An Upper Bound on the Capacity of Censored Channels

Gholamreza Alirezaei and Rudolf Mathar
Institute for Theoretical Information Technology
RWTH Aachen University, D-52056 Aachen, Germany
{alirezaei, mathar}@ti.rwth-aachen.de

***Abstract*—In this paper, we consider a channel which is linear over the interval $[0, 1]$ and is censored to the left by zero and to the right by one. Examples of this channel type are radio frequency amplifiers which amplify only up to certain thresholds. In the baseband, this channel is a model for censoring symbols whenever they exceed given thresholds. One-bit quantization may be seen as an extreme case when the right censoring bound converges to the left one. Determining mutual information and capacity of this channel is a fundamental information theoretic problem which seems to be unsolved in general. One reason seems to be that the output distribution has two mass points at the bounds of the censoring interval and can be continuous within the linear region. In this paper, we provide a compact formula for mutual information of this channel. Furthermore, an upper bound for the capacity of this channel is given. Finally, selected numerical results for additive uniformly distributed and Gaussian noise are presented to evaluate the accuracy of the bound.**

I. Introduction and Motivation

A channel is called *censored* if a noisy input signal is transmitted unaltered within certain bounds, and is clipped to a maximum or minimum value whenever the bounds are exceeded.*Truncation* should not be confused with censoring data or distribution. In the first case data outside a certain interval never occur while in the latter case data exceeding an interval are mapped onto the corresponding interval boundary, cf. [1]. To keep the formal apparatus low, in this paper we consider censoring at 0 and 1, respectively, leading to the function $Q(z)$, described in (1) and illustrated in Fig. 1.

This channel may serve as an approximate model for a nonlinear amplifier which is unable to amplify beyond certain boundaries and in this case simply generates maximum or minimum possible power. Moreover, the peak-to-average power ratio (PAPR) problem for orthogonal frequency-division multiplexing (OFDM) symbols may be analyzed by this approach, if clipping of the signal is applied not to exceed the limits of the system. Furthermore, if baseband signals, exceeding a certain threshold, cannot be properly decoded but are returned as threshold values, this model seems to be an appropriate description.

Mutual information and the capacity of this channel are of high interest from a practical and theoretical point of view, since the censored channel is a valuable member for the class of nonlinear channels, cf. [2]. However, mutual information is not easily accessible and to the best knowledge of the authors, there is neither a closed-form formula nor a systematic evaluation of the censored channel available in the literature. It is the main purpose of this paper to contribute for closing this gap.

This work was partially supported by the DFG grant Ma 1184/20-2 in the priority program InKomBio.

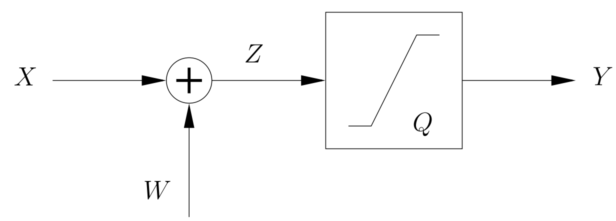

Fig. 1: The system model: some real input X is subject to additive noise W and is censored at 0 and 1 to yield output Y.

We derive a closed-form compact formula for the mutual information of the censored channel. Determining capacity is identified as a convex optimization problem, however, because of iterated integrals involved, an explicit solution seems to be out of reach. We hence deduce an upper bound for the capacity of this channel. Finally, for the cases of additive Gaussian and uniformly distributed noise selected numerical results are presented in order to evaluate the accuracy of the bound and to provide more insight.

II. Channel Model

We consider an additive noise channel, not necessarily Gaussian, although our concrete computational examples refer to the uniformly distributed and Gaussian cases. The input random variable X is assumed to be real-valued and is governed by the cumulative distribution function (CDF) $F(x)$. The input is subject to additive random noise W with density $\varphi(w)$ and corresponding CDF $\Phi(w)$. X and W are assumed to be stochastically independent. The noisy signal $Z = X+W$ is then censored at 0 and 1 by the function

$$Q(z) = \begin{cases} 0, & \text{if } z \leq 0, \\ z, & \text{if } 0 < z \leq 1, \\ 1, & \text{if } z > 1\,. \end{cases} \tag{1}$$

Hence, the output signal is represented as

$$Y = Q(X+W). \tag{2}$$

This channel model is depicted in Fig. 1.

In the following we aim at determining mutual information of this channel which may be written as

$$I(X;Y) = H(Y) - H(Y \mid X). \tag{3}$$

$H(Y)$ and $H(Y \mid X)$ denote the entropy of random variable Y and the conditional entropy of Y given X, respectively. These entropies are hard to determine since Y may have mass points with positive probabilities at 0 and 1.

We will also frequently use the weighted self-information

$$\rho(q) = -q \log q\,, \quad q \geq 0\,, \tag{4}$$

and the binary entropy function

$$\begin{aligned} h(p) &= -p \log p - (1-p)\log(1-p) \\ &= \rho(p) + \rho(1-p)\,, \qquad 0 \leq p \leq 1\,, \end{aligned} \tag{5}$$

where the logarithm is of a general base. It is well known that both $\rho(q)$ and $h(p)$ are strictly concave functions of their arguments q and p.

III. Entropy of a Mixture Distribution

In general the entropy of some random variable Y with density g and with respect to some dominating σ-finite measure μ on the real line may be written as

$$H(Y) = -\int g(y) \log g(y)\, d\mu(y), \tag{6}$$

see [3]. In the present case (2), however, we encounter a mixture of a two-point discrete and a continuous distribution, for which more must be said about the corresponding entropy, cf. [4].

To briefly discuss this case, random variables U, V and B are introduced. U is assumed to be absolutely-continuous with Lebesgue density $f_c(u)$, and V to be discrete with countably many support points v_i, probabilities p_i and discrete density $f_d(v) = p_i$, whenever $v = v_i$ and $f_d(v) = 0$, otherwise. B is a Bernoulli distributed random variabel with $P(B = 1) = \alpha$ and $P(B = 0) = 1 - \alpha$. Furthermore, U, V, B are assumed to be jointly stochastically independent. Then the random variable

$$Y = BU + (1-B)V$$

has density

$$g(y) = \alpha f_c(y) + (1-\alpha) f_d(y)$$

with respect to the measure $\mu = \lambda + \chi$, the sum of the Lebesgue measure λ and the counting measure χ on the support points of V.

From (6) it follows the identity

$$\begin{aligned} H(Y) &= -\alpha \int f_c(y) \log f_c(y)\, dy - \alpha \log \alpha \\ &\qquad - (1-\alpha) \sum_i p_i \log p_i - (1-\alpha)\log(1-\alpha) \\ &= H(B) + \alpha H(U) + (1-\alpha) H(V). \end{aligned} \tag{7}$$

The entropy of Y is hence a convex combination of the entropy of U and V with proportion α, plus the additional uncertainty introduced by switching random variable B, cf. [5, Thm. 3].

IV. Mutual Information of the Censored Channel

The censored channel $Y = Q(X + W)$ with Q defined by (1) is a typical example for a mixture of a continuous and a discrete distribution. Let $Z = X + W$. Then the distribution function of $Y = Q(Z)$ is given by

$$P\big(Q(Z) \leq z\big) = \begin{cases} 0, & \text{if } z < 0, \\ P(Z \leq z), & \text{if } 0 \leq z \leq 1, \\ 1, & \text{if } z > 1\,. \end{cases} \tag{8}$$

If Z is a continuous random variable, which is the case whenever the noise is and the interval $[0,1]$ is a subset of its support, then Y has two singleton mass points at 0 and 1.

Since X and W are stochastically independent, the conditional distribution function reads as

$$\begin{aligned} &P\big(Q(X+W) \leq z \mid X = x\big) \\ &\qquad = \begin{cases} 0, & \text{if } z < 0, \\ P(W \leq z - x), & \text{if } 0 \leq z \leq 1, \\ 1, & \text{if } z > 1\,. \end{cases} \end{aligned} \tag{9}$$

To achieve a compact form for the mutual information of the censored channel the following notation is used. Let f be a general (Lebesgue) density. For $x \in \mathbb{R}$ define

$$\begin{aligned} \ell(f,x) &= \int_{-\infty}^{-x} f(u) du \\ \alpha(f,x) &= \int_{-x}^{1-x} f(u) du \\ r(f,x) &= \int_{1-x}^{\infty} f(u) du \end{aligned}$$

Furthermore, let

$$\begin{aligned} G(f,x) &= h\big(\alpha(f,x)\big) \\ &\quad - \alpha(f,x) \int_{-x}^{1-x} \frac{f(u)}{\alpha(f,x)} \log \frac{f(u)}{\alpha(f,x)}\, du \\ &\quad + \big(1 - \alpha(f,x)\big) h\Big(\frac{\ell(f,x)}{1 - \alpha(f,x)}\Big), \end{aligned} \tag{10}$$

where $h(p)$ denotes the binary entropy function (5).

By formula (7), equation (10) represents the entropy of the mixture of two distributions, namely one with the density

$$\frac{1}{\alpha(f,x)} f(u), \quad -x < u < 1 - x,$$

and the other one by a two-point discrete distribution with support $\{0,1\}$ and corresponding probabilities

$$\frac{\ell(f,x)}{1 - \alpha(f,x)} \quad \text{and} \quad \frac{r(f,x)}{1 - \alpha(f,x)}\,.$$

The mixing parameter is $\alpha(f,x)$.

After some algebra an elegant and intuitively evident representation of $G(f,x)$ is achieved

$$\begin{aligned} G(f,x) &= -\int_{-x}^{1-x} f(u) \log f(u)\, du \\ &\qquad - \ell(f,x) \log \ell(f,x) - r(f,x) \log r(f,x) \\ &= \int_{-x}^{1-x} \rho\big(f(u)\big) du + \rho\big(\ell(f,x)\big) + \rho\big(r(f,x)\big)\,. \end{aligned} \tag{11}$$

Note that $G(f,x)$ is a concave function of density f for any $x \in \mathbb{R}$. This can be seen as follows.

It is well known that $\rho(p)$ is a concave function of $p \geq 0$. Furthermore, $\ell(f,x)$, $r(f,x)$ and $\alpha(f,x)$ are linear functions of f. Hence, for any two densities f and g and any $\lambda \in [0,1]$ it holds that

$$\begin{aligned}
&G\big(\lambda f + (1-\lambda)g, x\big) \\
&\quad = \int_{-x}^{1-x} \rho\big(\lambda f(u) + (1-\lambda)g(u)\big)du \\
&\quad + \rho\Big(\ell\big(\lambda f + (1-\lambda)g, x\big)\Big) + \rho\Big(r\big(\lambda f + (1-\lambda)g, x\big)\Big) \\
&\quad \geq \int_{-x}^{1-x} \big[\lambda\rho\big(f(u)\big) + (1-\lambda)\rho\big(g(u)\big)\big]du \\
&\quad + \lambda\rho\big(\ell(f,x)\big) + (1-\lambda)\rho\big(\ell(g,x)\big) \\
&\quad + \lambda\rho\big(r(f,x)\big) + (1-\lambda)\rho\big(r(g,x)\big) \\
&\quad = \lambda G(f,x) + (1-\lambda)G(g,x),
\end{aligned}$$

showing concavity of $G(f,x)$ as a function of f.

Turning back to channel model (2) and assuming Lebesgue densities

f_Z for input plus noise $Z = X + W$ and
f_W for noise only,

with the help of (10) the entropy and conditional entropy of $Y = Q(X+W)$ can be written as

$$H(Y) = G(f_Z, 0), \tag{12}$$

$$H(Y \mid X = x) = G(f_W, x), \tag{13}$$

$$H(Y \mid X) = \int G(f_W, x)\, dF(x)\,. \tag{14}$$

In summary, we have achieved a compact formula for the mutual information of the censored channel $Y = Q(X+W)$ from (2), setting $Z = X + W$, as

$$\begin{aligned}
I(X;Y) &= G(f_Z, 0) - \int G(f_W, x)\, dF(x) \\
&= \rho\Big(\int\int_{-\infty}^{0} f_W(u-x)du\, dF(x)\Big) \\
&- \int \rho\Big(\int_{-\infty}^{0} f_W(u-x)du\Big)dF(x) \\
&+ \rho\Big(\int\int_{1}^{\infty} f_W(u-x)du\, dF(x)\Big) \\
&- \int \rho\Big(\int_{1}^{\infty} f_W(u-x)du\Big)dF(x) \\
&+ \int_0^1 \rho\Big(\int f_W(u-x)dF(x)\Big)du \\
&- \int_0^1 \int \rho\big(f_W(u-x)\big)dF(x)\, du\,.
\end{aligned} \tag{15}$$

Since $f_Z(z) = \int f_W(z-x)\, dF(x)$ is linear in F and $G(f_Z, 0)$ is a concave function of f_Z, mutual information $I(X;Y)$ in (15) is a concave function of F. Thus, maximizing mutual information over F leads to a convex optimization problem. Achieving some explicit representation of the capacity seems to be an extremely hard variational problem. Deriving sharp bounds and accurate approximations might be possible by applying techniques used in [6] and [7].

Our next goal is to evaluate this formula for certain concrete cases and find simple upper and lower bounds for the corresponding capacity.

V. Bounds

In this section we will determine an upper bound on $H(Y)$ from (12) and a lower bound on (14) so that by merging both an upper bound on the mutual information is achieved. Thereafter, by maximizing over the input distribution an upper bound on the capacity is achieved.

We start from (12) and employ (11) at $x = 0$ to obtain

$$\begin{aligned}
H(Y) &= -\int_{-\infty}^{0} f_Z(u)du \cdot \log \int_{-\infty}^{0} f_Z(u)du \\
&\quad - \int_{1}^{\infty} f_Z(u)du \cdot \log \int_{1}^{\infty} f_Z(u)du \\
&\quad - \int_0^1 f_Z(u)\log f_Z(u)du \\
&= +\rho\Big(\int_{-\infty}^{0} f_Z(u)du\Big) + \rho\Big(\int_{1}^{\infty} f_Z(u)du\Big) \\
&\quad + \int_0^1 \rho\big(f_Z(u)\big)du \\
&\leq +\rho\Big(\int_{-\infty}^{0} f_Z(u)du\Big) + \rho\Big(\int_{1}^{\infty} f_Z(u)du\Big) \\
&\quad + \rho\Big(\int_0^1 f_Z(u)du\Big) \\
&= +\rho(p_\ell) + \rho(p_r) + \rho(p_m),
\end{aligned} \tag{16}$$

say, with $p_\ell + p_r + p_m = 1$. The inequality follows from the concavity of ρ.

In order to find a global upper bound independent of f_Z we solve the optimization problem

$$\begin{aligned}
\max_{p_\ell, p_r, p_m} \quad & \rho(p_\ell) + \rho(p_r) + \rho(p_m) \\
\text{subject to} \quad & p_\ell + p_r + p_m = 1.
\end{aligned} \tag{17}$$

This is a convex optimization problem. The objective function is even Schur-convex, cf. [8], such that the solution is given by

$$p_\ell = p_r = p_m = \frac{1}{3}$$

Hence,

$$H(Y) \leq 3\rho(1/3) = \log 3, \tag{18}$$

with equality if and only if $f_Z(z) = 1/3$ for $z \in [0,1]$ and $\int_{-\infty}^{0} f_Z(z)\, dz = \int_1^\infty f_Z(z)\, dz = 1/3$.

Now for the conditional entropy (14) the following chain of inequalities holds

$$
\begin{aligned}
&H(Y \mid X)\\
&\quad= \int \Bigg[+\rho\Big(\int_{-\infty}^{-x} f_W(u)\,du\Big) + \rho\Big(\int_{1-x}^{\infty} f_W(u)\,du\Big)\\
&\qquad\quad + \int_{-x}^{1-x} \rho\big(f_W(u)\big)\,du \Bigg] dF(x)\\
&\quad\geq \int \int_{-x}^{1-x} \rho\big(f_W(u)\big)\,du\,dF(x)\\
&\quad= \int \int_{0}^{1} \rho\big(f_W(v-x)\big)\,dv\,dF(x)\\
&\quad= -\int \int_{0}^{1} f_W(v-x)\log f_W(v-x)\,dv\,dF(x)\\
&\quad\geq -\int \int_{0}^{1} f_W(u-x)du \log\left[\frac{\int_0^1 f_W^2(v-x)dv}{\int_0^1 f_W(u-x)du}\right] dF(x)\\
&\quad\geq -\rho\big(\mathrm{e}^{-1}\big) \int \int_{0}^{1} f_W^2(v-x)\,dv\,dF(x),
\end{aligned}
\tag{19}
$$

where e denotes the Euler's number. Equality in this chain can never be achieved. The first inequality is due to the positivity of the first two expressions, while the second inequality is a consequence of the Jensen inequality [9]. The third inequality is because of the relation $p \log \frac{q}{p} \leq q\rho(\mathrm{e}^{-1})$ for any positive numbers p and q.

In summary, we obtain the following upper bound for the mutual information of the censored channel (2).

$$
\begin{aligned}
I(X;Y) &= H(Y) - H(Y \mid X)\\
&\leq \log 3 + \rho\big(\mathrm{e}^{-1}\big) \int \int_{0}^{1} f_W^2(v-x)\,dv\,dF(x)
\end{aligned}
\tag{20}
$$

Equation (20) still depends on the input distribution F. An upper bound for the capacity of the censored channel with noise density f_W is obtained by maximizing over F as given by

$$
\begin{aligned}
C &= \max_F I(X;Y)\\
&\leq \log 3 + \rho\big(\mathrm{e}^{-1}\big) \max_F \int \int_{0}^{1} f_W^2(v-x)\,dv\,dF(x)\\
&\leq \log 3 + \rho\big(\mathrm{e}^{-1}\big) \int \max_x \Big\{ \int_{0}^{1} f_W^2(v-x)\,dv \Big\}\,dF(x)\\
&= \log 3 + \rho\big(\mathrm{e}^{-1}\big) \max_x \Big\{ \int_{0}^{1} f_W^2(v-x)\,dv \Big\}.
\end{aligned}
\tag{21}
$$

The integral term in the last line above may be interpreted as the maximum energy/power of the noise density that can be found in a sliding window of length 1, or equivalently, the maximum of the autocorrelation function of the censored noise density at 0. An important insight revealed by this simple upper bound is that the capacity of the censored channel is limited for any signal energy (power) $\int x^2\,dF(x)$. This means that increasing signal energy increases the capacity only up to a specific finite value. Furthermore, as is expected, capacity decreases with increasing the noise variance.

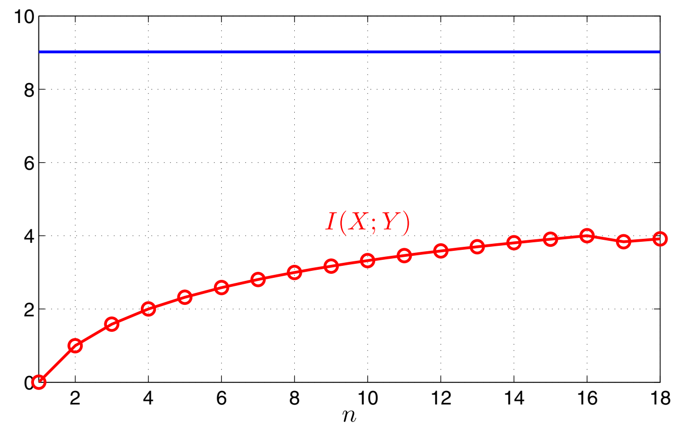

Fig. 2: Uniformly distributed noise.

VI. Numerical Investigations

In order to provide more insight and verify the quality of the upper bound (21), we numerically maximize mutual information (15) over the input distribution F. Note that a maximization of $I(X;Y)$ over specific classes of input distributions only provides a lower bound for the capacity. Hence, in all following results we will observe an upper bound (blue lines) by evaluation of (21) and a lower bound (red curves) by numerical maximization of (15) on the unknown capacity. It should be mentioned that the numerical maximization of (15) is computationally extremely demanding. Thus, only a few important cases are discussed in the following. All results are based on the binary logarithm.

A. Discrete Input with Uniformly Distributed Noise

We assume that the input distribution is a weighted dirac-train which results in the discrete distribution $F(x) = \sum_{i=1}^{n} p_i \varepsilon(x-x_i)$, where $\varepsilon(x)$ denotes the Heaviside (unit) step function. The mass-points x_i are equidistant on the interval $[a,b] = [-\frac{1}{28}, 1+\frac{1}{28}]$, i.e., $x_i = x_i(n) = a + (b-a)\frac{i-1}{n-1}$. This means that increasing the number n of signaling points, will lead to a reduced distance between the mass-points. For each maximization of $I(X;Y)$ the mass-points x_i and the number n are held fixed, while the probabilities p_i are numerically optimized. The number n is incremented for each new maximization of $I(X;Y)$ in order to investigate its effect. The noise is assumed to be uniformly distributed over the interval $[-\frac{1}{28}, \frac{1}{28}]$ such that the interval (support) length is $\Delta = \frac{1}{14}$.

The numerical solutions for this specific scenario are visualized in Fig. 2. It is shown that mutual information is increasing in n while n is an element of $\{1, 2, \ldots, 16\}$. This behavior occurs since for $n \leq 16$ the shifted uniform distributions, describing the distribution of $X + W$, are not overlapping. It is easy to prove that mutual information for $n \leq 16$ is given by $I(X,Y) = \sum_{i=1}^{n} \rho(p_i) \leq \log n$, where the upper bound is achieved for equal probabilities $p_i = \frac{1}{n}$. A maximum value of $I(X;Y) = 4$ is obtained at $n = 16$ which means that 4-bits can error-free be communicated over the censored channel.

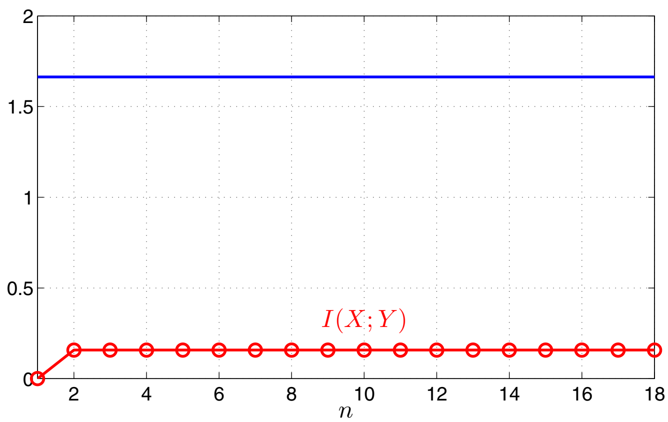

Fig. 3: Gaussian noise with standard deviation $\sigma = 1$.

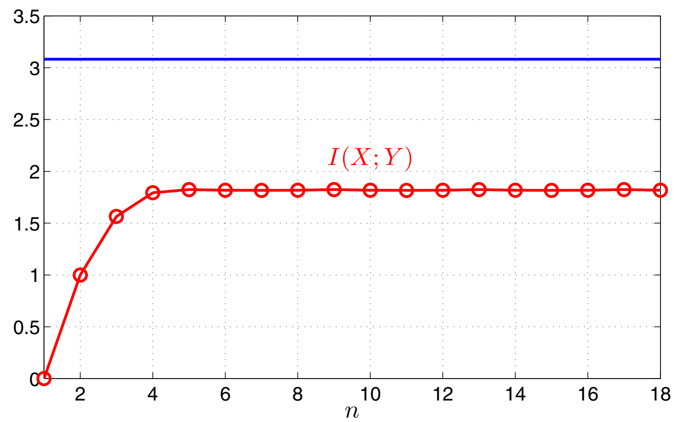

Fig. 4: Gaussian noise with standard deviation $\sigma = \frac{1}{10}$.

Hence, the maximum mutual information of the censored channel, given a uniformly distributed noise, is achieved for a discrete input distribution F when the distance between each two mass-points equals the support length Δ of the uniform noise distribution. For all $n > 16$ the distribution of $X + W$ is a convex combination of overlapping uniform distributions such that an error-free communication is not possible anymore. Thus, a maximization of mutual information can never achieve capacity and only provides a bad lower bound on the capacity. The upper bound in (21) is also depicted in Fig. 2. It is approximately equal to 9.02-bits which is much larger than the maximum value of 4-bits. However, this gap is highly dependent on the noise distribution and must specifically be analyzed for each considered noise distribution.

In summary, for uniform noise distributions with support lengths $\Delta > 0$, the number n of signaling points for achieving maximum mutual information is the greatest integer number for which the inequality $n \leq 2 + \frac{1}{\Delta}$ holds. The maximum mutual information for this number of signaling points is then equal to $\log n$ while the upper bound in (21) is equal to $\log 3 + \rho(\mathrm{e}^{-1})/\Delta$ and $\log 3 + \rho(\mathrm{e}^{-1})/\Delta^2$ for all $\Delta > 1$ and $\Delta \leq 1$, respectively. Both upper bounds can properly be lower bounded by the single expression $\log 3 + \rho(\mathrm{e}^{-1})(n-2)$ which grows linearly with n instead of logarithmically. The maximum-achieving signaling points x_i are distributed on the interval $[-\frac{(n-1)\Delta-1}{2}, 1 + \frac{(n-1)\Delta-1}{2}]$ with equal distances Δ and probabilities $p_i = \frac{1}{n}$. For example, the signal constellation $x_1 = -\frac{1}{2}$, $x_2 = \frac{1}{2}$ and $x_3 = \frac{3}{2}$ with equal probabilities and a uniformly distributed noise with unit support length yield the error-free rate $\log 3$ while the corresponding upper bound is $1 + \log 3$. Note that for this constellation, both expressions $H(Y) = \log 3$ and $H(Y \mid X) = 0$ hold.

B. Discrete Input with Gaussian Noise

For the second case, we consider the same setup as described in the last subsection with another type of noise distributions. We now assume zero-mean Gaussian noise distributions with standard deviations $\sigma \in \{\frac{1}{30}, \frac{1}{10}, 1\}$. We again maximize $I(X;Y)$ over the probabilities p_i while the mass-points x_i and the number n are kept fixed. The variation of n and σ yields the results which are shown in Fig. 3, 4 and 5. In addition, the upper bound in (21) is determined and depicted.

For $\sigma = 1$ a maximum mutual information of $I(X;Y) = 0.158$ is achieved with $n = 2$ signaling points, since the overlap of two Gaussians in the distribution of $X + W$ is lower than for more Gaussians. However, the overlap of both Gaussians is still too large such that only 0.158-bits can be communicated instead of 1-bit. The effect of censoring information is dominant and becomes even more distinct for larger values of σ. The mutual information (and consequently the capacity as well) of the censored channel behaves like the mutual information (capacity) of a Gaussian channel with a one-bit quantizer, cf. [10]. The value of the upper bound (21) is equal to 1.663-bits which significantly exceeds the theoretical value of $\log 2 =$1-bit for two signaling points.

For $\sigma = \frac{1}{10}$ a maximum mutual information of $I(X;Y) = 1.8248$ is obtained with $n = 5$ signaling points. Moreover, we observe the same maximum mutual information for all $n \in \{5, 9, 13, 17, \dots\}$, however, only five signalling points have positive and equal probabilities. Note that mutual information is increasing w.r.t. n only for all $n \leq 5$. For this choice of σ the effect of censoring becomes less important and the mutual information behaves increasingly like a standard Gaussian channel without censoring or quantization. The upper bound (21) is equal to 3.082.

Finally, for $\sigma = \frac{1}{30}$ a maximum mutual information of $I(X;Y) = 3.114$ is achieved with $n = 14$ signaling points. The upper bound yields the value 6.076-bits. It should be noted that for all $n \leq 8$ the signalling points have nearly equal probabilities while for all $n > 8$ the probability distribution is far from being uniform. The optimal probability distribution is shown in Fig. 6. Note that mutual information increases w.r.t. n only for $n \leq 14$.

In summary, the censored channel behaves like a one-bit quantized channel for sufficiently large noise variances. For sufficiently small variances, the censored channel behaves like a standard noisy channel without censoring or quantization. Furthermore, the optimal distribution of the input probabilities for equidistant signalling points is not uniform, if the noise variance is sufficiently small and the number of signalling

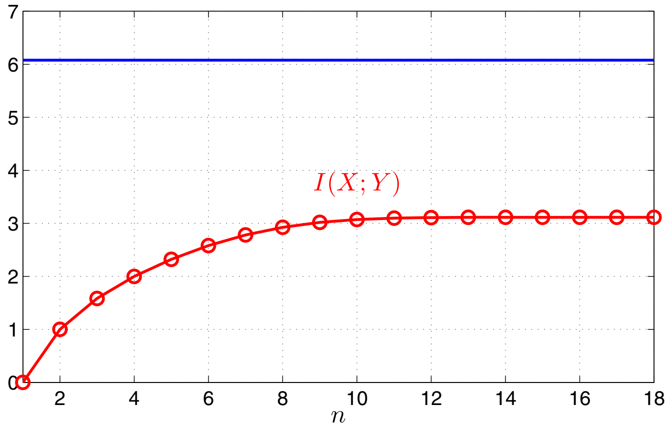

Fig. 5: Gaussian noise with standard deviation $\sigma = \frac{1}{30}$.

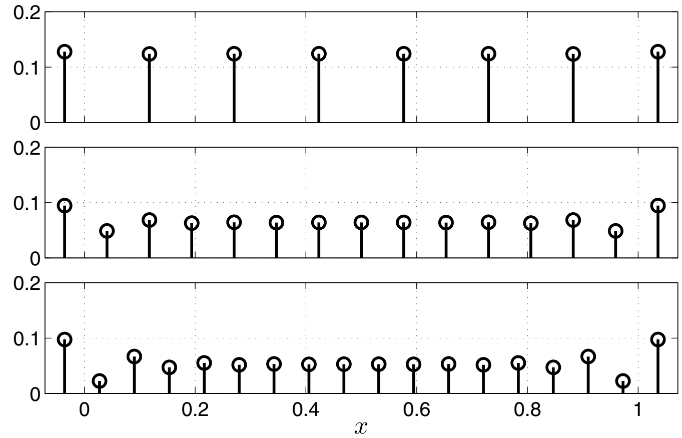

Fig. 6: Optimal probabilities for $n \in \{8, 15, 18\}$ under Gaussian noise.

points is sufficiently large.

C. Gaussian Input with Gaussian Noise

Again we assume a zero-mean Gaussian noise with standard deviation $\sigma = \frac{1}{30}$ and consider the convex combination of two Gaussians

$$F(x) = \frac{1}{2\sqrt{2\pi\tau^2}} \int_{-\infty}^{x} \left(\mathrm{e}^{-\frac{(t-1/2+\Delta)^2}{2\tau^2}} + \mathrm{e}^{-\frac{(t-1/2-\Delta)^2}{2\tau^2}} \right) dt$$

for the input distribution. The variance of the input signal is thus given by $\tau^2 + \Delta^2$ with $\tau > 0$, where $\Delta \geq 0$ describes the deviation of each Gaussian from the mean $1/2$. In order to change only the shape of the input distribution F while the energy (power) of the input signal is kept bounded for any variation of τ and Δ, we maximize the mutual information (15) over τ and Δ subject to $\tau^2 + \Delta^2 \leq \frac{1}{10}$. By this optimization problem, two limiting cases can be compared. The first limiting case is achieved when τ^2 tends to zero and $F(x)$ approaches to the distribution $\frac{1}{2}\varepsilon(x - \frac{1}{2} + \Delta) + \frac{1}{2}\varepsilon(x - \frac{1}{2} - \Delta)$ with $\Delta \leq \frac{1}{\sqrt{10}} = 0.316$. This limiting distribution is a special case of the already investigated scenario in Subsection VI-B. The second limiting case is specified when Δ^2 tends to zero and $F(x)$ approaches to the single Gaussian distribution $\frac{1}{\sqrt{2\pi\tau^2}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{(t-1/2)^2}{2\tau^2}} \, dt$ with $\tau^2 \leq \frac{1}{10}$. The second limiting case describes the optimal input distribution with limited signal energy for standard Gaussian channels without quantization and is thus an important case.

By numerical maximization of (15) over τ and Δ subject to $\tau^2 + \Delta^2 \leq \frac{1}{10}$, we observe the optimal values $I(X;Y) = 3.0865$, $\Delta = 0.2382 > 0$ and $\tau = 0.2080 > 0$. For the choice $\sigma = \frac{1}{30}$ and $\tau^2 + \Delta^2 \leq \frac{1}{10}$, this result shows that neither a single Gaussian for the input distribution nor a discrete input distribution can be capacity-achieving. In addition, by comparing this result with the capacity 3.2539 of a standard Gaussian channel without censoring or quantization, cf. [11], we observe only a small loss in the capacity, since the noise variance is sufficiently small. Note that by comparing this result with those in Fig. 5 and 6, one must be aware that the input signal X for the results in Fig. 5 has more power in average.

VII. Conclusion

We have investigated the mutual information of a channel, which is linear over the interval $[0, 1]$ and is censored to the left by zero and to the right by one. Determining mutual information and capacity of this channel is a fundamental information theoretic problem. We have provided a compact closed-form formula for the mutual information of this channel. Furthermore, a simple upper bound on the capacity of this channel has been developed, which shows that the capacity of the censored channel is bounded even if the energy of the input signal is arbitrarily high. Finally, selected numerical results for uniformly distributed noise as well as for Gaussian noise have been presented. Optimizing corresponding parameters has yielded insight into the behavior of the censored channel and the accuracy of the bound.

References

[1] J. P. Klein and M. L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer, 1997.

[2] J. Fahs and I. Abou-Faycal, "Using hermite bases in studying capacity-achieving distributions over awgn channels," *Information Theory, IEEE Transactions on*, vol. 58, no. 8, pp. 5302–5322, Aug. 2012.

[3] M. Pinsker, *Information and Information Stability of Random Variables and Processes*. San Francisco, California, USA: Holden-Day, 1964.

[4] D. N. Politis, "Maximum entropy modelling of mixture distributions," *Kybernetes*, vol. 23, no. 1, pp. 49–54, 1994.

[5] A. Rényi, "On the dimension and entropy of probability distributions," *Acta Mathematica Academiae Scientiarum Hungarica*, vol. 10, no. 1–2, pp. 193–215, 1959.

[6] G. Alirezaei and R. Mathar, "Scrutinizing the average error probability for nakagami fading channels," in *The IEEE International Symposium on Information Theory (ISIT'14)*, Honolulu, Hawai'i, USA, Jun. 2014, pp. 2884–2888.

[7] ——, "Analytical bounds on the average error probability for nakagami fading channels," in *Information Theory and Applications Workshop (ITA'15)*, San Diego, California, USA, Feb. 2015.

[8] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic Press, 1979.

[9] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, ser. Cambridge Mathematical Library. Cambridge University Press, 1952.

[10] G. Alirezaei and R. Mathar, "Optimum one-bit quantization," in *The IEEE Information Theory Workshop (ITW'15)*, Jeju Island, Korea, Oct. 2015.

[11] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York, NY, USA: Wiley-Interscience, 1991.