$

[3]Uniqueness relies on the strictness of the inequalities in (3).

$$x_{\min,\lambda,i}(\bar{x}) = 0, \qquad i \geq k+1, \tag{27c}$$

where ν is defined in (17). Furthermore, the one- and two-norms of $\mathbf{x}_{\min,\lambda}(\bar{x})$ are given respectively by

$$\begin{aligned}\|\mathbf{x}_{\min,\lambda}(\bar{x})\|_1 &= m(\lambda,\bar{x}) \\ &\triangleq (\ell-1)\mathsf{A} - (k-\ell+1)\nu\frac{\mathsf{A}}{2} \\ &\quad + \left(\frac{\bar{x} - \mathsf{A}\sum_{j=1}^{\ell-1} h_j + \nu\frac{\mathsf{A}}{2}\cdot\sum_{j=\ell}^{k} h_j}{\sum_{j=\ell}^{k} h_j^2}\right)\cdot\left(\sum_{i=\ell}^{k} h_i\right), \end{aligned} \tag{28}$$

and

$$\begin{aligned}\|\mathbf{x}_{\min,\lambda}(\bar{x})\|_2^2 &= v(\lambda,\bar{x}) \\ &\triangleq (\ell-1)\mathsf{A}^2 + (k-\ell+1)\nu^2\frac{\mathsf{A}^2}{4} \\ &\quad + \frac{\left(\bar{x} - \mathsf{A}\sum_{j=1}^{\ell-1} h_j + \nu\frac{\mathsf{A}}{2}\cdot\sum_{j=\ell}^{k} h_j\right)^2}{\sum_{j=\ell}^{k} h_j^2} \\ &\quad - \left(\frac{\bar{x} - \mathsf{A}\sum_{j=1}^{\ell-1} h_j + \nu\frac{\mathsf{A}}{2}\cdot\sum_{j=\ell}^{k} h_j}{\sum_{j=\ell}^{k} h_j^2}\right)\cdot\left(\sum_{i=\ell}^{k} h_i\right)\cdot\nu\mathsf{A}. \end{aligned} \tag{29}$$

If $\bar{x}$ lies in an interval $\mathcal{I}_{\ell,\emptyset}$ as defined in (25), then $\mathbf{x}_{\min,\lambda}(\bar{x})$ is given as in (27) with k replaced by κ_ℓ; its one- and two-norms are given by $m(\lambda,\bar{x})$ and $v(\lambda,\bar{x})$ as in (28) and (29), again with the parameter k replaced by κ_ℓ.

Proof: Omitted. ∎

Remark 10: Lemma 9 shows that $\mathbf{x}_{\min,\lambda}(\bar{x})$ sets the strongest $\ell-1$ antennas at the maximum value A and switches off the weakest $n_{\mathrm{T}} - k$ antennas; the remaining $k-\ell+1$ antennas are used in a shifted beamforming fashion, in the sense that the first term on the right-hand side of (27b) is the same as in standard beamforming subject to a second-moment constraint [4], while the second term there is a shifting constant that depends on the antenna index i.

To illustrate our minimum-energy solution, we present the following example.

Example 11: Consider a MISO channel with $n_{\mathrm{T}} = 3$ input antennas and channel vector

$$\mathbf{h} = (4,2,1)^{\mathsf{T}}. \tag{30}$$

The goal is to induce an 8-ary PAM on $\bar{X}$, where the eight points are $0, \mathsf{A}, 2\mathsf{A}, \ldots, 7\mathsf{A}$, with smallest weighted sum

$$\frac{\|\mathbf{x}\|_1}{2\mathsf{A}} + \frac{\|\mathbf{x}\|_2^2}{2\mathsf{A}^2}. \tag{31}$$

To determine the minimum-energy input vectors, we first compute the κ-values as defined in (18):

$$\kappa_1 = 2, \qquad \kappa_2 = 3, \qquad \kappa_3 = 3, \tag{32}$$

and the s- and t-values as defined in (20) and (19):

$$t_1 = 4.5\mathsf{A}, \qquad t_2 = 6.25\mathsf{A}, \qquad t_3 = 7\mathsf{A}, \tag{33}$$

and

$$s_1 = 0, \qquad s_2 = 2\mathsf{A}, \qquad s_3 = 5\mathsf{A}. \tag{34}$$

We thus have

$$s_1 < s_2 < t_1 < s_3 < t_2 < t_3. \tag{35}$$

Thus, when the weakest antenna is used for signaling, the strongest antenna is fixed to its maximum value A.

The corresponding intervals from Definition 7 are

$$\begin{aligned}&\mathcal{I}_{1,1} = [s_1, s_2] = [0, 2\mathsf{A}], \qquad \mathcal{I}_{1,2} = [s_2, t_1] = [2\mathsf{A}, 4.5\mathsf{A}], \\ &\mathcal{I}_{2,2} = [t_1, s_3] = [4.5\mathsf{A}, 5\mathsf{A}], \; \mathcal{I}_{2,3} = [s_3, t_2] = [5\mathsf{A}, 6.25\mathsf{A}], \\ &\mathcal{I}_{3,\emptyset} = [t_2, t_3] = [6.25\mathsf{A}, 7\mathsf{A}]. \end{aligned} \tag{36}$$

For the desired 8-ary PAM, the minimum-energy input vector $\mathbf{x}_\eta$ that induces the η-th modulation point

$$\bar{x}_\eta = (\eta-1)\mathsf{A}, \qquad \eta = 1, \ldots, 8, \tag{37}$$

is now given according to Lemma 9 as follows:

$$\begin{aligned}&\mathbf{x}_1 = \begin{pmatrix}0\\0\\0\end{pmatrix}, \qquad \mathbf{x}_2 = \begin{pmatrix}1/4\\0\\0\end{pmatrix}\mathsf{A}, \quad \mathbf{x}_3 = \begin{pmatrix}1/2\\0\\0\end{pmatrix}\mathsf{A}, \\ &\mathbf{x}_4 = \begin{pmatrix}7/10\\1/10\\0\end{pmatrix}\mathsf{A}, \quad \mathbf{x}_5 = \begin{pmatrix}9/10\\1/5\\0\end{pmatrix}\mathsf{A}, \quad \mathbf{x}_6 = \begin{pmatrix}1\\1/2\\0\end{pmatrix}\mathsf{A}, \\ &\mathbf{x}_7 = \begin{pmatrix}1\\9/10\\1/5\end{pmatrix}\mathsf{A}, \quad \mathbf{x}_8 = \begin{pmatrix}1\\1\\1\end{pmatrix}\mathsf{A}. \end{aligned} \tag{38}$$

The average one- and two-norms of the above solution are, respectively,

$$\mathsf{E}[\|\mathbf{X}\|_1] = 1.1562\,\mathsf{A}, \tag{39a}$$

$$\mathsf{E}\left[\|\mathbf{X}\|_2^2\right] = 0.9703\,\mathsf{A}^2. \tag{39b}$$

We can compare these to the one- and two-norms of the solution that minimizes the one-norm, which is given by (6), and the solution that minimizes the two-norm, which is beamforming with a peak constraint:

$$\mathsf{E}[\|\mathbf{X}_{\text{peak}}\|_1] = 1.125\,\mathsf{A}, \tag{40a}$$

$$\mathsf{E}\left[\|\mathbf{X}_{\text{peak}}\|_2^2\right] = 1.0156\,\mathsf{A}^2; \tag{40b}$$

$$\mathsf{E}[\|\mathbf{X}_{\text{beamforming}}\|_1] = 1.275\,\mathsf{A}, \tag{41a}$$

$$\mathsf{E}\left[\|\mathbf{X}_{\text{beamforming}}\|_2^2\right] = 0.9274\,\mathsf{A}^2. \tag{41b}$$

V. Characterization of All Achievable $\overline{X}$

Theorem 12: The target random variable $\overline{X}$ can be generated with a random input vector $\mathbf{X}$ satisfying (7) if, and only if, there exists a value $\lambda \in [0,1]$ such that the following two inequalities are satisfied:

$$\mathsf{E}_{\overline{X}}\left[m\left(\lambda,\overline{X}\right)\right] \leq \alpha_1\mathsf{A}, \tag{42a}$$

$$\mathsf{E}_{\overline{X}}\left[v\left(\lambda,\overline{X}\right)\right] \leq \alpha_2\mathsf{A}^2, \tag{42b}$$

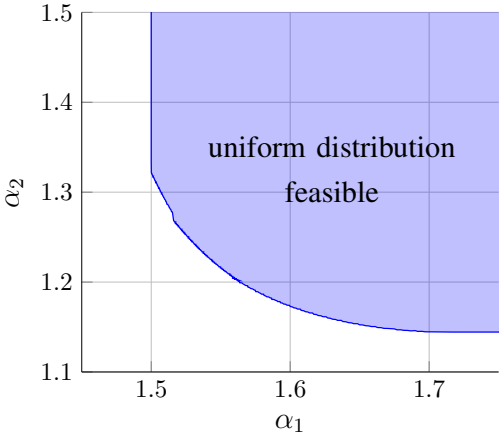

Fig. 1. The figure illustrates the set of all (α_1, α_2)-pairs that allow to induce a uniform distribution over $\overline{X}$ under a MISO channel vector $\mathbf{h} = (4, 3, 2, 1)^{\mathsf{T}}$.

where for $\lambda \in [0, 1)$ the functions $m(\cdot, \cdot)$ and $v(\cdot, \cdot)$ are defined in Lemma 9 and where for $i \in \{1, \ldots, n_{\mathrm{T}}\}$ and

$$\bar{x} \in \left[\mathsf{A} \sum_{k=1}^{i-1} h_k, \mathsf{A} \sum_{k=1}^{i} h_k \right], \tag{43}$$

we define

$$m(1, \bar{x}) \triangleq \sum_{k=1}^{i-1} \mathsf{A} + \frac{\bar{x} - \mathsf{A} \sum_{k=1}^{i-1} h_k}{h_i}, \tag{44}$$

$$v(1, \bar{x}) \triangleq \sum_{k=1}^{i-1} \mathsf{A}^2 + \left(\frac{\bar{x} - \mathsf{A} \sum_{k=1}^{i-1} h_k}{h_i} \right)^2. \tag{45}$$

Proof: The theorem follows immediately from Lemmas 2, 3, 4, and 9. ∎

Example 13: Consider a MISO channel with $n_{\mathrm{T}} = 4$ input antennas and channel vector

$$\mathbf{h} = (4, 3, 2, 1)^{\mathsf{T}}. \tag{46}$$

Figure 1 illustrates the set of (α_1, α_2)-pairs that permit the uniform distribution for $\overline{X} \in \overline{\mathcal{X}}$. The uniform distribution is optimal in the high-signal-to-noise-ratio (high-SNR) limit among all $\overline{X}$ taking values on $\overline{\mathcal{X}}$ without other constraints; see, e.g., [6]. Thus, we have identified the set of (α_1, α_2)-pairs for which the constraints (7) do not limit the capacity in the high-SNR limit. This set depends on the channel vector $\mathbf{h}$ and has a fundamentally different shape from the corresponding set for the single-input single-output channels, which is characterized by $\alpha_1 \geq 1/2$ and $\alpha_2 \geq 1/3$ and is rectangular [11].

Acknowledgments

M. Wigger acknowledges funding support from the European Research Council (ERC) under the European Union's Horizon 2020 program, grant agreement number 715111.

References

[1] Z. Ghassemlooy, L. N. Alves, S. Zvánovec, and M.-A. Khalighi (Eds.). "Visible Light Communications: Theory and Applications (1st ed.)," CRC Press, 2016.

[2] S. Arnon, J. Barry, G.K. Karagiannidis, R. Schober and M. Uysal (Eds.), "Advanced Optical Wireless Communication Systems," Cambridge University Press, 2012.

[3] E. Biglieri, A. R. Calderbank, A. Constantinides, A.J. Goldsmith, A. Paulraj, H. V. Poor, "MIMO Wireless Communications," Cambridge University Press, 2007.

[4] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.

[5] P. Algoet and J. Cioffi, "The capacity of a channel with Gaussian noise and intersymbol interference," in *Proc. IEEE Int. Symp. Inf. Theory*, Budapest, Hungary, June 1991, pp. 1–6.

[6] S. M. Moser, L. Wang, and M. Wigger, "Capacity results on multiple-input single-output wireless optical channels," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 6954–6966, Nov. 2018.

[7] L. Li, S. M. Moser, L. Wang, and M. Wigger, "On the capacity of MIMO optical wireless channels," *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5660–5682, Sept. 2020.

[8] X. Ling, J. Wang, X. Liang, Z. Ding, and C. Zhao, "Offset and power optimization for DCO-OFDM in visible light communication systems," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 349–363, Jan. 2016.

[9] X. Huang, J. Shi, J. Li, Y. Wang, and N. Chi, "A Gb/s VLC transmission using hardware preequalization circuit," *IEEE Photon. Technol. Lett.*, vol. 27, no. 18, pp. 1915–1918, Sep. 2015.

[10] F. Che, L. Wu, B. Hussain, X. Li, and C. P. Yue, "A fully integrated IEEE 802.15.7 visible light communication transmitter with on-chip 8-W 85% efficiency boost LED driver," *J. Lightw. Technol.*, vol. 34, no. 10, pp. 2419–2430, May, 2016.

[11] S. Ma and M. Wigger, "First- and second-moment constrained Gaussian channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Melbourne, Australia, Jul. 12–20, 2021, pp. 432–437.

[12] S. Ma, H. Li, Y. He, R. Yang, S. Lu, W. Cao, and S. Li, "Capacity bounds and interference management for interference channel in visible light communication networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 182–193, Jan. 2019.

[13] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Inf. Contr.*, vol. 18, no. 3, pp. 203–219, Feb. 1971.

[14] T. H. Chan, S. Hranilovic, and F. R. Kschischang, "Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2073–2088, June 2005.

[15] S. Hranilovic and F. R. Kschischang, "Capacity bounds for power- and band-limited optical intensity channels corrupted by Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 784–795, May 2004.

[16] A. A. Farid and S. Hranilovic, "Upper and lower bounds on the capacity of wireless optical intensity channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Nice, France, June 24–30, 2007.

[17] A. Dytso, M. Goldenbaum, S. Shamai (Shitz), and H. V. Poor, "Upper and lower bounds on the capacity of amplitude-constrained MIMO channels," in Proc. *IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[18] A. Chaaban, Z. Rezki, and M.-S. Alouini, "Capacity bounds and high-SNR capacity of MIMO intensity-modulation optical channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3003–3017, May 2018.

[19] A. Chaaban, Z. Rezki, and M.-S. Alouini, "Low-SNR asymptotic capacityof MIMO optical intensity channels with peak and average constraints," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4694–4705, Oct. 2018

[20] A. Dytso, M. Al, H. V. Poor, and S. Shamai (Shitz), "On the capacity of the peak power constrained vector Gaussian channel: An estimation theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, pp. 3907–3921, Jun. 2019.

[21] R.-H. Chen, L. Li, J. Zhang, W. Zhang, and J. Zhou, "On the capacity of MISO optical intensity channels with per-antenna intensity constraints." Available: https://arxiv.org/abs/2111.11637.

[22] R.-H. Chen, L. Li, J. Zhang, and L. Li, "Capacity results for MIMO optical wireless communication with per-antenna intensity constraints." Available: https://arxiv.org/abs/2112.02220.

[23] K. Keykhosravi, E. Agrell, M. Secondini, and M. Karlsson, "When to use optical amplification in moncoherent transmission: An information-theoretic approach," in *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2438–2445, April 2020.

[24] A. Lapidoth, S. M. Moser, and M. A. Wigger,"On the capacity of free-space optical intensity channels," in *IEEE Transactions on Information Theory*, vol. 55, no. 10, pp. 4449–4461, Oct. 2009.