Exact Expressions in Source and Channel Coding Problems Using Integral Representations

Neri Merhav Igal Sason
Andrew and Erna Viterbi Faculty of Electrical Engineering
Technion-Israel Institute of Technology
Technion City, Haifa 3200003, Israel
E-mails: {merhav, sason}@ee.technion.ac.il

***Abstract*—We explore known integral representations of the logarithmic and power functions, and demonstrate their usefulness for information-theoretic analyses. We obtain compact, easily–computable exact formulas for several source and channel coding problems that involve expectations and higher moments of the logarithm of a positive random variable and the moment of order $\rho > 0$ of a non-negative random variable (or the sum of i.i.d. positive random variables). These integral representations are used in a variety of applications, including the calculation of the degradation in mutual information between the channel input and output as a result of jamming, universal lossless data compression, Shannon and Rényi entropy evaluations, and the ergodic capacity evaluation of the single-input, multiple–output (SIMO) Gaussian channel with random parameters (known to both transmitter and receiver). The integral representation of the logarithmic function and its variants are anticipated to serve as a rigorous alternative to the popular (but non–rigorous) replica method (at least in some situations).**

I. Introduction

In analytic derivations pertaining to many problem areas in information theory, one frequently encounters the need to calculate expectations and higher moments of expressions that involve the logarithm of a positive–valued random variable, or more generally, the logarithm of the sum of several i.i.d. positive random variables. The common practice, in such situations, is either to resort to upper and lower bounds on the desired expression (e.g., using Jensen's inequality or any other well–known inequalities), or to apply the Taylor series expansion of the logarithmic function. A more modern approach is to use the replica method, which is a popular (but non–rigorous) tool that has been borrowed from the field of statistical physics with considerable success.

The purpose of this work is to point out to an alternative approach and to demonstrate its usefulness in some frequently–encountered situations. In particular, we consider the following integral representation of the logarithmic function (see [3, p. 363, Identity (3.434.2)]),

$$\ln x = \int_0^\infty \frac{e^{-u} - e^{-ux}}{u} \, \mathrm{d}u, \quad x > 0. \tag{1}$$

The immediate use of this representation is in situations where the argument of the logarithmic function is a positive–valued random variable, X, and we wish to calculate the expectation, $\mathbb{E}\{\ln X\}$. By commuting the expectation operator with the integration over u (assuming that this commutation is valid), the calculation of $\mathbb{E}\{\ln X\}$ is replaced by the (often easier) calculation of the moment–generating function (MGF) of X, as

$$\mathbb{E}\{\ln X\} = \int_0^\infty \left[e^{-u} - \mathbb{E}\left\{e^{-uX}\right\}\right] \frac{\mathrm{d}u}{u}. \tag{2}$$

Moreover, if $X_1, \ldots, X_n$ are positive i.i.d. random variables, then

$$\mathbb{E}\left\{\ln \sum_{i=1}^n X_i\right\} = \int_0^\infty \left(e^{-u} - \left[\mathbb{E}\left\{e^{-uX_1}\right\}\right]^n\right) \frac{\mathrm{d}u}{u}. \tag{3}$$

This simple idea is not quite new. It has been used in the physics literature, see, e.g., [2, Eq. (2.4) and onward], [7, Exercise 7.6, p. 140] and [9, Eq. (12) and onward]. With the exception of [8], we are not aware of any work in the information theory literature where it has been used. The purpose of this paper is to demonstrate additional information-theoretic applications, as the need to evaluate logarithmic expectations is not rare at all in many problem areas of information theory. Moreover, the integral representation (1) is useful also for evaluating higher moments of $\ln X$, most notably, the second moment or variance, in order to assess the statistical fluctuations around the mean.

We demonstrate the usefulness of this approach in several application areas. In some of these examples, we also demonstrate the calculation of variances associated with the relevant random variables of interest.

We also consider the utility of the related identity (see [3, p. 363, Identity (3.434.1)]),

$$x^\rho = 1 + \frac{\rho}{\Gamma(1-\rho)} \int_0^\infty \frac{e^{-u} - e^{-ux}}{u^{\rho+1}} \, \mathrm{d}u, \quad x \geq 0, \tag{4}$$

for $\rho \in (0, 1)$, where

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} \, \mathrm{d}t, \quad x > 0 \tag{5}$$

denotes Euler's Gamma function. Identity (4) is used for calculating the ρ-th moment of a non-negative random variable (or the ρ-th moment of the sum of such random variables). In [6], we also generalize (4) (in a non-trivial way) to all values $\rho > 0$, which then enable to obtain expressions for the moment of order ρ which are not restricted to $\rho \in (0,1)$. We note that (1) follows from (4) by the identity

$$\lim_{\rho \to 0} \frac{x^\rho - 1}{\rho} = \ln x, \quad x > 0. \tag{6}$$

It should be pointed out that most of our results remain in the form of a single– or double– definite integral of certain functions that depend on the parameters of the problem in question. Strictly speaking, such a definite integral may not be considered a closed–form expression, but nevertheless:

a) In most of our examples, the new expression is more compact, elegant, and insightful than the original quantity.
b) The resulting definite integrals can actually be considered a closed–form expression "for every practical purpose" since definite integrals in one or two dimensions can be calculated instantly using mathematical software tools.
c) At least in three of our examples, we show how to pass from an n–dimensional integral (with an arbitrarily large n) to one or two–dimensional integrals. This passage is in the spirit of the transition from a multi–letter expression to a single–letter expression.

In Sections II–IV, we provide some applications to Shannon and Rényi entropy calculations, and to source and channel coding problems, respectively. This conference paper presents in part our work in [5] and [6].

II. Applications to Shannon and Rényi Entropies

This section is focused on the calculation of the differential Shannon and Rényi entropies for generalized multivariate Cauchy densities.

Let $(X_1, \ldots, X_n)$ be a random vector whose probability density function is of the form

$$f(\underline{x}) = \frac{C_n}{\left[1 + \sum_{i=1}^n g(x_i)\right]^q}, \quad \forall\, \underline{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n, \tag{7}$$

for a certain non–negative function g and positive constant q such that

$$\int_{\mathbb{R}^n} \frac{1}{\left[1 + \sum_{i=1}^n g(x_i)\right]^q}\, \mathrm{d}\underline{x} < \infty. \tag{8}$$

We refer to this kind of density as *generalized multivariate Cauchy* because the multivariate Cauchy density is obtained as a special case where $g(x) = x^2$ and $q = \frac{1}{2}(n+1)$. Using the Laplace transform relation,

$$\frac{1}{s^q} = \frac{1}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-st}\, \mathrm{d}t, \quad \forall\, q > 0,\ \mathrm{Re}(s) > 0, \tag{9}$$

f can be represented as a mixture of product measures:

$$f(\underline{x}) = \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} \exp\left\{-t \sum_{i=1}^n g(x_i)\right\} \mathrm{d}t. \tag{10}$$

Defining

$$Z(t) := \int_{-\infty}^\infty e^{-t g(x)}\, \mathrm{d}x, \quad \forall\, t > 0, \tag{11}$$

it can be verified from (10) that

$$C_n = \frac{\Gamma(q)}{\displaystyle\int_0^\infty t^{q-1} e^{-t} Z^n(t)\, \mathrm{d}t}. \tag{12}$$

From (7), the calculation of the differential entropy of f is associated with evaluating $\mathbb{E}\big\{\ln\big[1+\sum_{i=1}^n g(X_i)\big]\big\}$. Using (1),

$$\begin{aligned} &\mathbb{E}\left\{\ln\left[1 + \sum_{i=1}^n g(X_i)\right]\right\} \\ &= \int_0^\infty \frac{e^{-u}}{u} \left(1 - \mathbb{E}\left\{\exp\left[-u \sum_{i=1}^n g(X_i)\right]\right\}\right) \mathrm{d}u. \end{aligned} \tag{13}$$

From (10) and by interchanging the integration,

$$\begin{aligned} &\mathbb{E}\left\{\exp\left[-u \sum_{i=1}^n g(X_i)\right]\right\} \\ &= \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} \int_{\mathbb{R}^n} \exp\left\{-(t+u) \sum_{i=1}^n g(x_i)\right\} \mathrm{d}\underline{x}\, \mathrm{d}t \\ &= \frac{C_n}{\Gamma(q)} \int_0^\infty t^{q-1} e^{-t} Z^n(t+u)\, \mathrm{d}t. \end{aligned} \tag{14}$$

In view of (10), (13) and (14), the differential entropy of $(X_1, \ldots, X_n)$ can be verified to be given by

$$\begin{aligned} &h(X_1, \ldots, X_n) \\ &= q\, \mathbb{E}\left\{\ln\left[1 + \sum_{i=1}^n g(X_i)\right]\right\} - \ln C_n \\ &= \frac{q C_n}{\Gamma(q)} \int_0^\infty \int_0^\infty \frac{t^{q-1} e^{-(t+u)}}{u} \Big[Z^n(t) - Z^n(t+u)\Big]\, \mathrm{d}t\, \mathrm{d}u \\ &\quad - \ln C_n. \end{aligned} \tag{15}$$

For $g(x) = |x|^\theta$, with a fixed $\theta > 0$, we obtain from (11) that

$$Z(t) = \frac{2\,\Gamma(1/\theta)}{\theta\, t^{1/\theta}}. \tag{16}$$

In particular, for $\theta = 2$ and $q = \frac{1}{2}(n+1)$, we get the multivariate Cauchy density from (7). In this case, it follows from (16) that $Z(t) = \sqrt{\frac{\pi}{t}}$ for $t > 0$, and from (12)

$$C_n = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{(n+1)/2}}. \tag{17}$$

Combining (15), (16) and (17) gives

$$\begin{aligned} &h(X_1, \ldots, X_n) \\ &= \frac{n+1}{2\pi^{(n+1)/2}} \int_0^\infty \int_0^\infty \frac{e^{-(t+u)}}{u\sqrt{t}} \left[1 - \left(\frac{t}{t+u}\right)^{n/2}\right] \mathrm{d}t\, \mathrm{d}u \\ &\quad + \frac{(n+1)\ln \pi}{2} - \ln \Gamma\left(\frac{n+1}{2}\right). \end{aligned} \tag{18}$$

Fig. 1 displays the normalized differential entropy, $\frac{1}{n}\, h(X_1, \ldots, X_n)$, for $1 \le n \le 100$.

We believe that the interesting point, conveyed in this application example, is that (15) provides a kind of a "single–letter

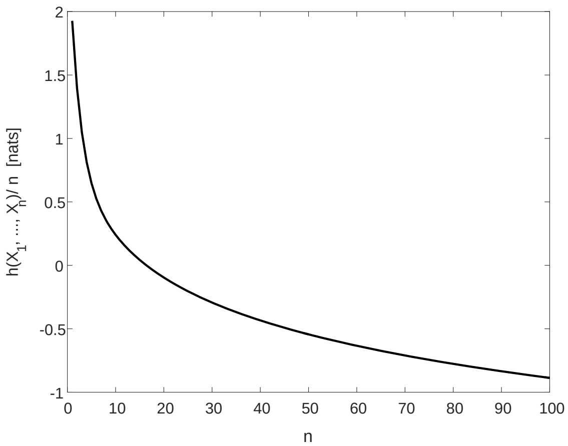

Fig. 1. The normalized differential entropy, $\frac{1}{n}\, h(X_1, \ldots, X_n)$ (see (18)), for a multivariate Cauchy density, $f(\underline{x}) = C_n/[1+\sum_{i=1}^n x_i^2]^{(n+1)/2}$, with C_n in (17).

expression"; the n–dimensional integral, associated with the original expression of the differential entropy $h(X_1, \ldots, X_n)$, is replaced by the two–dimensional integral in (15), independently of n. For $\alpha > 1$, it relies on the Laplace transform relation in (9), which after some calculations give

$$\begin{aligned} &H_\alpha(X_1, \ldots, X_n) \\ &= \frac{1}{1-\alpha} \log \mathbb{E}\bigl\{f^{\alpha-1}(\underline{X})\bigr\} \\ &= \frac{\alpha}{\alpha-1} \log \int_0^\infty t^{q-1} e^{-t} Z^n(t)\, \mathrm{d}t \\ &\quad - \frac{1}{\alpha-1} \log \int_0^\infty \int_0^\infty t^{q(\alpha-1)-1} u^{q-1} e^{-(t+u)} \\ &\qquad\qquad \cdot Z^n(t+u)\, \mathrm{d}u\, \mathrm{d}t \\ &\quad + \frac{1}{\alpha-1} \log \Gamma\bigl(q(\alpha-1)\bigr) - \log \Gamma(q), \end{aligned} \tag{19}$$

with $Z(\cdot)$ as given in (11). For $\alpha \in (0,1)$, we express the differential Rényi entropy of order α as a double-integral whose derivation relies on a non-trivial extension of Identity (4) for all $\rho > 0$. In particular, it can be obtained from (4) that, for all $\alpha \in \bigl(1-\frac{1}{q}, 1\bigr)$,

$$\begin{aligned} &H_\alpha(X_1, \ldots, X_n) \\ &= \log\Biggl(1 + \frac{q(1-\alpha)}{\Gamma\bigl(q(1-\alpha)\bigr)} \int_0^\infty \int_0^\infty \frac{e^{-(t+u)}}{u^{q(\alpha-1)+1}} \\ &\qquad \cdot \biggl(1 - \frac{C_n}{\Gamma(q)} \cdot t^{q-1} Z^n(t+u)\biggr) \mathrm{d}t\, \mathrm{d}u\Biggr). \end{aligned} \tag{20}$$

As a final note, we mention that a lower bound on the differential Shannon entropy of a different form of extended multivariate Cauchy distributions (cf. [4, Eq. (42)]) was derived in [4, Theorem 6]. The latter result relies on obtaining lower bounds on the differential entropy of random vectors whose densities are symmetric log-concave or γ-concave (i.e., densities f for which f^γ is concave for some $\gamma < 0$).

III. Applications to Source Coding

We refer here to universal source coding for binary arbitrarily-varying sources. Consider a source coding setting, where there are n binary DMS's, and let $x_i \in [0,1]$ denote the Bernoulli parameter of source no. $i \in \{1, \ldots, n\}$. Assume that a hidden memoryless switch selects uniformly at random one of these sources, and the data is then emitted by the selected source. Since it is unknown a-priori which source is selected at each instant, a universal lossless source encoder (e.g., a Shannon or Huffman code) is designed to match a binary DMS whose Bernoulli parameter is given by $\frac{1}{n}\sum_{i=1}^n x_i$. Neglecting integer length constraints, the average redundancy in the compression rate (measured in nats per symbol), due to the unknown realization of the hidden switch, is about

$$R_n = h_{\mathrm{b}}\Biggl(\frac{1}{n}\sum_{i=1}^n x_i\Biggr) - \frac{1}{n}\sum_{i=1}^n h_{\mathrm{b}}(x_i), \tag{21}$$

where $h_{\mathrm{b}}\colon [0,1] \to [0, \ln 2]$ is the binary entropy function (defined to the base e), and the redundancy is given in nats per source symbol. Now, let us assume that the Bernoulli parameters of the n sources are i.i.d. random variables, $X_1, \ldots, X_n$, all having the same density as that of some generic random variable X, whose support is the interval $[0,1]$. We wish to evaluate the expected value of the above defined redundancy, under the assumption that the realizations of $X_1, \ldots, X_n$ are known. We are then facing the need to evaluate

$$\overline{R}_n = \mathbb{E}\Biggl\{h_{\mathrm{b}}\Biggl(\frac{1}{n}\sum_{i=1}^n X_i\Biggr)\Biggr\} - \mathbb{E}\{h_{\mathrm{b}}(X)\}. \tag{22}$$

We now express the first and second terms on the right–hand side of (22) as a function of the MGF of X.

In view of (1), the binary entropy function h_{b} admits the integral representation

$$h_{\mathrm{b}}(x) = \int_0^\infty \Bigl[x e^{-ux} + (1-x) e^{-u(1-x)} - e^{-u}\Bigr] \frac{\mathrm{d}u}{u}, \tag{23}$$

for all $x \in [0,1]$, which implies that $\mathbb{E}\{h_{\mathrm{b}}(X)\}$ can be expressed as functionals of the MGF of X, $M_X(\nu) = \mathbb{E}\{e^{\nu X}\}$, and its derivative, for $\nu < 0$. For all $u \in \mathbb{R}$,

$$\mathbb{E}\bigl\{X e^{-uX}\bigr\} = M'_X(-u), \tag{24}$$

and

$$\mathbb{E}\bigl\{(1-X) e^{-u(1-X)}\bigr\} = e^{-u}\Bigl[M_X(u) - M'_X(u)\Bigr]. \tag{25}$$

From (23)–(25), we readily obtain

$$\begin{aligned} &\mathbb{E}\{h_{\mathrm{b}}(X)\} \\ &= \int_0^\infty \frac{1}{u}\Bigl\{M'_X(-u) + \bigl[M_X(u) - M'_X(u) - 1\bigr] e^{-u}\Bigr\}\, \mathrm{d}u. \end{aligned} \tag{26}$$

Define $Y_n := \frac{1}{n}\sum_{i=1}^n X_i$. Then,

$$M_{Y_n}(u) = M_X^n\bigl(\tfrac{u}{n}\bigr), \quad \forall\, u \in \mathbb{R}, \tag{27}$$

which yields, in view of (26), (27) and the change of integration variable, $t = \frac{u}{n}$, the following:

$$\mathbb{E}\left\{ h_{\mathrm{b}}\left(\tfrac{1}{n} \sum_{i=1}^{n} X_i \right) \right\} = \int_0^\infty \frac{1}{t} \left\{ M_X^{n-1}(-t)\, M_X'(-t) + e^{-nt} \cdot \left[M_X^n(t) - M_X^{n-1}(t)\, M_X'(t) - 1 \right] \right\} \mathrm{d}t. \tag{28}$$

Here too, we pass from an n-dimensional integral to a one-dimensional integral. In general, similar calculations can be carried out for higher integer moments, thus passing from n-dimensional integration for a moment of order s to an s-dimensional integral, independently of n.

For example, if $X_1, \ldots, X_n$ are i.i.d. and uniformly distributed on [0,1], then (28) gives that $\mathbb{E}\left\{ h_{\mathrm{b}}\left(\frac{1}{n} \sum_{i=1}^{n} X_i \right) \right\}$ is equal to $\frac{1}{2}$, 0.602, 0.634, 0.650, 0.659 nats for $n = 1, \ldots, 5$, respectively, with the limit $h_{\mathrm{b}}\left(\frac{1}{2}\right) = \ln 2 \approx 0.693$ as we let $n \to \infty$ (this is expected by the law of large numbers).

IV. Applications to Channel Coding

A. Mutual Information Calculations for Communication Channels with Jamming

Consider a channel with input $\underline{X} = (X_1, \ldots, X_n) \in \mathcal{X}^n$ and output $\underline{Y} = (Y_1, \ldots, Y_n) \in \mathcal{Y}^n$, having the transition probability law

$$p_n(\underline{y}|\underline{x}) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \prod_{j \neq i} q(y_j|x_j)\, p(y_i|x_i) \right\}, \tag{29}$$

for all $(\underline{x}, \underline{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$. This channel model refers to the case of a memoryless, stationary channel with a transition probability law $q_n(\underline{y}|\underline{x}) = \prod_{i=1}^{n} q(y_i|x_i)$, where one of its symbols is jammed uniformly at random, and the transition law of the randomly jammed symbol x_i is given by $p(y_i|x_i)$ instead of $q(y_i|x_i)$.

We wish to evaluate the mutual information $I(\underline{X}; \underline{Y})$, due to jamming as above, and see by how much it decreases as compared to the mutual information between the input and output vectors of the memoryless, stationary channel q_n (without the jamming). Let the input distribution be memoryless and stationary with $p_{\underline{X}}(\underline{x}) = \prod_{i=1}^{n} p_X(x_i)$ for all $\underline{x} \in \mathcal{X}^n$.

The mutual information (in nats) is given by

$$\begin{aligned} I(\underline{X}; \underline{Y}) &= h(\underline{Y}) - h(\underline{Y}|\underline{X}) \\ &= \int_{\mathcal{X}^n \times \mathcal{Y}^n} p_{\underline{X},\underline{Y}}(\underline{x}, \underline{y})\, \ln p_n(\underline{y}|\underline{x})\, \mathrm{d}\underline{x}\, \mathrm{d}\underline{y} \\ &\quad - \int_{\mathcal{Y}^n} p_{\underline{Y}}(\underline{y})\, \ln p_{\underline{Y}}(\underline{y})\, \mathrm{d}\underline{y}. \end{aligned} \tag{30}$$

For simplicity of notation, we omit the domains of integration when they are clear from the context. We have,

$$\begin{aligned} &\int p_{\underline{X},\underline{Y}}(\underline{x}, \underline{y})\, \ln p_n(\underline{y}|\underline{x})\, \mathrm{d}\underline{x}\, \mathrm{d}\underline{y} \\ &= \int p_{\underline{X},\underline{Y}}(\underline{x}, \underline{y})\, \ln \frac{p_n(\underline{y}|\underline{x})}{q_n(\underline{y}|\underline{x})}\, \mathrm{d}\underline{x}\, \mathrm{d}\underline{y} \\ &\quad + \int p_{\underline{X},\underline{Y}}(\underline{x}, \underline{y})\, \ln q_n(\underline{y}|\underline{x})\, \mathrm{d}\underline{x}\, \mathrm{d}\underline{y}. \end{aligned} \tag{31}$$

By using the logarithmic representation in (1), and based on the following equality (see (29)):

$$\frac{p_n(\underline{y}|\underline{x})}{q_n(\underline{y}|\underline{x})} = \frac{1}{n} \sum_{i=1}^{n} \frac{p(y_i|x_i)}{q(y_i|x_i)}, \tag{32}$$

it follows after some calculations (which involve changing the order of integration) that

$$\begin{aligned} &\int p_{\underline{X},\underline{Y}}(\underline{x}, \underline{y})\, \ln \frac{p_n(\underline{y}|\underline{x})}{q_n(\underline{y}|\underline{x})}\, \mathrm{d}\underline{x}\, \mathrm{d}\underline{y} \\ &= \int_0^\infty \frac{1}{u} \left[e^{-u} - f^{n-1}\left(\frac{u}{n}\right) g\left(\frac{u}{n}\right) \right] \mathrm{d}u, \end{aligned} \tag{33}$$

where, for $u \geq 0$,

$$f(u) := \int p_X(x)\, q(y|x)\, \exp\left(-\frac{u\, p(y|x)}{q(y|x)} \right) \mathrm{d}x\, \mathrm{d}y, \tag{34}$$

$$g(u) := \int p_X(x)\, p(y|x)\, \exp\left(-\frac{u\, p(y|x)}{q(y|x)} \right) \mathrm{d}x\, \mathrm{d}y. \tag{35}$$

Moreover, due to the product form of q_n, it can be verified that

$$\begin{aligned} &\int p_{\underline{X},\underline{Y}}(\underline{x}, \underline{y})\, \ln q_n(\underline{y}|\underline{x})\, \mathrm{d}\underline{x}\, \mathrm{d}\underline{y} \\ &= \int p_X(x)\, p(y|x)\, \ln q(y|x)\, \mathrm{d}x\, \mathrm{d}y \\ &\quad + (n-1) \int p_X(x)\, q(y|x)\, \ln q(y|x)\, \mathrm{d}x\, \mathrm{d}y. \end{aligned} \tag{36}$$

Combining (31)–(36) expresses the conditional (differential) entropy $h(\underline{Y}|\underline{X})$ as a double integral over $\mathcal{X} \times \mathcal{Y}$, independently of n (rather than an integration over $\mathcal{X}^n \times \mathcal{Y}^n$).

We next need to calculate the differential entropy of $\underline{Y}$ at the output of the channel whose transition probability law is given by p_n in (29), assuming the product input distribution as above. It can be verified that, for all $\underline{y} \in \mathcal{Y}^n$,

$$p_{\underline{Y}}(\underline{y}) = \prod_{j=1}^{n} v(y_j) \cdot \frac{1}{n} \sum_{i=1}^{n} \frac{w(y_i)}{v(y_i)}, \tag{37}$$

where, for all $y \in \mathcal{Y}$,

$$v(y) := \int q(y|x)\, p_X(x)\, \mathrm{d}x, \tag{38}$$

$$w(y) := \int p(y|x)\, p_X(x)\, \mathrm{d}x. \tag{39}$$

By the integral representation of the logarithmic function in (1), the following identity holds for an arbitrary positive random variable Z:

$$\mathbb{E}\{Z \ln Z\} = \int_0^\infty \frac{1}{u} \left[M_Z'(0)\, e^{-u} - M_Z'(-u) \right] \mathrm{d}u \tag{40}$$

where $M_Z(u) := \mathbb{E}\{e^{uZ}\}$ denotes the MGF of Z. By setting $Z := \frac{1}{n}\sum_{i=1}^{n} \frac{w(R_i)}{v(R_i)}$ where $\{R_i\}_{i=1}^{n}$ are i.i.d. random variables with the density v, some calculations finally give that

$$h(\underline{Y}) = \int_0^\infty \frac{1}{u}\left[t^{n-1}\left(\frac{u}{n}\right) s\left(\frac{u}{n}\right) - e^{-u}\right] \mathrm{d}u \qquad (41)$$
$$- \int w(y)\, \ln v(y)\, \mathrm{d}y - (n-1) \int v(y) \ln v(y)\, \mathrm{d}y,$$

where v and w are given in (38) and (39), respectively, and

$$s(u) := \int w(y)\, \exp\left(-\frac{u\, w(y)}{v(y)}\right) \mathrm{d}y, \quad u \geq 0, \qquad (42)$$

$$t(u) := \int v(y)\, \exp\left(-\frac{u\, w(y)}{v(y)}\right) \mathrm{d}y, \quad u \geq 0. \qquad (43)$$

Fig. 2 refers to the special case where transmission takes place over a binary symmetric channel (BSC) with crossover probability $q = 10^{-3}$, and p_n is the transition probability law of such a channel where the randomly selected bit in a transmitted block of $n = 128$ bits which is jammed has a new crossover probability $p \in (q, \frac{1}{2}]$. The mutual information for the jamming-free BSC with a symmetric and memoryless binary-input distribution is equal to

$$I(\underline{X}; \underline{Y}) = n\big(\ln 2 - h_{\mathrm{b}}(q)\big) = 87.71 \text{ nats},$$

where $h_{\mathrm{b}}(\cdot)$ denotes the binary entropy function (to the base e), and the value of the mutual information is decreased by 2.88 nats as a result of the random single-bit jamming with a crossover probability of $p = \frac{1}{2}$ (see Fig. 2).

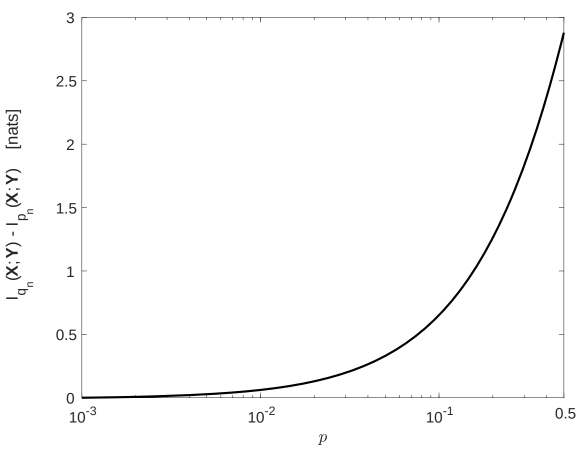

Fig. 2. The degradation in the mutual information $I(\underline{X}; \underline{Y})$ for blocks of $n = 128$ bits. The jamming-free channel q_n acts like a binary symmetric channel with crossover probability $q = 10^{-3}$, and p_n is the channel transition law in (29). The random bit which is jammed according to p_n has crossover probability $p \in \big(q, \frac{1}{2}\big]$. The binary-input distribution is memoryless and symmetric. The degradation in $I(\underline{X}; \underline{Y})$ (in units of nats) is plotted as a function of the parameter p.

The restriction to a single bit which is jammed is for simplicity.

B. Ergodic Capacity of Fading SIMO Channels

Consider the SIMO channel with L receive antennas and assume that the channel transfer coefficients, $h_1, \ldots, h_L$, are independent, zero–mean, circularly symmetric complex Gaussian random variables with variances $\sigma_1^2, \ldots, \sigma_L^2$. The ergodic capacity (in nats per channel use) of the channel is given by

$$C = \mathbb{E}\left\{\ln\left(1 + \rho \sum_{\ell=1}^{L} |h_\ell|^2\right)\right\}$$
$$= \mathbb{E}\left\{\ln\left(1 + \rho \sum_{\ell=1}^{L} \big(f_\ell^2 + g_\ell^2\big)\right)\right\}, \qquad (44)$$

where $f_\ell := \mathrm{Re}\{h_\ell\}$, $g_\ell := \mathrm{Im}\{h_\ell\}$, and $\rho := \frac{P}{N_0}$ is the signal–to–noise ratio (SNR) (see, e.g., [1] and [10]).

Paper [1] is devoted, among other things, to the exact evaluation of (44) by finding the density of the random variable defined by $\sum_{\ell=1}^{L}(f_\ell^2 + g_\ell^2)$, and then taking the expectation w.r.t. that density. Here, we show that the integral representation in (1) suggests a more direct approach to the evaluation of (44). It should also be pointed out that this approach is more flexible than the one in [1], as the latter strongly depends on the assumption that $\{h_i\}$ are Gaussian and statistically independent. The integral representation approach also allows other distributions of the channel transfer gains, as well as possible correlations between the coefficients and/or the channel inputs. Moreover, we are also able to calculate the variance of $\ln\left(1 + \rho \sum_{\ell=1}^{L} |h_\ell|^2\right)$, as a measure of the fluctuations around the mean, which is obviously related to the outage. Specifically, let $X := \rho \sum_{\ell=1}^{L}(f_\ell^2 + g_\ell^2)$. Then,

$$M_X(-u) = \mathbb{E}\big\{\exp(-uX)\big\} = \prod_{\ell=1}^{L} \frac{1}{1 + u\rho\sigma_\ell^2}, \; u > 0. \qquad (45)$$

From (1), (44) and (45), the ergodic capacity (in nats per channel use) is given by

$$C = \int_0^\infty \frac{e^{-x/\rho}}{x} \left(1 - \prod_{\ell=1}^{L} \frac{1}{1 + \sigma_\ell^2 x}\right) \mathrm{d}x. \qquad (46)$$

A similar approach appears in [8, Eq. (12)]. As for the variance, [5, Proposition 2] (relying on (1)) and (45) yield

$$\mathrm{Var}\left\{\ln\left(1 + \rho \sum_{\ell=1}^{L} [f_\ell^2 + g_\ell^2]\right)\right\}$$
$$= \int_0^\infty \int_0^\infty \frac{e^{-(x+y)/\rho}}{xy} \left\{\prod_{\ell=1}^{L} \frac{1}{1 + \sigma_\ell^2(x+y)} - \prod_{\ell=1}^{L} \frac{1}{(1 + \sigma_\ell^2 x)(1 + \sigma_\ell^2 y)}\right\} \mathrm{d}x\, \mathrm{d}y. \qquad (47)$$

Consider the example of $L = 2$, $\sigma_1^2 = \frac{1}{2}$ and $\sigma_2^2 = 1$. From (46), the ergodic capacity of the SIMO channel is given by

$$C = 2e^{1/\rho}\, E_1\left(\frac{1}{\rho}\right) - e^{2/\rho}\, E_1\left(\frac{2}{\rho}\right), \qquad (48)$$

where $E_1(\cdot)$ is the (modified) exponential integral function:

$$E_1(x) := \int_x^\infty \frac{e^{-s}}{s}\, \mathrm{d}s, \quad \forall\, x > 0. \qquad (49)$$

REFERENCES

[1] A. Dong, H. Zhang, D. Wu, and D. Yuan, “Logarithmic expectation of the sum of exponential random variables for wireless communication performance evaluation,” *Proc. 2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*, Boston, MA, USA, September 2015.

[2] S. E. Esipov and T. J. Newman, “Interface growth and Burgers turbulence: the problem of random initial conditions,” *Phys. Rev. E*, vol. 48, no. 2, pp. 1046–1050, August 1993.

[3] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, Eighth Edition, Elsevier, 2014.

[4] A. Marsiglietti and V. Kostina, “A lower bound on the differential entropy of log-concave random vectors with applications,” *Entropy*, vol. 20, no. 3, paper 185, March 2018.

[5] N. Merhav and I. Sason, “An integral representation of the logarithmic function with applications in information theory,” *Entropy*, vol. 22, no. 1, paper 51, pp. 1–22, January 2020.

[6] —, “Some useful integral representations for information-theoretic analyses,” *Entropy*, vol. 22, no. 6, paper 707, pp. 1–28, June 2020.

[7] M. Mézard and A. Montanari, *Information, Physics, and Computation*, Oxford University Press, New-York, USA, 2009.

[8] A. Rajan and C. Tepedelenlioğlu, “Stochastic ordering of fading channels through the Shannon transform,” *IEEE Trans. Inform. Theory*, vol. 61, no. 4, pp. 1619–1628, April 2015.

[9] J. Song, S. Still, R. D. H. Rojas, I. P. Castillo, and M. Marsili, “Optimal work extraction and mutual information in a generalized Szilárd engine,” arXiv:1910.0419v1, October 9, 2019.

[10] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.