# Exact Expressions in Source and Channel Coding Problems Using Integral Representations

Speaker: Igal Sason

Joint Work with Neri Merhav

EE Department, Technion - Israel Institute of Technology

2020 IEEE International Symposium on Information Theory

**ISIT 2020**

Online Virtual Conference

June 21-26, 2020

## Motivation

In information-theoretic analyses, one frequently needs to calculate:

- Expectations or, more generally, $\rho$-th moments, for some $\rho > 0$;
- Logarithmic expectations

of sums of i.i.d. positive random variables.

## Motivation

In information-theoretic analyses, one frequently needs to calculate:

- Expectations or, more generally, $\rho$-th moments, for some $\rho > 0$;
- Logarithmic expectations

of sums of i.i.d. positive random variables.

## Commonly Used Approaches

- Resorting to bounds (e.g., Jensen's inequality).
- A modern approach for logarithmic expectations is to use the replica method, which is a popular (but non–rigorous) tool, borrowed from statistical physics with considerable success.

## Motivation

In information-theoretic analyses, one frequently needs to calculate:

- Expectations or, more generally, $\rho$-th moments, for some $\rho > 0$;
- Logarithmic expectations

of sums of i.i.d. positive random variables.

## Commonly Used Approaches

- Resorting to bounds (e.g., Jensen's inequality).
- A modern approach for logarithmic expectations is to use the replica method, which is a popular (but non–rigorous) tool, borrowed from statistical physics with considerable success.

## Purpose of this Work

Pointing out an alternative approach, by using integral representations, and demonstrating its usefulness in information-theoretic analyses.

## Useful Integral Representation for the Logarithm

$$\ln z = \int_0^\infty \frac{e^{-u} - e^{-uz}}{u} \, \mathrm{d}u, \quad \mathrm{Re}(z) \geq 0.$$

## Useful Integral Representation for the Logarithm

$$\ln z = \int_0^\infty \frac{e^{-u} - e^{-uz}}{u} \, \mathrm{d}u, \quad \mathrm{Re}(z) \geq 0.$$

## Proof

$$\begin{aligned}
\ln z &= (z-1) \int_0^1 \frac{\mathrm{d}v}{1 + v(z-1)} \\
&= (z-1) \int_0^1 \int_0^\infty e^{-u[1+v(z-1)]} \, \mathrm{d}u \, \mathrm{d}v \\
&= (z-1) \int_0^\infty e^{-u} \int_0^1 e^{-uv(z-1)} \, \mathrm{d}v \, \mathrm{d}u \\
&= \int_0^\infty \frac{e^{-u}}{u} \left[1 - e^{-u(z-1)}\right] \mathrm{d}u \\
&= \int_0^\infty \frac{e^{-u} - e^{-uz}}{u} \, \mathrm{d}u.
\end{aligned}$$

## Useful Integral Representation for the Logarithm

$$\ln z = \int_0^\infty \frac{e^{-u} - e^{-uz}}{u} \, \mathrm{d}u, \quad \mathrm{Re}(z) \geq 0.$$

## Logarithmic Expectation

$$\mathbb{E}\{\ln X\} = \int_0^\infty \left[e^{-u} - M_X(-u)\right] \, \frac{\mathrm{d}u}{u},$$

where $M_X(u) := \mathbb{E}\left\{e^{uX}\right\}$ is the moment-generating function (MGF).

## Useful Integral Representation for the Logarithm

$$\ln z = \int_0^\infty \frac{e^{-u} - e^{-uz}}{u} \, \mathrm{d}u, \quad \mathrm{Re}(z) \geq 0.$$

## Logarithmic Expectation

$$\mathbb{E}\{\ln X\} = \int_0^\infty \left[e^{-u} - M_X(-u)\right] \frac{\mathrm{d}u}{u},$$

where $M_X(u) := \mathbb{E}\left\{e^{uX}\right\}$ is the moment-generating function (MGF).

## Logarithmic Expectation of a sum of i.i.d. random variables

Let $X_1, \ldots, X_n$ be i.i.d. random variables, then

$$\mathbb{E}\{\ln(X_1 + \ldots + X_n)\} = \int_0^\infty \left[e^{-u} - M_{X_1}^n(-u)\right] \frac{\mathrm{d}u}{u}.$$

## Example 1: Logarithms of Factorials

$$\ln(n!) = \sum_{k=1}^{n} \ln k$$
$$= \sum_{k=1}^{n} \int_0^\infty \left(e^{-u} - e^{-uk}\right) \frac{\mathrm{d}u}{u}$$
$$= \int_0^\infty e^{-u} \left(n - \frac{1 - e^{-un}}{1 - e^{-u}}\right) \frac{\mathrm{d}u}{u}.$$

## Example 1: Logarithms of Factorials

$$\begin{aligned}\ln(n!) &= \sum_{k=1}^{n} \ln k \\ &= \sum_{k=1}^{n} \int_0^\infty \left(e^{-u} - e^{-uk}\right) \frac{\mathrm{d}u}{u} \\ &= \int_0^\infty e^{-u} \left(n - \frac{1-e^{-un}}{1-e^{-u}}\right) \frac{\mathrm{d}u}{u}.\end{aligned}$$

## Example 2: Entropy of Poisson Random Variable $N \sim \mathrm{Poisson}(\lambda)$

$$\begin{aligned}H(N) &= \lambda - \mathbb{E}\{N\} \ln \lambda + \mathbb{E}\{\ln N!\} \\ &= \lambda \ln \frac{e}{\lambda} + \int_0^\infty e^{-u} \left(\lambda - \frac{1-e^{-\lambda(1-e^{-u})}}{1-e^{-u}}\right) \frac{\mathrm{d}u}{u}.\end{aligned}$$

### $\rho$-th moment for all $\rho \in (0, 1)$

$$\mathbb{E}\{X^\rho\} = 1 + \frac{\rho}{\Gamma(1-\rho)} \int_0^\infty \frac{e^{-u} - M_X(-u)}{u^{1+\rho}} \, \mathrm{d}u,$$

where $\Gamma(\cdot)$ denotes Euler's Gamma function:

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

## $\rho$-th moment for all $\rho \in (0,1)$

$$\mathbb{E}\{X^\rho\} = 1 + \frac{\rho}{\Gamma(1-\rho)} \int_0^\infty \frac{e^{-u} - M_X(-u)}{u^{1+\rho}} \, \mathrm{d}u,$$

where $\Gamma(\cdot)$ denotes Euler's Gamma function:

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

## $\rho$-th moment of the sum of i.i.d. RVs for all $\rho \in (0,1)$

If $\{X_i\}_{i=1}^n$ are i.i.d. nonnegative real-valued random variables, then

$$\mathbb{E}\left\{ \left( \sum_{i=1}^n X_i \right)^\rho \right\} = 1 + \frac{\rho}{\Gamma(1-\rho)} \int_0^\infty \frac{e^{-u} - M_{X_1}^n(-u)}{u^{1+\rho}} \, \mathrm{d}u.$$

## Passage to Logarithmic Expectations

Since

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

then, swapping limit and expectation (based on the Monotone Convergence Theorem) gives

$$\begin{aligned} \mathbb{E}\{\ln X\} &= \lim_{\rho \to 0^+} \frac{\mathbb{E}\{X^\rho\} - 1}{\rho} \\ &= \int_0^\infty \frac{e^{-u} - M_X(-u)}{u} \, \mathrm{d}u. \end{aligned}$$

## Extension to Fractional $\rho$-th moments with $\rho > 0$

$$\mathbb{E}\{X^\rho\} = \frac{1}{1+\rho} \sum_{\ell=0}^{\lfloor \rho \rfloor} \frac{\alpha_\ell}{B(\ell+1, \rho+1-\ell)}$$

$$+ \frac{\rho \sin(\pi\rho)\,\Gamma(\rho)}{\pi} \int_0^\infty \left( \sum_{j=0}^{\lfloor \rho \rfloor} \left\{ \frac{(-1)^j \, \alpha_j}{j!} \, u^j \right\} e^{-u} - M_X(-u) \right) \frac{\mathrm{d}u}{u^{\rho+1}},$$

where for all $j \in \{0, 1, \ldots, \}$

$$\alpha_j := \mathbb{E}\big\{(X-1)^j\big\} = \frac{1}{j+1} \sum_{\ell=0}^{j} \frac{(-1)^{j-\ell} \, M_X^{(\ell)}(0)}{B(\ell+1, j-\ell+1)},$$

and $B(\cdot, \cdot)$ denotes the Beta function:

$$B(u,v) := \int_0^1 t^{u-1} (1-t)^{v-1} \, \mathrm{d}t, \quad u, v > 0.$$

## Moments of Estimation Errors

Let $X_1, \ldots, X_n$ be i.i.d. random variables with an unknown expectation $\theta$ to be estimated, and consider the simple estimator,

$$\widehat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

## Moments of Estimation Errors

Let $X_1, \ldots, X_n$ be i.i.d. random variables with an unknown expectation $\theta$ to be estimated, and consider the simple estimator,

$$\widehat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Let

$$D_n := \left(\widehat{\theta}_n - \theta\right)^2$$

and

$$\rho' := \frac{\rho}{2}.$$

Then,

$$\mathbb{E}\left\{\left|\widehat{\theta}_n - \theta\right|^{\rho}\right\} = \mathbb{E}\left\{D_n^{\rho'}\right\}.$$

## Moments of Estimation Errors (Cont.)

By our formula, if $\rho > 0$ is a non–integral multiple of 2, then

$$\mathbb{E}\left\{\left|\widehat{\theta}_n - \theta\right|^\rho\right\} = \frac{2}{2+\rho} \sum_{\ell=0}^{\lfloor \rho/2 \rfloor} \frac{\alpha_\ell}{B\left(\ell+1, \rho/2+1-\ell\right)}$$

$$+ \frac{\rho \sin\left(\frac{\pi\rho}{2}\right) \Gamma\left(\frac{\rho}{2}\right)}{2\pi} \int_0^\infty \left( \sum_{j=0}^{\lfloor \rho/2 \rfloor} \left\{ \frac{(-1)^j \, \alpha_j}{j!} \, u^j \right\} e^{-u} - M_{D_n}(-u) \right) \frac{\mathrm{d}u}{u^{\rho/2+1}},$$

where

$$\alpha_j = \frac{1}{j+1} \sum_{\ell=0}^{j} \frac{(-1)^{j-\ell} \, M_{D_n}^{(\ell)}(0)}{B(\ell+1, j-\ell+1)}, \quad j \in \{0, 1, \ldots\},$$

$$M_{D_n}(-u) = \frac{1}{2\sqrt{\pi u}} \int_{-\infty}^{\infty} e^{-j\omega\theta} \, \phi_{X_1}^n\left(\frac{\omega}{n}\right) e^{-\omega^2/(4u)} \, \mathrm{d}\omega, \quad \forall\, u > 0,$$

and $\phi_{X_1}(\omega) := \mathbb{E}\{e^{j\omega X_1}\}$ $(\omega \in \mathbb{R})$ is the characteristic function of $X_1$.

## Moments of Estimation Errors: Example

Consider the case where $\{X_i\}_{i=1}^n$ are i.i.d. Bernoulli random variables with

$$\mathbb{P}\{X_1 = 1\} = \theta, \quad \mathbb{P}\{X_1 = 0\} = 1 - \theta$$

where the characteristic function is given by

$$\phi_X(u) := \mathbb{E}\Big\{e^{juX}\Big\} = 1 + \theta\Big(e^{ju} - 1\Big), \quad u \in \mathbb{R}.$$

## Moments of Estimation Errors: Example

Consider the case where $\{X_i\}_{i=1}^n$ are i.i.d. Bernoulli random variables with

$$\mathbb{P}\{X_1 = 1\} = \theta, \quad \mathbb{P}\{X_1 = 0\} = 1 - \theta$$

where the characteristic function is given by

$$\phi_X(u) := \mathbb{E}\left\{e^{juX}\right\} = 1 + \theta\left(e^{ju} - 1\right), \quad u \in \mathbb{R}.$$

## An Upper Bound via a Concentration Inequality

$$\mathbb{E}\left\{\left|\widehat{\theta}_n - \theta\right|^\rho\right\} \le K(\rho, \theta) \cdot n^{-\rho/2},$$

which holds for all $n \in \mathbb{N}$, $\rho > 0$ and $\theta \in [0, 1]$, with

$$K(\rho, \theta) := \rho\, \Gamma\left(\frac{\rho}{2}\right) \, \left(2\theta\,(1-\theta)\right)^{\rho/2}.$$

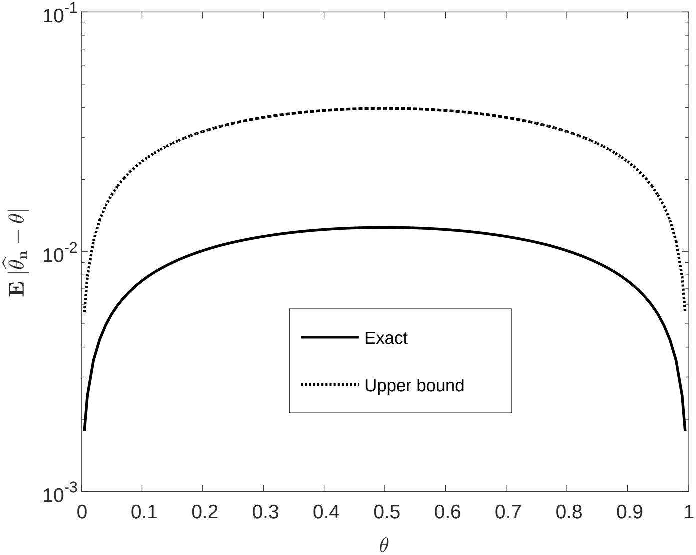


Figure: $\mathbb{E}\left|\widehat{\theta}_n - \theta\right|$ versus its upper bound as functions of $\theta$ with $n = 1000$.

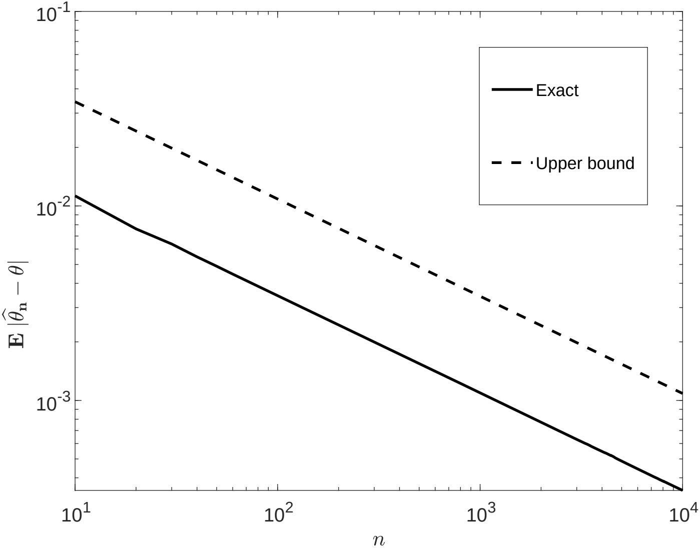


Figure: $\mathbb{E}\left|\widehat{\theta}_n - \theta\right|$ versus its upper bound as functions of $n$ with $\theta = \frac{1}{4}$.

## Channel Model

Consider a channel with

- input $\underline{X} = (X_1, \ldots, X_n) \in \mathcal{X}^n$ and output $\underline{Y} = (Y_1, \ldots, Y_n) \in \mathcal{Y}^n$;
- transition probability law:

$$p_{Y^n|X^n}(\underline{y}|\underline{x}) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \prod_{j \neq i} q_{Y|X}(y_j|x_j) \, r_{Y|X}(y_i|x_i) \right\}, \;\; (\underline{x}, \underline{y}) \in \mathcal{X}^n \times \mathcal{Y}^n.$$

This channel model refers to a DMC with a transition probability law

$$q_{Y^n|X^n}(\underline{y}|\underline{x}) = \prod_{i=1}^{n} q_{Y|X}(y_i|x_i),$$

where one of the transmitted symbols is jammed at a uniformly distributed random time, $i$, and the transition distribution of the jammed symbol is given by $r_{Y|X}(y_i|x_i)$ instead of $q_{Y|X}(y_i|x_i)$.

## Calculation of Mutual Information

We evaluate how the jamming affects the mutual information $I(X^n; Y^n)$.

For all $y \in \mathcal{Y}$, let

$$v(y) := \int_{\mathcal{X}} q_{Y|X}(y|x) \, p_X(x) \, \mathrm{d}x,$$

$$w(y) := \int_{\mathcal{X}} r_{Y|X}(y|x) \, p_X(x) \, \mathrm{d}x,$$

and let

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

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

## Calculation of Mutual Information

The integral representation of the logarithmic expectation give

$$
\begin{aligned}
& I_p(X^n; Y^n) \\
&= \int_0^\infty \frac{1}{u} \left[ t^{n-1}\left(\frac{u}{n}\right) s\left(\frac{u}{n}\right) - f^{n-1}\left(\frac{u}{n}\right) g\left(\frac{u}{n}\right) \right] \mathrm{d}u \\
&\quad + \int p_X(x)\, r_{Y|X}(y|x)\, \ln q_{Y|X}(y|x)\, \mathrm{d}x\, \mathrm{d}y - \int w(y)\, \ln v(y)\, \mathrm{d}y \\
&\quad + (n-1) \left[ \int p_X(x)\, q_{Y|X}(y|x)\, \ln q_{Y|X}(y|x)\, \mathrm{d}x\, \mathrm{d}y - \int v(y) \ln v(y)\, \mathrm{d}y \right].
\end{aligned}
$$

## Calculation of Mutual Information: Example

Let

- $q_{Y|X}$ be a BSC with crossover probability $\delta \in (0, \frac{1}{2})$;
- $r_{Y|X}$ be a BSC with a larger crossover probability, $\varepsilon \in (\delta, \frac{1}{2}]$;
- the input bits be i.i.d. and equiprobable.

## Calculation of Mutual Information: Example

Let

- $q_{Y|X}$ be a BSC with crossover probability $\delta \in (0, \frac{1}{2})$;
- $r_{Y|X}$ be a BSC with a larger crossover probability, $\varepsilon \in (\delta, \frac{1}{2}]$;
- the input bits be i.i.d. and equiprobable.

$$\begin{aligned} & I_p(X^n; Y^n) \\ & = n \ln 2 - d(\varepsilon \| \delta) - h_{\mathrm{b}}(\varepsilon) - (n-1) h_{\mathrm{b}}(\delta) \\ & \quad + \int_0^\infty \left\{ e^{-u} - \left[ (1-\delta) \exp\left( -\frac{(1-\varepsilon)u}{(1-\delta)n} \right) + \delta \exp\left( -\frac{\varepsilon u}{\delta n} \right) \right]^{n-1} \right. \\ & \qquad \left. \cdot \left[ (1-\varepsilon) \exp\left( -\frac{(1-\varepsilon)u}{(1-\delta)n} \right) + \varepsilon \exp\left( -\frac{\varepsilon u}{\delta n} \right) \right] \right\} \frac{\mathrm{d}u}{u}, \end{aligned}$$

where $h_{\mathrm{b}}(\cdot)$ and $d(\cdot \| \cdot)$ denote the binary entropy and binary relative entropy, respectively.

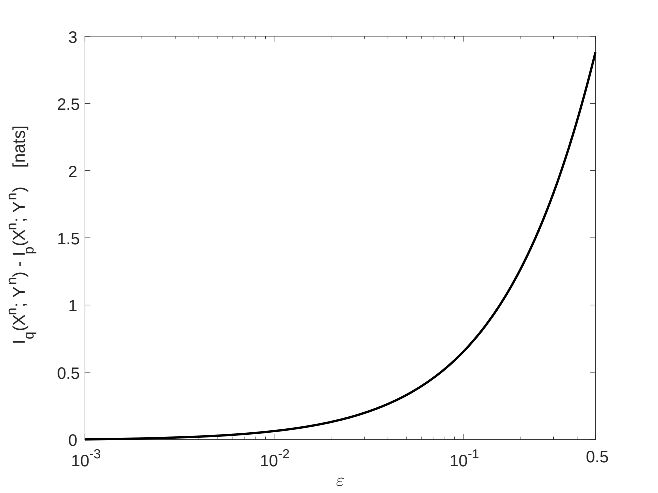


Figure: The degradation in mutual information for $n = 128$. The jammer–free channel $q$ is a BSC with crossover probability $\delta = 10^{-3}$, and $r$ is a BSC with crossover probability $\varepsilon \in \left(\delta, \frac{1}{2}\right]$.

## Summary

- We explore integral representations of the logarithmic and power functions.
- We demonstrate their usefulness for information-theoretic analyses.
- We obtain compact, easily-computable exact formulas for several source and channel coding problems.

## Journal Papers

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.

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