On MIMO Channel Capacity
with Output Quantization Constraints

Abbas Khalili
NYU
New York, USA
ako274@nyu.edu

Stefano Rini
NCTU
Hsinchu, Taiwan
stefano@nctu.edu.tw

Luca Barletta
Politecnico di Milano
Milano, Italy
luca.barletta@polimi.it

Elza Erkip
NYU
New York, USA
elza@nyu.edu

Yonina C. Eldar
Technion
Haifa, Israel
yonina@ee.technion.ac.il

Abstract—The capacity of a Multiple-Input Multiple-Output (MIMO) channel in which the antenna outputs are processed by an analog linear combining network and quantized by a set of threshold quantizers is studied. The linear combining weights and quantization thresholds are selected from a set of possible configurations as a function of the channel matrix. The possible configurations of the combining network model specific analog receiver architectures, such as single antenna selection, sign quantization of the antenna outputs or linear processing of the outputs. An interesting connection between the capacity of this channel and a constrained sphere packing problem in which unit spheres are packed in a hyperplane arrangement is shown. From a high-level perspective, this follows from the fact that each threshold quantizer can be viewed as a hyperplane partitioning the transmitter signal space. Accordingly, the output of the set of quantizers corresponds to the possible regions induced by the hyperplane arrangement corresponding to the channel realization and receiver configuration. This connection provides a number of important insights into the design of quantization architectures for MIMO receivers; for instance, it shows that for a given number of quantizers, choosing configurations which induce a larger number of partitions can lead to higher rates.[1]

Index Terms—MIMO, capacity, one-bit quantization, sphere packing, hybrid analog-digital receiver.

I. INTRODUCTION

As the coupling of multiple antennas and low-resolution quantization hold the promise of enabling millimeter-wave communication, the effect of finite-precision output quantization on the performance of MIMO systems has been widely investigated in recent literature. In [1], the authors propose a general framework to study the capacity of MIMO channels with various output quantization constraints and derive some initial results on the scaling of capacity in the number of available quantization levels. In this paper, we further our understanding of output quantization constraints in MIMO channels by drawing a connection between a constrained sphere packing problem and formulation in [1]. This connection suggest a rather insightful geometric-combinatorial approach to the design of receiver quantization strategies for MIMO channel with output quantization.

Literature Review: In [2], low resolution output quantization for MIMO channels is investigated through numerical evaluations. The authors are perhaps the first to note that the loss due to quantization can be relatively small. Quantization for the SISO channel is studied in detail where it is shown that, if the output is quantized using M bits, then the optimal input distribution need not have more than $M + 1$ points in its support. A cutting-plane algorithm is employed to compute this capacity and to generate optimum input support. In [3], the authors study the capacity of MIMO channels with sign quantization of the outputs and reveal a connection between a geometric-combinatorial problem and the capacity of this model at high SNR.

Contributions: In the model of [1], the output of a MIMO channel is processed by an analog combining network before being fed to N_{tq} threshold quantizers. The combining network is chosen among a set of possible configurations as a function of the channel matrix: these configurations represent analog operations that can be performed by the receiver analog front-end. Through the problem formulation in [1], it is possible to study the performance of different receiver architectures as a function of the available quantization bits N_{tq} and transmit power.

In this paper, we show that the capacity of the model in [1] can be approximately characterized using the solution of a geometric-combinatorial problem. Each threshold quantizer in effect obtains a linear combination of the noisy channel inputs and can thus be viewed as partitioning the transmit signal space with a hyperplane. The output of the set of quantizers identifies a region among those induced by the hyperplane arrangement corresponding to the channel matrix and receiver configuration. Transmitted points can be reliably distinguished at the receiver when they are separated by a hyperplane in the transmit space. Our result generalizes those of [1], [3] and provides an intuitive approach to design effective, and sometimes surprising, quantization strategies. For example, one would expect that, for a receiver able to perform linear combination before quantization, the optimal transmission strategy is to perform Singular Value Decomposition (SVD) followed by multilevel quantization of each sub-channel. We show that this scheme is actually sub-optimal at high SNR as receiver configurations which induce a larger number of partitions may lead to higher transmission rates.

Organization: The channel model is introduced in Sec. II. Combinatorial notions are presented in Sec. III. Prior results and a motivating example given in Sec. IV. The main result is presented in Sec. V. Sec. VI concludes the paper.
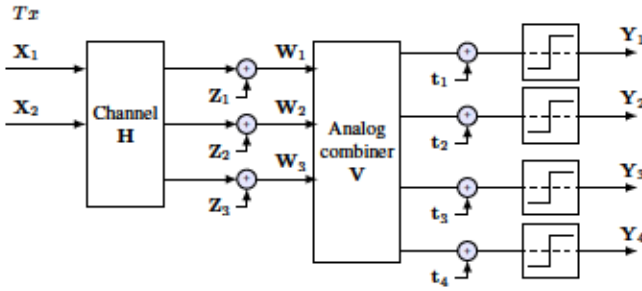

Fig. 1: System model for $N_t = 2$, $N_r = 3$, and $N_{tq} = 4$.

Notation: All logarithms are taken in base two. The vector $\text{diag}\{\mathbf{M}\}$ is the diagonal of the matrix $\mathbf{M}$ while $\lambda(\mathbf{M})$ is the vector of eigenvalues of $\mathbf{M}$. The identity matrix of size n is $\mathbf{I}_n$, the all–zero/all–one matrix of size $n \times m$ is $\mathbf{0}_{n \times m}/\mathbf{1}_{n \times m}$, respectively. Dimensions for these matrices are omitted when implied by the context. We adopt the convelntion that $\binom{n}{i} = 0$ for $i > n$.

II. CHANNEL MODEL

Consider the discrete-time MIMO channel with N_t/N_r transmit/receive antennas in which an input vector $\mathbf{X}_n = [X_{1,n} \ldots X_{N_t,n}]^T$ results in the output vector $\mathbf{W}_n = [W_{1,n} \ldots W_{N_r,n}]^T$ according to the relationship

$$\mathbf{W}_n = \mathbf{H}\mathbf{X}_n + \mathbf{Z}_n, \quad 1 \leq n \leq N, \tag{1}$$

where $\mathbf{Z}_n$ is an i.i.d. Gaussian vector of size N_r with zero mean and unit variance and $\mathbf{H}$ is a full rank matrix of size $N_r \times N_t$ (i.e. $\text{rank}(\mathbf{H}) = \min(N_t, N_r)$), fixed through the transmission block-length and known at the receiver and transmitter.[2] The input is subject to the power constraint $\sum_{n=1}^{N} \mathbb{E}[|\mathbf{X}_n|_2^2] \leq NP$ where $|\mathbf{X}_n|_2$ is the 2-norm. We study a variation of the model in (1) shown in Fig.1 in which the output vector $\mathbf{W}_n$ is processed by a receiver analog front-end and fed to N_{tq} threshold one-bit quantizers. This results in the vector $\mathbf{Y}_n = [Y_{1,n} \ldots Y_{N_{tq},n}]^T \in \{-1, +1\}^{N_{tq}}$ given by

$$\mathbf{Y}_n = \text{sign}(\mathbf{V}\mathbf{W}_n + \mathbf{t}), \quad 1 \leq n \leq N, \tag{2}$$

where $\mathbf{V}$ is the analog combining matrix of size $N_{tq} \times N_r$, $\mathbf{t}$ is a threshold vector of length N_{tq} and $\text{sign}(\mathbf{u})$ is the function producing the sign of each component of the vector $\mathbf{u}$ as plus or minus one. The matrix $\mathbf{V}$ and the vector $\mathbf{t}$ are selected among a set of available configurations $\mathcal{F}$ [1]:

$$\{\mathbf{V}, \mathbf{t}\} \in \mathcal{F} \subseteq \{\mathbb{R}^{N_{tq} \times N_r}, \mathbb{R}^{N_{tq}}\}. \tag{3}$$

For a fixed receiver configuration, $\{\mathbf{V}, \mathbf{t}\}$, the capacity of the channel in (2) is obtained as

$$C(\mathbf{V}, \mathbf{t}) = \max_{P_{\mathbf{X}}(\mathbf{x}), \; \mathbb{E}[|\mathbf{X}|_2^2] \leq P} I(\mathbf{X}; \mathbf{Y}). \tag{4}$$

We are interested in determining the largest attainable performance over all possible configurations, leading to

$$C(\mathcal{F}) = \max_{\{\mathbf{V}, \mathbf{t}\} \in \mathcal{F}} C(\mathbf{V}, \mathbf{t}). \tag{5}$$

[2]The full rank assumption is justified for richly scattering environments.

In the following, we provide an approximate characterization of the solution of the maximization in (5) under the assumption that $\text{diag}\{\mathbf{H}\mathbf{H}^T\} = \text{diag}\{\mathbf{V}\mathbf{V}^T\} = \mathbf{1}_{1 \times N_r}$. Under this assumption, the derivation of the results is particularly succinct and thus fitting to the available space. The more general case of any arbitrary channel matrix $\mathbf{H}$ and any combining matrix $\mathbf{V}$ is presented in [4].

III. COMBINATORIAL INTERLUDE

This section briefly introduces a few combinatorial concepts useful in the remainder of the paper.

A *hyperplane arrangement* $\mathcal{A}$ is a finite set of n affine hyperplanes in $\mathbb{R}^m$ for some $n, m \in \mathbb{N}$. A hyperplane arrangement $\mathcal{A} = \{\mathbf{x} \in \mathbb{R}^m, \; \mathbf{a}_i^T \mathbf{x} = b_i\}_{i=1}^n$ can be expressed as $\mathcal{A} = \{\mathbf{x}, \mathbf{A}\mathbf{x} = \mathbf{b}\}$ where $\mathbf{A}$ is obtained by letting each row i correspond to $\mathbf{a}_i^T$ and defining $\mathbf{b} = [b_1 \ldots b_n]^T$. A plane arrangement is said to be in General Position (GP) if and only if every $n \times n$ sub-matrix of $\mathbf{A}$ has non zero determinant [5].

Lemma III.1. *A hyperplane arrangement of size n in $\mathbb{R}^m$ divides $\mathbb{R}^m$ into at most $\text{r}(m, n) = \sum_{i=0}^{m} \binom{n}{i} \leq 2^n$ regions.*

Lemma III.2. *A hyperplane arrangement of size n in $\mathbb{R}^m$ where all the hyperplanes pass through the origin divides $\mathbb{R}^m$ into at most $\text{r}_0(n, m) = 2 \sum_{i=0}^{m} \binom{n-1}{i}$ regions.*

Lemma III.3. *Let $\mathcal{A}$ be a hyperplane arrangement of size l in $\mathbb{R}^m$ and consider a hyperplane arrangement $\mathcal{B}$ of size dl with $d \in \mathbb{N}$ hyperplanes parallel to each of the hyperplanes in $\mathcal{A}$. Then $\mathcal{B}$ divides $\mathbb{R}^m$ into at most $\text{r}_p(m, n, d) = \sum_{i=0}^{m} \binom{l}{i} d^i \leq (1 + d)^l$ regions.*

A necessary condition to attain the qualities in Lem. III.1, Lem. III.2 and Lem. III.3 is for the hyperplane arrangement $\mathcal{A}$ to be in GP.

A unitary sphere packing in $\mathbb{R}^m$ is defined as

$$\mathcal{P} = \bigcup_{t=1}^{N} \mathcal{S}^m(c_t, 1), \tag{6}$$

where $\mathcal{S}^m(c, r)$ is the m-dimensional hyper-sphere with center c and radius r. A hyperplane separates two spheres if each sphere belongs to one of the half-spaces induced by the hyperplane. A sphere packing $\mathcal{P}$ is said to be separable by the hyperplane arrangement $\mathcal{A}$ if any two spheres are separated by at least one hyperplane in $\mathcal{A}$. A sphere packing in a sphere is a packing $\mathcal{P}$ for which $\mathcal{P} \subseteq \mathcal{S}(c, r)$ for some c, r.

Our aim is to show a connection between the capacity in (5) and the following sphere packing problem.

Definition III.4. *Separable sphere packing in a sphere:* Given a hyperplane arrangement $\mathcal{A}$ and a constant $r \in \mathbb{R}^+$, define $\text{r}_{\text{ssps}}(\mathcal{A}, r)$ as the largest number of unit spheres in a packing $\mathcal{P}$ contained in the sphere $\mathcal{S}(0, r)$ and separable by $\mathcal{A}$.

IV. PRIOR RESULTS AND A MOTIVATING EXAMPLE

The maximization in (5) yields the optimal performance for any set of possible receiver configurations. One is often interested in studying and comparing the performance for

specific classes of receiver configurations: three such classes are studied in detail in [1]: single antenna selection and multilevel quantization, sign quantization of the outputs and linear combining and multilevel quantization.

A. Prior Results

The simplest receiver architecture of interest is perhaps the one in which a single antenna output is selected by the receiver and quantized through a high-resolution quantizer. This is obtained in the model of Sec. II by setting

$$\mathcal{F} = \left\{ \mathbf{V} = \left[\mathbf{0}_{N_{tq} \times i} \; \mathbf{1}_{N_{tq} \times 1} \; \mathbf{0}_{N_{tq} \times N_r - i - 1} \right], \; 0 \leq i \leq N_r - 1 \right.$$
$$\left. \mathbf{t} \in \mathbb{R}^{N_{tq}} \right\}, \tag{7}$$

where the term $\mathbf{1}_{N_{t_q} \times 1}$ selects the antenna with the highest channel gain.

Proposition 1. [1, Prop. 2]. *The capacity of the MIMO channel with single antenna selection and multilevel quantization is upper–bounded as*

$$C_{\text{select}} \leq \frac{1}{2} \log \left(\min \left\{ 1 + |\mathbf{h}_{\max}^T|_2^2 P, (N_{tq} + 1)^2 \right\} \right), \tag{8}$$

where $\mathbf{h}_{\max}^T$ is the row of $\mathbf{H}$ with the largest 2-norm. The upper bound in (8) can be attained to within 2 bits-per-channel-use (bpcu).

Our main result, discussed in detail in Sec. V, is inspired by an intriguing connection between Lem. III.2 and the infinite SNR capacity of the MIMO channel with sign quantization of the outputs [3]. Note that the model in [3] is obtained by setting $N_{tq} = N_r$ and letting $\mathcal{F}$ be the set of all matrices obtained by permuting the rows of $[\mathbf{I}, \mathbf{0}]$.

Proposition 2. [3, Prop. 3]. *The capacity of the MIMO channel with sign quantization of the outputs in which $\mathbf{H}$ is in GP at infinite SNR is bounded as*

$$\log(r_0(N_r, N_t)) \leq C_{\text{sign}}^{\text{SNR} \to \infty} \leq \log(r_0(N_r, N_t) + 1).$$

Recall that the most general architecture in Sec. II has

$$\mathcal{F} = \left\{ \mathbf{V} \in \mathbb{R}^{N_{tq} \times N_r}, \; \mathbf{t} \in \mathbb{R}^{N_{tq}} \right\}, \tag{9}$$

and corresponds to a receiver analog front-end which is able to perform any linear combination of the antenna outputs before quantization.

Proposition 3. [1, Prop. 6]. *The capacity of a MIMO channel with linear combining and multilevel quantization is upper–bounded as*

$$C_{\text{linear}} \leq R^\star(\lambda(\mathbf{H}), P, N_{tq}) + K. \tag{10}$$

The capacity is to within a gap of $3K$ bpcu from the upper bound in (10) for

$$R^\star(\lambda(\mathbf{H}), P, N_{tq}) =$$
$$\begin{cases} \sum_{i=1}^K \frac{1}{2} \log(1 + \lambda_i P_i) \\ \qquad \text{if } \sum_{i=1}^K \left(\sqrt{1 + \lambda_i^2 P_i} - 1 \right) \leq N_{tq} \\ K \log \left(\frac{N_{tq}}{K} + 1 \right) \\ \qquad \text{otherwise,} \end{cases} \tag{11}$$

with $K = \max\{N_t, N_r\}$, $P_i = (\mu - \lambda_i^{-2})^+$ and $\mu \in \mathbb{R}^+$ is *the smallest value for which $\sum_i P_i = P$.*

To establish the achievability of Prop. 3, the SVD can be used to transform the channel into $K = \min\{N_t, N_r\}$ parallel sub-channels with independent unit-variance additive noise and gains $\lambda(\mathbf{H})$. After SVD, the quantization strategy is chosen depending on whether the performance is bounded by the effect of the additive noise or by the quantization noise.

B. Motivating Example for the Combinatorial Approach

Let us consider the three architectures in Propositions 1-3 for the case of $N_t = 2$, $N_r = 3$ and $N_{tq} = 4$, also shown in Fig. 1, and provide some high-level intuition on the relationship between capacity and the sphere packing problem in Def. III.4.

Prop. 1: Since the threshold quantizers are used to sample the same antenna output, the number of possible outputs is at most $N_r + 1$ so that the performance in Prop. 1 is bounded by $\log(N_{tq} + 1) = \log 5 \approx 2.32$ bpcu at high SNR. This receiver configuration can be interpreted as follow: an antenna output represents a line in the two-dimensional transmit signal space; each threshold quantizer corresponds to a translation of this line and these N_{tq} parallel lines partition the signal space into at most $N_{tq} + 1$ subregions.

Prop. 2: Sign quantization of the outputs corresponds to the hyperplane arrangement in which all hyperplanes pass through the origin: the number of regions induced by this arrangement is obtained through Lem. III.2. There are $r_0 = 8$ partitions, as also shown in Fig. 2b, yielding a maximum rate of 3 bpcu, attainable at high SNR.

Prop. 3: When the receiver can perform linear combining before quantization, the SVD can be used to transform the channel into two parallel sub-channels. This strategy corresponds to the hyperplane arrangement in Lem. III.3 and the number of partitions induced is 9, as also shown in Fig. 2c.

Lem. III.1: This lemma actually indicates that the largest number of regions is 11 so that the rate $\log(11) = 3.46$ bpcu can be obtained through the receiver configuration in Fig. 2d at high SNR.[3]

Given the above interpretation of the capacity at high SNR, a feasible finite SNR strategy is the one in which, for a given receiver configuration, the channel inputs are chosen as the center of the spheres with sufficiently large radius inside each partition subject to the power constraint. The average achievable rate of the four strategies discussed above is plotted in Fig. 3. Each line in Fig. 3 corresponds to one of the sphere packing configurations in Fig. 2. For a given channel realization, $\mathbf{V}$ and $\mathbf{t}$ are chosen to result in the partitionings of the transmitter space corresponding to each of the subfigures in Fig. 2, appropriately scaled by the available transmit power. Note that the configurations are not optimized. The channel inputs are then chosen as uniformly distributed over the center of the spheres packed in the partitionings.

---

[3]Note that this does not contradict the result of Prop. 3 since the inner bound is 2 bpcu from the outer bound.
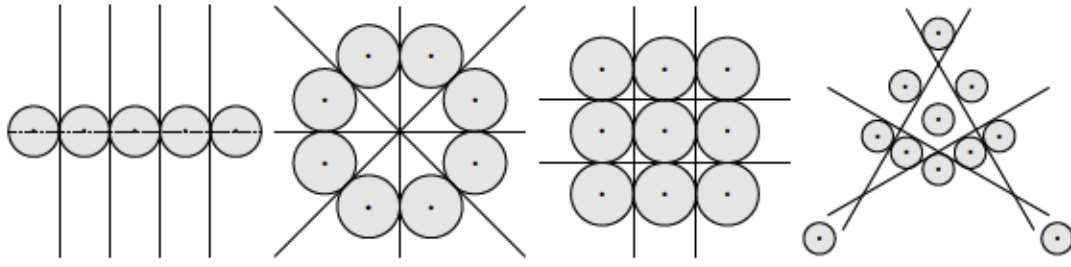
(a) Configuration corresponding to Prop.1 (b) Configuration corresponding to Prop.2 (c) Configuration corresponding to Prop.3 (d) Configuration inspired by Lem. III.1

Fig. 2: Different receiver output quantization strategies.
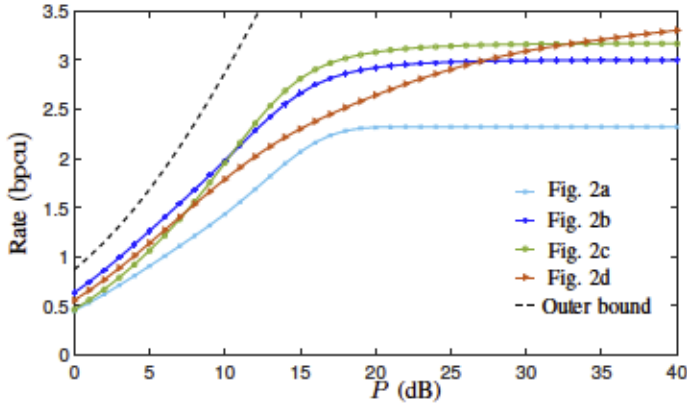
Fig. 3: Simulation results for $N_{tq} = 4$, $N_r = 3$, and $N_t = 2$ discussed in Sec. IV-B.

The average performance is calculated over real i.i.d. zero-mean, unitary variance Gaussian channel gains, further scaled to guarantee that each row has unitary 2-norm. The capacity of the channel without quantization constraint is also provided as a reference.

From Fig. 3 we see that, at high SNR, the best performance is attained by the configuration corresponding to Lem. III.1, since at high SNR the performance is determined by the number of transmitted points. As the SNR decreases, configurations with less transmitted points perform better.

V. MAIN RESULT

Sec. IV-B provides a geometric-combinatorial interpretation of the capacity of the model in (1)-(2) for the receiver architectures considered in [1]. The main result of the paper is to make such interpretation more rigorous and more general.

Theorem 1. *The capacity expression in (5) when* $\mathrm{diag}\{\mathbf{HH}^T\} = \mathrm{diag}\{\mathbf{VV}^T\} = \mathbf{1}_{1 \times N_r}$, *is upper bounded as*

$$C(\mathcal{F}) \leq \max_{\mathcal{A}} \ \log r_{ssps}(\mathcal{A}, \sqrt{P}) + \frac{3}{2}K + 3, \quad (12)$$

for

$$\mathcal{A} \in \{\mathbf{x}, \ \mathbf{VHx} = \mathbf{t}, \ (\mathbf{V}, \mathbf{t}) \in \mathcal{F}\}, \quad (13)$$

and $K = \max\{N_t, N_r\}$. *The capacity is within* $2.5N_t$ *bpcu from the outer bound in* (12).

Proof: Only the converse proof is presented here while the achievability proof is provided in [4]. Let us choose the input

and output alphabets as $\mathcal{X} = \mathcal{Y} = [0 : r(N_t, N_{tq})]$ and let the channel transition probability be determined by the channel input support and the receiver analog configuration. Also, let us define $\mathrm{sign}^*(x)$ as

$$\mathrm{sign}^*(x) = \begin{cases} x & |x| < 1 \\ \mathrm{sign}(x) & |x| \geq 1, \end{cases}$$

and the set $\overline{\mathbb{N}}^m$ as

$$\overline{\mathbb{N}}^m = \{\overline{\mathbf{n}}_{\mathbf{n}}, \ \overline{\mathbf{n}}_{\mathbf{n}} \in \mathcal{S}^m(\mathbf{n}, 1), \ \forall \ \mathbf{n} \in \mathbb{N}^m\}, \quad (14)$$

that is $\overline{\mathbb{N}}^m$ is composed of a set of points selected from the unit sphere around the integer points in $\mathbb{N}^m$. Finally, let $Q_{\overline{\mathbb{N}}^m}(\mathbf{x})$ be the mapping which assigns each point in $\mathbb{R}^m$ to the closest point in $\overline{\mathbb{N}}^m$ and

$$\overline{\mathbf{W}}^N = \mathbf{H} Q_{\overline{\mathbb{N}}^N_t}(\mathbf{X}^N), \quad \overline{\mathbf{Y}}^N = \mathrm{sign}^* \left(\mathbf{V} \overline{\mathbf{W}}^N + \mathbf{t} \right),$$

$$\mathbf{E}^N = \mathbf{W}^N - \overline{\mathbf{W}}^N.$$

Using Fano's inequality, we write

$$N(R - \epsilon_N) \leq I(\overline{\mathbf{Y}}^N, \mathbf{E}^N; \mathbf{X}^N) \quad (15a)$$

$$\leq I(\overline{\mathbf{Y}}^N; \mathbf{X}^N) + H(\mathbf{E}^N) - H(\mathbf{Z}^N)$$

$$= I(\overline{\mathbf{Y}}^N; \mathbf{X}^N) + \frac{N N_r}{2} \log \frac{3}{2}, \quad (15b)$$

where, in (15a), we have used the fact that we can reconstruct $\mathbf{Y}^N$ from $\overline{\mathbf{Y}}^N$ and the value of $\mathbf{E}^N$. In (15b), we used the fact that since $\mathrm{diag}\{\mathbf{HH}^T\} = 1$, components of $H(\mathbf{X} - Q_{\overline{\mathbb{N}}^m_t}(\mathbf{X}))$ have support at most $[-1, +1]$. The largest variance of a random variable with finite support is for the case in which the probability distribution is evenly distributed at the end points, so that $\mathrm{Var}[\mathbf{E}_i] \leq 3/2$. Using the "Gaussian maximizes entropy" property, we obtain $H(\mathbf{E}_i) \leq 1/2 \log(\pi e 3)$.

From a high-level perspective, (15) shows that the capacity of the channel in (1)-(2) is close to the capacity of the channel with no additive noise but in which the input is mapped to $\overline{\mathbb{N}}^m$. Next, we show that restricting the input to a peak power constraint, instead of an average power constraint, has a bounded effect on the capacity.

Let us represent $\mathbf{X}_i$ in hyper-geometric coordinates as $\mathbf{X}_i = \phi_i |\mathbf{X}_i|_2$ for $\phi_i \in \mathcal{S}^{N_t}(0, 1)$ and $|\phi_i|_2 = 1$ and define $\hat{\mathbf{X}}^N$ as

$$\hat{\mathbf{X}}_i = \phi_i \left(|\mathbf{X}_i|_2 \mod \lceil \sqrt{P} \rceil \right), \ 1 \leq i \leq N \quad (16)$$

where $\mod(x)$ indicates the modulus operation; in other words, $\hat{\mathbf{X}}_i$ has the same direction as $\mathbf{X}_i$ but its modulus is folded over $\lceil\sqrt{P}\rceil$. Accordingly, define

$$\widehat{\mathbf{W}}^N = \mathbf{H}Q_{\overline{\mathbf{N}}_i^N}(\hat{\mathbf{X}}^N), \quad \hat{\mathbf{Y}}^N = \text{sign}^* \left(\mathbf{V}\widehat{\mathbf{W}}^N + \mathbf{t}\right),$$

and use these definitions to further bound the term $I(\overline{\mathbf{Y}}^N; \mathbf{X}^N)$ in (15b) as

$$I(\overline{\mathbf{Y}}^N; \mathbf{X}^N) \le I(\hat{\mathbf{Y}}^N, \overline{\mathbf{Y}}^N; \hat{\mathbf{X}}^N, \mathbf{X}^N) \quad (17)$$

$$= I(\hat{\mathbf{Y}}^N; \hat{\mathbf{X}}^N) + I(\hat{\mathbf{Y}}^N; \mathbf{X}^N | \hat{\mathbf{X}}^N) \quad (18)$$

$$+ I(\overline{\mathbf{Y}}^N; \mathbf{X}^N, \hat{\mathbf{X}}^N | \hat{\mathbf{Y}}^N).$$

Note that $I(\hat{\mathbf{Y}}^N; \mathbf{X}^N | \hat{\mathbf{X}}^N) = 0$ because of the Markov chain $\hat{\mathbf{Y}}^N - \hat{\mathbf{X}}^N - \mathbf{X}^N$. For the term $I(\overline{\mathbf{Y}}^N; \mathbf{X}^N, \hat{\mathbf{X}}^N | \hat{\mathbf{Y}}^N)$ we write

$$I(\overline{\mathbf{Y}}^N; \mathbf{X}^N, \hat{\mathbf{X}}^N | \hat{\mathbf{Y}}^N) \le H(\overline{\mathbf{Y}}^N - \hat{\mathbf{Y}}^N) \quad (19a)$$

$$\le H(\mathbf{H}(\overline{\mathbf{X}}^N - \hat{\mathbf{X}}^N)) \quad (19b)$$

$$\le H(\overline{\mathbf{X}}^N - \hat{\mathbf{X}}^N), \quad (19c)$$

where (19a) follows from the fact that $\overline{\mathbf{Y}}^N$ is a discrete random variable, (19b) from the fact that $\overline{\mathbf{X}}$ and $\hat{\mathbf{X}}$ are also discrete random variables, and (19c) from the fact that $\mathbf{H}$ is full rank by assumption.

Next, to bound the term $H(\overline{\mathbf{X}}^N - \hat{\mathbf{X}}^N)$, we can write

$$\overline{\mathbf{X}}_i - \hat{\mathbf{X}}_i = \phi_i \left(|\overline{\mathbf{X}}_i|_2 / \lceil\sqrt{P}\rceil\right), \quad (20)$$

where $/$ indicates the quotient of the modulus operation. The entropy of this random variable can then be rewritten as

$$H(\overline{\mathbf{X}}^N - \hat{\mathbf{X}}^N) \le H(\phi^N) + H(|\overline{\mathbf{X}}^N|_2 / \lceil\sqrt{P}\rceil)$$

$$\le N(N_r - 1) + N \max_i H(|\overline{\mathbf{X}}_i|_2 / \lceil\sqrt{P}\rceil).$$

It can be shown that $H(|\overline{\mathbf{X}}_i|_2 / \lceil\sqrt{P}\rceil) \le 3$ bpcu: the proof follows from the fact that the power constraint can be violated only a finite number of times, which leads to the fact that $|\overline{\mathbf{X}}_i|_2 / \lceil\sqrt{P}\rceil$ is concentrated around small integer values. By combining the bounds in (15) and (17) we obtain

$$N(R - \epsilon_N) \le I(\hat{\mathbf{Y}}^N; \hat{\mathbf{X}}^N) + \frac{3}{2}NK + 3N$$

$$\le N \max_{P_{\hat{\mathbf{X}}}} I(\hat{\mathbf{Y}}; \hat{\mathbf{X}}) + \frac{3}{2}NK + 3N. \quad (21)$$

Let us now evaluate the mutual information term $I(\hat{\mathbf{Y}}^N; \hat{\mathbf{X}}^N)$, $\hat{\mathbf{Y}}^N$ is a deterministic function of $\hat{\mathbf{X}}^N$ and can be interpreted as the membership function indicating to which partition of the hyperplane arrangement the input belongs to. For this reason, $I(\hat{\mathbf{Y}}^N; \hat{\mathbf{X}}^N)$ is maximized by choosing an input support as the subset of $\overline{\mathbb{N}}^m$ in which a single point is contained in each partition induced by $\{\mathbf{VHx} = \mathbf{t}\}$ and letting the input distribution be uniformly distributed over this set. As a final step of the proof, we note that the upper bound in (21) can be minimized over the choice of the set $\overline{\mathbb{N}}^m$ in (14). In other words, by varying the choice of $\overline{\mathbf{n}}_{\mathbf{n}}$ in (14), the points in $\overline{\mathbb{N}}^m$ are moved outside the corresponding partition, thus

tightening the bound in (21). Accordingly, unless a partition contains a unit ball centered around a value $\mathbf{n} \in \mathbb{R}^{N_r}$, a value $\overline{\mathbf{n}}_{\mathbf{n}}$ can be chosen so that $\overline{\mathbb{N}}^m$ does not contain a value in such partition. It then follows that $I(\hat{\mathbf{Y}}^N; \hat{\mathbf{X}}^N) \le \log r_{\text{ssps}}(\mathcal{A}, \sqrt{P})$ which is the desired result. ∎

Remark V.1. Th. 1 extends the results in Sec. IV-A as it holds for any set of possible receiver configurations $\mathcal{F}$ in (3). The results in Sec. IV-A only hold when $\mathcal{F}$ has a specific form as in (7) or (9). On the other hand, Th. 1 does not provide a closed-form characterization of capacity as it involves the solution of a packing problem. In particular, letting $\mathcal{F}$ in (13) have the form of (7) or (9) does not immediately recover the capacity characterization in Sec. IV-A as Th. 1 follows a different approach than [1] to bound capacity.

Remark V.2. When considering the model with any $\mathbf{H}$ and $\mathbf{V}$, the result in Th. 1 generalizes as follows. The channel model in (2) is reduced to model where $\mathbf{V}$ and $\mathbf{H}$ are such that $\text{diag}\{\mathbf{HH}^T\} = \text{diag}\{\mathbf{VV}^T\} = \mathbf{1}_{1 \times N_r}$ by letting the additive noise $\mathbf{Z}_n$ have a general covariance matrix. For a channel model under this normalization, consider the additive noise after combining, $\tilde{\mathbf{Z}}_n = \mathbf{VZ}_n$: the variance of the ith entry of $\tilde{\mathbf{Z}}_n$, $\tilde{Z}_{i,n}$, determines the uncertainty in the output of the ith quantizer, $Y_{i,n}$. Accordingly, the capacity is approximatively equal to the number of separable points which can be fitted in the sphere of radius $\sqrt{P}$ such that each point is at distance at least $(\mathbb{V}\text{ar}[\tilde{Z}_{i,n}])^{1/2}$ from the ith hyperplane. The complete derivation can be found in [4].

VI. CONCLUSION

In this paper, the capacity of a MIMO channel with output quantization constraints for receivers equipped with analog combiners and one-bit threshold quantizers is investigated. The connection between the capacity of the system and a constrained sphere packing problem is showed by arguing that the threshold quantizers can be interpreted as hyperplanes partitioning the transmit signal space. This connection reveals, for example, that the infinite SNR capacity of a channel with linear combiner is attained by a receiver configuration which partitions the transmit signal space in the largest number of regions.

REFERENCES

[1] S. Rini, L. Barletta, Y. C. Eldar, and E. Erkip, "A general framework for MIMO receivers with low-resolution quantization," *Proc. IEEE Inf. Theory Workshop*, Nov. 2017.

[2] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, 2009.

[3] J. Mo and R. W. Heath, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, 2015.

[4] A. Khalili, S. Rini, L. Barletta, Y. Eldar, and E. Erkip, "A general framework for low-resolution receivers for MIMO channels," *under preparation*.

[5] T. Cover, "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition," *IEEE Trans. Electron. Comput.*, vol. EC-14, no.3, pp. 326–334, Jun. 1965.