J. Appl. Prob. **53,** 360–368 (2016)
*doi:*10.1017/jpr.2016.5

EXTENSION OF DE BRUIJN'S IDENTITY TO DEPENDENT NON-GAUSSIAN NOISE CHANNELS

NAYEREH BAGHERI KHOOLENJANI * ** AND
MOHAMMAD HOSSEIN ALAMATSAZ,* *** *University of Isfahan*

Abstract

De Bruijn's identity relates two important concepts in information theory: Fisher information and differential entropy. Unlike the common practice in the literature, in this paper we consider general additive non-Gaussian noise channels where more realistically, the input signal and additive noise are not independently distributed. It is shown that, for general dependent signal and noise, the first derivative of the differential entropy is directly related to the conditional mean estimate of the input. Then, by using Gaussian and Farlie–Gumbel–Morgenstern copulas, special versions of the result are given in the respective case of additive normally distributed noise. The previous result on independent Gaussian noise channels is included as a special case. Illustrative examples are also provided.

Keywords: Differential entropy; Fisher information; Gaussian copula; Farlie–Gumbel–Morgenstern copula

2010 Mathematics Subject Classification: Primary 54C70; 62H20

1. Introduction

Differential entropy and Fisher information are two performance measures in information theory and estimation theory, respectively. They are of fundamental importance and have been extensively studied. Recall that the differential entropy of a continuous random variable (RV) Y is defined as

$$h(Y) = -\int_{-\infty}^{\infty} f_Y(y;\gamma)\log f_Y(y;\gamma)\,\mathrm{d}y, \tag{1}$$

where γ is a deterministic parameter in the probability density function f_Y. Fisher information with respect to location parameter is defined as

$$J(Y) = \mathbb{E}_Y\bigg[\bigg(\frac{\partial}{\partial Y}\log f_Y(Y;\gamma)\bigg)^2\bigg],$$

which is equivalently expressed as

$$J(Y) = -\mathbb{E}_Y\bigg[\bigg(\frac{\partial^2}{\partial Y^2}\log f_Y(Y;\gamma)\bigg)\bigg].$$

The importance of Fisher information as a measure of information is well known. It has many implications in estimation theory, as exemplified by the Cramer–Rao lower bound which

Received 15 May 2015; revision received 16 August 2015.
* Postal address: Department of Statistics, University of Isfahan, Isfahan, 81746-73441, Iran.
** Email address: n.b.khoolenjani@gmail.com
*** Email address: alamatho@sci.ui.ac.ir

is a fundamental limit on the variance of an estimator. De Bruijn's identity, stated in Theorem 1 below, relates these two concepts [5]. It was first considered by Stam [16] and, since then, it has been widely used by many researchers. Recently, a renewed interest was manifested in the applications of de Bruijn's identity in estimation and turbo (iterative) decoding schemes, and in relating the input-output mutual information with the minimum mean-square error for additive Gaussian and non-Gaussian noise channels [3], [12].

Consider additive noise channels of the form

$$Y = X + \sqrt{\gamma} Z, \tag{2}$$

where X, the input signal and Z, the additive noise, are arbitrary RVs and the parameter γ is nonnegative. When the additive noise Z is a Gaussian RV with zero mean and unit variance independent of the input RV X, an elegant algebraic connection between the differential entropy of output signal Y and Fisher information, known as de Bruijn's identity, is as stated below.

Theorem 1. (De Bruijn's identity [1], [14].) *Let X be an arbitrary RV with finite second-order moment and Z be normally distributed independent of X with zero mean and unit variance. Then, we have*

$$\frac{\partial}{\partial \gamma} h(X + \sqrt{\gamma} Z) = \frac{1}{2} J(X + \sqrt{\gamma} Z).$$

A new version of de Bruijn's identity for stable RVs was stated by Johnson [6]. He obtained expressions for the derivative of the differential entropy of output signal Y as an inner product of score functions. Recently, an extension of de Bruijn's identity to non-Gaussian Z was provided by Park *et al.* [11]. Their results, however, are based on an independence assumption between input signal and additive noise. In the real world, there are signal-dependent noise channels in which the noise characteristics depend highly on the transmitted signal; see [2], [9], and [13]. Applications for such dependency appear also in radar and sonar systems [7]. Therefore, a natural question arises as to how de Bruijn's identity can be extended to deal with dependent RVs. In this paper, an extension of de Bruijn's identity is presented in the case when input signal X and additive noise Z are two arbitrary dependent RVs. It is found that the first derivative of the differential entropy of output signal Y can be expressed as a function of the conditional mean estimate associated with the input signal. Assuming the dependence, between input signal X and additive noise Z, is modeled by some copula functions, relationships between the first derivative of the differential entropy of Y and the Fisher information are derived.

The remainder of this paper is organized as follows. By removing the independence condition, an extension of de Bruijn's identity is provided in Section 2. In Section 3, using Gaussian and Farlie–Gumbel–Morgenstern copulas, special versions of the result are given and some illustrative examples are provided. Finally, conclusions are made in Section 4.

2. De Bruijn's identity for dependent RVs

Consider the general additive noise channel described by (2) in which the input signal X and the additive noise Z are dependent RVs with a differentiable joint probability density function (PDF) $f_{X,Z}(x, z)$. Then, the conditional distribution of the output signal Y given the input signal X is given by

$$f_{Y \mid X}(y \mid x; \gamma) = \frac{f_{X,Z}(x, (y - x)/\sqrt{\gamma})}{\sqrt{\gamma} f_X(x)}, \tag{3}$$

and, thus, the marginal unconditional PDF of the output is

$$f_Y(y; \gamma) = \mathbb{E}_X[f_{Y \mid X}(y \mid X; \gamma)]. \tag{4}$$

https://doi.org/10.1017/jpr.2016.5 Published online by Cambridge University Press

Between the two PDFs $f_{Y\mid X}(y \mid X; \gamma)$ and $f_{X,Z}(x, z)$, there exists a general relationship that can be established as follows.

First, by using (3), we obtain

$$\frac{\partial}{\partial y} f_{Y\mid X}(y \mid x; \gamma) = \frac{1}{\sqrt{\gamma} f_X(x)} \left(\frac{\partial}{\partial y} f_{X,Z}\left(x, \frac{y-x}{\sqrt{\gamma}}\right)\right),$$

and

$$\begin{aligned}
\frac{\partial}{\partial \gamma} & f_{Y\mid X}(y \mid x; \gamma) \\
&= \frac{\partial}{\partial \gamma}\left(\frac{f_{X,Z}(x, (y-x)/\sqrt{\gamma})}{\sqrt{\gamma} f_X(x)}\right) \\
&= -\frac{1}{2\gamma}\left\{\frac{f_{X,Z}(x, (y-x)/\sqrt{\gamma})}{\sqrt{\gamma} f_X(x)} + \frac{y-x}{\sqrt{\gamma} f_X(x)}\left(\frac{1}{\sqrt{\gamma}} \frac{\partial}{\partial z} f_{X,Z}(x, z)|_{z=(y-x)/\sqrt{\gamma}}\right)\right\} \\
&= -\frac{1}{2\gamma}\left\{\left(\frac{\partial}{\partial y}(y-x)\right) f_{Y\mid X}(y \mid x; \gamma) + (y-x)\frac{\partial}{\partial y} f_{Y\mid X}(y \mid x; \gamma)\right\} \\
&= -\frac{1}{2\gamma}\frac{\partial}{\partial y}((y-x) f_{Y\mid X}(y \mid x; \gamma)). \qquad (5)
\end{aligned}$$

Now, using (4) and (5), we have

$$\frac{\partial}{\partial \gamma} f_Y(y; \gamma) = \mathbb{E}_X\left(\frac{\partial}{\partial \gamma} f_{Y\mid X}(y \mid X; \gamma)\right) = -\frac{1}{2\gamma}\mathbb{E}_X\left[\frac{\partial}{\partial y}((y-X) f_{Y\mid X}(y \mid X; \gamma))\right]. \qquad (6)$$

Our main result concerns the first derivative of the differential entropy of the channel model (2) which is presented in the following theorem.

Theorem 2. *Let X and Z in (2) be two arbitrary dependent RVs with a joint PDF $f_{X,Z}(x, z)$. Then, under the following conditions:*

$$\lim_{y\to\pm\infty} y\sqrt{f_Y(y; \gamma)} = 0, \qquad \left|\frac{\mathbb{E}_X[X f_{Y\mid X}(y \mid X; \gamma)]}{\sqrt{f_{Y(y;\gamma)}}}\right| < \infty, \qquad (7)$$

the first derivative of the differential entropy can be expressed as

$$\frac{\partial}{\partial \gamma} h(Y) = \frac{1}{2\gamma}\mathbb{E}_Y\{1 + \varrho_Y(Y; \gamma)\mathbb{E}_{X\mid Y}[X \mid Y]\}, \qquad (8)$$

where $\varrho_Y = (\partial/\partial Y)\log f_Y$ is the Fisher score with respect to location parameter.

Proof. From (1), we know that

$$\frac{\partial}{\partial \gamma} h(Y) = -\int_{-\infty}^{\infty} \log f_Y(y; \gamma)\frac{\partial}{\partial \gamma} f_Y(y; \gamma)\, \mathrm{d}y. \qquad (9)$$

Using (6), (9) can be expressed as

$$\frac{\partial}{\partial \gamma} h(Y) = \frac{1}{2\gamma}\int_{-\infty}^{\infty} \log f_Y(y; \gamma)\frac{\partial}{\partial y}\mathbb{E}_X[(y-X) f_{Y\mid X}(y \mid X; \gamma)]\, \mathrm{d}y. \qquad (10)$$

https://doi.org/10.1017/jpr.2016.5 Published online by Cambridge University Press

Integrating by parts with respect to y and using (7), (10) can be written as

$$\begin{aligned}\frac{\partial}{\partial\gamma}h(Y) &= -\frac{1}{2\gamma}\int_{-\infty}^{\infty}\biggl(\frac{\partial}{\partial y}\log f_Y(y;\gamma)\biggr)\mathbb{E}_X[(y-X)f_{Y\mid X}(y\mid X;\gamma)]\,\mathrm{d}y \\ &= -\frac{1}{2\gamma}\int_{-\infty}^{\infty}\frac{\partial}{\partial y}f_Y(y;\gamma)y\,\mathrm{d}y + \frac{1}{2\gamma}\int_{-\infty}^{\infty}\frac{\partial}{\partial y}f_Y(y;\gamma)\mathbb{E}_X\biggl[X\frac{f_{Y\mid X}(y\mid X;\gamma)}{f_Y(y;\gamma)}\biggr]\mathrm{d}y \\ &= \frac{1}{2\gamma} + \frac{1}{2\gamma}\int_{-\infty}^{\infty} f_Y(y)\varrho_Y(y;\gamma)\mathbb{E}_{X\mid Y}[X\mid y]\,\mathrm{d}y. \end{aligned} \tag{11}$$

Thus, the proof is completed. □

Remark 1. This result is an extension of Park *et al.* [11], due to the signal-dependent nature of the noise. It is also interesting to note that (8) can be expressed as a covariance, in a form reminiscent of [6, Equation (18)], as follows:

$$\frac{\partial}{\partial\gamma}h(Y) = \frac{1}{2\gamma}\int_{-\infty}^{\infty} f_Y(y)\biggl(\varrho_Y(y;\gamma) + \frac{y-\mathbb{E}[Y]}{\operatorname{cov}(X,Y)}\biggr)\mathbb{E}_{X\mid Y}[X\mid y]\,\mathrm{d}y.$$

The equivalence of these forms arises from

$$\int_{-\infty}^{\infty} f_Y(y)\biggl(\frac{y-\mathbb{E}[Y]}{\operatorname{cov}(X,Y)}\biggr)\mathbb{E}_{X\mid Y}[X\mid y]\,\mathrm{d}y = 1.$$

Remark 2. Note that using (4), we have

$$\frac{\partial}{\partial y}f_Y(y;\gamma) = \int_{-\infty}^{\infty}\biggl(\frac{\partial}{\partial y}f_{Y\mid X}(y\mid x;\gamma)\biggr)f_X(x)\,\mathrm{d}x \tag{12}$$

$$\begin{aligned} &= \int_{-\infty}^{\infty}\biggl(\frac{\partial}{\partial y}\log f_{Y\mid X}(y\mid x;\gamma)\biggr)f_{Y\mid X}(y\mid x;\gamma)f_X(x)\,\mathrm{d}x \\ &= f_Y(y;\gamma)\mathbb{E}_{X\mid Y}\biggl[\frac{\partial}{\partial Y}\log f_{Y\mid X}(Y\mid X;\gamma)\mid Y\biggr]. \end{aligned} \tag{13}$$

Thus, substituting (13) into (11), the statement of Theorem 2 can also be written as

$$\frac{\partial}{\partial\gamma}h(Y) = \frac{1}{2\gamma}\mathbb{E}_Y\biggl\{1 + \mathbb{E}_{X\mid Y}[X\mid Y]\mathbb{E}_{X\mid Y}\biggl[\frac{\partial}{\partial Y}\log f_{Y\mid X}(Y\mid X;\gamma)\mid Y\biggr]\biggr\}. \tag{14}$$

An alternative perspective on this result, in the independent case, is given by Guo *et al.* [3]. Instead of using the de Bruijn's identity (Theorem 1), the authors show that the derivative of a certain mutual information quantity can be expressed in terms of the minimum mean-squared error of the corresponding noisy channel.

Remark 3. Most signal processing algorithms are designed and based on prior knowledge of signal and noise characteristics. It is therefore natural (and useful) to view them as Bayesian inference strategies. In this case, the result of Theorem 2 may be viewed as a function of posterior mean.

Remark 4. In the above proof, we have exchanged integration and differentiation in (6), (9), (10), and (12). Strict justification of these exchanges requires application of the dominated convergence theorem.

3. Using copula functions

The study of copula functions gives a fully developed mathematical theory for multivariate distribution analysis [10]. A copula is a function that links univariate distribution functions to generate a multivariate distribution function and thus represents the dependency structure of RVs. In other words, copulas enable us to extract the dependence structure from the joint distribution function of a set of RVs and, at the same time, to isolate the dependence structure from the univariate marginal behavior [8]. In recent years, there has been a revival of interest in copula in applications where the matter of dependency between RVs is of importance [4].

The foundation theorem for copula was introduced by Sklar [15] and states that for a given joint multivariate PDF and relevant marginal PDFs, there exists a copula function that relates them. In a bivariate case, Sklar's theorem is stated as follows.

Theorem 3. *Suppose that $F_{X,Y}$ is a joint cumulative distribution function (CDF) with margins F_X and F_Y, then there exists a function $C\colon [0,1]^2 \longrightarrow [0,1]$ such that*

$$F_{X,Y}(x,y) = C(F_X(x), F_Y(y)). \tag{15}$$

If F_X and F_Y are continuous then C is unique; otherwise, C is uniquely determined on the (range of F_X) $\times$ (range of F_Y). Conversely, if C is a copula and F_X and F_Y are CDFs, then the function $F_{X,Y}$ defined by (15) is a joint CDF with margins F_X and F_Y.

For any copula function, there is a corresponding copula density function, which is the mixed partial derivative of function C, given by

$$c(F_X(x), F_Y(y)) = \frac{f_{X,Y}(x,y)}{f_X(x) f_Y(y)}, \tag{16}$$

where $f_{X,Y}$, f_X, and f_Y are the joint and marginal PDFs of X and Y, respectively. Equation (16) can be expressed in an equivalent and more suitable form:

$$f_{X,Y}(x,y) = c(u,v) f_X(x) f_Y(y), \tag{17}$$

where u, v are related to x, y through the marginal CDFs $u = F_X(x)$, $v = F_Y(y)$.

Two of the most popular parametric families of copulas, which are considered in this paper, are the Farlie–Gumbel–Morgenstern (FGM) and the Gaussian families. The key advantage of these copulas is that one can specify different levels of dependency between the margins.

Definition 1. The Gaussian copula is defined as

$$C_\rho(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)), \tag{18}$$

where ρ is the Gaussian copula parameter, Φ_ρ is the bivariate standard Gaussian distribution function, and Φ^{-1} is the inverse of the univariate standard Gaussian distribution function.

The corresponding Gaussian copula density is

$$\begin{aligned} c_\rho(u,v) = {} & \frac{1}{\sqrt{1-\rho^2}} \exp\biggl[-\frac{[\Phi^{-1}(u)]^2 - 2\rho[\Phi^{-1}(u)][\Phi^{-1}(v)] + [\Phi^{-1}(v)]^2}{2(1-\rho^2)}\biggr] \\ & \times \exp\biggl(\frac{[\Phi^{-1}(u)]^2}{2}\biggr) \exp\biggl(\frac{[\Phi^{-1}(v)]^2}{2}\biggr), \end{aligned} \tag{19}$$

where $-1 < \rho < 1$, $0 \le u, v \le 1$.

https://doi.org/10.1017/jpr.2016.5 Published online by Cambridge University Press

Definition 2. The FGM copula, which can be considered when the dependent structure between the variables are not very strong, is defined as $C_\theta(u, v) = uv[1 + \theta(1 - u)(1 - v)]$, where $\theta \in [-1, 1]$ is the FGM copula parameter. The corresponding FGM copula density function is given by

$$c_\theta(u, v) = [1 + \theta(1 - 2u)(1 - 2v)]. \tag{20}$$

As revealed in Theorem 2, for dependent signal and noise, the first derivative of the differential entropy of output signal Y can be expressed by the function of the conditional mean estimate $\mathbb{E}_{X \mid Y}(X \mid Y)$. In Corollaries 1 and 2 below, special versions of this result are given in the case when the dependence between input signal X and additive noise Z are modeled by assuming the Gaussian and FGM copulas, for the case of weaker dependence, for the joint distribution of signal and noise.

Corollary 1. *In (2), let X be an arbitrary RV with a finite second-order moment depending on the RV Z. Assume that Z is normally distributed with zero mean and unit variance and the dependent structure of X and Z can be described by the Gaussian copula (19). If the expectations $\mathbb{E}_X[\Phi^{-1}(F_X(X))]$ and $\mathbb{E}_X[X\Phi^{-1}(F_X(X))]$ exist and the conditions in (7) hold, then*

$$\frac{\partial}{\partial\gamma}h(Y) = \bigg(\frac{1-\rho^2}{2}\bigg)J(Y) - \frac{\rho}{2\sqrt{\gamma}}\mathbb{E}_Y\{\varrho_Y(Y)\mathbb{E}_{X \mid Y}[\Phi^{-1}(F_X(X)) \mid Y]\}. \tag{21}$$

Proof. When Z is normally distributed with zero mean and unit variance, the joint PDF of X and Z can be written as

$$f_{X,Z}(x, z) = \exp\bigg\{-\frac{1}{2(1-\rho^2)}[[\Phi^{-1}(F_X(x))]^2 - 2\rho z\Phi^{-1}(F_X(x)) + z^2]\bigg\} \times \frac{1}{\sqrt{2\pi(1-\rho^2)}}f_X(x)\exp\bigg\{\frac{[\Phi^{-1}(F_X(x))]^2}{2}\bigg\}. \tag{22}$$

Differentiating both sides of (22) with respect to y yields

$$-\bigg(\frac{y-x}{\gamma(1-\rho^2)}\bigg)f_{X,Z}\bigg(x, \frac{y-x}{\sqrt{\gamma}}\bigg) = \frac{\partial}{\partial y}f_{X,Z}\bigg(x, \frac{y-x}{\sqrt{\gamma}}\bigg) - \bigg(\frac{\rho\Phi^{-1}(F_X(x))}{\sqrt{\gamma}(1-\rho^2)}\bigg)f_{X,Z}\bigg(x, \frac{y-x}{\sqrt{\gamma}}\bigg).$$

Therefore, we have

$$\begin{aligned}\mathbb{E}_{X \mid Y}[X \mid y] &= \frac{1}{\sqrt{\gamma}f_Y(y)}\int xf_{X,Z}\bigg(x, \frac{y-x}{\sqrt{\gamma}}\bigg)\,\mathrm{d}x \\ &= \frac{1}{f_Y(y)}\int\bigg\{\frac{y}{\sqrt{\gamma}}f_{X,Z}\bigg(x, \frac{y-x}{\sqrt{\gamma}}\bigg) + \sqrt{\gamma}(1-\rho^2)\frac{\partial}{\partial y}f_{X,Z}\bigg(x, \frac{y-x}{\sqrt{\gamma}}\bigg) \\ &\qquad - \rho\sqrt{\gamma}\Phi^{-1}(F_X(x))\frac{1}{\sqrt{\gamma}}f_{X,Z}\bigg(x, \frac{y-x}{\sqrt{\gamma}}\bigg)\bigg\}\,\mathrm{d}x \\ &= y + \gamma(1-\rho^2)\varrho_Y(y) - \rho\sqrt{\gamma}\mathbb{E}_{X \mid Y}[\Phi^{-1}(F_X(X)) \mid y].\end{aligned}$$

Thus, the right-hand side of (8) becomes

$$\frac{1}{2\gamma} + \frac{1}{2\gamma}\int_{-\infty}^{\infty}f_Y(y)\varrho_Y(y)\{y + \gamma(1-\rho^2)\varrho_Y(y) - \rho\sqrt{\gamma}\mathbb{E}_{X \mid Y}[\Phi^{-1}(F_X(X)) \mid y]\}\,\mathrm{d}y. \tag{23}$$

Using integration by parts, the first two terms in (23) vanish, completing the proof. □

https://doi.org/10.1017/jpr.2016.5 Published online by Cambridge University Press

Corollary 2. *Given (2), let X be an arbitrary RV with a finite second-order moment depending on a normally distributed RV Z with zero mean and unit variance. Assume the dependence between X and Z is modeled by the FGM copula (20) and the conditions in (7) hold. Then, we have*

$$\frac{\partial}{\partial\gamma}h(Y)=\frac{1}{2}J(Y)+\frac{\theta}{2\pi\gamma}\mathbb{E}_Y\bigg\{\frac{\varrho_Y(Y)}{f_Y(Y;\gamma)}\mathbb{E}_X\bigg[(1-2F_x(X))\exp\bigg(-\frac{(Y-X)^2}{\gamma}\bigg)\bigg]\bigg\}. \quad (24)$$

Proof. By using (17), the joint PDF of X and Z can be written as

$$f_{X,Z}(x,z)=\frac{1}{\sqrt{2\pi}}f_X(x)[1+\theta(1-2F_x(X))(1-2\Phi(z))]\exp\bigg\{-\frac{z^2}{2}\bigg\}.$$

Therefore, we have

$$-\bigg(\frac{y-x}{\gamma}\bigg)f_{X,Z}\bigg(x,\frac{y-x}{\sqrt{\gamma}}\bigg)$$
$$=\frac{\partial}{\partial y}f_{X,Z}\bigg(x,\frac{y-x}{\sqrt{\gamma}}\bigg)+\frac{\theta\sqrt{\gamma}}{\pi}f_X(x)(1-2F_X(x))\exp\bigg(-\frac{(y-x)^2}{\gamma}\bigg).$$

The remainder of the proof is similar to the proof of Corollary 1 and is omitted for brevity. □

Remark 5. As a corollary, setting $\rho=0$ in (21) and $\theta=0$ in (24), both cases reduce to the conventional de Bruijn's identity (Theorem 1) for the case where input signal X and additive noise Z are independent RVs.

It should be noted that for additive non-Gaussian noise channels, the differential entropy cannot be expressed in terms of the Fisher information. Instead, in such cases, the differential entropy is expressed by the posterior mean. If two conditional mean estimates $\mathbb{E}_{X\mid Y}[X\mid Y]$ and $\mathbb{E}_{X\mid Y}[(\partial/\partial Y)\log f_{Y\mid X}(Y\mid X;\gamma)\mid Y]$ are expressed by polynomial functions of Y, then (8) and (14) can be expressed in simpler forms.

Example 1. Consider the channel model (2), where X and Z are two dependent RVs distributed according to a bivariate standard normal distribution with PDF

$$f_{X,Z}(x,z)=\frac{1}{2\pi\sqrt{(1-\rho^2)}}\exp\bigg\{-\frac{1}{2(1-\rho^2)}[x^2-2\rho xz+z^2]\bigg\}.$$

Then, we have

$$\mathbb{E}_{X\mid Y}[X\mid Y=y]=\frac{1+\rho\sqrt{\gamma}}{1+\gamma+2\rho\sqrt{\gamma}}y$$

and

$$\mathbb{E}_Y\{\varrho_Y(Y;\gamma)\mathbb{E}_{X\mid Y}[X\mid Y]\}=-\frac{1+\rho\sqrt{\gamma}}{1+\gamma+2\rho\sqrt{\gamma}}.$$

Therefore, (8) can be expressed as

$$\frac{\partial}{\partial\gamma}h(Y)=\frac{1}{2\gamma}\mathbb{E}_Y\{1+\varrho_Y(Y;\gamma)\mathbb{E}_{X\mid Y}[X\mid Y]\}=\frac{1}{2(1+\gamma+2\rho\sqrt{\gamma})}\bigg(1+\frac{\rho}{\sqrt{\gamma}}\bigg).$$

Note that in this case, the output signal Y is also Gaussian with zero mean and variance $1+\gamma+2\rho\sqrt{\gamma}$. Thus, this result can also be alternatively obtained using the derivative of the entropy of output signal Y directly.

Example 2. Let X and Z in (2) be two dependent RVs which are distributed uniformly on (0, 1) and the dependence between them is modeled by the FGM copula (20). Then, the joint PDF of X and Y becomes

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{\gamma}} f_{X,Z}\left(x, \frac{y-x}{\sqrt{\gamma}}\right),$$

where $f_{X,Z}(x,z) = [1+\theta(1-2x)(1-2z)]$, $0 < x < 1$, $0 < z < 1$. Thus, the marginal PDF of Y is

$$f_Y(y;\gamma) = \begin{cases} \dfrac{1}{\sqrt{\gamma}}\left[\dfrac{2\theta}{3\sqrt{\gamma}}y^3 - \theta y^2\left(1+\dfrac{1}{\sqrt{\gamma}}\right) + y(1+\theta)\right], & 0 < y < 1, \\ \dfrac{1}{\sqrt{\gamma}}\left[1 - \dfrac{\theta}{3\sqrt{\gamma}}\right], & 1 \le y < \sqrt{\gamma}, \\ \dfrac{1}{\sqrt{\gamma}}\left[-\dfrac{2\theta}{3\sqrt{\gamma}}y^3 + \theta y^2\left(1+\dfrac{1}{\sqrt{\gamma}}\right)\right. & \\ \qquad \left. - y(\theta+1) - \theta\left(\dfrac{\gamma}{3} + \dfrac{1}{3\sqrt{\gamma}}\right) + \sqrt{\gamma} + 1\right], & \sqrt{\gamma} \le y < 1+\sqrt{\gamma} \end{cases}$$

for $\gamma > 1$, and

$$f_Y(y;\gamma) = \begin{cases} \dfrac{1}{\sqrt{\gamma}}\left[\dfrac{2\theta}{3\sqrt{\gamma}}y^3 - \theta y^2\left(1+\dfrac{1}{\sqrt{\gamma}}\right) + y(1+\theta)\right], & 0 < y < \sqrt{\gamma}, \\ 1 - \dfrac{\theta\sqrt{\gamma}}{3}, & \sqrt{\gamma} \le y < 1, \\ \dfrac{1}{\sqrt{\gamma}}\left[-\dfrac{2\theta}{3\sqrt{\gamma}}y^3 + \theta y^2\left(1+\dfrac{1}{\sqrt{\gamma}}\right)\right. & \\ \qquad \left. - y(\theta+1) - \theta\left(\dfrac{\gamma}{3} + \dfrac{1}{3\sqrt{\gamma}}\right) + \sqrt{\gamma} + 1\right], & 1 \le y < 1+\sqrt{\gamma} \end{cases}$$

for $0 < \gamma \le 1$. But, it is important to note that for any $\gamma > 0$, we obtain

$$\begin{aligned} \mathbb{E}_{X\mid Y}\left[\frac{\partial}{\partial Y}\log f_{Y\mid X}(Y \mid X;\gamma) \mid Y\right] &= \frac{1}{f_Y(y;\gamma)}\int_0^1 \frac{\partial}{\partial y} f_{X,Y}(x,y)\,\mathrm{d}x \\ &= \frac{1}{f_Y(y;\gamma)}\int_0^1 \frac{-2\theta}{\gamma}(1-2x)\,\mathrm{d}x \\ &= 0. \end{aligned}$$

Therefore, by using (14), the derivative of the differential entropy of output signal Y becomes

$$\frac{\partial}{\partial\gamma}h(Y) = \frac{1}{2\gamma}.$$

Remark 6. It is worth mentioning here that from Example 2, it follows that if X and Z are two dependent uniform(0, 1) RVs whose dependence structure can be described by an FGM copula, then the PDF of their sum $Y = X + Z$ is given by

$$f_Y(y) = \begin{cases} \left[\dfrac{2\theta}{3}y^3 - 2\theta y^2 + (1+\theta)y\right], & 0 < y < 1, \\ \left[-\dfrac{2\theta}{3}y^3 + 2\theta y^2 - (1+\theta)y - \dfrac{2}{3}\theta + 2\right], & 1 \le y < 2. \end{cases}$$

https://doi.org/10.1017/jpr.2016.5 Published online by Cambridge University Press

4. Conclusion

In this paper we unveiled an information-estimation relationship which holds in general for dependent noise channels. Considering dependent noise and signal, the first derivative of the differential entropy of output signal was expressed by the conditional mean estimate associated with the input signal. Special versions of the result were given in the cases where input signal and additive noise are jointly distributed according to either Gaussian or FGM copula functions. De Bruijn's identity for independent Gaussian channels followed as a special case.

Acknowledgements

The authors are grateful to an anonymous referee whose valuable comments and suggestions led to an improvement of the manuscript. The first author is also grateful to the office of Graduate Studies of the University of Isfahan for their support.

References

[1] Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*, 2nd edn. John Wiley, Hoboken, NJ.
[2] Gholizadeh, M. H., Amindavar, H. and Ritcey, J. A. (2013). Analytic Nakagami fading parameter estimation in dependent noise channel using copula. *EURASIP J. Adv. Signal Process.* **129,** 10.1186/1687-6180-2013-129.
[3] Guo, D., Shamai, S. and Verdu, S. (2005). Additive non-Gaussian noise channels: mutual information and conditional mean estimation. In *Proc. IEEE Internat. Symp. Inf. Theory*, IEEE, New York, pp. 719–723.
[4] Joe, H. (1997). *Multivariate Models and Dependence Concepts* (Monogr. Statist. Appl. Prob. **73**). Chapman & Hall, London.
[5] Johnson, O. (2004). *Information Theory and the Central Limit Theorem*. Imperial College Press, London.
[6] Johnson, O. (2013). A de Bruijn identity for symmetric stable laws. Preprint. Available at http://arxiv.org/abs/1310.2045v1.
[7] Kay, S. (2009). Waveform design for multistatic radar detection. *IEEE Trans. Aerospace Electron. Systems* **45,** 1153–1166.
[8] Li, H. and Sun, Y. (2009). Tail dependence for heavy-tailed scale mixtures of multivariate distributions. *J. Appl. Prob.* **46,** 925–937.
[9] Moon, J. and Park, J. (2001). Pattern-dependent noise prediction in signal-dependent noise. *IEEE J. Selected Areas Commun.* **19,** 730–743.
[10] Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd edn. Springer, New York.
[11] Park, S., Serpedin, E. and Qaraqe, K. (2012). On the equivalence between Stein and de Bruijn identities. *IEEE Trans. Inf. Theory* **58,** 7045–7067.
[12] Payaro, M. and Palomar, D. P. (2009). Hessian and concavity of mutual information, differential entropy, and entropy power in linear vector Gaussian channels. *IEEE Trans. Inf. Theory* **55,** 3613–3628.
[13] Pham, D.-T. (2005). Entropy of a variable slightly contaminated with another. *IEEE Signal Process. Lett.* **12,** 536–539.
[14] Rioul, O. (2011). Information theoretic proofs of entropy power inequalities. *IEEE Trans. Inf. Theory* **57,** 33–55.
[15] Sklar, M. (1959). Fonctions de répartition à *n* dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8,** 229–231.
[16] Stam, A. J. (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information Control* **2,** 101–112.