Some Results on the Generalized Gaussian Distribution

Alex Dytso[1], Ronit Bustin[2], H. Vincent Poor[1], Daniela Tuninetti[3], Natasha Devroye[3], and Shlomo Shamai (Shitz)[2]

***Abstract*—The paper considers information-theoretic applications of a broad class of distributions termed generalized Gaussian (GG). The flexible parametric form of the probability density function of the GG family makes it an excellent choice for many modeling scenarios. Well-know examples of this distribution include Laplace, Gaussian, and uniform.**

The first part of the paper explores properties of the GG distribution. In particular, it is shown that a GG random variable can be decomposed into a product of a Gaussian random variable and a independent positive random variable. The properties of this decomposition are carefully examined.

The second part of paper consideres a rate-distortion problem of GG sources under the L_p error distortion. For example, conditions are derived under which Shannon's lower bound is tight.

I. Introduction

A classical rate-distortion problem, first formulated by Shannon in [1], considers a source with the distribution P_X on $\mathcal{X}$, a reconstruction alphabet $\hat{\mathcal{X}}$ and a rate distortion measure $d : \mathcal{X} \times \hat{\mathcal{X}} \to \mathbb{R}^+$. One of the crowning achievements of Shannon is an exact expression for the rate-distortion function given by

$$R(D) = \inf_{P_{\hat{X}|X}: \mathbb{E}\left[d(X,\hat{X})\right] \le D} I(X; \hat{X}). \tag{1}$$

For continuous sources the rate-distortion function has been found for a Laplace source with an absolute error distortion $d(x, \hat{x}) = |x - \hat{x}|$ [2], exponential source with absolute error distortion [3] and Gaussian source with square error distortion $d(x, \hat{x}) = |x - \hat{x}|^2$ [1]. In this paper, we will enlarge this set of known cases by considering the rate-distortion problem for generalized Gaussian (GG) distributions.

[1]A. Dytso and H.V. Poor are with the Department of Electrical Engineering, Princeton University, NJ, Princeton 08544, USA (email: adytso, poor@princeton.edu).

[2]R. Bustin and S. Shamai (Shitz) are with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Technion City, Haifa 32000, Israel (e-mail: bustin@technion.ac.il, sshlomo@ee.technion.ac.il).

[3]D. Tuninetti and N. Devroye are with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, IL, Chicago 60607, USA (e-mail: danielat, devroye@uic.edu).

The work of A. Dytso and H.V. Poor was supported by the National Science Foundation under Grants CCF-1420575 and CNS-1456793. The work of R. Bustin and S. Shamai has been supported by the European Union's Horizon 2020 Research and Innovation Programme, grant agreement no. 694630. The work of Daniela Tuninetti and Natasha Devroye was supported in part by the National Science Foundation under Grant 1422511.

A. Problem Formulation

We shall refer to X with GG distribution given by the probability density function (pdf)

$$f_X(x) = \frac{c_q}{\alpha} \mathrm{e}^{-\frac{|x-\mu|^q}{2\alpha^q}}, \tag{2a}$$

$$c_q = \frac{q}{2^{\frac{q+1}{q}} \Gamma\left(\frac{1}{q}\right)}, \ x \in \mathbb{R}, \tag{2b}$$

as $X \sim \mathcal{N}_q(\mu, \alpha^q)$. The well-know examples of this family of distributions include: the Laplace distribution for $q = 1$; the Gaussian distribution for $q = 2$; and the uniform distribution on $[-\beta, \beta]$ for $q = \infty$ and $\alpha = \lim_{q\to\infty} \left(\frac{1}{2}\right)^{\frac{1}{q}} \beta$.

We consider the rate distortion problem with a GG source and a distortion measure that corresponds an ℓ_p-norm

$$d(x, \hat{x}) = |x - \hat{x}|^p, p \ge 1. \tag{3}$$

Formally, we seek to solve

$$R_{p,q}(\alpha, D) = \inf_{P_{\hat{X}|X}: \mathbb{E}\left[|X-\hat{X}|^p\right] \le \frac{2D^p}{p}} I(X; \hat{X}), \tag{4}$$

where $X \sim \mathcal{N}_q(\mu, \alpha^q)$. We also define

$$R_{\infty,q}(\alpha, D) = \inf_{P_{\hat{X}|X}: |X-\hat{X}| \le D \text{ a.s. } P_X} I(X; \hat{X}). \tag{5}$$

It should be noted that our analysis carries over to distributions on $\mathbb{R}^+$ such as an exponential distribution, and thus our analysis will subsume the known cases mentioned above.

B. Why Generalized Gaussian Sources?

The flexible parametric form of the pdf of GG distributions allows for tails that are either heavier than Gaussian ($p < 2$) or lighter than Gaussian ($p > 2$) which makes it an excellent choice for many modeling scenarios.

From the information theoretic perspective the GG distribution is interesting because it maximizes the entropy and Réyni entropy under a p-th absolute moment constraint [4], [5].

Theorem 1. *Let $X \in \mathbb{R}$ such that $\mathbb{E}[|X|^q] \le \frac{2\alpha^q}{q}$. Then*

$$h(X) \le \frac{1}{q} \log\left(\frac{q\mathrm{e}}{2c_q^q} \cdot \mathbb{E}[|X|^q]\right) \le \frac{1}{q} \log\left(\frac{\alpha^q \mathrm{e}}{c_q^q}\right). \tag{6}$$

The inequality in (6) *is attained with equality iff $X \sim \mathcal{N}_q(0, \alpha^q)$.*

Proof: This result can be proved via a method outlined in [4, Chapter 12]. ■

C. Shannon's Lower Bound

In [1] Shannon developed a technique for constructing a lower bound on the rate distortion which we now explore in our context.

Theorem 2. *For $X \sim \mathcal{N}_q(0, \alpha^q)$ and $(p, q) \in \mathbb{R}_+^2$*

$$R_{p,q}(\alpha, D) \geq \left[\log\left(\frac{\alpha}{D}\right) + \log\left(\frac{c_p}{c_q} \mathrm{e}^{\frac{1}{q}-\frac{1}{p}}\right)\right]^+. \tag{7}$$

Proof:

$$\begin{aligned} I(X;\hat{X}) &= h(X) - h(X|\hat{X}) \\ &= h(X) - h(X-\hat{X}|\hat{X}) \\ &\overset{a)}{\geq} h(X) - h(X-\hat{X}) \qquad (8a) \\ &\overset{b)}{\geq} \frac{1}{q}\log\left(\frac{\mathbb{E}[|X|^q]\cdot q\mathrm{e}}{2c_q^q}\right) \\ &\quad - \frac{1}{p}\log\left(\frac{\mathbb{E}[|X-\hat{X}|^p]\cdot p\mathrm{e}}{2c_p^p}\right) \qquad (8b) \\ &\overset{c)}{\geq} \log\left(\frac{\alpha}{D}\right) + \log\left(\frac{c_p}{c_q}\mathrm{e}^{\frac{1}{q}-\frac{1}{p}}\right), \qquad (8c) \end{aligned}$$

where the inequalities follow from: a) the fact that conditioning reduces entropy; b) the maximum entropy principle from Theorem 1; and c) the distortion constraint $\mathbb{E}\left[|X-\hat{X}|^p\right] \leq \frac{2D^p}{p}$. This concludes the proof. ■

The inequalities in (8) are tight if there exists a *backward test channel* for some random variable $\hat{X}$ such that

$$X = \hat{X} + Z, \tag{9}$$

and where $Z \sim \mathcal{N}_p(0, D^p)$ and independent of $\hat{X}$. Moreover, denoting by $\phi_r(t)$ the characteristic function of $X \sim \mathcal{N}_r(0,1)$, showing the existence of a test channel in (9) is equivalent to showing that the function

$$h(t) = \frac{\phi_q(\alpha t)}{\phi_p(Dt)}, \tag{10}$$

is a valid characteristic function of some random variable $\hat{X}$. W.l.o.g. in what follows we set $D = 1$ and $\alpha \geq 1$ and define

$$\phi_{(q,p,\alpha)}(t) = \frac{\phi_q(\alpha t)}{\phi_p(t)}. \tag{11}$$

Therefore, showing the existence of the test channel in (9) amounts to showing that $\phi_{(q,p,\alpha)}(t)$ is a valid characteristic function for $\alpha \geq 1$.

Remark 1. *Note that using an additive backward test channel is not the only way of achieving equalities in* (8). *However, this is one of the most commonly used techniques and understanding its limitations can be very valuable.*

D. Examples of Gaussian and Laplace Sources

The problem of analyzing $\phi_{(q,p,\alpha)}(t)$ in (11) can be best demonstrated by studying the following four cases:

- $(q,p) = (2,2)$ Gaussian source with a square-error distortion;
- $(q,p) = (1,1)$ Laplace source with an absolute-error distortion;
- $(q,p) = (2,1)$ Gaussian source with an absolute-error distortion; and
- $(q,p) = (1,2)$ Laplace source with a square-error distortion.

Recall, that the characteristic functions of Gaussian and Laplace random variables are given by $\phi_2(t) = \mathrm{e}^{-\frac{t^2}{2}}$ and $\phi_1(t) = \frac{1}{1+4t^2}$, respectively. Therefore, for the above four mentioned cases the function $\phi_{(q,p,\alpha)}(t)$ is given by

$$\phi_{(2,2,\alpha)}(t) = \mathrm{e}^{-\frac{(\alpha^2-1)t^2}{2}}, \quad \phi_{(2,1,\alpha)}(t) = (1+4t^2)\mathrm{e}^{-\frac{\alpha^2 t^2}{2}},$$
$$\phi_{(1,1,\alpha)}(t) = \frac{1+4t^2}{1+4\alpha^2t^2}, \quad \phi_{(1,2,\alpha)}(t) = \frac{\mathrm{e}^{\frac{t^2}{2}}}{1+4\alpha^2t^2}.$$

On the one hand, note that $\phi_{(2,2,\alpha)}(t) = \mathrm{e}^{-\frac{(\alpha^2-1)t^2}{2}}$ is a valid characteristic function and corresponds to $\hat{X} \sim \mathcal{N}_2(0, \alpha^2-1)$, and

$$\phi_{(1,1,\alpha)}(t) = \frac{1+4t^2}{1+4\alpha^2t^2} = \frac{1}{\alpha^2} + \left(1-\frac{1}{\alpha^2}\right)\frac{1}{1+4\alpha^2t^2}, \tag{12}$$

is a valid characteristic function of $\hat{X}$ with pdf given by

$$f_{\hat{X}}(x) = \frac{1}{\alpha^2}\delta(x) + \left(1-\frac{1}{\alpha^2}\right)\frac{c_1}{\alpha}\mathrm{e}^{-\frac{|x|}{2\alpha}}, \tag{13}$$

which is a pdf of a random variable $\hat{X} = B \cdot Y$ where $B \sim \mathsf{Bernoulli}\left(1-\frac{1}{\alpha^2}\right)$ and $Y \sim \mathcal{N}_1(0,\alpha)$ and where B and Y are independent.

On the other hand, recall that all characteristic functions satisfy the inequality $|\phi(t)| \leq 1$, which is not satisfied by $\phi_{(1,2,\alpha)}(t) = \frac{\mathrm{e}^{\frac{t^2}{2}}}{1+4\alpha^2t^2}$. Moreover, observe that $\phi_{(2,1,\alpha)}(t) = (1+4t^2)\mathrm{e}^{-\frac{\alpha^2t^2}{2}}$ has an inverse Fourier transform given by

$$\mathcal{F}^{-1}\{\phi_{(2,1,\alpha)}(t)\}(x) = \frac{1}{2\pi}\frac{(\alpha^2-x)\mathrm{e}^{-\frac{x^2}{2\alpha^2}}}{\alpha^5}, \tag{14}$$

which takes on negative values for $\alpha^2 < x$ and therefore cannot be a valid pdf.

Therefore, $\phi_{(1,2,\alpha)}(t)$ and $\phi_{(2,1,\alpha)}(t)$ are not valid characteristic functions and the test channel in (9) does not exists for $(p,q) = (1,2)$ and $(p,q) = (2,1)$.

Remark 2. *From the Gaussian and Laplace examples, we see that the existence of the test channel is not always guaranteed. This motivates answering the question for which values of (p,q) the test channel can be formed. However, answering such a question can be challenging since there exists no closed form expression for the characteristic function $\phi_r(t)$ for $r \neq 1, 2$. In fact, very little is known about properties* *of $\phi_r(t)$ for $r \neq 1,2$. Section III is, therefore, devoted to studying properties of the characteristic function of the GG random variable, which are of independent interest.*

II. Main Results

Our main results are given next.

Theorem 3. *Let $Z \sim \mathcal{N}_p(0,1)$ and for $(p,q) \in \mathbb{R}_+^2$ let*

$$\begin{aligned}\mathbb{S} = \{(q,p) : q \in (0,2),\ p>0,\ p \neq q\} \\ \cup \{(q,p) : q \geq 2, 0<p<q\}.\end{aligned} \tag{15}$$

Then we have:

- *for $(q,p) \in \mathbb{S}$ and any $\alpha \geq 1$ the test channel in* (9) cannot *be formed. In other words, for $(p,q) \in \mathbb{S}$ and any $\alpha \geq 1$ there exists no $\hat{X}$ independent of Z such that* (9) *holds with $X \sim \mathcal{N}_q(0,\alpha^q)$.*
- *for $\{(q,p) : 2 < q \leq p\}$ and* almost all[1] $\alpha \geq 1$ *the test channel in* (9) cannot *be formed. In other words, for $\{(p,q) : 2 < q \leq p\}$ and* almost all $\alpha \geq 1$ *there exists no $\hat{X}$ independent of Z such that* (9) *holds with $X \sim \mathcal{N}_q(0,\alpha^q)$.*
- *for $\{(p,q) : 1 \leq p = q \leq 2\}$ and all $\alpha \geq 1$ the test channel in* (9) *exists and Shannon's lower bound in Theorem 2 is attainable. Moreover, the distribution of $\hat{X}$ is given by*

$$f_{\hat{X}}(x) = \frac{\mathrm{e}^{-\frac{x^2}{2(\alpha^2-D^2)-1)}}}{\sqrt{2\pi((\alpha^2-D^2)-1)}}, \tag{16}$$

for $p=2$ and

$$f_{\hat{X}}(x) = \left(\frac{D}{\alpha}\right)^{p+1}\delta(x) + \left(1 - \left(\frac{D}{\alpha}\right)^{p+1}\right) g(x), \tag{17}$$

for $1 \leq p < 2$, and where $g(x)$ is the pdf of a random variable with characteristic function given by

$$\phi(t) = \frac{\phi_p(\alpha t) - \left(\frac{D}{\alpha}\right)^{p+1}\phi_p(Dt)}{\left(1-\left(\frac{D}{\alpha}\right)^{p+1}\right)\phi_p(Dt)}. \tag{18}$$

Note that for the regime $2 < q \leq p$ the results holds for almost all $\alpha \geq 1$, but we would like to point out that there are cases in this regime when a test channel can be established and Shannon's lower bound is tight as demonstrated next for the case of $(p,q) = (\infty,\infty)$.

Theorem 4. *For $X \sim \mathsf{Unif}[-\alpha,\alpha]$ and $\frac{\alpha}{D} \in \mathbb{N}$ (set of measure zero)*

$$R_{\infty,\infty}(\alpha,D) = \log^+\left(\frac{\alpha}{D}\right), \tag{19}$$

where the equality in (19) *is attained by a discrete random variable $\hat{X}$ uniformly distributed on*

$$\mathsf{supp}(\hat{X}) = \left\{\pm i \cdot D : i \in \left[1 : \frac{N}{2}\right]\right\},\ \textit{if } N \textit{ is even},$$
$$\mathsf{supp}(\hat{X}) = \left\{\pm i \cdot D : i \in \left[0 : \frac{N-1}{2}\right]\right\},\ \textit{if } N \textit{ is odd}.$$

where on $N = \frac{\alpha}{D}$.

Proof: Let $Z \sim \mathsf{Unif}[-D,D]$ independent of $\hat{X}$ and define a test channel as in (9).

Note, that by this construction we have that $X \sim \mathsf{Unif}[-\alpha,\alpha]$, $|X-\hat{X}| \leq D$ a.s., and

$$\begin{aligned} I(X;\hat{X}) &= h(X) - h(X|\hat{X}) \\ &= h(X) - h(Z) \\ &= \log\left(\frac{\alpha}{D}\right). \end{aligned}$$

This concludes the proof. ■

III. Properties of the Generalized Gaussian Distribution

In this section we study properties of the GG distribution.

A. Moments and Mellin Transform

The moments and absolute moments of the GG distribution are given next.

Proposition 1. *(Moments [6].) For any $p>0$ and $k>-1$ the moments of $X \sim \mathcal{N}_p(0,\alpha^p)$ are given by*

$$\mathbb{E}[X^k] = \frac{2^{\frac{k}{p}}\alpha^k}{\Gamma\left(\frac{1}{p}\right)}\Gamma\left(\frac{k+1}{p}\right),\ \textit{for } k = \textit{even}, \tag{20}$$

$$\mathbb{E}[X^k] = 0,\ \textit{for } k = \textit{odd}. \tag{21}$$

Definition 1. *The* Mellin transform *of a positive random variable X is defined as*

$$m_X(s) = \mathbb{E}[X^{s-1}], \tag{22}$$

for $s \in \mathbb{C}$ and $\mathsf{Re}(s) > 0$.

The Mellin transform emerges as a major tool in characterizing products of positive independent random variables since

$$m_{X \cdot Y}(s) = m_X(s) \cdot m_Y(s). \tag{23}$$

Proposition 2. *(Mellin Transform of $|X|$.) For any $p>0$ and $X \sim \mathcal{N}_p(0,1)$*

$$\mathbb{E}[|X|^{s-1}] = \frac{2^{\frac{s-1}{p}}\alpha^k}{\Gamma\left(\frac{1}{p}\right)}\Gamma\left(\frac{s}{p}\right) \tag{24}$$

where $s \in \mathbb{C}$ such that $\mathsf{Re}(s) > 0$. Moreover, for any $p>0$ and $k>-1$ the absolute moments are given by

$$\mathbb{E}[|X|^k] = \frac{2^{\frac{k}{p}}\alpha^k}{\Gamma\left(\frac{1}{p}\right)}\Gamma\left(\frac{k+1}{p}\right). \tag{25}$$

[1]In other words, the set of α for which the statement does not hold has Lebesgue measure zero.

Proof: The Mellin transform can be easily computed by using the integral

$$\int_0^\infty x^{s-1} e^{-x^p} dx = \frac{1}{p}\Gamma\left(\frac{s}{p}\right), \tag{26}$$

and a change of variable, and where the above integral is finite if $\mathsf{Re}(s) > 0$ and $p > 0$. ■

Note that the p-th absolute moment of $X \sim \mathcal{N}_p(0, \alpha^p)$ is given by

$$\mathbb{E}[|X|^p] = \frac{2\alpha^p}{p}. \tag{27}$$

The following corollary, which relates k-th moments of two GG distributions of a different order, is useful in many proofs.

Corollary 1. *Let $X_q \sim \mathcal{N}_q(0,1)$ and $X_p \sim \mathcal{N}_p(0,1)$ then for $q \geq p$*

$$\mathbb{E}[|X_q|^k] \leq \mathbb{E}[|X_p|^k], \tag{28}$$

for any $k \in \mathbb{R}^+$. Moreover, for $q > p$

$$\lim_{k\to\infty} \left(\frac{\mathbb{E}[|X_p|^k]}{\mathbb{E}[|X_q|^k]}\right)^{\frac{1}{k}} = \infty. \tag{29}$$

Proof: The proof follows by using Proposition 2 and Stirling's approximation for the gamma function. ■

B. Relation to Positive Definite Functions

As will be observed throughout this paper, the GG distribution exhibits different properties depending whether $p \leq 2$ or $p > 2$. At the heart of this behavior is the concept of positive-definite functions.

Definition 2. *A function $f : \mathbb{R} \to \mathbb{C}$ is called* positive definite *if for every positive integer n and all real numbers $x_1, x_2, ..., x_n$ the $n \times n$ matrix*

$$A = (a_{i,j})_{i,j=1}^n, \ a_{i,j} = f(x_i - x_j), \tag{30}$$

is positive semi-definite.

Our first result relates the pdf of the GG distribution to the class of positive definite functions.

Theorem 5. *The function $\mathrm{e}^{-\frac{|x|^p}{2}}$ is*

- *not positive definite for $p > 2$; and*
- *positive definite for $0 < p \leq 2$. Moreover, for $x > 0$*

$$\mathrm{e}^{-\frac{x^p}{2}} = \int_0^\infty e^{-\frac{tx^2}{2}} d\mu(t), \tag{31}$$

where $d\mu(t)$ (independent of x) is a finite non-negative Borel measure on $t \in [0, \infty]$.

Proof: See Appendix B. ■

The concept of positive-definite functions will also play an important role in examining properties of the characteristic function of the GG distribution.

C. Product Decomposition of the GG Random Variable

As a consequence of Theorem 5 we have the following decompositional representation of the GG random variable.

Proposition 3. *For any $0 < p \leq 2$ and $X \sim \mathcal{N}_p(0,1)$ we have that*

$$X = V \cdot Z, \tag{32}$$

where V (dependent on p) is a positive random variable independent of $Z \sim \mathcal{N}_2(0,1)$. Moreover, V has the following properties:

- *V is an unbounded random variable for $p < 2$ and $V = 1$ for $p = 2$; and*
- *for $p < 2$ V is a continuous random variable with pdf given by*

$$f_V(v) = \frac{1}{2\pi}\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{p}\right)} \int_{\mathbb{R}} v^{-it-1} \frac{2^{\frac{it}{p}}\Gamma\left(\frac{it+1}{p}\right)}{2^{\frac{it}{2}}\Gamma\left(\frac{it+1}{2}\right)} dt, \ v > 0. \tag{33}$$

Proof: See Appendix A. ■

The pdf of V can be also be represented as a power series expansion.

Proposition 4. *For $p < 2$ the pdf of V in (33) is given by*

$$f_V(v) = \frac{p}{\sqrt{\pi}\Gamma\left(\frac{1}{p}\right)} \sum_{k=0}^\infty a_k v^{kp}, \ v > 0, \tag{34}$$

where

$$a_k = \frac{(-1)^{k+1}\sin\left(\frac{\pi kp}{2}\right)\Gamma\left(\frac{kp}{2}+1\right)2^{(kp+1)\left(\frac{1}{2}-\frac{1}{p}\right)}}{k!}. \tag{35}$$

Proof: The proof follows by using the residue theorem on the integral in (33). ■

Remark 3. *For the case of $p = 1$, the random variable V is distributed according to the Rayleigh distribution.*

D. Characteristic Function

The focus of this section is on the characteristic function of the GG distribution which will play an important role in analyzing the output distribution of additive channels with GG distributed noise. The characteristic function of the GG distribution can be given in the following integral form.

Theorem 6. *For any $p > 0$ the characteristic function of $X \sim \mathcal{N}_p(0, \alpha^p)$ is given by*

$$\phi_p(t) = 2c_p \int_0^\infty \cos(t\alpha x) e^{-\frac{x^p}{2}} dx. \tag{36}$$

Proof: The proof follows from the fact that $e^{-\frac{|x|^p}{2}}$ is an even function and therefore its Fourier transform is equivalent to the cosine transform. ■

Examples of characteristic functions of $X \sim \mathcal{N}_p(0,1)$ for several values of p are given in Fig. 1.

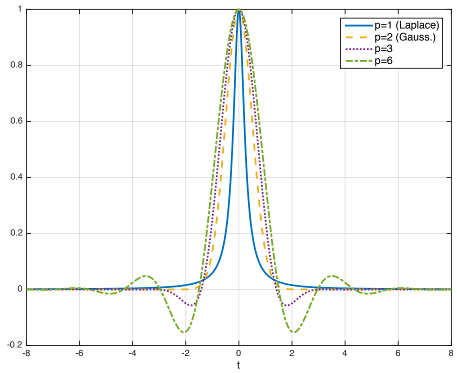

Fig. 1: Plot of the characteristic function of $X \sim \mathcal{N}_p(0,2)$ for several values of p.

E. On the Distribution of Zeros of the Characteristic Function

As can be seen from Fig. 1 the characteristic function of the GG distribution can have zeros. The following theorem gives a somewhat surprising result on the distribution of the zeros of $\phi_p(t)$ in (36).

Theorem 7. *(*Distribution of Zeros.*) The characteristic function of $\phi_p(t)$ in* (36) *has the following properties:*

- *for $p > 2$, $\phi_p(t)$ has at least one positive to negative zero crossing; and*
- *for $0 \le p \le 2$, $\phi_p(t)$ is non-negative and can be expressed as*

$$\phi_p(t) = \mathbb{E}\left[\mathrm{e}^{-\frac{t^2 V^2}{2}}\right], \tag{37}$$

where the random variable V (dependent on p) is defined in (32)*.*

Proof: See Appendix B. ∎

Remark 4. *Note that for $0 < p \le 2$ the function $\phi_p(\sqrt{2t})$ can be thought of as a Laplace transform of a random variable V^2. This observation will be useful in several of the proofs.*

F. Asymptotics of $\phi_p(t)$

By using the representation of $\phi_p(t)$ in Theorem 7, we can give an exact tail behavior for $\phi_p(t)$ in (36).

Proposition 5. *For $0 < p < 2$ we have that*

$$\lim_{t\to\infty} \phi_p(t) \cdot \frac{t^{p+1}}{A_1} = 1, \tag{38}$$

where $A_1 = \frac{p 2^{\frac{p-1}{2}} \Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{1}{p}\right)} a_1$ and a_1 is defined in (35)*. Moreover,*

$$\lim_{t\to\infty} \phi_{(q,p,\alpha)}(t) = \begin{cases} 0, & q > p \\ \frac{1}{\alpha^{q+1}}, & q = p \\ \infty, & q < p \end{cases}. \tag{39}$$

Proof: See Appendix C. ∎

G. Deconvolution Results

We are now in a position to answer whether $\phi_{(q,p,\alpha)}(t)$ in (11) is a characteristic function.

Theorem 8. *The function $\phi_{(q,p,\alpha)}(t)$ in* (11) *has the following properties*

- *for $(q,p) \in \mathbb{S}$, $\phi_{(q,p,\alpha)}(t)$ is not a characteristic function for any $\alpha \ge 1$; and*
- *for every $\{(p,q) : 2 < q \le p\}$ there exists an $\alpha \ge 1$ such that $\phi_{(q,p,\alpha)}(t)$ is not a characteristic function.*

Proof: See Appendix D. ∎

We would like to point out that for $2 < q \le p$ there are cases when $\phi_{(q,p,\alpha)}(t)$ is a characteristic function. Specifically, we can find an $\alpha \ge 1$ such that $\phi_{(q,p,\alpha)}(t)$ is a characteristic function for $p > 2$ and $q > 2$. The most trivial case is of $p = q$ and $\alpha = 1$, for which

$$\phi_{(q,p,\alpha)}(t) = 1, \tag{40}$$

is a characteristic function. A less trivial example is when $p = q = \infty$ in which case $\phi_\infty(t) = \mathrm{sinc}(t)$ and

$$\phi_{(\infty,\infty,\alpha)}(t) = \frac{\mathrm{sinc}(\alpha t)}{\mathrm{sinc}(t)}. \tag{41}$$

For example, when $\alpha = 2$ we have that $\phi_{(\infty,\infty,\alpha)}(t) = \frac{1}{2}\cos(2t)$, which corresponds to the characteristic function of the random variable $\hat{X} = \pm 1$ equally likely. Note that in the above example, because zeros of $\phi_p(t)$ occur periodically, we can select α such that the poles and zeros of $\phi_{(q,p,\alpha)}(t)$ cancel. However, we conjecture that such examples are only possible for $p = \infty$, and for $2 < p < \infty$ zeros of $\phi_p(t)$ do not appear periodically leading to the following conjecture

Conjecture 1. *For $2 < q \le p < \infty$, $\phi_{(q,p,\alpha)}(t)$ is not a characteristic function for any $\alpha > 1$.*

It is not difficult to check, by using the property that convolution with an analytic function is again analytic, that Conjecture 1 is true if p is an even integer and q is any non-even real number.

In view of Theorem 8 it remains to answer what happens to $\phi_{(q,p,\alpha)}(t)$ when $0 < q = p < 2$.

Theorem 9. *For $1 \le q = p < 2$, $\phi_{(p,p,\alpha)}(t)$ in* (11) *is a characteristic function. Moreover,*

$$\phi_{(p,p,\alpha)}(t) = \frac{1}{\alpha^{p+1}} + \left(1 - \frac{1}{\alpha^{p+1}}\right) G(t), \tag{42}$$

where $G(t)$ is a characteristic function of a continuous random variable.

Proof: See Appendix E. ∎

Note that for the case of $0 < q = p < 1$ whether $\phi_{(p,p,\alpha)}(t)$ is a characteristic remains an open question.

H. Proof of Theorem 3

The proof the main results now follows easily by noting that the test channel in (9) does not exist under the circumstances described in Theorem 8, and exists (i.e., Shannon's lower bound is achievable) under the circumstances described in Theorem 9.

IV. Conclusion

In this work, we have examined a problem of lossy compression for the generalized Gaussian family of source distributions. For the case when the distortion and the source distribution are matched (i.e., $p = q$), a closed form expression for the rate-distortion function has been given and shown to be attained by Shannon's lower bound. Moreover, for cases when the distortion and the source distribution are not matched, conditions under which no backward test channel exists have been given.

Appendix A
Proof of Proposition 3

To show that $X = V \cdot Z$, first observe that $d\nu(t) = c_p\sqrt{\frac{2\pi}{t}}d\mu(t)$ is a probability measure where $d\mu(t)$ was defined in Theorem 5.

$$\begin{aligned}
1 = \mathbb{P}(X \in \mathbb{R}) &= \int_{\mathbb{R}} c_p \mathrm{e}^{-\frac{|x|^p}{2}} dx \\
&\overset{a)}{=} \int_{\mathbb{R}} c_p \int_0^\infty e^{-\frac{tx^2}{2}} d\mu(t) dx \\
&\overset{b)}{=} c_p \int_0^\infty \int_{\mathbb{R}} e^{-\frac{tx^2}{2}} dx d\mu(t) \\
&= c_p \int_0^\infty \sqrt{\frac{2\pi}{t}} d\mu(t) \\
&= \int_0^\infty d\nu(t),
\end{aligned}$$

where the equalities follow from: a) using the representation of $\mathrm{e}^{-\frac{|x|^p}{2}}$ in Theorem 5; and b) switching of the order of integration as justified by Tonelli's theorem for positive functions.

The above implies that $d\nu(t) = c_p\sqrt{\frac{2\pi}{t}}d\mu(t)$ is a probability measure on $[0,\infty)$. Moreover, for any measurable set $\mathcal{S} \subset \mathbb{R}$ we have

$$\begin{aligned}
\mathbb{P}(X \in \mathcal{S}) &= \int_{\mathcal{S}} c_p \mathrm{e}^{-\frac{|x|^p}{2}} dx \\
&\overset{a)}{=} \int_{\mathcal{S}} c_p \int_0^\infty e^{-\frac{tx^2}{2}} d\mu(t) dx \\
&= \int_0^\infty \int_{\mathcal{S}} \sqrt{\frac{t}{2\pi}} e^{-\frac{tx^2}{2}} dx c_p \sqrt{\frac{2\pi}{t}} d\mu(t) \\
&= \int_0^\infty \mathbb{P}\left(\frac{1}{\sqrt{t}} Z \in \mathcal{S}\right) c_p \sqrt{\frac{2\pi}{t}} d\mu(t) \\
&\overset{b)}{=} \int_0^\infty \mathbb{P}\left(\frac{1}{\sqrt{T}} Z \in \mathcal{S} \mid T = t\right) c_p \sqrt{\frac{2\pi}{t}} d\mu(t) \\
&\overset{c)}{=} \mathbb{E}\left[\mathbb{P}\left(\frac{1}{\sqrt{T}} Z \in \mathcal{S} \mid T\right)\right] \\
&= \mathbb{P}\left(\frac{1}{\sqrt{T}} \cdot Z \in \mathcal{S}\right) \\
&\overset{d)}{=} \mathbb{P}\left(V \cdot Z \in S\right),
\end{aligned}$$

where the equalities follow from: a) the representation of $\mathrm{e}^{-\frac{|x|^p}{2}}$ in Theorem 5; b) the fact that $c_p\sqrt{\frac{2\pi}{t}}d\mu(t)$ is a probability measure; c) because Z is independent of t; and d) renaming $V = \frac{1}{\sqrt{T}}$.

Therefore, it follows that X can be decomposed into $V \cdot Z$.

To find the pdf of V we use the Mellin transform approach by observing that

$$\mathbb{E}[|X|^{it}] = \mathbb{E}[|V \cdot Z|^{it}] = \mathbb{E}[V^{it}] \cdot \mathbb{E}[|Z|^{it}]. \tag{43}$$

Therefore, by using Proposition 2 we have that the Mellin transform of V is given by

$$\mathbb{E}[V^{it}] = \frac{\mathbb{E}[|X|^{it}]}{\mathbb{E}[|Z|^{it}]} = \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{p}\right)} \frac{2^{\frac{it}{p}}\Gamma\left(\frac{it+1}{p}\right)}{2^{\frac{it}{2}}\Gamma\left(\frac{it+1}{2}\right)}. \tag{44}$$

Finally, the pdf of V is computed by using the inverse Mellin transform of (44)

$$f_V(v) = \frac{1}{2\pi} \frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{p}\right)} \int_{\mathbb{R}} v^{-it-1} \frac{2^{\frac{it}{p}}\Gamma\left(\frac{it+1}{p}\right)}{2^{\frac{it}{2}}\Gamma\left(\frac{it+1}{2}\right)} dt, \ v > 0.$$

This concludes the proof.

Appendix B
Proof of Theorem 7

We first show that for $p > 2$ there is at least one zero. We use the approach of [7]. Towards a contradiction assume that $\phi_p(t) \geq 0$ for all $t \geq 0$, then for all $t \geq 0$ we have that

$$\begin{aligned}
0 &\leq \frac{1}{2c_p}\frac{1}{2\pi} \int_0^\infty \phi_p(x)(1-\cos(xt))^2 dx \\
&\overset{a)}{=} \frac{1}{2c_p}\frac{1}{2\pi} \int_0^\infty \phi_p(x) \frac{1}{2}\left(3 - 4\cos(tx) + \cos(2tx)\right) dx \\
&\overset{b)}{=} 3 - 4\mathrm{e}^{-\frac{t^p}{2}} + \mathrm{e}^{-\frac{(2t)^p}{2}},
\end{aligned}$$

where the equalities follow from: a) using $(1-\cos(xt))^2 = \frac{1}{2}\left(3 - 4\cos(tx) + \cos(2tx)\right)$; and b) using the inverse Fourier transform and Parseval's identity. For small x we can write $\mathrm{e}^{-x} = 1 - x + O(x^2)$ and therefore we have that

$$\begin{aligned}
0 &\leq 3 - 4\left(1 - \frac{t^p}{2}\right) + \left(1 - \frac{(2t)^p}{2}\right) + O(t^{2p}) \\
&= (4 - 2^p)\frac{t^p}{2} + O(t^{2p}).
\end{aligned}$$

Therefore, for $p > 2$ we reach a contradiction since $4 - 2^p < 0$ for $p > 2$. This concludes the proof for the case of $p > 2$.

For $0 < p \leq 2$ the proof goes as follows:

$$\begin{aligned}
\phi_p(t) &= \mathbb{E}[\mathrm{e}^{itX}] \\
&\overset{a)}{=} \mathbb{E}[\mathrm{e}^{itVZ}] \\
&= \mathbb{E}\left[\mathbb{E}[\mathrm{e}^{itVZ}|V]\right] \\
&\overset{b)}{=} \mathbb{E}\left[\mathrm{e}^{-\frac{t^2V^2}{2}}\right] \\
&\overset{c)}{>} 0,
\end{aligned}$$

where the (in)-equalities follow from: a) the decomposition in Proposition 3; b) the independence of V and Z and the fact that the characteristic function of Z is $\mathrm{e}^{-\frac{t^2}{2}}$; and c) the positivity of $\mathrm{e}^{-\frac{t^2 V^2}{2}}$.

This concludes the proof.

Appendix C
Proof of Proposition 5

The proof follows by observing that

$$\begin{aligned}\phi_p(t) &= \int_0^\infty \mathrm{e}^{-\frac{t^2 v^2}{2}} f_V(v) dv \\ &\overset{a)}{\leq} \frac{p}{\sqrt{\pi}\Gamma\left(\frac{1}{p}\right)} a_1 \int_0^\infty \mathrm{e}^{-\frac{t^2 v^2}{2}} v^p dv \\ &\overset{b)}{=} \frac{p}{\sqrt{\pi}\Gamma\left(\frac{1}{p}\right)} a_1 2^{\frac{p-1}{2}} \Gamma\left(\frac{p+1}{2}\right) \frac{1}{t^{p+1}} \\ &= A_1 \frac{1}{t^{p+1}},\end{aligned} \tag{45}$$

where the (in)-equalities follow from: a) using the first term of the series in Proposition 4 to bound $f_V(v)$; and b) using the p-th absolute moment of a Gaussian random variable from Proposition 2.

The lower bound on $\phi_p(t)$ follows similarly and is given by

$$A_1 \frac{1}{t^{p+1}} - O\left(\frac{1}{t^{2p+1}}\right) \geq \phi_p(t). \tag{46}$$

The upper bound in (45) and the lower bound in (46) imply that

$$\lim_{t\to\infty} \phi_p(t) \cdot \frac{t^{p+1}}{A_1} = 1. \tag{47}$$

This concludes the proof.

Appendix D
Proof of Theorem 8

A. *Case of $q > p > 0$*

We start by looking at the regime $q > p > 0$. We want to show that there exists no random variable X independent of $Z \sim \mathcal{N}_p(0,1)$ such that

$$\alpha X = \hat{X} + Z, \tag{48}$$

where $X \sim \mathcal{N}_q(0,1)$ for all $\alpha \geq 1$. Note that, since X and Z are symmetric and have finite moments, if such an $\hat{X}$ exists it must also be symmetric with finite moments. Then for all $k \geq 1$

$$\begin{aligned}\alpha^k \mathbb{E}[|X|^k] &= \mathbb{E}[|\hat{X} + Z|^k] \\ &= \mathbb{E}[\mathbb{E}[|\hat{X} + Z|^k \mid Z]] \\ &\overset{a)}{\geq} \mathbb{E}[|\mathbb{E}[\hat{X} + Z \mid Z]|^k]] \\ &\overset{b)}{=} \mathbb{E}[|\mathbb{E}[\hat{X}] + Z|^k]] \\ &= \mathbb{E}[|Z|^k],\end{aligned} \tag{49}$$

where the (in)-equalities follow from: a) Jensen's inequality; and b) the independence of $\hat{X}$ and Z.

This in turn implies that, in order for inequality in (49) to be true, we must have that

$$\alpha \geq \left(\frac{\mathbb{E}[|Z|^k]}{\mathbb{E}[|X|^k]}\right)^{\frac{1}{k}}, \tag{50}$$

for all $k \geq 1$. However, by Corollary 1 for $p < q$

$$\alpha \geq \lim_{k\to\infty} \left(\frac{\mathbb{E}[|Z|^k]}{\mathbb{E}[|X|^k]}\right)^{\frac{1}{k}} = \infty;$$

therefore, there exists no $\alpha \geq 1$ that can satisfy (50) for all $k \geq 1$. This concludes the proof for $p < q$.

B. *Case of $p = 2$ and $q < 2$*

Note that in the case of $p = 2$ and $q < 2$ we want to show that there is no random variable $\hat{X}$ such that the convolution leads to

$$f_X(y) = c_2 \mathbb{E}\left[\mathrm{e}^{-\frac{(y-\hat{X})^2}{2}}\right], \text{ where } f_X(y) = \frac{c_q}{\alpha} \mathrm{e}^{-\frac{|y|^q}{2\alpha^q}}.$$

Such an $\hat{X}$ does not exist since the convolution preserves analyticity. That is the convolution with an analytic pdf must result in an analytic pdf. Noting that $f_X(y)$ is not analytic for $q < 2$ (i.e., the derivative at zero is not defined) leads to the desired conclusion.

C. *Case of $p > 2$ and $q \leq 2$*

Now for $p > 2$ and $q \leq 2$ the function $\phi_{(q,p,\alpha)}(t)$ has a pole but no zeros by Theorem 7. Therefore, for the case of $p > 2$ and $q \leq 2$ there exists a t_0, namely the pole of $\phi_{(q,p,\alpha)}(t)$, such that $\phi_{(q,p,\alpha)}(t)$ is not continuous at $t = t_0$. This violates the condition that the characteristic function is always a continuous function of t and therefore $\phi_{(q,p,\alpha)}(t)$ is not a characteristic function for all $\alpha \geq 1$.

D. *Case of $p \geq q > 2$.*

For the case of $p > 2$ and $q > 2$ the function

$$\phi_{(q,p,\alpha)}(t) = \frac{\phi_q(\alpha t)}{\phi_p(t)}, \tag{51}$$

has both poles and zeros by Theorem 7. Moreover, let t_1 be such that is $\phi_p(t_1) = 0$ and we can always choose an α such that $\phi_q(\alpha t_1) \neq 0$ and

$$\phi_{(q,p,\alpha)}(t_1) = \infty. \tag{52}$$

In other words, we choose an α such that the poles do not cancel the zeros. This implies that there exists an α such that $\phi_{(p,q,\alpha)}(t)$ is not a continuous function of t and is not a characteristic function. Moreover, it is not difficult to show that the above can be done for almost all $\alpha \geq 1$.

E. Case of $q < p < 2$

Finally, for the regime $q < p < 2$ the result follows from Proposition 5 where it is shown that

$$\lim_{n \to \infty} \phi_{(q,p,\alpha)}(t) = \infty, \tag{53}$$

which violates the condition that the characteristic function is bounded. This concludes the proof.

APPENDIX E
PROOF OF THEOREM 9

To show that $\phi_{(p,p,\alpha)}$ is a characteristic function we use Cramer's theorem [8] which requires verification that

$$\frac{2}{x^2} \int_0^{xA} (xA - u) \phi_{(p,p,\alpha)} \left(\frac{u}{x} \right) \cos(u) du \geq 0, \tag{54}$$

for $x > 0$ and $A \geq 0$.

First, we focus on $xA \leq \pi$. By using the fact that $\phi_{(p,p,\alpha)}(t)$ is monotonically decreasing, by the *second mean value theorem for integration* we have that

$$\begin{aligned} &\int_0^{xA} (xA - u) \phi_{(p,p,\alpha)} \left(\frac{u}{x} \right) \cos(u) du \\ &= \phi_{(p,q,\alpha)} (0) \int_0^{b} (A - \frac{u}{x}) \cos(u) du, \end{aligned} \tag{55}$$

for some $b \in [0, xA]$. Note that the integral $\int_0^b (xA - u) \cos(u) du \geq 0$ for all $b \in [0, xA]$.

For $xA > \pi$ let K be the largest integer such that $xA \geq K\pi$; then

$$\begin{aligned} &\int_0^{xA} (xA - u) \phi_{(p,p,\alpha)} \cos(u) du \\ &= \sum_{k=0}^{K-1} \int_{k\pi}^{(k+1)\pi} (xA - u) \phi_{(p,p,\alpha)} \cos(u) du \\ &+ \int_{K\pi}^{xA} (xA - u) \phi_{(p,p,\alpha)} \cos(u) du, \end{aligned}$$

and the proof then follows similarly to the case of $xA \leq \pi$ by using second mean value theorem for integration on each of the domains. For further details see [9].

Next, we show that

$$\phi_{(p,p,\alpha)}(t) = \frac{1}{\alpha^{p+1}} + \left(1 - \frac{1}{\alpha^{p+1}} \right) G(t), \tag{56}$$

where $G(t) = \frac{\left(\phi_{(p,p,\alpha)}(t) - \frac{1}{\alpha^{p+1}} \right)}{\left(1 - \frac{1}{\alpha^{p+1}} \right)}$ is the characteristic function of a continuous distribution.

First, observe that $f(t) = 1$ is a characteristic function of $\delta(x)$. Second, recall that the set of characteristic functions is closed under convex combinations, and since $\phi_{(p,p,\alpha)}(t)$ and $f(t) = 1$ are characteristic function so is $G(t)$.

To show that $G(t)$ is the characteristic function of a continuous distribution it is enough to show that $G(t)$ is integrable (i.e., we can perform Fourier inversion). The integrability of $G(t)$ follows from the fact that $G(t)$ is bounded and

$$\lim_{t \to \infty} t^2 G(t) = 0; \tag{57}$$

for details see [9]. This concludes the proof.

REFERENCES

[1] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 379-423, 623-656, Jul., Oct. 1948.

[2] H. Si, O. O. Koyluoglu, and S. Vishwanath, "Lossy compression of exponential and laplacian sources using expansion coding," in *Proc. IEEE Int. Symp. Inf. Theory*. IEEE, 2014, pp. 3052–3056.

[3] S. Verdú, "The exponential distribution in information theory," *Problemy Peredachi Informatsii*, vol. 32, no. 1, pp. 100–111, 1996.

[4] T. Cover and J. Thomas, *Elements of Information Theory: Second Edition*. Wiley, 2006.

[5] 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, April 2007.

[6] S. Nadarajah, "A generalized normal distribution," *Journal of Applied Statistics*, vol. 32, no. 7, pp. 685–694, 2005.

[7] N. Elkies, A. Odlyzko, and J. Rush, "On the packing densities of superballs and other bodies," *Inventiones Mathematicae*, vol. 105, no. 1, pp. 613–639, 1991.

[8] N. G. Ushakov, *Selected Topics in Characteristic Functions*. Walter de Gruyter, 1999.

[9] A. Dytso, R. Bustin, H. V. Poor, D. Tuninetti, N. Devroye, and S. Shamai, "The generalized Gaussian distribution in information theory: Rate distortion," www.princeton.edu/ ∼ adytso/papers/GGInfoTheoreSourceCoding.pdf, 2017.