

A Generalized Ozarow-Wyner Capacity Bound with Applications

Alex Dytso*, Mario Goldenbaum*, H. Vincent Poor*, and Shlomo Shamai (Shitz)[†]

*Department of Electrical Engineering, Princeton University

[†]Department of Electrical Engineering, Technion – Israel Institute of Technology

Abstract—In this paper, a generalized Ozarow-Wyner capacity bound is presented that holds for arbitrary noise channels. The bound is then used to approximate the capacity of a large class of additive noise channels that are subject to a p -th moment input constraint, where p is some positive real number, as well as to the Cauchy noise channel with a logarithmic moment constraint. For both channel models the gap to the capacity is precisely specified.

Index Terms—Ozarow-Wyner bound, capacity approximation, Cauchy noise, additive noise channel.

I. INTRODUCTION

One of the crown jewels of Shannon’s seminal paper [1] has been the characterization of the fundamental limit of reliable communication over an arbitrary memoryless communication channel called its capacity. The capacity is given by the mutual information between the channel input and the output maximized over a feasible set of channel input distributions. Up to the writing of this paper, however, a closed form solution to this maximization problem is known only for a handful of channel models such as the additive white Gaussian noise (AWGN) channel. While finding the optimal solution might even be impossible for certain channels, there is hope of finding a solution that is at least approximately optimal. The objective of this work is therefore to provide a simple and yet powerful tool that may lead to “good” capacity approximations. At the heart of our technique is a generalization of the Ozarow-Wyner capacity bound given in [2].

A. Related Work

The original bound of [2] was established by Ozarow and Wyner in order to demonstrate that pulse-amplitude modulation (PAM) signaling can achieve the capacity of an AWGN channel that is subject to an average transmit power constraint to within one bit. This result analytically confirmed an observation made by Ungerboeck in [3]. In [4], the bound of [2] has been sharpened and was then extended to AWGN channels with more general input constraints in [5]. The bound has also been proven to be useful for other problems such as two-user Gaussian interference channels [4], [6], communication with a disturbance constraint [7], energy harvesting problems [8], [9], and information-theoretic security [10].

This work was supported in part by the U. S. National Science Foundation under Grants CCF-1420575 and CNS-1456793, by the German Research Foundation under Grant GO 2669/1-1, and by the European Union’s Horizon 2020 Research And Innovation Programme, grant agreement no. 694630.

For the AWGN channel there exist a number of other bounds that use discrete inputs as well (see [11]–[14] and references therein). The advantage of using Ozarow-Wyner type of bounds, however, lies in their simplicity as they only depend on the number of signal constellation points and the minimum distance of the constellation.

B. Contribution and Paper Outline

In Section II, we present an Ozarow-Wyner bound that holds for arbitrary channels (not necessarily additive ones). In Section III, we use this bound to approximate the capacity of additive noise channels that are subject to a p -th moment input constraint, where p is any positive real number. In particular, we provide an approximation that is based on discrete uniformly distributed inputs and precisely determine the gap to the capacity. In Section IV, we show that our technique can also be applied to more exotic additive noise channels such as the Cauchy channel with a logarithmic moment constraint (recall that Cauchy noise has infinite variance). Finally, Section V concludes the paper.

C. Notational Remarks

Deterministic scalar/vector quantities are denoted by lowercase normal/bold letters, random variables by uppercase letters, and random vectors by bold uppercase letters. The trace of a matrix is denoted as $\text{Tr}(\cdot)$. The support of a random vector \mathbf{X} is written as $\text{supp}(\mathbf{X})$. For $x \in \mathbb{R}$, $\sec(x)$ denotes the secant and $\Gamma(x)$ the gamma function. The entropy of any discrete X is denoted as $H(X)$, whereas $h(X)$ is used whenever X is continuous. All logarithms are taken to the base e and entropies and mutual informations are measured in nats.

II. A GENERALIZED OZAROW-WYNER BOUND

Let $\mathbf{U} \in \mathbb{R}^n$ be some random vector. Then, for any $p > 0$

$$\|\mathbf{U}\|_p := \left(\frac{1}{n} \mathbb{E} \left[\text{Tr}(\mathbf{U}\mathbf{U}^T)^{\frac{p}{2}} \right] \right)^{\frac{1}{p}}. \quad (1)$$

For $n = 1$, we simply have $\|U\|_p^p = \mathbb{E}[|U|^p]$ and so we refer to $\|\mathbf{U}\|_p^p$ as the p -th moment of \mathbf{U} in what follows.

The following bound on the conditional differential entropy will be useful throughout the rest of the paper.

Theorem 1. *Let $p \in (0, \infty)$ be arbitrary and $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a measurable function. Then, for any pair of random vectors*

$\mathbf{U} \in \mathbb{R}^n$ and $\mathbf{V} \in \mathbb{R}^n$ such that $h(\mathbf{U}) < \infty$ and $\|\mathbf{U}\|_p < \infty$, we have

$$\frac{1}{n}h(\mathbf{U}|\mathbf{V}) \leq \log\left(k_{n,p} \cdot n^{\frac{1}{p}} \cdot \|\mathbf{U} - g(\mathbf{V})\|_p\right),$$

where

$$k_{n,p} := \frac{\sqrt{\pi} e^{\frac{1}{p}} \left(\frac{p}{n}\right)^{\frac{1}{p}} \Gamma\left(\frac{n}{p} + 1\right)^{\frac{1}{n}}}{\Gamma\left(\frac{n}{2} + 1\right)^{\frac{1}{n}}}. \quad (2)$$

Proof: The proof follows from the proof of [5, Th. 1]. ■

Now, consider any time-discrete channel determined by the conditional probability distribution $P_{\mathbf{Y}|\mathbf{X}}$, where $\mathbf{X} \in \mathbb{R}^n$ denotes the length- n channel input sequence and $\mathbf{Y} \in \mathbb{R}^n$ the corresponding sequence of channel outputs.¹ We make use of Theorem 1 in order to obtain a novel bound on the capacity of $P_{\mathbf{Y}|\mathbf{X}}$ if the inputs are restricted to be discrete, which is a generalization of the Ozarow-Wyner bound given in [2].

Theorem 2. (Generalized Ozarow-Wyner Bound) *Let $\mathbf{X}_D \in \text{supp}(\mathbf{X}_D) \subset \mathbb{R}^n$ be a discrete random vector of finite entropy, $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a measurable function, and $p > 0$. Furthermore, let \mathcal{K}_p be a set of continuous random vectors, independent of \mathbf{X}_D , such that for every $\mathbf{U} \in \mathcal{K}_p$, $h(\mathbf{U}), \|\mathbf{U}\|_p < \infty$, and*

$$\text{supp}(\mathbf{U} + \mathbf{x}_i) \cap \text{supp}(\mathbf{U} + \mathbf{x}_j) = \emptyset \quad (3)$$

for all $\mathbf{x}_i, \mathbf{x}_j \in \text{supp}(\mathbf{X}_D)$, $i \neq j$. Then, for every channel $P_{\mathbf{Y}|\mathbf{X}}$ we have

$$[H(\mathbf{X}_D) - \text{gap}^*]^+ \leq I(\mathbf{X}_D; \mathbf{Y}) \leq H(\mathbf{X}_D),$$

where

$$\text{gap}^* := \inf_{\substack{\mathbf{U} \in \mathcal{K}_p \\ g \text{ measurable} \\ p > 0}} \underbrace{(G_{1,p}(\mathbf{U}, \mathbf{X}_D, g) + G_{2,p}(\mathbf{U}))}_{=:\text{gap}(\mathbf{U}, g, p)} \quad (4)$$

with

$$G_{1,p}(\mathbf{U}, \mathbf{X}_D, g) := n \log\left(\frac{\|\mathbf{U} + \mathbf{X}_D - g(\mathbf{Y})\|_p}{\|\mathbf{U}\|_p}\right), \quad (5)$$

$$G_{2,p}(\mathbf{U}) := n \log\left(\frac{k_{n,p} \cdot n^{\frac{1}{p}} \cdot \|\mathbf{U}\|_p}{e^{\frac{1}{n}h(\mathbf{U})}}\right). \quad (6)$$

Proof: Let \mathbf{U} and \mathbf{X}_D be statistically independent. Then, the mutual information $I(\mathbf{X}_D; \mathbf{Y})$ can be lower bounded as

$$\begin{aligned} I(\mathbf{X}_D; \mathbf{Y}) &\stackrel{a)}{\geq} I(\mathbf{X}_D + \mathbf{U}; \mathbf{Y}) \\ &= h(\mathbf{X}_D + \mathbf{U}) - h(\mathbf{X}_D + \mathbf{U}|\mathbf{Y}) \\ &\stackrel{b)}{=} H(\mathbf{X}_D) + h(\mathbf{U}) - h(\mathbf{X}_D + \mathbf{U}|\mathbf{Y}). \end{aligned} \quad (7)$$

Here, *a)* follows from the data processing inequality as $\mathbf{X}_D + \mathbf{U} \rightarrow \mathbf{X}_D \rightarrow \mathbf{Y}$ forms a Markov chain in that order, and *b)* from the assumption in (3).

By using Theorem 1, we have that the last term in (7) can be bounded from above as

$$h(\mathbf{X}_D + \mathbf{U}|\mathbf{Y}) \leq n \log\left(k_{n,p} \cdot n^{\frac{1}{p}} \cdot \|\mathbf{X}_D + \mathbf{U} - g(\mathbf{Y})\|_p\right).$$

¹For Theorem 2 to be valid we do not need \mathbf{Y} to be from \mathbb{R}^n . Actually, \mathbf{Y} can be an element of any space.

Combining this expression with (7) results in

$$I(\mathbf{X}_D; \mathbf{Y}) \geq H(\mathbf{X}_D) - (G_{1,p}(\mathbf{U}, \mathbf{X}_D, g) + G_{2,p}(\mathbf{U})),$$

with $G_{1,p}$ and $G_{2,p}$ as defined in (5) and (6), respectively. Maximizing the right-hand side over all $\mathbf{U} \in \mathcal{K}_p$, measurable functions $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$, and $p > 0$ provides the bound. ■

Remark 1. Note that the bound in Theorem 2 is very general in the sense that no assumptions have been made about $P_{\mathbf{Y}|\mathbf{X}}$.

Remark 2. In the context of the AWGN channel, the term $G_{2,p}$ corresponds to the *shaping loss*, which vanishes as n tends to infinity (i.e., $G_{2,p} \in \mathcal{O}(\log(n)/n)$).

III. ADDITIVE NOISE CHANNELS WITH p -TH MOMENT INPUT CONSTRAINT

In this section, we demonstrate the usefulness of the generalized Ozarow-Wyner bound of Theorem 2 by applying it to additive noise channels whose inputs are constrained in their p -th moment. More precisely, let $p > 0$ and $A \geq 0$ be real numbers and consider the channel model

$$Y = X + Z, \quad \mathbb{E}[|X|^p] \leq A \quad (8)$$

for some noise variable Z such that $|h(Z)| < \infty$ and where Z is independent of X .² Based on Theorems 1 and 2, in the following two subsections we provide an upper bound on the capacity of (8) as well as a bound on the gap to the capacity.

A. Capacity Upper Bound

Theorem 3. *Let \mathcal{F} be a convex and compact set of channel input distributions that fulfill the p -th moment constraint $\mathbb{E}[|X|^p] \leq A$ and let C denote the capacity of the additive noise channel given in (8). Then,*

$$C \leq \inf_{p>0} \frac{1}{p} \log\left(k_{1,p}^p \cdot e^{-ph(Z)} \mathbb{E}[|X^* + Z|^p]\right)$$

for some $X^* \in \mathcal{F}$ (dependent on p) and $k_{1,p}$ as defined in (2).

Proof: By means of Theorem 1, we can upper bound the mutual information between the channel input and output as

$$\begin{aligned} I(X; Y) &= h(X + Z) - h(Z) \\ &\leq \frac{1}{p} \log(k_{1,p}^p \cdot \mathbb{E}[|X + Z|^p]) - h(Z) \\ &= \frac{1}{p} \log\left(k_{1,p}^p \cdot e^{-ph(Z)} \mathbb{E}[|X + Z|^p]\right). \end{aligned}$$

This leads immediately to an upper bound on the capacity:

$$\begin{aligned} C &\leq \max_{X \in \mathcal{F}} \inf_{p>0} \frac{1}{p} \log\left(k_{1,p}^p \cdot e^{-ph(Z)} \mathbb{E}[|X + Z|^p]\right) \\ &\stackrel{a)}{\leq} \inf_{p>0} \max_{X \in \mathcal{F}} \frac{1}{p} \log\left(k_{1,p}^p \cdot e^{-ph(Z)} \mathbb{E}[|X + Z|^p]\right) \\ &\stackrel{b)}{=} \inf_{p>0} \frac{1}{p} \log\left(k_{1,p}^p \cdot e^{-ph(Z)} \mathbb{E}[|X^* + Z|^p]\right). \end{aligned}$$

²For $p = 2$, the input constraint in (8) is simply an average transmit power constraint.

Note that *a*) follows from the fact that $\max \inf \leq \inf \max$ and *b*) since \mathcal{F} is convex and compact and $\log(k_{1,p}^p \cdot e^{-ph(Z)} \mathbb{E}[|X+Z|^p])$ is concave in the distribution of X , we have that for a given p the maximizer X^* exists. ■

B. Gap to Capacity

As evaluating the capacity bound of Theorem 3 is notoriously difficult, in the following we determine the gap to the capacity under some simplifying assumptions. In doing so, let

$$C(A, p, q) := \max_{X: \mathbb{E}[|X|^p] \leq A} I(X; Y),$$

where (if existent)

$$q := \sup\{q' > 0 : \mathbb{E}[|Z|^{q'}] < \infty\}. \quad (9)$$

An important ingredient to specify the gap will be PAM.

Definition 1. Let N be a positive integer and P some non-negative real number. Then,

$$\text{PAM}(N, P) := \begin{cases} 0, \pm 2\Delta, \pm 4\Delta, \dots, \pm(N-1)\Delta, & N \text{ odd} \\ \pm\Delta, \pm 3\Delta, \pm 5\Delta, \dots, \pm(N-1)\Delta, & N \text{ even} \end{cases}$$

with

$$\Delta := \sqrt{3P/(N^2 - 1)}, \quad (10)$$

refers to a *PAM constellation* with N equally spaced points and average power P . Moreover, we use $X_D \sim \text{PAM}(N, P)$ to denote that the discrete random variable X_D is uniformly distributed over $\text{PAM}(N, P)$.

Remark 3. Notice that the Euclidean distance between adjacent points in $\text{PAM}(N, P)$ is 2Δ . It can be easily verified that $\mathbb{E}[X_D^2] = P$ if $X_D \sim \text{PAM}(N, P)$.

Theorem 4. Let the noise, Z , be distributed such that (9) exists and let $X_D \sim \text{PAM}(N, P)$ with

$$N = \left\lfloor \left(k_{1,r}^r e^{-rh(Z)} (\mathbb{E}[|X_D|^r] + \mathbb{E}[|Z|^r]) \right)^{\frac{1}{r}} \right\rfloor \quad (11)$$

and P chosen such that $\mathbb{E}[|X_D|^r] = \mathbb{E}[|X^*|^r]$, where X^* as defined in Theorem 3. Then, for $A \geq 1$ and any $p > 0$

$$C(A, p, q) - I(X_D; X_D + Z) \leq \inf_{\substack{0 < r \leq \min(p, q) \\ 0 < s < \infty}} \text{Gap}(r, s)$$

with Gap as defined in (12) at the top of the next page.

Proof: Fix r such that $0 < r \leq \min(p, q)$, choose X^* according to Theorem 3, and note that by Jensen's inequality we have that

$$\mathbb{E}[|X^*|^r] \leq \mathbb{E}[|X^*|^p]^{\frac{r}{p}} \leq A^{\frac{r}{p}}.$$

Let $X_D \sim \text{PAM}(N, P)$ with N as in (11) and P chosen in the way that $\mathbb{E}[|X_D|^r] = \mathbb{E}[|X^*|^r]$. Observe (see Remark 3)

that the squared Euclidean distance between any given pair of adjacent points in $\text{PAM}(N, P)$ can be bounded from below as

$$\begin{aligned} (2\Delta)^2 &= \frac{12P}{N^2 - 1} \\ &= \frac{12\mathbb{E}[X_D^2]}{\left[\left(k_{1,r}^r e^{-rh(Z)} (\mathbb{E}[|X_D|^r] + \mathbb{E}[|Z|^r]) \right)^{\frac{1}{r}} \right]^2 - 1} \\ &\stackrel{a)}{\geq} 12k_{1,r}^{-2} e^{2h(Z)} \mathbb{E}[X_D^2] (\mathbb{E}[|X_D|^r] + \mathbb{E}[|Z|^r])^{-\frac{2}{r}} \\ &\stackrel{b)}{\geq} 12k_{1,r}^{-2} e^{2h(Z)} \mathbb{E}[|X_D|^r]^{\frac{2}{r}} (\mathbb{E}[|X_D|^r] + \mathbb{E}[|Z|^r])^{-\frac{2}{r}} \\ &= 12k_{1,r}^{-2} e^{2h(Z)} \left(1 + \frac{\mathbb{E}[|Z|^r]}{\mathbb{E}[|X_D|^r]} \right)^{-\frac{2}{r}} \\ &= 12k_{1,r}^{-2} e^{2h(Z)} \left(1 + \frac{\mathbb{E}[|Z|^r]}{A^r} \right)^{-\frac{2}{r}} \\ &\stackrel{c)}{\geq} 12k_{1,r}^{-2} e^{2h(Z)} \underbrace{\left(1 + \mathbb{E}[|Z|^r] \right)^{-\frac{2}{r}}}_{=: \mu_r(Z)^{-\frac{2}{r}}}, \end{aligned} \quad (13)$$

where the inequalities follow from: *a*) dropping the floor function and the minus one; *b*) using the bound $\mathbb{E}[|X_D|^r] \leq \mathbb{E}[X_D^2]^{\frac{r}{2}}$; and *c*) using that $A \geq 1$.

Now, let $\xi_r := \max(2^{r-1}, 1)$ and observe that

$$\begin{aligned} I(X_D; X_D + Z) + \text{gap}(U, g, r) &\geq \log(N) \\ &= \log \left(\left[\left(k_{1,r}^r e^{-rh(Z)} (\mathbb{E}[|X_D|^r] + \mathbb{E}[|Z|^r]) \right)^{\frac{1}{r}} \right] \right) \\ &\stackrel{a)}{\geq} \frac{1}{r} \log \left(k_{1,r}^r e^{-rh(Z)} (\mathbb{E}[|X_D|^r] + \mathbb{E}[|Z|^r]) \right) - \frac{1}{r} \log(2) \\ &\stackrel{b)}{=} \frac{1}{r} \log \left(k_{1,r}^r e^{-rh(Z)} (\mathbb{E}[|X^*|^r] + \mathbb{E}[|Z|^r]) \right) - \frac{1}{r} \log(2) \\ &\stackrel{c)}{\geq} \frac{1}{r} \log \left(k_{1,r}^r e^{-rh(Z)} \mathbb{E}[|X^* + Z|^r] \right) - \frac{1}{r} \log(2\xi_r) \\ &\stackrel{d)}{\geq} C(A, p, q) - \frac{1}{r} \log(2\xi_r). \end{aligned} \quad (14)$$

Here, *a*) follows from the fact that $\lfloor x \rfloor \geq \frac{x}{2}$ for every $x \geq 1$, *b*) from $\mathbb{E}[|X_D|^r] = \mathbb{E}[|X^*|^r]$, *c*) by using the bound $\mathbb{E}[|X^* + Z|^r] \leq \xi_r (\mathbb{E}[|X^*|^r] + \mathbb{E}[|Z|^r])$, and *d*) by means of Theorem 3.

We proceed by letting U be a continuous random variable that is uniformly distributed over the interval $[-\Delta, \Delta]$ and independent of X_D and by choosing $g(y) \equiv y$. Then, based on Theorem 2, we upper bound the gap term as follows:³

$$\begin{aligned} \text{gap}(U, g, s) &= \log \left(\frac{k_{1,s} \|U\|_s}{e^{h(U)}} \right) + \log \left(\frac{\|X_D + U - Y\|_s}{\|U\|_s} \right) \\ &\stackrel{a)}{\leq} \log \left(\frac{k_{1,s} \mathbb{E}[|U|^s]^{\frac{1}{s}}}{e^{h(U)}} \right) + \frac{1}{s} \log \left(\xi_s + \frac{\mathbb{E}[|X_D - Y|^s]}{\mathbb{E}[|U|^s]} \right) \\ &= \log \left(\frac{k_{1,s} \mathbb{E}[|U|^s]^{\frac{1}{s}}}{e^{h(U)}} \right) + \frac{1}{s} \log \left(\xi_s + \frac{\mathbb{E}[|Z|^s]}{\mathbb{E}[|U|^s]} \right) \end{aligned}$$

³The parameter s can be chosen independently of r and therefore provides an additional degree of freedom in order to minimize the gap to capacity.

$$\text{Gap}(r, s) := \log\left(\frac{k_{1,s}}{2(s+1)^{\frac{1}{s}}}\right) + \frac{1}{s} \log\left(\max(2^{s-1}, 1) + \frac{k_{1,r}^s 2^s (s+1) \mathbb{E}[|Z|^s] (1 + \mathbb{E}[|Z|^r])^{\frac{s}{r}}}{12^{\frac{s}{2}} e^{sh(Z)}}\right) + \frac{1}{r} \log(2 \max(2^{r-1}, 1)) \quad (12)$$

$$\begin{aligned} & \stackrel{b)}{=} \log\left(\frac{k_{1,s}}{2(s+1)^{\frac{1}{s}}}\right) + \frac{1}{s} \log\left(\xi_s + \frac{2^s (s+1) \mathbb{E}[|Z|^s]}{(2\Delta)^s}\right) \\ & \stackrel{c)}{\leq} \log\left(\frac{k_{1,s}}{2(s+1)^{\frac{1}{s}}}\right) + \frac{1}{s} \log\left(\xi_s + \frac{k_{1,r}^s 2^s (s+1) \mathbb{E}[|Z|^s]}{12^{\frac{s}{2}} e^{sh(Z)} \mu_r(Z)^{-\frac{s}{r}}}\right), \end{aligned} \quad (15)$$

where the (in)equalities follow from: *a)* using $\mathbb{E}[|V+U|^s] \leq \max(2^{s-1}, 1)(\mathbb{E}[|V|^s] + \mathbb{E}[|U|^s])$, which holds for any U and V ; *b)* using the facts that $\mathbb{E}[|U|^s] = \frac{(2\Delta)^s}{2^s(s+1)}$ and $h(U) = \log(2\Delta)$; and *c)* using the bound in (13).

Combining (14) with (15) and taking the infimum over r and s proves the theorem. ■

C. Discussion

In the special case of $Z \sim \mathcal{N}(0, \sigma^2)$ and $p = 2$ (i.e., an average power constraint $\mathbb{E}[|X|^2] \leq P$) the number of constellation points (i.e., codewords) has to be chosen as

$$N \approx \sqrt{\frac{\mathbb{E}[|X|^2] + \mathbb{E}[|Z|^2]}{\mathbb{E}[|Z|^2]}} = \sqrt{\frac{P + \sigma^2}{\sigma^2}} = e^{C(P, \sigma^2)},$$

where $C(P, \sigma^2)$ refers to the AWGN channel capacity. For arbitrary Z and p , however, the number of constellation points has to be chosen such that

$$N \approx \left(k_{1,r}^r e^{-rh(Z)} (\mathbb{E}[|X+Z|^r])\right)^{\frac{1}{r}} \approx e^{C(A, p, q)},$$

$0 < r \leq \min(p, q)$, which is the reasoning behind (11). The product $k_{1,r}^r e^{-rh(Z)}$ can be interpreted as the r -th *entropy power* of Z . In [15] it has been shown that it is related to $\mathbb{E}[|Z|^r]$ in the sense that

$$\mathbb{E}[|Z|^r] \geq k_{1,r}^{-r} e^{rh(Z)}. \quad (16)$$

If a one dimensional signal constellation is used for the AWGN channel, the asymptotic gap to the capacity is

$$\varrho_L(2, q) := \lim_{A \rightarrow \infty} (C(A, 2, q) - I(X_D; Y)) = \frac{1}{2} \log\left(\frac{\pi e}{3}\right).$$

This high power loss can be decomposed into the *shaping loss* $\frac{1}{2} \log\left(\frac{\pi e}{6}\right)$ and the *coding loss* $\frac{1}{2} \log(2)$. Similarly, we may define the asymptotic loss of an arbitrary additive noise channel with a p -th moment constraint as

$$\varrho_L(p, q) := \lim_{A \rightarrow \infty} (C(A, p, q) - I(X_D; Y)).$$

With the techniques used in the proof of Theorem 4 one can show that

$$\begin{aligned} \varrho_L(p, q) & \leq \inf_{\substack{0 < r \leq \min(p, q) \\ 0 < s < \infty}} \left[\log\left(\frac{k_{1,s}}{2(s+1)^{\frac{1}{s}}}\right) \right. \\ & \quad \left. + \frac{1}{s} \log\left(\max(2^{s-1}, 1) + \frac{k_{1,r}^s 2^s (s+1) \mathbb{E}[|Z|^s]}{12^{\frac{s}{2}} e^{sh(Z)}}\right) \right]. \end{aligned}$$

Moreover, for channels where we approximately have equality in (16) and where $2 \leq \min(p, q)$ it follows that

$$\varrho_L(p, q) \leq \log\left(\frac{k_{1,2}}{2\sqrt{3}}\right) + \frac{1}{2} \log(2) = \frac{1}{2} \log\left(\frac{\pi e}{3}\right),$$

showing that the AWGN channel with a second moment constraint has the worst asymptotic loss.

IV. CAUCHY-NOISE CHANNEL WITH A LOGARITHMIC MOMENT CONSTRAINT

Now, we apply the generalized Ozarow-Wyner bound of Theorem 2 to an additive noise channel with a different type of constraint: the *Cauchy-noise channel with a logarithmic moment constraint*.⁴ The output of this channel is given by $Y = X + Z$ with Z being a Cauchy random variable of density

$$p_Z(z) = \frac{1}{\pi\gamma} \left(\frac{\gamma^2}{z^2 + \gamma^2} \right), \quad z \in \mathbb{R},$$

$\gamma > 0$, whereas the channel input is restricted to the space

$$\mathcal{F} := \left\{ X : \mathbb{E} \left[\log \left(\left(\frac{A+\gamma}{A} \right)^2 + \left(\frac{X}{A} \right)^2 \right) \right] \leq \log(4) \right\}. \quad (17)$$

In [16], it has been shown that for $A \geq \gamma$, the capacity of the Cauchy-noise channel with logarithmic moment constraint is

$$C(A, \gamma) := \max_{X \in \mathcal{F}} I(X; Y) = \log\left(\frac{A}{\gamma}\right),$$

which is achieved by a Cauchy-distributed input. Under a mild assumption on the ratio A/γ , however, using $X_D \sim \text{PAM}(N, P)$ as the channel input results in a gap to the capacity of at most 2.5 nats.

Theorem 5. *Let $X_D \sim \text{PAM}(N, P)$ with $N = \lfloor A/\gamma \rfloor$ and*

$$P = A^2 \left(4 - \left(\frac{A+\gamma}{A} \right)^2 \right). \quad (18)$$

Then, $X_D \in \mathcal{F}$ and for $\frac{A}{\gamma} \geq \frac{1}{\sqrt{2}-1}$

$$C(A, \gamma) - I(X_D; Y) \leq \min_{p \in (0, 1)} \text{Gap}^C(p) \approx 2.5 \text{ nats},$$

with Gap^C as defined in (19) at the top of the next page.

Proof: The fact that $X_D \in \mathcal{F}$ follows immediately by using Jensen's inequality:

$$\begin{aligned} & \mathbb{E} \left[\log \left(\left(\frac{A+\gamma}{A} \right)^2 + \left(\frac{X_D}{A} \right)^2 \right) \right] \\ & \leq \log \left(\left(\frac{A+\gamma}{A} \right)^2 + \frac{\mathbb{E}[X_D^2]}{A^2} \right) = \log \left(\left(\frac{A+\gamma}{A} \right)^2 + \frac{P}{A^2} \right). \end{aligned}$$

⁴The Cauchy-noise channel with a p -th moment input constraint is covered by Theorem 4.

$$\text{Gap}^C(p) := \log\left(\frac{(pe)^{\frac{1}{p}} \Gamma(\frac{1}{p} + 1)}{(p+1)^{\frac{1}{p}}}\right) + \frac{1}{p} \log\left(1 + \frac{\sec(\frac{\pi p}{2})(p+1)}{6^{\frac{p}{2}}}\right) - \log(2 - \sqrt{2}) \quad (19)$$

Hence, choosing P as in (18) satisfies the constraint in (17).

With this choice of the channel input let U be a random variable that is independent of X_D and uniformly drawn from $[-\Delta, \Delta)$. Note that for the Cauchy noise Z , only fractional absolute moments exist (i.e., $\mathbb{E}[|Z|^p]$ exists only for $p \in (0, 1)$). Thus, following along similar lines as in the proof of Theorem 4, we have that

$$\begin{aligned} & \text{gap}(U, g, p) \\ & \leq \log\left(\frac{k_{1,p} \mathbb{E}[|U|^p]^{\frac{1}{p}}}{e^{h(U)}}\right) + \frac{1}{p} \log\left(1 + \frac{\mathbb{E}[|Z|^p]}{\mathbb{E}[|U|^p]}\right) \\ & \stackrel{a)}{=} \log\left(\frac{k_{1,p}}{2(p+1)^{\frac{1}{p}}}\right) + \frac{1}{p} \log\left(1 + \frac{\gamma^p \sec(\frac{\pi p}{2}) 2^p (p+1)}{(2\Delta)^p}\right) \\ & \stackrel{b)}{\leq} \log\left(\frac{k_{1,p}}{2(p+1)^{\frac{1}{p}}}\right) + \frac{1}{p} \log\left(1 + \frac{\gamma^p \sec(\frac{\pi p}{2})(p+1)N^p}{(3P)^{\frac{p}{2}}}\right) \\ & \stackrel{c)}{\leq} \log\left(\frac{k_{1,p}}{2(p+1)^{\frac{1}{p}}}\right) + \frac{1}{p} \log\left(1 + \frac{\sec(\frac{\pi p}{2})(p+1)A^p}{(3P)^{\frac{p}{2}}}\right) \\ & \stackrel{d)}{=} \log\left(\frac{k_{1,p}}{2(p+1)^{\frac{1}{p}}}\right) + \frac{1}{p} \log\left(1 + \frac{\sec(\frac{\pi p}{2})(p+1)}{3^{\frac{p}{2}} \left(4 - \left(\frac{A+\gamma}{A}\right)^2\right)^{\frac{p}{2}}}\right) \\ & \stackrel{e)}{\leq} \log\left(\frac{k_{1,p}}{2(p+1)^{\frac{1}{p}}}\right) + \frac{1}{p} \log\left(1 + \frac{\sec(\frac{\pi p}{2})(p+1)}{6^{\frac{p}{2}}}\right), \quad (20) \end{aligned}$$

where $a)$ follows from the facts $\mathbb{E}[|Z|^p] = \gamma^p \sec(\frac{\pi p}{2})$, $\mathbb{E}[|U|^p] = \frac{(2\Delta)^p}{2^p(p+1)}$ and $h(U) = \log(2\Delta)$, $b)$ from (10), $c)$ from using $N = \lfloor A/\gamma \rfloor$ and dropping the floor function, $d)$ from inserting (18), and $e)$ from $4 - \left(\frac{A+\gamma}{A}\right)^2 \geq 2$ for $\frac{A}{\gamma} \geq \frac{1}{\sqrt{2}-1}$.

Furthermore, we have

$$\begin{aligned} I(X_D; Y) + \text{gap}(U, g, p) & \geq \log(\lfloor A/\gamma \rfloor) \\ & \stackrel{a)}{\geq} C(A, \gamma) + \log\left(1 - \frac{\gamma}{A}\right) \\ & \stackrel{b)}{\geq} C(A, \gamma) + \log(2 - \sqrt{2}). \quad (21) \end{aligned}$$

Here, $a)$ follows from $\lfloor x \rfloor \geq x - 1$, which is valid for all $x \in \mathbb{R}$, and $b)$ from the assumption $\frac{A}{\gamma} \geq \frac{1}{\sqrt{2}-1}$. Inserting (2) and combining (20) with (21) results in (19) at the top of the page. It can be shown that Gap^C is a strictly convex function attaining its minimum at $p \approx 0.6993$. This proves the result. ■

V. CONCLUSIONS

While finding exact capacity expressions can be difficult, in this work we have provided a very general yet simple achievable scheme that results in approximate capacity expressions valid for a large class of additive noise channels under very general input constraints. The results are based on

a novel generalization of the Ozarow-Wyner bound. As the generalization does not impose assumptions on the structure of the channel, we believe it can also be applied to non-additive noise channels (e.g., the Poisson channel, the optical intensity channel), channels with memory, fading channels, intersymbol interference channels, as well as to channels with side information. For instance, we have recently shown that the bound is very useful in obtaining bounds on the capacity of *amplitude-constrained* multiantenna channels [17].

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3 & 4, pp. 379–423 & pp. 623–656, 1948.
- [2] L. Ozarow and A. Wyner, "On the capacity of the Gaussian channel with a finite number of input levels," *IEEE Trans. Inf. Theory*, vol. 36, no. 6, pp. 1426–1428, 1990.
- [3] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, 1982.
- [4] A. Dytso, D. Tuninetti, and N. Devroye, "Interference as noise: Friend or foe?" *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3561–3596, 2016.
- [5] A. Dytso, R. Bustin, D. Tuninetti, N. Devroye, S. Shamai, and H. V. Poor, "On the minimum mean p -th error in Gaussian noise channels and its applications." Submitted to *IEEE Trans. Inf. Theory*, <https://arxiv.org/abs/1607.01461>, 2016.
- [6] A. Dytso, D. Tuninetti, and N. Devroye, "On the two-user interference channel with lack of knowledge of the interference codebook at one receiver," *IEEE Trans. Inf. Theory*, vol. 61, no. 3, pp. 1257–1276, 2015.
- [7] A. Dytso, R. Bustin, D. Tuninetti, N. Devroye, S. Shamai, and H. V. Poor, "New bounds on MMSE and applications to communication with the disturbance constraint." Submitted to *IEEE Trans. Inf. Theory*, <https://arxiv.org/pdf/1603.07628>, 2016.
- [8] Y. Dong, F. Farnia, and A. Özgür, "Near optimal energy control and approximate capacity of energy harvesting communication," *IEEE J. Select. Areas Commun.*, vol. 33, no. 3, pp. 540–557, 2015.
- [9] D. Shaviv, P.-M. Nguyen, and A. Özgür, "Capacity of the energy-harvesting channel with a finite battery," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6436–6458, 2016.
- [10] M. Bloch, J. Barros, M. R. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, 2008.
- [11] A. Alvarado, F. Brannstrom, E. Agrell, and T. Koch, "High-SNR asymptotics of mutual information for discrete constellations with applications to BICM," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1061–1076, 2014.
- [12] Y. Wu and S. Verdú, "The impact of constellation cardinality on Gaussian channel capacity," in *Proc. 48th Annu. Allerton Conf. Commun., Control and Comp.*, Monticello, IL, Sep. 2010, pp. 620–628.
- [13] A. Dytso, D. Tuninetti, and N. Devroye, "On discrete alphabets for the two-user Gaussian interference channel with one receiver lacking knowledge of the interfering codebook," in *Proc. Inf. Theory and Appl. Workshop*, San Diego, CA, 2014, pp. 1–8.
- [14] K. Moshksar, A. Ghasemi, and A. Khandani, "An alternative to decoding interference or treating interference as Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 305–322, 2015.
- [15] E. Lutwak, D. Yang, and G. Zhang, "Moment-entropy inequalities for a random vector," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1603–1607, 2007.
- [16] J. Fahn and I. Abou-Faycal, "A Cauchy input achieves the capacity of a Cauchy channel under a logarithmic constraint," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, 2014, pp. 3077–3081.
- [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.*, Singapore, 2017, submitted.