The Capacity of Fading Vector Gaussian Channels Under Amplitude Constraints on Antenna Subsets

Antonino Favano*†, Marco Ferrari†, Maurizio Magarini*, and Luca Barletta*
*Politecnico di Milano, Milano, Italy, {antonino.favano, maurizio.magarini, luca.barletta}@polimi.it
†CNR-IEIIT, Milano, Italy, marco.ferrari@ieiit.cnr.it

***Abstract*—Upper bounds on the capacity of vector Gaussian channels affected by fading are derived under peak amplitude constraints at the input. The focus is on constraint regions that can be decomposed in a Cartesian product of sub-regions. This constraint models a transmitter configuration employing a number of power amplifiers less than or equal to the total number of transmitting antennas. In general, the power amplifiers feed distinct subsets of the transmitting antennas and partition the input in independent subspaces. Two upper bounds are derived: The first one is suitable for high signal-to-noise ratio (SNR) values and, as we prove, it is tight in this regime; The second upper bound is accurate at low SNR. Furthermore, the derived upper bounds are applied to the relevant case of amplitude constraints induced by employing a distinct power amplifier for each transmitting antenna.**

I. Introduction

Amplitude constraints accurately model the main limitation induced by power amplifiers due to their nonlinear behavior. For this reason, the evaluation of the channel capacity under peak amplitude constraints is a research topic of great practical interest. One of the first contributions in this field is thanks to Smith [1]. In his work, he investigates the capacity of scalar Gaussian channels and the capacity-achieving input distribution. He proves that the optimal input distribution is discrete and composed of a finite number of mass points. In [2], the authors extend Smith's findings to quadrature Gaussian channel under amplitude constraints on the norm of the input, proving that the capacity-achieving input distribution is again discrete, made of a finite number of mass points, and also uniformly distributed in its phase. A further generalization to vector Gaussian channels is presented in [3]. Other significant results on the discreteness of the optimal input distribution are presented in [4]–[6].

In [7], McKellips presents a tight upper bound on the capacity of scalar Gaussian channels under peak amplitude constraints. The authors of [8] rederive the McKellips' upper bound through a dual capacity expression and generalize it to higher dimensions. Furthermore, they improve on McKellips' result and define a more accurate upper bound, which they refer to as *refined* upper bound. In [9], the present authors define a numerical algorithm to evaluate an arbitrarily precise estimate of the channel capacity and of its capacity-achieving distribution.

In the aforementioned works, the amplitude constraint is set on the norm of the input vector, which correctly models the limitation induced by a single power amplifier common to all the transmitting antennas. Furthermore, the considered channel matrix is assumed to be an identity matrix.

The authors of [10] evaluate capacity bounds for 2×2 multiple input multiple output (MIMO) systems under rectangular peak amplitude constraints and any arbitrary channel matrix. In [11], the authors further generalize the investigation to higher dimensional vector Gaussian channels and derive bounds for arbitrary constraint regions. In [12], interesting insights on the capacity-achieving input distribution for low signal-to-noise ratio (SNR) levels are presented. Finally, in [13], [14] the present authors derive an upper bound for arbitrary convex constraint regions that, together with the entropy power inequality (EPI) lower bound, provides a vanishing capacity gap at high SNR.

In [15, Appendix F], the authors use a duality-based upper-bounding technique and a suitable auxiliary product output distribution to derive an upper bound that is given by a sum of upper bounds on independent sub-spaces. They derive their upper bound for a system with number of transmitting antennas N_T strictly larger than the number of receiving antennas N_R and for an input constraint region defined as a Cartesian product of N_R one-dimensional sub-regions.

Contributions

In this paper, we adapt the result in [15, Appendix F] to the case of $\mathsf{N} \times \mathsf{N}$ MIMO systems and generalize their approach to input constraint regions defined as the Cartesian product of an arbitrary number of sub-regions $\mathsf{K} \leq \mathsf{N}$. In addition to the mentioned transmitter configuration using a single power amplifier, another configuration of practical interest is that of employing separate power amplifiers for each transmitting antenna. For this latter case, the resulting constraint region turns out to be a Cartesian product of the constraint imposed by each amplifier, which we refer to as *per-antenna* constraint.

In this work, we further generalize the constraint region as a Cartesian product of sub-regions lying in sub-spaces of the MIMO system. This generalization can model the transmitter configuration employing multiple power amplifiers, each one feeding a given subset of the transmitting antennas. We propose two upper bounds targeting peak amplitude constraints that can be decomposed into a Cartesian product of sub-regions. The first upper bound that we derive is suitable for high SNR values, and we prove that it converges to the EPI lower bound for increasing SNR. As for the low SNR regime, we propose an upper bound based on a Gaussian maximum-entropy argument. Finally, we apply our bounds to the practical scenario of the per-antenna constraint, which can be seen as a special case of the considered Cartesian constraint regions.

Paper Organization

In Sec. II we define the channel model, while in Sec. III we present our main results. We provide high and low SNR regime upper bounds and we investigate their asymptotic behavior. Furthermore, in Sec. IV we specialize the derived upper bounds to the per-antenna constraint an provide numerical results verifying the predicted asymptotic behavior. Finally, Sec. V concludes the paper.

Notation

We use bold letters for vectors ($\mathbf{x}$), uppercase letters for random variables (X), and calligraphic uppercase letters for subsets of vector spaces ($\mathcal{X}$). We represent the $n \times 1$ vector of zeros by $\mathbf{0}_n$ and the $n \times n$ identity matrix by I_n. We denote by $\mathcal{CN}(\boldsymbol{\mu}, \Sigma)$ a multivariate complex Gaussian distribution and by $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$ a multivariate real Gaussian distribution, both with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ. For a given matrix H, we define by $\lambda_i(\mathsf{H})$ the ith singular value of H. By $\mathcal{B}_n(\mathsf{R})$ we denote the n-dimensional closed ball of radius R and we define the n-dimensional box of sides R as $\mathrm{Box}_n(\mathsf{R}) \triangleq \{\mathbf{x} : |x_i| \leq \mathsf{R}/2,\ i = 1, \dots, n\}$. Finally, the n-dimensional volume of a set $\mathcal{X}$ is denoted by $\mathrm{Vol}_n(\mathcal{X})$.

II. Channel Model

Let us consider an $\mathsf{N} \times \mathsf{N}$ real MIMO system with input-output relationship given by

$$\mathbf{Y} = \mathsf{H}\mathbf{X} + \mathbf{Z}, \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{\mathsf{N}}$ is the output vector, H is any full rank channel fading matrix, $\mathbf{X} \in \mathcal{X} \subset \mathbb{R}^{\mathsf{N}}$ is the input vector, with $\mathcal{X}$ being the input constraint region, and $\mathbf{Z} \in \mathbb{R}^{\mathsf{N}}$ is a noise vector such that $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}_{\mathsf{N}}, \sigma_z^2 \mathsf{I}_{\mathsf{N}})$. Let us assume H to be constant over all channel uses and known both at the transmitter and at the receiver.

Throughout this paper we consider input constraint regions that can be decomposed into the Cartesian product of sub-regions. Let us denote by K the number of sub-regions in $\mathcal{X}$. We define

$$\mathcal{X} \triangleq \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_{\mathsf{K}}, \tag{2}$$

where the operator $\times$ denotes the Cartesian product and $\mathcal{X}_i \subset \mathbb{R}^{\mathsf{N}_i}$ is the ith sub-region of $\mathcal{X}$ with dimension N_i. For convenience in the indexing notation, we also define $\mathsf{N}_0 = 0$. It is worth observing that $\forall i > 0$, $\mathsf{N}_i \in \mathbb{N}^+$ and that $\sum_{i=1}^{\mathsf{K}} \mathsf{N}_i = \mathsf{N}$. Let us define the maximum radius of each sub-region as

$$\mathsf{R}_i \triangleq \sup_{\mathbf{x} \in \mathcal{X}_i} \|\mathbf{x}\|, \tag{3}$$

for all $i = 1, \dots, \mathsf{K}$. For the sake of simplicity we will assume $\mathsf{R}_i = \mathsf{R},\ i = 1, \dots, \mathsf{K}$. Note that setting all the R_i's to R can be done without loss of generality by scaling the related columns of H accordingly.

Since we consider peak amplitude-constrained input distributions, we resort to the following SNR definition

$$\mathsf{SNR} = \frac{\mathsf{R}^2 \mathsf{K}}{\mathsf{N}\sigma_z^2}. \tag{4}$$

Finally, we define the channel capacity as

$$\mathsf{C} \triangleq \max_{P_{\mathbf{X}} :\, \mathsf{supp}(P_{\mathbf{X}}) \subseteq \mathcal{X}} \mathrm{I}(\mathbf{X}; \mathbf{Y}), \tag{5}$$

where $P_{\mathbf{X}}$ is the input distribution law.

III. Main Results

In this section we derive two upper bounds. The first upper bound is suitable for the high SNR regime and since in this SNR range the input signal is predominant over the noise, the MIMO capacity can be approximated, broadly speaking, by a sum of capacities, each induced by the sub-spaces where the $\mathcal{X}_i$'s lie.

Furthermore, we introduce a second upper bound, targeting the low SNR regime. For this latter SNR range, we assume the Gaussian noise to be the dominant component and, therefore, we upper-bound the capacity by using a Gaussian output distribution.

A. High SNR regime

To derive an upper bound on the channel capacity, suitable for the high SNR regime, we consider an equivalent output multiplied by the inverse of the channel matrix H. Note that the receiver can compute H^{-1} because the matrix H is full rank and it is known at the receiver. We have

$$\begin{aligned} \mathsf{H}^{-1}\mathbf{Y} &= \mathsf{H}^{-1}\mathsf{H} \cdot \mathbf{X} + \mathsf{H}^{-1}\mathbf{Z} & (6)\\ &= \mathbf{X} + \mathsf{H}^{-1}\mathbf{Z} & (7)\\ &= \mathbf{X} + \mathbf{Z}_{\mathsf{D}}, & (8) \end{aligned}$$

where $\mathbf{Z}_{\mathsf{D}} = \mathsf{H}^{-1}\mathbf{Z}$ is the resulting noise vector with $\mathbf{Z}_{\mathsf{D}} \sim \mathcal{N}(\mathbf{0}_{\mathsf{N}}, \mathsf{D})$ and $\mathsf{D} = \sigma_z^2 \mathsf{H}^{-1}\mathsf{H}^{-T}$. Let us denote by $d_{k,l}$ the element (k, l) of the matrix D and define the main-diagonal block submatrices D_i's of D as

$$\mathsf{D}_i \triangleq [d_{k,l}]_{k,l=m_i+1}^{m_i+\mathsf{N}_i}, \quad i = 1, \dots, \mathsf{K}, \tag{9}$$

where $m_i = \sum_{j=1}^{\mathsf{N}_{i-1}} \mathsf{N}_j$ and $m_1 = 0$. Furthermore, let us denote by $\mathbf{X}_i$ the $\mathsf{N}_i \times 1$ vector $\mathbf{X}_i = (X_{m_i+1}, X_{m_i+2}, \dots, X_{m_i+\mathsf{N}_i})^T$ and $\mathbf{Z}_{\mathsf{D},i}$ analogously.

Theorem 1. *Given the input constraint region $\mathcal{X}$ defined in (2), the channel capacity is upper-bounded by*

$$\mathsf{C} \leq \overline{\mathsf{C}}_1 \triangleq \left(\sum_{i=1}^{\mathsf{K}} \mathsf{C}_i\right) + \frac{1}{2}\log\frac{\prod_{j=1}^{\mathsf{K}} \det(\mathsf{D}_j)}{\det(\mathsf{D})}, \tag{10}$$

where

$$\mathsf{C}_i \triangleq \max_{P_{\mathbf{X}_i} :\, \mathbf{X}_i \in \mathcal{X}_i} \mathsf{h}(\mathbf{X}_i + \mathbf{Z}_{\mathsf{D},i}) - \mathsf{h}(\mathbf{Z}_{\mathsf{D},i}) \tag{11}$$

and $\mathbf{Z}_{\mathsf{D},i} \sim \mathcal{N}(\mathbf{0}_{\mathsf{N}_i}, \mathsf{D}_i)$.

Proof.

$$\mathsf{C} = \max_{P_{\mathbf{X}}:\,\mathbf{X}\in\mathcal{X}} \mathrm{I}(\mathbf{X}\,;\mathsf{H}\mathbf{X}+\mathbf{Z}) \tag{12}$$
$$= \max_{P_{\mathbf{X}}:\,\mathbf{X}\in\mathcal{X}} \mathrm{I}(\mathbf{X}\,;\mathbf{X}+\mathbf{Z}_{\mathsf{D}}) \tag{13}$$
$$= \max_{P_{\mathbf{X}}:\,\mathbf{X}\in\mathcal{X}} \mathsf{h}(\mathbf{X}+\mathbf{Z}_{\mathsf{D}}) - \mathsf{h}(\mathbf{Z}_{\mathsf{D}}) \tag{14}$$
$$\overset{(a)}{\le} \left(\max_{P_{\mathbf{X}}:\,\mathbf{X}\in\mathcal{X}} \sum_{i=1}^{\mathsf{K}} \mathsf{h}(\mathbf{X}_i+\mathbf{Z}_{\mathsf{D},i})\right) + \frac{1}{2}\log\frac{1}{\det(\mathsf{D})} - \frac{\mathsf{N}}{2}\log(2\pi e) \tag{15}$$
$$= \left(\sum_{i=1}^{\mathsf{K}} \max_{P_{\mathbf{X}_i}:\,\mathbf{X}_i\in\mathcal{X}_i} \mathsf{h}(\mathbf{X}_i+\mathbf{Z}_{\mathsf{D},i})\right) + \frac{1}{2}\log\frac{1}{\det(\mathsf{D})} - \frac{\mathsf{N}}{2}\log(2\pi e) + \frac{1}{2}\log\frac{\prod_{j=1}^{\mathsf{K}}\det(\mathsf{D}_j)}{\prod_{k=1}^{\mathsf{K}}\det(\mathsf{D}_k)} \tag{16}$$
$$= \left(\sum_{i=1}^{\mathsf{K}} \max_{P_{\mathbf{X}_i}:\,\mathbf{X}_i\in\mathcal{X}_i} \mathsf{h}(\mathbf{X}_i+\mathbf{Z}_{\mathsf{D},i}) - \mathsf{h}(\mathbf{Z}_{\mathsf{D},i})\right) + \frac{1}{2}\log\frac{\prod_{j=1}^{\mathsf{K}}\det(\mathsf{D}_j)}{\det(\mathsf{D})}, \tag{17}$$

where (a) holds because of the sub-additivity of the differential entropy and $\mathbf{Z}_{\mathsf{D},i}$ is obtained by marginalizing $\mathbf{Z}_{\mathsf{D}}$ on the ith sub-space. Note that, since $\mathbf{Z}_{\mathsf{D}}$ is a multivariate Gaussian with zero mean and covariance matrix D, it holds $\mathbf{Z}_{\mathsf{D},i} \sim \mathcal{N}(\mathbf{0}_{\mathsf{N}_i}, \mathsf{D}_i)$. Because of the structure of $\mathcal{X}$, we know that the optimal $P_{\mathbf{X}}$ is a product distribution of the K terms $P_{\mathbf{X}_i}$. Consequently, we can decompose and evaluate the maximization in (15) over the K independent subspaces corresponding to each $P_{\mathbf{X}_i}$. In turns, this decomposition allows us to swap the maximization and the sum operators of (15) without introducing any further upper bound. Finally, in (16) we add and subtract $\frac{1}{2}\log\prod_i\det(\mathsf{D}_i)$, to obtain the term $\mathsf{h}(\mathbf{Z}_{\mathsf{D},i})$ in (17). □

Remark 1. Since the C_i's are defined implicitly via a maximization over the corresponding $P_{\mathbf{X}_i}$'s, $\overline{\mathsf{C}}_1$ cannot be directly evaluated. Nonetheless, we can obtain explicit results by evaluating upper bounds on the C_i's via suitable techniques, like those presented in [8], [11], [14]. In turn, the upper bounds on the C_i's can be plugged into (10) to obtain an explicit upper bound on $\overline{\mathsf{C}}_1$.

Remark 2. Since D is positive-semidefinite, by Fischer's inequality [16], we have that $\det(\mathsf{D}) \le \prod_{j=1}^{\mathsf{K}}\det(\mathsf{D}_j)$. Therefore, it holds

$$\log\frac{\prod_{j=1}^{\mathsf{K}}\det(\mathsf{D}_j)}{\det(\mathsf{D})} \ge 0. \tag{18}$$

Remark 3. Intuitively, the logarithmic term in (10) accounts for the inaccuracy introduced by considering the noise vector $\mathbf{Z}_{\mathsf{D}}$ to be independent on each of the K sub-spaces. Indeed, whenever H is diagonal we have that $\det(\mathsf{D}) = \prod_{i=1}^{\mathsf{K}}\det(\mathsf{D}_i)$, then $\log\frac{\det(\mathsf{D})}{\prod_j\det(\mathsf{D}_j)}$ goes to zero and inequality (10) becomes an equality.

Let us introduce the EPI lower bound [11] for the channel in (1) as

$$\mathsf{C} \ge \underline{\mathsf{C}} \triangleq \frac{\mathsf{N}}{2}\log\left(1+\frac{(\mathrm{Vol}_{\mathsf{N}}(\mathsf{H}\mathcal{X}))^{\frac{2}{\mathsf{N}}}}{2\pi e\sigma_z^2}\right). \tag{19}$$

In the following lemma, we show that the capacity gap between the EPI lower bound and the upper bound in Theorem 1 is vanishing when the SNR tends to infinity.

Lemma 1. *When $\sigma_z^2 \to 0$, we have*

$$\lim_{\sigma_z^2\to 0} \overline{\mathsf{C}}_1 - \underline{\mathsf{C}} = 0. \tag{20}$$

Proof. Let us consider the mutual information for the ith sub-channel $\mathrm{I}(\mathbf{X}_i\,;\mathbf{X}_i+\mathbf{Z}_{\mathsf{D},i}) = \mathsf{h}(\mathbf{X}_i+\mathbf{Z}_{\mathsf{D},i}) - \mathsf{h}(\mathbf{Z}_{\mathsf{D},i})$. Let us denote by M_i the $\mathsf{N}_i\times\mathsf{N}_i$ matrix such that $\mathsf{D}_i = \sigma_z^2\mathsf{M}_i^{-1}\mathsf{M}_i^{-T}$. We can derive such matrix M_i because D_i is a covariance matrix and therefore it is positive-semidefinite. We have that

$$\mathrm{I}(\mathbf{X}_i\,;\mathbf{X}_i+\mathbf{Z}_{\mathsf{D},i}) = \mathrm{I}(\mathbf{X}_i\,;\mathsf{M}_i\mathbf{X}_i+\mathbf{Z}_i), \tag{21}$$

where $\mathbf{Z}_i \sim \mathcal{N}(\mathbf{0}_{\mathsf{N}_i}, \sigma_z^2\mathsf{I}_{\mathsf{N}_i})$. Then, we have

$$\lim_{\sigma_z^2\to 0}\sum_{i=1}^{\mathsf{K}}\mathsf{C}_i = \sum_{i=1}^{\mathsf{K}} \max_{P_{\mathbf{X}_i}:\,\mathbf{X}_i\in\mathcal{X}_i} \mathsf{h}(\mathsf{M}_i\mathbf{X}_i) - \lim_{\sigma_z^2\to 0}\mathsf{h}(\mathbf{Z}_i) \tag{22}$$
$$= \left(\sum_{i=1}^{\mathsf{K}}\log\mathrm{Vol}_{\mathsf{N}_i}(\mathsf{M}_i\mathcal{X}_i)\right) - \lim_{\sigma_z^2\to 0}\mathsf{h}(\mathbf{Z}) \tag{23}$$
$$= \left(\sum_{i=1}^{\mathsf{K}}\log\det(\mathsf{M}_i)\mathrm{Vol}_{\mathsf{N}_i}(\mathcal{X}_i)\right) - \lim_{\sigma_z^2\to 0}\mathsf{h}(\mathbf{Z}), \tag{24}$$

where (23) holds because $\mathsf{h}(\mathsf{M}_i\mathbf{X}_i)$ is maximized by the uniform distribution over $\mathcal{X}_i$. Notice also that

$$\lim_{\sigma_z^2\to 0}\frac{1}{2}\log\frac{\prod_{i=1}^{\mathsf{K}}\det(\mathsf{D}_i)}{\det(\mathsf{D})} \tag{25}$$
$$= \lim_{\sigma_z^2\to 0}\frac{1}{2}\log\frac{\prod_{i=1}^{\mathsf{K}}\det\left(\sigma_z^2\mathsf{M}_i^{-1}\mathsf{M}_i^{-T}\right)}{\det\left(\sigma_z^2\mathsf{H}^{-1}\mathsf{H}^{-T}\right)} \tag{26}$$
$$= \log\det(\mathsf{H}) - \sum_{i=1}^{\mathsf{K}}\log\det(\mathsf{M}_i), \tag{27}$$

where we used the fact that $\det(\mathsf{H}^{-1}\mathsf{H}^{-T}) = (\det(\mathsf{H}^{-1}))^2 = 1/(\det(\mathsf{H}))^2$ and similarly for the M_i's.

Furthermore, for the lower bound it holds that

$$\lim_{\sigma_z^2\to 0}\underline{\mathsf{C}} = \lim_{\sigma_z^2\to 0}\frac{\mathsf{N}}{2}\log\left(\frac{(\mathrm{Vol}_{\mathsf{N}}(\mathsf{H}\mathcal{X}))^{\frac{2}{\mathsf{N}}}}{2\pi e\sigma_z^2}\right) \tag{28}$$
$$= \lim_{\sigma_z^2\to 0}\log(\mathrm{Vol}_{\mathsf{N}}(\mathsf{H}\mathcal{X})) - \mathsf{h}(\mathbf{Z}) \tag{29}$$
$$= \lim_{\sigma_z^2\to 0}\log\det(\mathsf{H}) + \log\left(\prod_{i=1}^{\mathsf{K}}\mathrm{Vol}_{\mathsf{N}_i}(\mathcal{X}_i)\right) - \mathsf{h}(\mathbf{Z}). \tag{30}$$

Notice that, since $\mathcal{X}$ is defined by a Cartesian product, it holds that $\mathrm{Vol}_{\mathsf{N}}(\mathcal{X}) = \prod_i\mathrm{Vol}_{\mathsf{N}_i}(\mathcal{X}_i)$.

Finally, by putting everything together we get

$$\lim_{\sigma_z^2 \to 0} \overline{\mathsf{C}}_1 - \underline{\mathsf{C}} = 0. \tag{31}$$

□

Remark 4. As mentioned in Remark 1, one could be interested in upper-bounding the C_i's to evaluate explicit bounds on $\overline{\mathsf{C}}_1$. The results of Lemma 1 can always be extended to these explicit upper bounds on $\overline{\mathsf{C}}_1$, as long as the bounding techniques used on the C_i's guarantee asymptotic convergence to the true value as the SNR increases.

B. Low SNR regime

At low SNR, i.e., when the Gaussian noise is dominant, the upper bound in Theorem 1 is loose. Indeed, even as the SNR decreases, the logarithmic term in (10) remains larger than zero, as mentioned in Remark 2. Therefore, aside from the special case in which H is diagonal (see Remark 3), we have that $\overline{\mathsf{C}}_1$ is bounded away from zero. Let us define an upper bound that is more suitable than $\overline{\mathsf{C}}_1$ in the low SNR regime and vanishes as the SNR decreases. Intuitively, as R goes to zero, $\mathcal{X}$ becomes smaller and the output distribution becomes closer to a Gaussian. Therefore, we can derive an upper bound for the low SNR regime by using a Gaussian maximum-entropy argument.

Let us consider the singular value decomposition of H, i.e., $\mathsf{H} = \mathsf{U}\Lambda\mathsf{V}^T$. Given (1), we can consider the equivalent model

$$\begin{aligned} \Lambda^{-1}\bar{\mathbf{Y}} &= \bar{\mathbf{X}} + \Lambda^{-1}\bar{\mathbf{Z}} &(32)\\ &= \bar{\mathbf{X}} + \bar{\mathbf{Z}}_{\bar{\mathsf{D}}}, &(33)\end{aligned}$$

where $\bar{\mathbf{Y}} = \mathsf{U}^{-1}\mathbf{Y}$, the input is $\bar{\mathbf{X}} = \mathsf{V}^T\mathbf{X}$, and the noise vector is $\bar{\mathbf{Z}}_{\bar{\mathsf{D}}} = \Lambda^{-1}\bar{\mathbf{Z}} = \Lambda^{-1}\mathsf{U}^{-1}\mathbf{Z}$. Notice that, since $\mathbf{Z}$ has a rotationally symmetric distribution, we still have $\bar{\mathbf{Z}} \sim \mathcal{N}(\mathbf{0}_\mathsf{N}, \sigma_z^2 \mathsf{I}_\mathsf{N})$ and $\bar{\mathbf{Z}}_{\bar{\mathsf{D}}} \sim \mathcal{N}(\mathbf{0}_\mathsf{N}, \bar{\mathsf{D}})$ with $\bar{\mathsf{D}} = \sigma_z^2 \Lambda^{-1}\Lambda^{-T}$.

Theorem 2. *Given the input constraint region $\mathcal{X}$ defined in (2), the channel capacity is upper-bounded by*

$$\mathsf{C} \le \overline{\mathsf{C}}_2 \triangleq \left(\sum_{i=1}^{\mathsf{N}} \frac{1}{2}\log(\mathsf{P}_i + \lambda_i(\mathsf{D})) \right) - \frac{1}{2}\log\det(\mathsf{D}), \tag{34}$$

where P_i is the power allocation given by the water-filling algorithm, for a total available average power $\mathsf{R}^2\mathsf{K}$ and N parallel channels with noise variances $\lambda_i(\mathsf{D})$'s.

Proof. Since $\mathcal{X}$ is a Cartesian product of K sub-regions, each one contained in a ball of radius R, we have that $\sup_{\mathbf{x}\in\mathcal{X}}\{\|\mathbf{x}\|\} = \mathsf{R}\sqrt{\mathsf{K}}$. Given the constraint imposed by $\mathcal{X}$, the looser constraint $\mathsf{E}\left[\mathbf{X}^T\mathbf{X}\right] \le \mathsf{R}^2\mathsf{K}$ is always satisfied. We have

$$\begin{aligned}
\mathsf{C} &= \max_{P_\mathbf{X}:\, \mathbf{X}\in\mathcal{X}} \mathrm{I}(\mathbf{X}; \mathsf{H}\mathbf{X} + \mathbf{Z}) &(35)\\
&= \max_{\substack{P_\mathbf{X}:\, \mathbf{X}\in\mathcal{X},\\ \mathsf{E}[\mathbf{X}^T\mathbf{X}]\le \mathsf{R}^2\mathsf{K}}} \mathrm{I}(\mathbf{X}; \mathsf{H}\mathbf{X} + \mathbf{Z}) &(36)\\
&\le \max_{P_\mathbf{X}:\, \mathsf{E}[\mathbf{X}^T\mathbf{X}]\le \mathsf{R}^2\mathsf{K}} \mathrm{I}(\mathbf{X}; \mathsf{H}\mathbf{X} + \mathbf{Z}) &(37)\\
&= \max_{P_{\bar{\mathbf{X}}}:\, \mathsf{E}[\bar{\mathbf{X}}^T\bar{\mathbf{X}}]\le \mathsf{R}^2\mathsf{K}} \mathrm{I}(\bar{\mathbf{X}}; \bar{\mathbf{X}} + \bar{\mathbf{Z}}_{\bar{\mathsf{D}}}) &(38)\\
&= \max_{P_{\bar{\mathbf{X}}}:\, \mathsf{E}[\bar{\mathbf{X}}^T\bar{\mathbf{X}}]\le \mathsf{R}^2\mathsf{K}} \mathrm{h}(\Lambda^{-1}\bar{\mathbf{Y}}) - \mathrm{h}(\bar{\mathbf{Z}}_{\bar{\mathsf{D}}}) &(39)\\
&\le \max_{P_{\bar{\mathbf{X}}}:\, \mathsf{E}[\bar{\mathbf{X}}^T\bar{\mathbf{X}}]\le \mathsf{R}^2\mathsf{K}} \mathrm{h}\left(\widetilde{\mathbf{Y}}\right) - \mathrm{h}(\bar{\mathbf{Z}}_{\bar{\mathsf{D}}}) &(40)\\
&\le \max_{P_{\bar{\mathbf{X}}}:\, \mathsf{E}[\bar{\mathbf{X}}^T\bar{\mathbf{X}}]\le \mathsf{R}^2\mathsf{K}} \sum_{i=1}^{\mathsf{N}} \frac{1}{2}\log\left(2\pi e\left(\mathsf{E}\left[\left|\bar{X}_i\right|^2\right] + \lambda_i(\bar{\mathsf{D}})\right)\right)\\
&\qquad - \frac{1}{2}\log\det\left(2\pi e\bar{\mathsf{D}}\right) &(41)\\
&= \left(\sum_{i=1}^{\mathsf{N}} \frac{1}{2}\log(\mathsf{P}_i + \lambda_i(\mathsf{D})) \right) - \frac{1}{2}\log\det(\mathsf{D}), &(42)
\end{aligned}$$

where the upper bound in (37) holds because we removed the constraint imposed by $\mathcal{X}$, in (38) we used the equivalent model defined in (33). Since V^T is a unitary matrix, we have that $\mathsf{E}\left[\mathbf{X}^T\mathbf{X}\right] = \mathsf{E}\left[\bar{\mathbf{X}}^T\bar{\mathbf{X}}\right]$. Let us define the normally distributed vector $\widetilde{\mathbf{Y}} \sim \mathcal{N}(\mathbf{0}_\mathsf{N}, \Sigma)$, with $\Sigma = \mathsf{E}\left[\bar{\mathbf{X}}\bar{\mathbf{X}}^T\right] + \bar{\mathsf{D}}$. For the upper bound in (40) we used a Gaussian maximum-entropy bound $\mathrm{h}\left(\Lambda^{-1}\bar{\mathbf{Y}}\right) \le \mathrm{h}(\widetilde{\mathbf{Y}})$, and in (41) we used $\mathrm{h}(\widetilde{\mathbf{Y}}) \le \sum_i \mathrm{h}(\widetilde{Y}_i)$. Finally, to obtain (42) we notice that $\lambda_i\left(\bar{\mathsf{D}}\right) = \lambda_i(\mathsf{D})$ for any i and we apply the water-filling algorithm. □

The following trivial lemma shows that the upper bound $\overline{\mathsf{C}}_2$ is vanishing in the low SNR regime.

Lemma 2. *The capacity upper bound $\overline{\mathsf{C}}_2$ tends to zero for $\sigma_z^2 \to \infty$*

$$\lim_{\sigma_z^2 \to \infty} \overline{\mathsf{C}}_2 = 0. \tag{43}$$

Proof. In Theorem 2, when $\sigma_z^2 \to \infty$ the P_i's tend to be negligible compared to the $\lambda_i(\mathsf{D})$, which are proportional to σ_z^2. Therefore, we have that

$$\lim_{\sigma_z^2 \to \infty} \overline{\mathsf{C}}_2 = \left(\sum_{i=1}^{\mathsf{N}} \frac{1}{2}\log(\lambda_i(\mathsf{D})) \right) - \frac{1}{2}\log\det(\mathsf{D}) = 0. \tag{44}$$

□

IV. Per-Antenna Constraint

The proposed upper bounds can be applied to a common and practical constraint, namely the *per-antenna* constraint. A transmitter configuration of practical interest in MIMO systems is that of a single power amplifier for each transmitting antenna. We model the transmitted signal on each antenna as a complex signal. Let us consider a MIMO system with $\mathsf{N}/2$ complex dimensions

$$\mathbf{Y}' = \mathsf{H}'\mathbf{X}' + \mathbf{Z}', \tag{45}$$

where $\mathbf{Y}' \in \mathbb{C}^{\mathsf{N}/2}$ is the output vector, H' is any full rank complex channel fading matrix, $\mathbf{X}' \in \mathcal{X}' = \mathrm{Box}_{\mathsf{N}/2}(2\mathsf{R}) \subset \mathbb{C}^{\mathsf{N}/2}$ is the input vector, with $\mathcal{X}'$ being the input constraint region, and $\mathbf{Z}' \in \mathbb{C}^{\mathsf{N}/2}$ is a noise vector such that $\mathbf{Z}' \sim \mathcal{CN}(\mathbf{0}_\mathsf{N}, 2\sigma_z^2\mathsf{I}_\mathsf{N})$. Note that we can still refer to the model in (1) simply by vectorizing the system in (45): We need to define $\mathsf{H} = \mathrm{Re}\{\mathsf{H}'\} \otimes \mathsf{I}_2 + \mathrm{Im}\{\mathsf{H}'\} \otimes \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$, where the operator $\otimes$ is the Kronecker product, while the output vector is such that $\mathbf{Y} = [\mathrm{Re}(Y_1'), \mathrm{Im}(Y_1'), \dots, \mathrm{Re}(Y_\mathsf{N}'), \mathrm{Im}(Y_\mathsf{N}')]^T$, and analogously for $\mathbf{X}$ and $\mathbf{Z}$. Notice also that $\mathsf{K} = \mathsf{N}/2$. Therefore, for any $i = 1, \dots, \mathsf{N}/2$, the constraint $|X_i'| \leq \mathsf{R}$ is equivalent to

$$\mathbf{X}_i \in \mathcal{X}_i = \mathcal{B}_2(\mathsf{R}), \;\; i = 1, \dots, \mathsf{N}/2, \tag{46}$$

where $\mathbf{X}_i = \begin{pmatrix} \mathrm{Re}(X_i') \\ \mathrm{Im}(X_i') \end{pmatrix}$. While the upper bound in Theorem 2 can be applied directly, the upper bound $\overline{\mathsf{C}}_1$ of Theorem 1 has to be specialized for the per-antenna case. Let us consider the following equivalent expression of C_i defined in the proof of Lemma 1

$$\mathsf{C}_i = \max_{P_{\mathbf{X}_i}:\, \mathbf{X}_i \in \mathcal{X}_i} \mathsf{h}(\mathsf{M}_i\mathbf{X}_i) - \mathsf{h}(\mathbf{Z}_i). \tag{47}$$

Since H is obtained by vectorizing H', the singular values of H are equal 2-by-2, i.e., $\lambda_{2i}(\mathsf{H}) = \lambda_{2i-1}(\mathsf{H})$, $i = 1, \dots, \mathsf{N}/2$. The same is true for the singular values of D, D_i's, and M_i's. To simplify the notation, we define

$$\lambda(\mathsf{M}_i) \triangleq \lambda_1(\mathsf{M}_i) = \lambda_2(\mathsf{M}_i), \quad i = 1, \dots, \mathsf{N}/2. \tag{48}$$

In the per-antenna case, suitable upper bounds for each C_i are defined in [8], [17]. The McKellips-Type upper bound, derived in [8, Eq. (32)], gives the following simple closed form expression

$$\mathsf{C}_i \leq \overline{\mathsf{C}}_{\mathrm{McK},i} \triangleq \log\left(1 + \sqrt{\frac{\pi}{2}}\frac{\lambda(\mathsf{M}_i)\mathsf{R}}{\sigma_z} + \frac{(\lambda(\mathsf{M}_i)\mathsf{R})^2}{2e\sigma_z^2}\right), \tag{49}$$

for $i = 1, \dots, \mathsf{N}/2$. Therefore, $\overline{\mathsf{C}}_1$ of Theorem 1 can be upper-bounded by

$$\overline{\mathsf{C}}_1 \leq \overline{\mathsf{C}}_{\mathrm{PA},1} \triangleq \left(\sum_{i=1}^{\mathsf{N}/2} \overline{\mathsf{C}}_{\mathrm{McK},i}\right) + \frac{1}{2}\log\frac{\prod_{j=1}^{\mathsf{K}} \det(\mathsf{D}_j)}{\det(\mathsf{D})}. \tag{50}$$

Furthermore, the authors of [8] derive an additional upper bound, tighter than (49), that however has to be computed via a numerical optimization. For a given C_i, let us denote by $\overline{\mathsf{C}}_{\mathrm{Ref},i} \geq \mathsf{C}_i$ this *refined* upper bound [8, Eq. (82)]. Then, by plugging the $\overline{\mathsf{C}}_{\mathrm{Ref},i}$'s into (10), we can define the following upper bound on $\overline{\mathsf{C}}_1$

$$\overline{\mathsf{C}}_1 \leq \overline{\mathsf{C}}_{\mathrm{PA},2} \triangleq \left(\sum_{i=1}^{\mathsf{N}/2} \overline{\mathsf{C}}_{\mathrm{Ref},i}\right) + \frac{1}{2}\log\frac{\prod_{j=1}^{\mathsf{K}} \det(\mathsf{D}_j)}{\det(\mathsf{D})}. \tag{51}$$

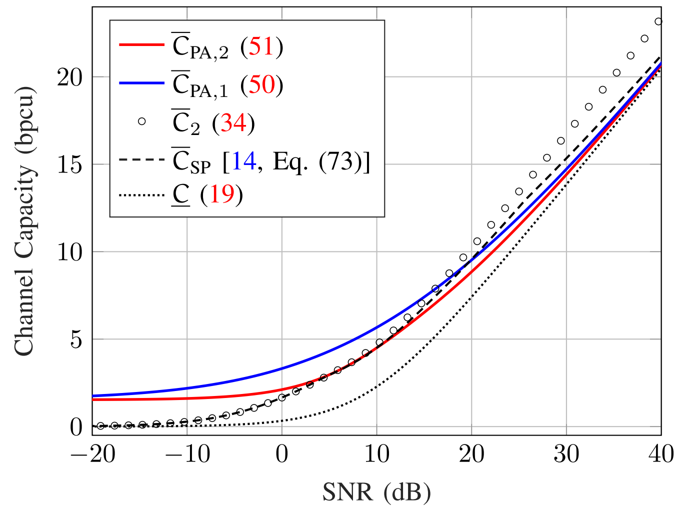

Figure 1. Capacity bounds in bit per channel use (bpcu) versus SNR, for $\mathsf{N} = 4$, $\lambda(\mathsf{M}_1) = 0.52$, and $\lambda(\mathsf{M}_2) = 0.37$.

Note that the $\overline{\mathsf{C}}_{\mathrm{McK},i}$'s and the $\overline{\mathsf{C}}_{\mathrm{Ref},i}$'s asymptotically converge to the C_i's as the SNR increases. Therefore, Lemma 1 can be extended to both $\overline{\mathsf{C}}_{\mathrm{PA},1}$ and $\overline{\mathsf{C}}_{\mathrm{PA},2}$ (see Remark 4).

A. Numerical Results

For the per-antenna case, let us now evaluate numerically $\underline{\mathsf{C}}$, $\overline{\mathsf{C}}_2$, and both the specialized upper bounds on $\overline{\mathsf{C}}_1$. We evaluate the bounds for a random realization of H. If we consider the compound upper bound given by $\min\left(\overline{\mathsf{C}}_2, \overline{\mathsf{C}}_{\mathrm{PA},2}\right)$ we see that, as predicted by Lemma 1 and Lemma 2, the capacity gap between upper and lower bounds is indeed vanishing both at high SNR, thanks to $\overline{\mathsf{C}}_{\mathrm{PA},2}$, and at low SNR, thanks to $\overline{\mathsf{C}}_2$. Moreover, we compare the proposed bounds to the previous best in the existing literature, which we proposed in [14]. Specifically, let us denote by $\overline{\mathsf{C}}_{\mathrm{SP}}$ the upper bound based on a sphere packing argument [14, Eq. (73)]. Notice that the sphere packing bound is also vanishing at both low and high SNR, but as seen in Fig. 1, the upper bounds $\overline{\mathsf{C}}_{\mathrm{PA},1}$ and $\overline{\mathsf{C}}_{\mathrm{PA},2}$ can be closer to the lower bound $\underline{\mathsf{C}}$ at finite SNR levels of practical interest.

V. Conclusion

We have derived two upper bounds on the channel capacity of peak amplitude-constrained vector Gaussian channels affected by fading. We considered constraint regions that can be decomposed into a Cartesian product, reflecting the fact that each power amplifier feeds a subset of the transmitting antennas. We have proved that the first upper bound, suitable for the high signal-to-noise (SNR) regime, has vanishing capacity gap when compared to the entropy power inequality lower bound. The second proposed upper bound is based on a Gaussian maximum-entropy argument and is used for the low SNR regime. Finally, for a transmitter that employs separate power amplifiers for each antenna, we have shown an example where the proposed upper bounds are tighter than the best known upper bounds at any SNR.

REFERENCES

[1] J. G. Smith, "The information capacity of amplitude- and variance-constrained scalar Gaussian channels," *Information and Control*, vol. 18, no. 3, pp. 203–219, April 1971.

[2] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Transactions on Information Theory*, vol. 41, no. 4, pp. 1060–1071, July 1995.

[3] B. Rassouli and B. Clerckx, "On the capacity of vector Gaussian channels with bounded inputs," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 6884–6903, December 2016.

[4] A. Tchamkerten, "On the discreteness of capacity-achieving distributions," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2773–2778, Nov. 2004.

[5] T. Chan, S. Hranilovic, and F. Kschischang, "Capacity-achieving probability measure for conditionally Gaussian channels with bounded inputs," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 2073–2088, Jun. 2005.

[6] B. Mamandipoor, K. Moshksar, and A. K. Khandani, "Capacity-achieving distributions in Gaussian multiple access channel with peak power constraints," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6080–6092, Oct. 2014.

[7] A. L. McKellips, "Simple tight bounds on capacity for the peak-limited discrete-time channel," in *International Symposium on Information Theory, 2004. ISIT 2004. Proceedings.*, June 2004, pp. 348–348.

[8] A. Thangaraj, G. Kramer, and G. Böcherer, "Capacity bounds for discrete-time, amplitude-constrained, additive white Gaussian noise channels," *IEEE Transactions on Information Theory*, vol. 63, no. 7, pp. 4172–4182, Jul. 2017.

[9] A. Favano, M. Ferrari, M. Magarini, and L. Barletta, "The capacity of the amplitude-constrained vector Gaussian channel," in *2021 IEEE International Symposium on Information Theory (ISIT)*, Jul. 2021, pp. 426–431.

[10] A. ElMoslimany and T. M. Duman, "On the capacity of multiple-antenna systems and parallel Gaussian channels with amplitude-limited inputs," *IEEE Transactions on Communications*, vol. 64, no. 7, pp. 2888–2899, Jul. 2016.

[11] A. Dytso, M. Goldenbaum, S. Shamai, and H. V. Poor, "Upper and lower bounds on the capacity of amplitude-constrained MIMO channels," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, December 2017, pp. 1–6.

[12] 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 Transactions on Information Theory*, vol. 65, no. 6, pp. 3907–3921, January 2019.

[13] A. Favano, M. Ferrari, M. Magarini, and L. Barletta, "A sphere packing bound for AWGN MIMO fading channels under peak amplitude constraints," in *2020 IEEE Information Theory Workshop (ITW)*, 2021, pp. 1–5.

[14] ——, "A sphere packing bound for vector Gaussian fading channels under peak amplitude constraints," *arXiv preprint arXiv:2111.13179*, Nov 2021.

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

[16] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge university press, 2012.

[17] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Transactions on Information Theory*, vol. 41, no. 4, pp. 1060–1071, 1995.