J. Appl. Probab. **60,** 1136–1156 (2023)
*doi:*10.1017/jpr.2022.128

COSTA'S CONCAVITY INEQUALITY FOR DEPENDENT VARIABLES BASED ON THE MULTIVARIATE GAUSSIAN COPULA

FATEMEH ASGARI,* ** AND
MOHAMMAD HOSSEIN ALAMATSAZ,* *** *University of Isfahan*

Abstract

An extension of Shannon's entropy power inequality when one of the summands is Gaussian was provided by Costa in 1985, known as Costa's concavity inequality. We consider the additive Gaussian noise channel with a more realistic assumption, i.e. the input and noise components are not independent and their dependence structure follows the well-known multivariate Gaussian copula. Two generalizations for the first- and second-order derivatives of the differential entropy of the output signal for dependent multivariate random variables are derived. It is shown that some previous results in the literature are particular versions of our results. Using these derivatives, concavity of the entropy power, under certain mild conditions, is proved. Finally, special one-dimensional versions of our general results are described which indeed reveal an extension of the one-dimensional case of Costa's concavity inequality to the dependent case. An illustrative example is also presented.

Keywords: Differential entropy; Fisher information; Gaussian copula

2020 Mathematics Subject Classification: Primary 62P30
Secondary 94A17; 94A15

1. Introduction

Let $h(\mathbf{Y}) = -\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y}; t) \log f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y}$ denote the differential entropy of a random vector $\mathbf{Y}$ with probability density function (PDF) $f_{\mathbf{Y}}(\mathbf{y}; t)$ depending on a real parameter t. The entropy power of an m-variate random vector $\mathbf{Y}$ is defined by

$$N(\mathbf{Y}) = \frac{\mathrm{e}^{(2/m)h(\mathbf{Y})}}{2\pi \mathrm{e}},$$

which was first introduced by Shannon [13]. One of the most important inequalities in information theory is the entropy power inequality (EPI), which gives a lower bound for the differential entropy of the sum of the independent random vectors $\mathbf{X}$ and $\mathbf{Y}$ as $N(\mathbf{X} + \mathbf{Y}) \geq N(\mathbf{X}) + N(\mathbf{Y})$. The first complete proof of the EPI was given in [15]; in its development, [15] proved an equality called de Bruijn's identity. This identity links Fisher information with Shannon's differential entropy (see [5]). Consider the additive Gaussian noise channel model

$$\mathbf{Y} = \mathbf{X} + \mathbf{W}_t, \tag{1}$$

Received 31 May 2021; revision received 12 November 2022; accepted 16 November 2022.
* Postal address: Department of Statistics, Faculty of Mathematics and Statistics, University of Isfahan, Isfahan 81746-73441, Iran.
** Email address: ft.asgari@sci.ui.ac.ir
*** Email address: alamatho@sci.ui.ac.ir

in which the input signal $\mathbf{X} = (X_1, \ldots, X_m)^\top$ and the additive noise $\mathbf{W}_t = (W_{t,1}, \ldots, W_{t,m})^\top$ are two m-variate random vectors and $\mathbf{W}_t$ is normally distributed with mean vector $\mathbf{0}$ and covariance matrix

$$\boldsymbol{\Sigma}_{\mathbf{W}_t} = \begin{pmatrix} t & \sigma_{12} & \ldots & \sigma_{1m} \\ \sigma_{21} & t & \ldots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \ldots & t \end{pmatrix}, \tag{2}$$

where the σ_{ij}, $i, j = 1, 2, \ldots, m$, are real numbers. De Bruijn's identity, generalized by Costa [7] to multivariate random variables, is given by

$$\frac{\partial}{\partial t} h(\mathbf{Y}) = \frac{1}{2} J(\mathbf{Y}), \tag{3}$$

in which $\mathbf{X}$ and $\mathbf{W}_t$ are independent random vectors and $J(\mathbf{Y})$ stands for the Fisher information of $f_{\mathbf{Y}}(\mathbf{y}; t)$, defined by

$$\begin{aligned} J(\mathbf{Y}) &= \int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y}; t) \|\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\|^2 \, \mathrm{d}\mathbf{y}, \\ &= \int_{\mathbb{R}^m} \frac{\|\nabla f_{\mathbf{Y}}(\mathbf{y}; t)\|^2}{f_{\mathbf{Y}}(\mathbf{y}; t)} \, \mathrm{d}\mathbf{y}. \end{aligned} \tag{4}$$

There are several applications of the EPI, such as in bounding the capacity of certain kinds of channels and proving converses of channel or source coding theorems; see, e.g., [6, 18]. Considering the channel model (1), [7] presented an extension of the EPI for the case in which $\mathbf{W}_t$ is independent of $\mathbf{X}$ with $\Sigma_{\mathbf{W}_t} = t\mathbf{I}_m$, where $\mathbf{I}_m$ is the $m \times m$ identity matrix. That is,

$$N(\mathbf{X} + \mathbf{W}_t) \geq (1 - t) N(\mathbf{X}) + t N(\mathbf{X} + \mathbf{W}_1),$$

or, equivalently, $N(\mathbf{X} + \mathbf{W}_t)$ is concave in t, i.e.

$$\frac{\partial^2}{\partial t^2} N(\mathbf{X} + \mathbf{W}_t) \leq 0. \tag{5}$$

Later, [8] provided another simple proof for the Costa's concavity inequality (5) via the Stam Fisher information inequality [15] defined by

$$\frac{1}{J(X + W)} \geq \frac{1}{J(X)} + \frac{1}{J(W)},$$

where X and W are independent random variables. Also, [17] used some advanced methods to simplify Costa's proof of the inequality (5).

As mentioned before, in all of the above results the assumption of independence between the input signal $\mathbf{X}$ and the additive noise $\mathbf{W}_t$ has been required. However, there are several real situations, such as in radar and sonar systems, in which the noise is highly dependent on the transmitted signal [11]. It was illustrated in [16] that, under some assumptions, Shannon's EPI can hold for weakly dependent random variables; [3] extended the EPI to dependent random variables with arbitrary distibutions; and [10] provided certain conditions under which the conditional EPI can hold for dependent summands as well.

One of the best methods for describing the dependency structure among random variables is by copula functions. Copula theory was first introduced in [14] in order to achieve the connection between a joint PDF and its marginals. In [4], the authors extended two inequalities based on the Fisher information when the input signal and noise components are dependent and their dependence structure is modeled by several well-known copulas. There are several families of copulas with different dependence structures. The Gaussian copula is one of the most usable, and describes different levels of dependence between marginal components. In the present paper, by considering the additive Gaussian noise channel model (1) where the input signal $\mathbf{X}$ and noise $\mathbf{W}_t$ are dependent random vectors obeying the multivariate Gaussian copula, first, an extension of de Bruijn's identity (3) is derived, and then Costa's concavity inequality (5) is proved, under some mild conditions.

The rest of the paper is organized as follows. In Section 2 we recall the copula theory concept and the basic definition of the multivariate Gaussian copula function, along with one of its particular cases. In Section 3 we provide a generalization of the first-order derivatives of the differential entropy and Fisher information, provided that the input signal and noise components are dependent variables. Thus, based on these derivatives, Costa's concavity inequality for the case that the random vector $\mathbf{X}$ is composed of independent coordinates is extended. Finally, we illustrate the one-dimensional versions of our results in Section 4.

Let us first establish the fundamental definitions and notation used in this paper. Let $\phi(\mathbf{y})$ and $\psi(\mathbf{y})$ be twice continuously differentiable functions on $\mathbb{R}^m$, and V be any closed and simply connected m-dimensional region in $\mathbb{R}^m$ bounded by a piecewise smooth, closed, and oriented surface S. We recall Green's identity [1], which is stated as

$$\int_V \phi \Delta \psi \, \mathrm{d}V = \int_S \phi \nabla \psi . \mathbf{n}_S \, \mathrm{d}S - \int_V \nabla \phi . \nabla \psi \, \mathrm{d}V, \tag{6}$$

in which $\nabla\phi$ and $\nabla\psi$ are the gradients of ϕ and ψ, respectively, $\mathbf{n}_S$ denotes the unit vector normal to the surface S, and $\nabla\psi.\mathbf{n}_S$ is the inner product of the two vectors. Now, the m-dimensional Stokes' theorem is recalled: it states that if $\mathbf{F}\colon \mathbb{R}^m \to \mathbb{R}^m$ is a vector field over $\mathbb{R}^m$, then

$$\int_V \nabla . \mathbf{F} \, \mathrm{d}V = \int_{\partial V} \mathbf{F} . \mathbf{n}_S \, \mathrm{d}S, \tag{7}$$

where $\partial V = S$ is the boundary of V.

We denote the PDF and cumulative distribution function (CDF) of a random variable X by $f_X(x)$ and $F_X(x)$, respectively.

2. Copula background

Copula theory is popular in multivariate distribution analysis as copulas allow easy modeling of the distribution of a random vector by its marginals. A copula is a multivariate CDF with standard uniform marginal distributions which couples univariate distribution functions to generate a multivariate CDF and indicates the dependency structure of the random variables. Copulas are important parts of the study of dependency between variables since they allow us to separate the effect of dependency from the effects of the marginal distributions [9]. In recent years, there has been a revival of copulas in applications where the matter of dependency between random variables is of great importance [2].

The fundamental theorem for copulas was introduced by Sklar [14] and illustrates the role that copulas play in the relationship between multivariate CDFs and their univariate marginals.

In an n-dimensional multivariate case, Sklar's theorem states that if $F_{T_1,T_2,\ldots,T_n}$ is an n-dimensional CDF with marginals $F_{T_1}, F_{T_2}, \ldots, F_{T_n}$, then there exists an n-copula $C\colon I^n \longrightarrow I$ such that

$$F_{T_1,T_2,\ldots,T_n}(t_1, t_2, \ldots, t_n) = C(F_{T_1}(t_1), F_{T_2}(t_2), \ldots, F_{T_n}(t_n)), \tag{8}$$

where $I = [0, 1]$. If $F_{T_1}, F_{T_2}, \ldots, F_{T_n}$ are continuous, the n-copula C is unique; otherwise, C is uniquely determined on the range of F_{T_1} × the range of F_{T_2} × ··· × the range of F_{T_n}. Conversely, if C is an n-copula and $F_{T_1}, F_{T_2}, \ldots, F_{T_n}$ are univariate distribution functions, then $F_{T_1,T_2,\ldots,T_n}$ is a joint CDF with marginals $F_{T_1}, F_{T_2}, \ldots, F_{T_n}$.

For any n-copula function C, there exists a corresponding copula density function c:

$$c(u_1, u_2, \ldots, u_n) = \frac{\partial^n}{\partial u_1 \partial u_2 \cdots \partial u_n} C(u_1, u_2, \ldots, u_n). \tag{9}$$

Therefore, if $f_{T_1,T_2,\ldots,T_n}$, $f_{T_1}, f_{T_2}, \ldots f_{T_n}$, and c are the density functions of $F_{T_1,T_2,\ldots,T_n}$, $F_{T_1}, F_{T_2}, \ldots F_{T_n}$, and C, respectively, the relation in (8) yields

$$f_{T_1,T_2,\ldots,T_n}(t_1, t_2, \ldots, t_n) = c(u_1, u_2, \ldots, u_n) f_{T_1}(t_1) f_{T_2}(t_2) \cdots f_{T_n}(t_n), \tag{10}$$

where $u_1, u_2, \ldots, u_n$ are related to $t_1, t_2, \ldots, t_n$ through the marginal distribution functions $u_1 = F_{T_1}(t_1)$, $u_2 = F_{T_2}(t_2)$, ..., $u_n = F_{T_n}(t_n)$.

Let us recall the definition of one of the most popular copulas, the multivariate Gaussian copula, which we consider here.

Definition 1. The n-dimensional Gaussian copula with covariance matrix $\boldsymbol{\Sigma}$ is defined by

$$C_{\boldsymbol{\Sigma}}(u_1, u_2, \ldots, u_n) = \Phi_{\boldsymbol{\Sigma}}(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)), \tag{11}$$

where $\Phi_{\boldsymbol{\Sigma}}$ denotes the CDF of the n-variate normal random vector with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$, Φ^{-1} is the inverse of the univariate standard Gaussian CDF, and $0 \le u_1, u_2, \ldots, u_n \le 1$.

In this paper we consider the special version of the n-dimensional Gaussian copula with

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho & \ldots & \rho \\ \rho & 1 & \ldots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \ldots & 1 \end{pmatrix} = (1-\rho)\mathbf{I}_n + \rho \mathbf{1}_n \mathbf{1}_n^\top,$$

and $-1/(n-1) < \rho < 1$ in which $\mathbf{1}_n = (1, 1, \ldots, 1)^\top_{1\times n}$. Thus, from (9), the n-dimensional Gaussian copula density is given by

$$\begin{aligned} c_{\boldsymbol{\Sigma}}(u_1, u_2, \ldots, u_n) &= \prod_{i=1}^{n} \frac{\partial}{\partial u_i} \Phi^{-1}(u_i) \phi_{\boldsymbol{\Sigma}}(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_n)) \\ &= (2\pi)^{n/2} \exp\left[\frac{1}{2}\sum_{i=1}^{n} z_i^2\right] \phi_{\boldsymbol{\Sigma}}(z_1, z_2, \ldots, z_n), \end{aligned} \tag{12}$$

where $\phi_{\boldsymbol{\Sigma}}$ is the PDF of the n-variate Gaussian distribution, and $z_i = \Phi^{-1}(u_i)$, $i = 1, 2, \ldots, n$. Since

$$|\boldsymbol{\Sigma}| = (1 + (n-1)\rho)(1-\rho)^{n-1}, \qquad \boldsymbol{\Sigma}^{-1} = \frac{1}{1-\rho}\left(\mathbf{I}_n - \frac{\rho}{1+(n-1)\rho}\mathbf{1}_n\mathbf{1}_n^\top\right),$$

we have

$$\phi_{\mathbf{\Sigma}}(z_1, z_2, \ldots, z_n) = \frac{(2\pi)^{-n/2}}{\sqrt{(1+(n-1)\rho)(1-\rho)^{n-1}}} \times \exp\left\{\frac{-1}{2(1-\rho)}\sum_{i=1}^{n} z_i^2 + \frac{\rho}{2(1+(n-1)\rho)(1-\rho)}\left(\sum_{i=1}^{n} z_i\right)^2\right\}. \quad (13)$$

Now, due to the fact that $\left(\sum_{i=1}^{n} z_i\right)^2 = \sum_{i=1}^{n} z_i^2 + \sum_{i\neq j} z_i z_j$, substituting (13) into (12) yields

$$c_{\mathbf{\Sigma}}(u_1, u_2, \ldots, u_n) = \alpha(\rho, n) \times \exp\left\{\beta(\rho, n)\left(\sum_{i=1}^{n}[\Phi^{-1}(u_i)]^2 - \frac{1}{(n-1)\rho}\sum_{i\neq j}\Phi^{-1}(u_i)\Phi^{-1}(u_j)\right)\right\}, \quad (14)$$

where

$$\alpha(\rho, n) = \frac{1}{\sqrt{(1+(n-1)\rho)(1-\rho)^{n-1}}}, \qquad \beta(\rho, n) = \frac{-(n-1)\rho^2}{2(1-\rho)(1+(n-1)\rho)}.$$

Remark 1. Note that setting $\mathbf{\Sigma} = \mathbf{I}_n$, i.e. $\rho = 0$, in (11) leads to the independent copula $C_{\mathbf{I}_n}(u_1, u_2, \ldots, u_n) = u_1 u_2 \cdots u_n$, which is equivalent to the random variables $T_1, T_2, \ldots, T_n$ being independent.

A particular case of the n-dimensional Gaussian copula is the bivariate Gaussian copula. If we put $n = 2$ and

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

with $-1 < \rho < 1$, then the bivariate Gaussian copula is defined by

$$C_\rho(u_1, u_2) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho),$$

where $\rho \in (-1, 1)$ is the Gaussian copula parameter and Φ_2 is the bivariate standard Gaussian CDF. The Gaussian copula density for $-1 < \rho < 1$ is obtained as

$$c_\rho(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{\rho^2}{2(1-\rho^2)}\left([\Phi^{-1}(u_1)]^2 - \frac{2}{\rho}\Phi^{-1}(u_1)\Phi^{-1}(u_2) + [\Phi^{-1}(u_2)]^2\right)\right\}. \quad (15)$$

3. The general case

Consider the additive Gaussian noise channel model (1). Let $\mathbf{X}$ and $\mathbf{W}_t$ be two dependent random vectors with a differentiable joint PDF $f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{w}_t)$. Then, for the PDF of $\mathbf{Y}$, we obtain

$$f_{\mathbf{Y}}(\mathbf{y}; t) = \int_{\mathbb{R}^m} f_{\mathbf{X}}(\mathbf{x}) f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}; t)\, \mathrm{d}\mathbf{x} = \int_{\mathbb{R}^m} f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x}, \quad (16)$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

where

$$f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y} \mid \mathbf{x}; t) = \frac{f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})}.$$

First, recall that assuming $\mathbf{X}$ and $\mathbf{W}_t$ are independent random vectors and $\Sigma_{\mathbf{W}_t} = t\mathbf{I}_m$, [7, 17] used the heat equation given by

$$\frac{\partial}{\partial t} f_{\mathbf{Y}}(\mathbf{y}; t) = \frac{1}{2} \sum_{j=1}^{m} \frac{\partial^2}{\partial y_j^2} f_{\mathbf{Y}}(\mathbf{y}; t)$$

in their proofs. We now need to generalize this heat equation to the case of multivariate random vectors, as below.

Lemma 1. *Suppose that $\mathbf{W}_t$ in channel model (1) has the covariance matrix (2), and let $\mathbf{X}$ and $\mathbf{W}_t$ be two dependent random vectors whose dependence structure is modeled by the multivariate Gaussian copula (14). Then, we have*

$$\begin{aligned} &f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) \\ &= \frac{\alpha(\rho, 2m)}{(2\pi t)^{m/2}} \exp\Bigg\{\gamma(\rho, m)\frac{\|\mathbf{y}-\mathbf{x}\|^2}{t} + \beta(\rho, 2m)\Bigg[\sum_{i=1}^{m} [\Phi^{-1}(F_{X_i}(x_i))]^2 \\ &\qquad - \frac{2}{(2m-1)\rho}\Bigg(\sum_{i<j} \Phi^{-1}(F_{X_i}(x_i))\Phi^{-1}(F_{X_j}(x_j)) + \sum_{k<l} \frac{(y_k - x_k)(y_l - x_l)}{t} \\ &\qquad + \sum_{i,k} \Phi^{-1}(F_{X_i}(x_i))\frac{(y_k - x_k)}{\sqrt{t}}\Bigg)\Bigg]\Bigg\} \prod_{i=1}^{m} f_{X_i}(x_i), \end{aligned} \tag{17}$$

where

$$\gamma(\rho, m) = \frac{2(1-m)\rho - 1}{2(1-\rho)(1+(2m-1)\rho)}.$$

Proof. Using (10) and (14), by setting $\mathbf{T} = (\mathbf{X}, \mathbf{W}_t)$ and $n = 2m$, we have

$$f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{w}) = c_{\boldsymbol{\Sigma}}(F_{X_1}(x_1), \ldots, F_{X_m}(x_m), F_{W_{t,1}}(w_1), \ldots, F_{W_{t,m}}(w_m)) \prod_{i=1}^{m} f_{X_i}(x_i) \prod_{k=1}^{m} f_{W_{t,k}}(w_k),$$

where

$$\begin{aligned} &c_{\boldsymbol{\Sigma}}(F_{X_1}(x_1), \ldots, F_{X_m}(x_m), F_{W_{t,1}}(w_1), \ldots, F_{W_{t,m}}(w_m)) \\ &= \alpha(\rho, 2m) \exp\Bigg\{\beta(\rho, 2m)\Bigg[\sum_{i=1}^{m} z_{x_i}^2 + \sum_{k=1}^{m} z_{w_k}^2 - \frac{2}{(2m-1)\rho} \\ &\qquad \Bigg(\sum_{i<j} z_{x_i} z_{x_j} + \sum_{k<l} z_{w_k} z_{w_l} + \sum_{i,k} z_{x_i} z_{w_k}\Bigg)\Bigg]\Bigg\}, \end{aligned}$$

in which

$$z_{x_i} = \Phi^{-1}(F_{X_i}(x_i)), \qquad z_{w_k} = \Phi^{-1}(F_{W_{t,k}}(w_k)) = \Phi^{-1}\left(\Phi\left(\frac{w_k}{\sqrt{t}}\right)\right) = \frac{w_k}{\sqrt{t}},$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

because $W_{t,k}$, $k = 1, 2, \ldots, m$, are normally distributed with zero mean and variance t. Thus,

$$\begin{aligned} f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{w}) &= \frac{\alpha(\rho, 2m)}{(2\pi t)^{\frac{m}{2}}} \exp\left\{\beta(\rho, 2m)\left[\sum_{i=1}^{m}[\Phi^{-1}(F_{X_i}(x_i))]^2 + \frac{\|\mathbf{w}\|^2}{t}\right.\right. \\ &\quad - \frac{2}{(2m-1)\rho}\left(\sum_{i<j}\Phi^{-1}(F_{X_i}(x_i))\Phi^{-1}(F_{X_j}(x_j)) + \sum_{k<l}\frac{w_k w_l}{t}\right. \\ &\quad \left.\left.\left. + \sum_{i,k}\Phi^{-1}(F_{X_i}(x_i))\frac{w_k}{\sqrt{t}}\right)\right] - \frac{\|\mathbf{w}\|^2}{2t}\right\} \prod_{i=1}^{m} f_{X_i}(x_i). \end{aligned}$$

By some easy calculations, this expression can be rewritten as (17). □

Lemma 2. *Based on the same assumptions as in Lemma 1, we have*

$$\frac{\partial}{\partial t} f_{\mathbf{Y}}(\mathbf{y}; t) = \delta(\rho, m)\Delta f_{\mathbf{Y}}(\mathbf{y}; t) - \lambda(\rho, m)\nabla . q(\mathbf{y}; t), \tag{18}$$

in which $q(\mathbf{y}; t) = (q_1(\mathbf{y}; t), q_2(\mathbf{y}; t), \ldots, q_m(\mathbf{y}; t))$ *and*

$$\begin{aligned} \delta(\rho, m) &= \frac{[(1-\rho)(1+(2m-1)\rho)]}{2(1-2(1-m)\rho)}, \qquad \lambda(\rho, m) = \frac{\rho}{2\sqrt{t}(1-2(1-m)\rho)}, \\ \Delta f_{\mathbf{Y}}(\mathbf{y}; t) &= \sum_{j=1}^{m} \frac{\partial^2}{\partial y_j^2} f_{\mathbf{Y}}(\mathbf{y}; t), \\ q_j(\mathbf{y}; t) &= \int_{\mathbb{R}^m}\left[\sum_{i=1}^{m}\Phi^{-1}(F_{X_i}(x_i)) + \frac{1}{\sqrt{t}}\sum_{k\neq j}(y_k - x_k)\right] f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{w}_t)\,\mathrm{d}\mathbf{x} \\ &= p_j(\mathbf{y}; t) f_{\mathbf{Y}}(\mathbf{y}; t), \qquad j = 1, 2, \ldots, m, \end{aligned}$$

where $p_j(\mathbf{y}; t) = \mathbf{E}_{\mathbf{X}|\mathbf{Y}}\big[\sum_{i=1}^{m}\Phi^{-1}(F_{X_i}(X_i)) + (1/\sqrt{t})\sum_{k\neq j}(Y_k - X_k) \mid \mathbf{Y} = \mathbf{y}\big]$.

Proof. According to Lemma 1, differentiating (17) with respect to t and y_j yields

$$\begin{aligned} \frac{\partial}{\partial t} f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) &= \left[\frac{-\gamma(\rho, m)}{t^2}\|\mathbf{y}-\mathbf{x}\|^2 + \frac{\beta(\rho, 2m)}{(2m-1)\rho t^2}\left(\sum_{k\neq l}(y_k - x_k)(y_l - x_l)\right.\right. \\ &\quad \left.\left. + \sqrt{t}\sum_{i,k}\Phi^{-1}(F_{X_i}(x_i))(y_k - x_k)\right) - \frac{m}{2t}\right] f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) \end{aligned} \tag{19}$$

$$\begin{aligned} \frac{\partial}{\partial y_j} f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) &= \left[\frac{2\gamma(\rho, m)}{t}(y_j - x_j) - \frac{2\beta(\rho, 2m)}{(2m-1)\rho t}\left(\sum_{k\neq j}(y_k - x_k)\right.\right. \\ &\quad \left.\left. + \sqrt{t}\sum_{i=1}^{m}\Phi^{-1}(F_{X_i}(x_i))\right)\right] f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}), \quad j = 1, 2, \ldots, m, \end{aligned} \tag{20}$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

respectively. Thus, for the second-order derivative of (17) with respect to y_j, we obtain

$$\frac{\partial^2}{\partial y_j^2} f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) = \left[\frac{2\gamma(\rho, m)}{t} + \left[\frac{2\gamma(\rho, m)}{t}(y_j - x_j) - \frac{2\beta(\rho, 2m)}{(2m-1)\rho t} \left(\sum_{k\neq j} (y_k - x_k) + \sqrt{t} \sum_{i=1}^{m} \Phi^{-1}(F_{X_i}(x_i)) \right) \right]^2 \right] f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}). \tag{21}$$

Now, according to (16) and (19), we have

$$\begin{aligned} \frac{\partial}{\partial t} f_{\mathbf{Y}}(\mathbf{y}; t) &= \int_{\mathbb{R}^m} \frac{\partial}{\partial t} f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x} \\ &= \frac{-m}{2t} f_{\mathbf{Y}}(\mathbf{y}; t) - \frac{\gamma(\rho, m)}{t^2} \int_{\mathbb{R}^m} \|\mathbf{y}-\mathbf{x}\|^2 f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x} \\ &\quad + \frac{\beta(\rho, 2m)}{(2m-1)\rho t^2} \int_{\mathbb{R}^m} \left(\sum_{k\neq l} (y_k - x_k)(y_l - x_l) \right. \\ &\qquad \left. + \sqrt{t} \sum_{i,k} \Phi^{-1}(F_{X_i}(x_i))(y_k - x_k) \right) f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x}, \end{aligned} \tag{22}$$

$$\begin{aligned} \frac{\partial^2}{\partial y_j^2} f_{\mathbf{Y}}(\mathbf{y}; t) &= \int_{\mathbb{R}^m} \frac{\partial^2}{\partial y_j^2} f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x} \\ &= \frac{2\gamma(\rho, m)}{t} f_{\mathbf{Y}}(\mathbf{y}; t) + \frac{4\gamma^2(\rho, m)}{t^2} \int_{\mathbb{R}^m} (y_j - x_j)^2 f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x} \\ &\quad + \frac{4\beta^2(\rho, 2m)}{(2m-1)^2\rho^2 t^2} \int_{\mathbb{R}^m} \left(\sum_{k\neq j} (y_k - x_k) + \sqrt{t} \sum_{i=1}^{m} \Phi^{-1}(F_{X_i}(x_i)) \right)^2 f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x} \\ &\quad - \frac{8\gamma(\rho, m)\beta(\rho, 2m)}{(2m-1)\rho t^2} \int_{\mathbb{R}^m} (y_j - x_j) \left(\sum_{k\neq j} (y_k - x_k) \right. \\ &\qquad \left. + \sqrt{t} \sum_{i=1}^{m} \Phi^{-1}(F_{X_i}(x_i)) \right) f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x}. \end{aligned} \tag{23}$$

Thus, due to (19), by combining (22) with (23), we obtain

$$\begin{aligned} \frac{\partial}{\partial t} f_{\mathbf{Y}}(\mathbf{y}; t) &= \frac{-1}{4\gamma(\rho, m)} \sum_{j=1}^{m} \frac{\partial^2}{\partial y_j^2} f_{\mathbf{Y}}(\mathbf{y}; t) \\ &\quad - \frac{\beta(\rho, 2m)}{2\gamma(\rho, m)(2m-1)\rho\sqrt{t}} \sum_{j=1}^{m} \int_{\mathbb{R}^m} \left[\sum_{i=1}^{m} \Phi^{-1}(F_{X_i}(x_i)) \right. \\ &\qquad \left. + \frac{1}{\sqrt{t}} \sum_{k\neq j} (y_k - x_k) \right] \frac{\partial}{\partial y_j} f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})\, \mathrm{d}\mathbf{x}, \end{aligned}$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

where
$$\frac{-1}{4\gamma(\rho, m)} = \delta(\rho, m), \qquad \frac{\beta(\rho, 2m)}{2\gamma(\rho, m)(2m-1)\rho\sqrt{t}} = \lambda(\rho, m).$$
Therefore, using (20), the proof is complete. □

Now, we need to derive the first- and second-order derivatives of the differential entropy $h(\mathbf{Y})$ that are key instruments in establishing our main result.

Theorem 1. *Based on Lemma 2, the first-order derivative of the entropy $h(\mathbf{Y})$ is derived as*
$$\frac{\partial}{\partial t} h(\mathbf{Y}) = \delta(\rho, m) J(\mathbf{Y}) + A_t, \tag{24}$$
where
$$A_t = -\lambda(\rho, m) \sum_{j=1}^{m} \mathbf{E}_{\mathbf{Y}}\left[p_j(\mathbf{Y}; t) \frac{\partial}{\partial Y_j} \log f_{\mathbf{Y}}(\mathbf{Y}; t)\right].$$

Proof. Using (18), we obtain
$$\begin{aligned}\frac{\partial}{\partial t} h(\mathbf{Y}) &= -\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y}; t) \frac{\partial}{\partial t} \log f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y} - \int_{\mathbb{R}^m} \log f_{\mathbf{Y}}(\mathbf{y}; t) \frac{\partial}{\partial t} f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y}, \\ &= 0 - \delta(\rho, m) \int_{\mathbb{R}^m} \Delta f_{\mathbf{Y}}(\mathbf{y}; t) \log f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y} + \lambda(\rho, m) \int_{\mathbb{R}^m} \nabla . q(\mathbf{y}; t) \log f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y}.\end{aligned} \tag{25}$$

To apply Green's identity (6) to the second term in (25), we assume that V_r is the $m-$sphere of radius r centered at the origin with boundary $S_r = \partial V_r$. Now, we apply Green's identity to the second term in (25) with $\phi(\mathbf{y}) = \log f_{\mathbf{Y}}(\mathbf{y}; t)$ and $\psi(\mathbf{y}) = f_{\mathbf{Y}}(\mathbf{y}; t)$, and then take the limit on both sides as $r \to +\infty$. Thus,
$$\begin{aligned}\int_{\mathbb{R}^m} \nabla .(\nabla f_{\mathbf{Y}}(\mathbf{y}; t)) \log f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y} &= \lim_{r \to +\infty} \int_{S_r} \log f_{\mathbf{Y}}(\mathbf{y}; t) \nabla f_{\mathbf{Y}}(\mathbf{y}; t) . \mathbf{n}_{S_r}(\mathbf{y})\, \mathrm{d}S_r \\ &\quad - \int_{\mathbb{R}^m} \nabla f_{\mathbf{Y}}(\mathbf{y}; t) . \nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y}, \\ &= 0 - \int_{\mathbb{R}^m} \frac{\|\nabla f_{\mathbf{Y}}(\mathbf{y}; t)\|^2}{f_{\mathbf{Y}}(\mathbf{y}; t)}\, \mathrm{d}\mathbf{y},\end{aligned} \tag{26}$$
where $\mathbf{n}_{S_r}$ is the unit vector normal in the surface S_r. Consider the identity
$$\nabla .(\phi \mathbf{F}) = \nabla \phi . \mathbf{F} + \phi \nabla . \mathbf{F}, \tag{27}$$
where $\mathbf{F} \colon \mathbb{R}^m \to \mathbb{R}^m$. We set $\mathbf{F}(\mathbf{y}) = q(\mathbf{y}; t)$ and $\phi(\mathbf{y}) = \log f_{\mathbf{Y}}(\mathbf{y}; t)$, and then, using Stokes' theorem (8) and taking limits on both sides as $r \to +\infty$, we get
$$\begin{aligned}\int_{\mathbb{R}^m} \nabla . q(\mathbf{y}; t) \log f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y} &= \lim_{r \to +\infty} \int_{S_r} f_{\mathbf{Y}}(\mathbf{y}; t) \log f_{\mathbf{Y}}(\mathbf{y}; t) q(\mathbf{y}; t) . \mathbf{n}_{S_r}(\mathbf{y})\, \mathrm{d}S_r \\ &\quad - \int_{\mathbb{R}^m} q(\mathbf{y}; t) . \nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\, \mathrm{d}\mathbf{y}, \\ &= 0 - \sum_{j=1}^{m} \mathbf{E}_{\mathbf{Y}}\left[p_j(\mathbf{Y}; t) \frac{\partial}{\partial Y_j} \log f_{\mathbf{Y}}(\mathbf{Y}; t)\right].\end{aligned} \tag{28}$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

In Appendix A, the surface integrals in (26) and (28) over the surface S_r are shown to vanish as r approaches $+\infty$. Therefore, by substituting (26) and (28) into (25), the theorem is proved. □

Remark 2. Note that, in Theorem 1, from (24) with $\rho = 0$, we obtain

$$\frac{\partial}{\partial t} h(\mathbf{Y}) = \frac{1}{2} \int_{\mathbb{R}^m} \frac{\|\nabla f_{\mathbf{Y}}(\mathbf{y}; t)\|^2}{f_{\mathbf{Y}}(\mathbf{y}; t)} \, d\mathbf{y}.$$

That is, the first-order derivative of the entropy $h(\mathbf{Y})$ reduces to the case when $\mathbf{X}$ and $\mathbf{W}_t$ are independent random vectors with $\Sigma_{\mathbf{W}_t} = tI_m$ as in [7].

According to Theorem 1, to provide the second-order derivative of $h(\mathbf{Y})$, it is sufficient to derive the first-order derivative of the Fisher information $J(\mathbf{Y})$. First, we need the following lemma.

Lemma 3. *According to Lemma 2, the following two equations hold:*

$$\frac{\partial}{\partial t} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\|^2 = 2\delta(\rho, m) \nabla \log f_{\mathbf{Y}}(\mathbf{y}; t).\nabla \left(\frac{\Delta f_{\mathbf{Y}}(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \right) - 2\lambda(\rho, m) \nabla \log f_{\mathbf{Y}}(\mathbf{y}; t).\nabla \left(\frac{\nabla.q(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \right), \tag{29}$$

where

$$\frac{\Delta f_{\mathbf{Y}}(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} = \Delta \log f_{\mathbf{Y}}(\mathbf{y}; t) + \|\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\|^2, \tag{30}$$

$$\nabla \left(\frac{\nabla.q(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \right) = \frac{\nabla(\nabla.q(\mathbf{y}; t))}{f_{\mathbf{Y}}(\mathbf{y}; t)} - \frac{\nabla.q(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \nabla \log f_{\mathbf{Y}}(\mathbf{y}; t), \tag{31}$$

and

$$2\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t).\nabla(\Delta \log f_{\mathbf{Y}}(\mathbf{y}; t)) - \Delta \|\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\|^2 = -2 \sum_{i,j} \left(\frac{\partial^2}{\partial y_i \partial y_j} \log f_{\mathbf{Y}}(\mathbf{y}; t) \right)^2. \tag{32}$$

Proof. Simply, we know that

$$\frac{\partial}{\partial t} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\|^2 = 2 \sum_{j=1}^{m} \frac{\partial}{\partial y_j} \log f_{\mathbf{Y}}(\mathbf{y}, t) \frac{\partial^2}{\partial y_j \partial t} \log f_{\mathbf{Y}}(\mathbf{y}; t).$$

Also, from (18), we can write

$$\frac{\partial^2}{\partial y_j \partial t} \log f_{\mathbf{Y}}(\mathbf{y}; t) = \frac{\partial}{\partial y_j} \left(\frac{\delta(\rho, m)\Delta f_{\mathbf{Y}}(\mathbf{y}; t) - \lambda(\rho, m) \sum_{j=1}^{m} \frac{\partial}{\partial y_j} q_j(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \right),$$

which implies (29). To prove (30), we have

$$\begin{aligned}
\Delta \log f_{\mathbf{Y}}(\mathbf{y}; t) &= \sum_{j=1}^{m} \frac{\partial^2}{\partial y_j^2} \log f_{\mathbf{Y}}(\mathbf{y}; t) \\
&= \sum_{j=1}^{m} \left[\frac{\frac{\partial^2}{\partial y_j^2} f_{\mathbf{Y}}(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} - \left(\frac{\frac{\partial}{\partial y_j} f_{\mathbf{Y}}(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \right)^2 \right] = \frac{\Delta f_{\mathbf{Y}}(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} - \|\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\|^2.
\end{aligned}$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

Also, since $\nabla.q(\mathbf{y};t)=\sum_{j=1}^{m}(\partial/\partial y_j)q_j(\mathbf{y};t)$, (31) is obtained. Now, to prove (32), we obtain

$$\begin{aligned}\Delta\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2&=\sum_{i=1}^{m}\frac{\partial^2}{\partial y_i^2}\sum_{j=1}^{m}\left(\frac{\partial}{\partial y_i}\log f_{\mathbf{Y}}(\mathbf{y};t)\right)^2\\&=2\sum_{i,j}\left(\frac{\partial^2}{\partial y_i\partial y_j}\log f_{\mathbf{Y}}(\mathbf{y};t)\right)^2+2\sum_{i,j}\frac{\partial}{\partial y_i}\log f_{\mathbf{Y}}(\mathbf{y};t)\frac{\partial^3}{\partial y_i\partial y_j^2}\log f_{\mathbf{Y}}(\mathbf{y};t),\end{aligned}\tag{33}$$

where

$$\begin{aligned}2\sum_{i,j}\frac{\partial}{\partial y_i}\log f_{\mathbf{Y}}(\mathbf{y};t)\frac{\partial^3}{\partial y_i\partial y_j^2}\log f_{\mathbf{Y}}(\mathbf{y};t)&=2\sum_{i=1}^{m}\frac{\partial}{\partial y_i}\log f_{\mathbf{Y}}(\mathbf{y};t)\frac{\partial}{\partial y_i}\sum_{j=1}^{m}\frac{\partial^2}{\partial y_j^2}\log f_{\mathbf{Y}}(\mathbf{y};t)\\&=2\nabla\log f_{\mathbf{Y}}(\mathbf{y};t).\nabla(\Delta\log f_{\mathbf{Y}}(\mathbf{y};t));\end{aligned}$$

together with (33), this completes the proof. □

Theorem 2. *Under the conditions of Lemma 2, the first-order derivative of the Fisher information $J(\mathbf{Y})$ is as follows:*

$$\frac{\partial}{\partial t}J(\mathbf{Y})=-2\delta(\rho,m)\sum_{i,j}\mathbf{E}_{\mathbf{Y}}\left[\frac{\partial^2}{\partial Y_i\partial Y_j}\log f_{\mathbf{Y}}(\mathbf{Y};t)\right]^2+2\lambda(\rho,m)D_t,\tag{34}$$

where

$$\begin{aligned}D_t=&\sum_{j=1}^{m}\mathbf{E}_{\mathbf{Y}}\left[\frac{\partial}{\partial Y_j}p_j(\mathbf{Y};t)\Delta\log f_{\mathbf{Y}}(\mathbf{Y};t)\right]\\&+\sum_{i\neq j}\mathbf{E}_{\mathbf{Y}}\left[p_j(\mathbf{Y};t)\frac{\partial}{\partial Y_j}\log f_{\mathbf{Y}}(\mathbf{Y};t)\left(\frac{\partial^2}{\partial Y_i^2}\log f_{\mathbf{Y}}(\mathbf{Y};t)-\frac{\partial^2}{\partial Y_i\partial Y_j}\log f_{\mathbf{Y}}(\mathbf{Y};t)\right)\right].\end{aligned}$$

Proof. According to the Fisher information (4), we know that

$$\frac{\partial}{\partial t}J(\mathbf{Y})=\int_{\mathbb{R}^m}\frac{\partial}{\partial t}f_{\mathbf{Y}}(\mathbf{y};t)\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}+\int_{\mathbb{R}^m}f_{\mathbf{Y}}(\mathbf{y};t)\frac{\partial}{\partial t}\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}.\tag{35}$$

Based on Lemma 2, the first term in (35) is expressed as

$$\begin{aligned}\int_{\mathbb{R}^m}\frac{\partial}{\partial t}f_{\mathbf{Y}}(\mathbf{y};t)\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}=&\ \delta(\rho,m)\int_{\mathbb{R}^m}\Delta f_{\mathbf{Y}}(\mathbf{y};t)\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}\\&-\lambda(\rho,m)\sum_{j=1}^{m}\int_{\mathbb{R}^m}\frac{\partial}{\partial y_j}q_j(\mathbf{y};t)\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}.\end{aligned}\tag{36}$$

By applying Green's identity (6) to the first term in (36) and taking the limit as r tends to $+\infty$, we obtain

$$\begin{aligned}\int_{\mathbb{R}^m}\Delta f_{\mathbf{Y}}(\mathbf{y};t)\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}=&\lim_{r\to+\infty}\int_{S_r}\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\nabla f_{\mathbf{Y}}(\mathbf{y};t).\mathbf{n}_{S_r}(\mathbf{y})\,\mathrm{d}S_r\\&-\int_{\mathbb{R}^m}\nabla f_{\mathbf{Y}}(\mathbf{y};t).\nabla\|\nabla\log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}.\end{aligned}\tag{37}$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

Similarly, using Green's identity for the second term in (37) and taking the limit, we have

$$-\int_{\mathbb{R}^m} \nabla f_{\mathbf{Y}}(\mathbf{y};t).\nabla\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y} = -\lim_{r\to+\infty}\int_{S_r} f_{\mathbf{Y}}(\mathbf{y};t)\nabla\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2.\mathbf{n}_{S_r}(\mathbf{y})\,\mathrm{d}S_r + \int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\Delta\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}. \tag{38}$$

The first terms in (37) and (38) can be shown to vanish (see Appendix B), and therefore, by comparing (37) with (38), we can write

$$\int_{\mathbb{R}^m} \Delta f_{\mathbf{Y}}(\mathbf{y};t)\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y} = \int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\Delta\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}.$$

Substituting this into (36) yields

$$\int_{\mathbb{R}^m} \frac{\partial}{\partial t} f_{\mathbf{Y}}(\mathbf{y};t)\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y} = \delta(\rho,m)\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\Delta\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y} - \lambda(\rho,m)\int_{\mathbb{R}^m} \nabla.q(\mathbf{y};t)\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}. \tag{39}$$

Also, by using (29) in Lemma 3, the second term in (35) can be rewritten as

$$\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\frac{\partial}{\partial t}\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y} = 2\delta(\rho,m)\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\nabla \log f_{\mathbf{Y}}(\mathbf{y};t).\nabla\left(\frac{\Delta f_{\mathbf{Y}}(\mathbf{y};t)}{f_{\mathbf{Y}}(\mathbf{y};t)}\right)\mathrm{d}\mathbf{y} - 2\lambda(\rho,m)\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\nabla \log f_{\mathbf{Y}}(\mathbf{y};t).\nabla\left(\frac{\nabla.q(\mathbf{y};t)}{f_{\mathbf{Y}}(\mathbf{y};t)}\right)\mathrm{d}\mathbf{y}. \tag{40}$$

Now, according to (30), we obtain

$$\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\nabla \log f_{\mathbf{Y}}(\mathbf{y};t).\nabla\left(\frac{\Delta f_{\mathbf{Y}}(\mathbf{y};t)}{f_{\mathbf{Y}}(\mathbf{y};t)}\right)\mathrm{d}\mathbf{y} = \int_{\mathbb{R}^m} \nabla f_{\mathbf{Y}}(\mathbf{y};t).\nabla(\Delta \log f_{\mathbf{Y}}(\mathbf{y};t))\,\mathrm{d}\mathbf{y} + \int_{\mathbb{R}^m} \nabla f_{\mathbf{Y}}(\mathbf{y};t).\nabla\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}.$$

Therefore, from this and (38), we have

$$\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\nabla \log f_{\mathbf{Y}}(\mathbf{y};t).\nabla\left(\frac{\Delta f_{\mathbf{Y}}(\mathbf{y};t)}{f_{\mathbf{Y}}(\mathbf{y};t)}\right)\mathrm{d}\mathbf{y} = \int_{\mathbb{R}^m} \nabla f_{\mathbf{Y}}(\mathbf{y};t).\nabla(\Delta \log f_{\mathbf{Y}}(\mathbf{y};t))\,\mathrm{d}\mathbf{y} - \int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\Delta\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}. \tag{41}$$

Thanks to the identity (31), for the second term in (40) we obtain

$$\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};t)\nabla \log f_{\mathbf{Y}}(\mathbf{y};t).\nabla\left(\frac{\nabla.q(\mathbf{y};t)}{f_{\mathbf{Y}}(\mathbf{y};t)}\right)\mathrm{d}\mathbf{y} = \int_{\mathbb{R}^m} \nabla \log f_{\mathbf{Y}}(\mathbf{y};t).\nabla(\nabla.q(\mathbf{y};t))\,\mathrm{d}\mathbf{y} - \int_{\mathbb{R}^m} \nabla.q(\mathbf{y};t)\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};t)\|^2\,\mathrm{d}\mathbf{y}. \tag{42}$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

Using Green's identity, we arrive at

$$\int_{\mathbb{R}^m} \nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t).\nabla(\nabla.q(\mathbf{y};\, t))\,\mathrm{d}\mathbf{y} = \lim_{r\to+\infty}\int_{S_r} \nabla.q(\mathbf{y};\, t)\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y})\,\mathrm{d}S_r - \int_{\mathbb{R}^m} \Delta \log f_{\mathbf{Y}}(\mathbf{y};\, t)\nabla.q(\mathbf{y};\, t)\,\mathrm{d}\mathbf{y}, \tag{43}$$

whose first term becomes zero (see Appendix B). Using the identity

$$\nabla.q(\mathbf{y};\, t) = \sum_{j=1}^{m} p_j(\mathbf{y};\, t)\frac{\partial}{\partial y_j} f_{\mathbf{Y}}(\mathbf{y};\, t) + f_{\mathbf{Y}}(\mathbf{y};\, t)\sum_{j=1}^{m}\frac{\partial}{\partial y_j} p_j(\mathbf{y};\, t),$$

the second term in (43) is rewritten as

$$-\int_{\mathbb{R}^m} \Delta \log f_{\mathbf{Y}}(\mathbf{y};\, t)\nabla.q(\mathbf{y};\, t)\,\mathrm{d}\mathbf{y} = -\sum_{j=1}^{m}\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\, t)p_j(\mathbf{y};\, t)\frac{\partial}{\partial y_j}\log f_{\mathbf{Y}}(\mathbf{y};\, t)\Delta \log f_{\mathbf{Y}}(\mathbf{y};\, t)\,\mathrm{d}\mathbf{y} - \sum_{j=1}^{m}\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\, t)\frac{\partial}{\partial y_j}p_j(\mathbf{y};\, t)\Delta \log f_{\mathbf{Y}}(\mathbf{y};\, t)\,\mathrm{d}\mathbf{y}.$$

By combining this with (42) and (43), we get

$$\begin{aligned}
&\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\, t)\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t).\nabla\left(\frac{\nabla.q(\mathbf{y};\, t)}{f_{\mathbf{Y}}(\mathbf{y};\, t)}\right)\mathrm{d}\mathbf{y}\\
&= -\sum_{j=1}^{m}\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\, t)p_j(\mathbf{y};\, t)\frac{\partial}{\partial y_j}\log f_{\mathbf{Y}}(\mathbf{y};\, t)\Delta \log f_{\mathbf{Y}}(\mathbf{y};\, t)\,\mathrm{d}\mathbf{y}\\
&\quad -\sum_{j=1}^{m}\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\, t)\frac{\partial}{\partial y_j}p_j(\mathbf{y};\, t)\Delta \log f_{\mathbf{Y}}(\mathbf{y};\, t)\,\mathrm{d}\mathbf{y} - \int_{\mathbb{R}^m}\nabla.q(\mathbf{y};\, t)\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2\,\mathrm{d}\mathbf{y}.
\end{aligned} \tag{44}$$

Also, we have

$$\int_{\mathbb{R}^m}\nabla.q_j(\mathbf{y};\, t)\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2\,\mathrm{d}\mathbf{y} = \lim_{r\to+\infty}\int_{S_r}\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 q(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y})\,\mathrm{d}S_r, - \int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\, t)p(\mathbf{y};\, t).\nabla\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2\,\mathrm{d}\mathbf{y}, \tag{45}$$

whose first term vanishes (see Appendix B), and $p(\mathbf{y};\, t) = (p_1(\mathbf{y};\, t), p_2(\mathbf{y};\, t)\, \ldots, p_m(\mathbf{y};\, t))$. From (45), combining (35), (39), (40), (41), and (44), we obtain

$$\begin{aligned}
&\frac{\partial}{\partial t}J(\mathbf{Y})\\
&= \delta(\rho, m)\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\, t)[2\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t).\nabla(\Delta \log f_{\mathbf{Y}}(\mathbf{y};\, t)) - \Delta\|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2]\,\mathrm{d}\mathbf{y}\\
&\quad + 2\lambda(\rho, m)\sum_{j=1}^{m}\mathbf{E}_{\mathbf{Y}}\left[\frac{\partial}{\partial Y_j}p_j(\mathbf{Y};\, t)\Delta \log f_{\mathbf{Y}}(\mathbf{Y};\, t)\right]\\
&\quad + 2\lambda(\rho, m)\sum_{i\neq j}\mathbf{E}_{\mathbf{Y}}\left[p_j(\mathbf{Y};\, t)\frac{\partial}{\partial Y_j}\log f_{\mathbf{Y}}(\mathbf{Y};\, t)\left(\frac{\partial^2}{\partial Y_i^2}\log f_{\mathbf{Y}}(\mathbf{Y};\, t) - \frac{\partial^2}{\partial Y_i\partial Y_j}\log f_{\mathbf{Y}}(\mathbf{Y};\, t)\right)\right].
\end{aligned}$$

Hence, based on the relation (32) in Lemma 3, the proof is complete. □

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

Remark 3. It is interesting to see that, if we put $\rho = 0$ in (34), it reduces to

$$\frac{\partial}{\partial t} J(\mathbf{Y}) = -\sum_{i,j} \mathbf{E}_{\mathbf{Y}} \left[\frac{\partial^2}{\partial Y_i \partial Y_j} \log f_{\mathbf{Y}}(\mathbf{Y}; t) \right]^2.$$

That is, Theorem 2 results in the case where $\mathbf{X}$ and $\mathbf{W}_t$ are independent random variables as a special case. Hence, Theorem 2 encompasses the result of [17] as a corollary.

Now, we can establish our main result of this manuscript.

Theorem 3. *Let* $\mathbf{X}$ *and* $\mathbf{W}_t$ *in channel model* (1) *be two dependent random variables whose dependence structure is modeled by the multivariate Gaussian copula. For any* $\rho > -1/(2m - 1)$*, under the conditions*

$$m \frac{\partial A_t}{\partial t} + 2A_t^2 + 4\delta(\rho, m) A_t J(\mathbf{Y}) \leq 0, \tag{46a}$$

$$\rho D_t \leq 0, \tag{46b}$$

the entropy power $N(\mathbf{X} + \mathbf{W}_t)$ *is concave in t. i.e.*

$$\frac{\partial^2}{\partial t^2} N(\mathbf{X} + \mathbf{W}_t) \leq 0.$$

Proof. Simply, we have

$$\frac{\partial^2}{\partial t^2} N(\mathbf{Y}) = \frac{2}{m} N(\mathbf{Y}) \left[\frac{\partial^2}{\partial t^2} h(\mathbf{Y}) + \frac{2}{m} \left(\frac{\partial}{\partial t} h(\mathbf{Y}) \right)^2 \right].$$

Since the entropy power is nonnegative, to show that $(\partial^2/\partial t^2)N(\mathbf{Y}) \leq 0$, it is sufficient to prove that

$$-\frac{\partial^2}{\partial t^2} h(\mathbf{Y}) \geq \frac{2}{m} \left(\frac{\partial}{\partial t} h(\mathbf{Y}) \right)^2.$$

Based on Theorem 1, this is equivalent to

$$-\delta(\rho, m) \frac{\partial}{\partial t} J(\mathbf{Y}) - \frac{\partial}{\partial t} A_t \geq \frac{2\delta^2(\rho, m)}{m} J^2(\mathbf{Y}) + \frac{4\delta(\rho, m)}{m} A_t J(\mathbf{Y}) + \frac{2}{m} A_t^2.$$

Thus, since $\rho > -1/(2m - 1)$ and $\delta(\rho, m) > 0$, due to the condition (46a), we must prove that

$$-\frac{\partial}{\partial t} J(\mathbf{Y}) \geq \frac{2\delta(\rho, m)}{m} J^2(\mathbf{Y}).$$

According to proof of the proposition in [17, p. 3], we have

$$\sum_{i,j} \mathbf{E}_{\mathbf{Y}} \left[\frac{\partial^2}{\partial Y_i \partial Y_j} \log f_{\mathbf{Y}}(\mathbf{Y}; t) \right]^2 \geq \frac{J^2(\mathbf{Y})}{m}. \tag{47}$$

Hence, according to Theorem 3, (47), and assumption (46b), the proof is complete. □

4. The one-dimensional case

In this section, by considering the channel model (1) with $m = 1$, we describe special versions of our main results.

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

Corollary 1. *Let X and W_t in the channel model $Y = X + W_t$ be dependent one-dimensional random variables, and let W_t be normally distributed with mean zero and variance t. If their dependence structure is modeled by the bivariate Gaussian copula* (15)*, then*

$$\frac{\partial}{\partial t}h(Y) = \left(\frac{1-\rho^2}{2}\right)J(Y) + A'_t,$$

where

$$A'_t = -\frac{\rho}{2\sqrt{t}}\mathbf{E}_Y\left[p'(Y;\, t)\frac{\partial}{\partial Y}\log f_Y(Y;\, t)\right], \tag{48}$$

in which $p'(y;\, t) = \mathbf{E}_{X|Y}[\Phi^{-1}(F_X(X)) \mid Y = y]$.

Proof. Since W_t is normally distributed with mean zero and variance t, from (15),

$$f_{X,W_t}(x,\, y-x) = \frac{1}{\sqrt{2\pi(1-\rho^2)}}\exp\left\{-\frac{\rho^2}{2(1-\rho^2)}\left[(\Phi^{-1}(F_X(x)))^2 - \frac{2\rho}{\sqrt{t}}(y-x)\Phi^{-1}(F_X(x)) + \frac{(y-x)^2}{\rho^2 t}\right]\right\}f_X(x).$$

Thus, by some simple calculations, we obtain

$$\begin{aligned}\frac{\partial}{\partial t}f_Y(y;\, t) &= \int_{-\infty}^{+\infty}\frac{\partial}{\partial t}f_{X,W_t}(x,\, y-x)\,\mathrm{d}x\\ &= \int_{-\infty}^{+\infty}\frac{1}{2t}\left(-\frac{\rho\Phi^{-1}(F_X(x))(y-x)}{\sqrt{t}(1-\rho^2)} + \frac{(y-x)^2}{t(1-\rho^2)} - 1\right)f_{X,W_t}(x,\, y-x)\,\mathrm{d}x,\end{aligned} \tag{49}$$

$$\begin{aligned}\frac{\partial^2}{\partial y^2}f_Y(y;\, t) &= \int_{-\infty}^{+\infty}\frac{\partial^2}{\partial y^2}f_{X,W_t}(x,\, y-x)\,\mathrm{d}x,\\ &= \int_{-\infty}^{+\infty}\frac{1}{t(1-\rho^2)}\left[\frac{\rho^2(\Phi^{-1}(F_X(x)))^2}{(1-\rho^2)} + \frac{(y-x)^2}{t(1-\rho^2)}\right.\\ &\qquad\left. - \frac{2\rho\Phi^{-1}(F_X(x))(y-x)}{\sqrt{t}(1-\rho^2)} - 1\right]f_{X,W_t}(x,\, y-x)\,\mathrm{d}x.\end{aligned} \tag{50}$$

Now, by comparing (49) with (50), we obtain

$$\frac{\partial}{\partial t}f_Y(y;\, t) = \left(\frac{1-\rho^2}{2}\right)\frac{\partial^2}{\partial y^2}f_Y(y;\, t) - \frac{\rho}{2\sqrt{t}}\frac{\partial}{\partial y}q'(y;\, t), \tag{51}$$

in which

$$q'(y;\, t) = \int_{-\infty}^{+\infty}\Phi^{-1}(F_X(x))f_{X,W_t}(x,\, y-x)\,\mathrm{d}x = p'(y;\, t)f_Y(y;\, t), \tag{52}$$

where $p'(y;\, t) = \mathbf{E}_{X|Y}[\Phi^{-1}(F_X(X)) \mid Y = y]$. Hence, $q_j(\mathbf{y};\, t)$ and $p_j(\mathbf{y};\, t)$ in Lemma 2 reduce to $q'(y;\, t)$ and $p'(y;\, t)$, respectively. Now, since X and W_t are one-dimensional, it is sufficient to set $m = 1$ and $p_j(\mathbf{y};\, t) = p'(y;\, t)$ in (24). Therefore, the proof is complete. □

Remark 4. Corollary 1 is equivalent to a result in [12].

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

Now, under the same conditions as in Corollary 1, according to the relations (51) and (52), the first-order derivative of the Fisher information,

$$\frac{\partial}{\partial t}J(Y) = -(1-\rho^2)\mathbf{E}_Y\left(\frac{\partial^2}{\partial Y^2}\log f_Y(Y;t)\right)^2 + \frac{\rho}{\sqrt{t}}\mathbf{E}_Y\left[\frac{\partial}{\partial Y}p'(Y;t)\frac{\partial^2}{\partial Y^2}\log f_Y(Y;t)\right], \quad (53)$$

simply follows by setting $m = 1$ and $p_j(\mathbf{y}; t) = p'(y; t)$ in (34). This coincides with the result in [4], where a direct proof of (53) is provided.

Using the first-order derivatives of the entropy and Fisher information of the output signal Y, in what follows the concavity of Shannon's entropy power for the special one-dimensional case is obtained.

Corollary 2. *Given the channel model* (1)*, assume that* X *and* W_t *are dependent random variables modeled by the bivariate Gaussian copula* (14)*. Based on the assumptions*

$$\frac{\partial A'_t}{\partial t} + 2A_t'^2 + 2(1-\rho^2)J(Y)A'_t \le 0, \quad (54a)$$

$$\rho\mathbf{E}_Y\left[\frac{\partial}{\partial Y}p'(Y;t)\frac{\partial^2}{\partial Y^2}\log f_Y(Y;t)\right] \le 0, \quad (54b)$$

the entropy power $N(X + W_t)$ *is concave in* t*.*

Example 1. Consider the channel model $Y = X + W_t$ with $W_t = \sqrt{t}W$. Let X be standard Gaussian and suppose that X and W_t are jointly distributed according to the bivariate Gaussian copula, i.e. X and W are two dependent random variables distributed according to a bivariate standard Gaussian distribution with the PDF

$$f_{X,W}(x, w) = \frac{1}{2\pi\sqrt{(1-\rho^2)}}\exp\left\{-\frac{1}{2(1-\rho^2)}[x^2 - 2\rho xw + w^2]\right\}.$$

We know that Y is normally distributed with mean zero and variance $1 + t + 2\sqrt{t}\rho$. Thus, since $(X, Y) \sim N_2(\mathbf{0}, \Sigma_{X,Y})$ with

$$\Sigma_{X,Y} = \begin{pmatrix} 1 & 1+\sqrt{t}\rho \\ 1+\sqrt{t}\rho & 1+t+2\sqrt{t}\rho \end{pmatrix},$$

we have

$$p'(y;t) = \mathbf{E}_{X|Y}(X \mid Y = y) = \frac{1+\sqrt{t}\rho}{1+t+2\sqrt{t}\rho}y.$$

Further, we observe that

$$\frac{\partial}{\partial y}\log f_Y(y;t) = -\frac{1}{1+t+2\sqrt{t}\rho}y.$$

Thus, by (48), we can write

$$A'(t) = \frac{\rho(1+\sqrt{t}\rho)}{2\sqrt{t}(1+t+2\rho\sqrt{t})}.$$

As we can see, both conditions (54a) and (54b) are satisfied when $\rho > 0$. Thus, based on Corollary 2, $N(X + W_t)$ is concave in t.

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

5. Conclusions

In this paper, based on the multivariate Gaussian copula dependence structure, we have derived the first- and second-order derivatives of differential entropy of the output signal in the m-dimensional additive Gaussian noise channel model. Then, by using these derivatives, we have generalized Costa's concavity inequality for the particular case where the coordinates of the input signal and noise are dependent according to a multivariate Gaussian copula model. In particular, we have studied our results in the one-dimensional case and have provided an illustrative example.

Appendix A. Vanishing surface integrals of Theorem 1

We need to prove that

$$\lim_{r\to+\infty}\int_{S_r}\log f_{\mathbf{Y}}(\mathbf{y};\,t)\nabla f_{\mathbf{Y}}(\mathbf{y};\,t).\mathbf{n}_{S_r}(\mathbf{y})\,\mathrm{d}S_r=0. \tag{55}$$

We first assume that $h(\mathbf{Y})$ is finite. Next, we integrate the surface integral in (55) over $r\geq 0$ and then, by applying the identity (27) and Stokes' theorem, we obtain

$$\begin{aligned}\int_0^{+\infty}\int_{S_r}\log f_{\mathbf{Y}}(\mathbf{y};\,t)\nabla f_{\mathbf{Y}}(\mathbf{y};\,t).\mathbf{n}_{S_r}(\mathbf{y})\,\mathrm{d}S_r &= \int_0^{+\infty}\int_{S_r}\nabla f_{\mathbf{Y}}(\mathbf{y};\,t).(\log f_{\mathbf{Y}}(\mathbf{y};\,t)\mathbf{n}_{S_r}(\mathbf{y}))\,\mathrm{d}S_r\,\mathrm{d}r\\ &= \int_{\mathbb{R}^m}\nabla f_{\mathbf{Y}}(\mathbf{y};\,t).(\log f_{\mathbf{Y}}(\mathbf{y};\,t)\mathbf{n}_{S_r}(\mathbf{y}))\,\mathrm{d}\mathbf{y}\\ &= \lim_{r\to+\infty}\int_{S_r} f_{\mathbf{Y}}(\mathbf{y};\,t)\log f_{\mathbf{Y}}(\mathbf{y};\,t)\,\mathrm{d}S_r\\ &\quad -\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\,t)\nabla.(\log f_{\mathbf{Y}}(\mathbf{y};\,t)\mathbf{n}_{S_r}(\mathbf{y}))\,\mathrm{d}\mathbf{y}.\end{aligned} \tag{56}$$

Since the limit in the first part of (56) exists, due to

$$\left|\int_0^{+\infty}\int_{S_r} f_{\mathbf{Y}}(\mathbf{y};\,t)\log f_{\mathbf{Y}}(\mathbf{y};\,t)\,\mathrm{d}S_r\,\mathrm{d}r\right| = |h(\mathbf{Y})|<+\infty,$$

the first term in (56) vanishes. Now, since

$$|\nabla.(\log f_{\mathbf{Y}}(\mathbf{y};\,t)\mathbf{n}_{S_r}(\mathbf{y}))| = \frac{|\nabla f_{\mathbf{Y}}(\mathbf{y};\,t).\mathbf{n}_{S_r}(\mathbf{y})|}{f_{\mathbf{Y}}(\mathbf{y};\,t)}\leq\frac{\|\nabla f_{\mathbf{Y}}(\mathbf{y};\,t)\|}{f_{\mathbf{Y}}(\mathbf{y};\,t)},$$

for the second term in (56) we can write

$$\begin{aligned}\left|\int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\,t)\nabla.(\log f_{\mathbf{Y}}(\mathbf{y};\,t)\mathbf{n}_{S_r}(\mathbf{y}))\,\mathrm{d}\mathbf{y}\right| &\leq \int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\,t)|\nabla.(\log f_{\mathbf{Y}}(\mathbf{y};\,t)\mathbf{n}_{S_r}(\mathbf{y}))|\,\mathrm{d}\mathbf{y}\\ &\leq \mathbf{E}_{\mathbf{Y}}\left(\frac{\|\nabla f_{\mathbf{Y}}(\mathbf{Y};\,t)\|}{f_{\mathbf{Y}}(\mathbf{Y};\,t)}\right).\end{aligned} \tag{57}$$

Further, we know that

$$\mathbf{E}_{\mathbf{Y}}\left(\frac{\|\nabla f_{\mathbf{Y}}(\mathbf{Y};\,t)\|}{f_{\mathbf{Y}}(\mathbf{Y};\,t)}\right) = \mathbf{E}_{\mathbf{Y}}\left\{\left[\sum_{j=1}^{m}\left(\frac{\frac{\partial}{\partial y_j}f_{\mathbf{Y}}(\mathbf{Y};\,t)}{f_{\mathbf{Y}}(\mathbf{Y};\,t)}\right)^2\right]^{\frac{1}{2}}\right\}\leq\left\{\sum_{j=1}^{m}\mathbf{E}_{\mathbf{Y}}\left(\frac{\frac{\partial}{\partial y_j}f_{\mathbf{Y}}(\mathbf{Y};\,t)}{f_{\mathbf{Y}}(\mathbf{Y};\,t)}\right)^2\right\}^{\frac{1}{2}}. \tag{58}$$

On the other hand, from (20), we have

$$\begin{aligned}\mathbf{E}(W_{t,j} \mid \mathbf{Y}=\mathbf{y}) &= \mathbf{E}_{\mathbf{X}|\mathbf{Y}}[(Y_j - X_j) \mid \mathbf{Y}=\mathbf{y}] \\ &= \int_{\mathbb{R}^m} (y_j - x_j) \frac{f_{\mathbf{X},\mathbf{W}_t}(x, \mathbf{y}-\mathbf{x})}{f_{\mathbf{Y}}(\mathbf{y};\, t)} \,\mathrm{d}\mathbf{x} \\ &= -2\delta(\rho, m) t \left(\frac{\frac{\partial}{\partial Y_j} f_{\mathbf{Y}}(\mathbf{y};\, t)}{f_{\mathbf{Y}}(\mathbf{y};\, t)} \right) + 2t\lambda(\rho, m) p_j(\mathbf{y};\, t). \end{aligned} \tag{59}$$

Now, since for all $j = 1, 2, \ldots, m$, $|E(W_{t,j} \mid \mathbf{Y}=\mathbf{y})| < +\infty$, the first and second terms in (59) must be finite too. Therefore, we have

$$\mathbf{E}_{\mathbf{Y}} \left(\frac{\frac{\partial}{\partial Y_j} f_{\mathbf{Y}}(\mathbf{Y};\, t)}{f_{\mathbf{Y}}(\mathbf{Y};\, t)} \right)^2 < +\infty, \qquad j = 1, 2, \ldots, m, \tag{60}$$

and, due to (58), the right-hand side of inequality (57) is finite. Hence, the integral in (56) is finite and, since the limit in (55) exists, the desired result (55) is proved.

Now, we need to prove that

$$\lim_{r \to +\infty} \int_{S_r} f_{\mathbf{Y}}(\mathbf{y};\, t) \log f_{\mathbf{Y}}(\mathbf{y};\, t) q(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y}) \,\mathrm{d}S_r = 0, \tag{61}$$

in which the integral is taken from $r = 0$ to $r = +\infty$ on the surface integral. Thus, we have

$$\begin{aligned} &\left| \int_0^{+\infty} \int_{S_r} \log f_{\mathbf{Y}}(\mathbf{y};\, t) q(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y}) \,\mathrm{d}S_r \,\mathrm{d}r \right| \\ &\quad \leq \int_0^{+\infty} \int_{S_r} f_{\mathbf{Y}}(\mathbf{y};\, t) |\log f_{\mathbf{Y}}(\mathbf{y};\, t)| \, \|p(\mathbf{y};\, t)\| \, \|\mathbf{n}_{S_r}(\mathbf{y})\| \,\mathrm{d}S_r \,\mathrm{d}r \\ &\quad = \sqrt{m} \int_{\mathbb{R}^m} f_{\mathbf{Y}}(\mathbf{y};\, t) |\log f_{\mathbf{Y}}(\mathbf{y};\, t)| \, \|p(\mathbf{y};\, t)\| \,\mathrm{d}\mathbf{y}, \\ &\quad = \sqrt{m} \mathbf{E}_{\mathbf{Y}} \big| \log f_{\mathbf{Y}}(\mathbf{Y};\, t) \|p(\mathbf{Y};\, t)\| \big|. \end{aligned} \tag{62}$$

Since $f_{\mathbf{Y}}(\mathbf{y};\, t)$ converges to zero as $\mathbf{y}$ approaches $\pm\infty$, we have $f_{\mathbf{Y}}(\mathbf{y};\, t) \log f_{\mathbf{Y}}(\mathbf{y};\, t) \to 0$ as $\mathbf{y} \to \pm\infty$. Therefore, $\log f_{\mathbf{Y}}(\mathbf{y};\, t)$ is finite and, due to (59), the right-hand side of (62) becomes finite. Hence, since the limit in (61) exists, we can conclude the relation in (61).

Appendix B. Vanishing surface integrals of Theorem 2

We intend to prove that

$$\begin{aligned} u_1 &= \lim_{r \to +\infty} \int_{S_r} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 \nabla f_{\mathbf{Y}}(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y}) \,\mathrm{d}S_r = 0, \\ u_2 &= \lim_{r \to +\infty} \int_{S_r} f_{\mathbf{Y}}(\mathbf{y};\, t) \nabla \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2.\mathbf{n}_{S_r}(\mathbf{y}) \,\mathrm{d}S_r = 0, \end{aligned} \tag{63}$$

$$\begin{aligned} u_3 &= \lim_{r \to +\infty} \int_{S_r} \nabla.q(\mathbf{y};\, t) \nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y}) \,\mathrm{d}S_r = 0, \\ u_4 &= \lim_{r \to +\infty} \int_{S_r} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 q(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y}) \,\mathrm{d}S_r = 0. \end{aligned} \tag{64}$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

First, we consider the integral of the surface integral in (63) over $r \geq 0$;

$$\begin{aligned}
&\left| \int_0^{+\infty} \int_{S_r} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 \nabla f_{\mathbf{Y}}(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y})\, \mathrm{d}S_r\, \mathrm{d}r \right| \\
&\quad \leq \int_0^{+\infty} \int_{S_r} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 \nabla \|f_{\mathbf{Y}}(\mathbf{y};\, t)\| \, \|\mathbf{n}_{S_r}(\mathbf{y})\|\, \mathrm{d}S_r\, \mathrm{d}r = \mathbf{E}_{\mathbf{Y}} \left(\frac{\|\nabla f_{\mathbf{Y}}(\mathbf{Y};\, t)\|}{f_{\mathbf{Y}}(\mathbf{Y};\, t)} \right)^3. \qquad (65)
\end{aligned}$$

Simply, based on (58) and (60), the right-hand side of (65) becomes finite and, since the limit u_1 exists, this proves that $u_1 = 0$.

To show that $u_2 = 0$, we write

$$\begin{aligned}
&\int_0^{+\infty} \int_{S_r} f_{\mathbf{Y}}(\mathbf{y};\, t) \nabla \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 .\mathbf{n}_{S_r}(\mathbf{y})\, \mathrm{d}S_r\, \mathrm{d}r \\
&= \int_0^{+\infty} \int_{S_r} \nabla \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 .(f_{\mathbf{Y}}(\mathbf{y};\, t)\mathbf{n}_{S_r}(\mathbf{y}))\, \mathrm{d}S_r\, \mathrm{d}r, \\
&= \int_{\mathbb{R}^m} \nabla \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 .(f_{\mathbf{Y}}(\mathbf{y};\, t)\mathbf{n}_{S_r}(\mathbf{y}))\, \mathrm{d}\mathbf{y}, \\
&= \lim_{r \to +\infty} \int_{S_r} f_{\mathbf{Y}}(\mathbf{y};\, t) \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2\, \mathrm{d}S_r - \int_{\mathbb{R}^m} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 \nabla .(f_{\mathbf{Y}}(\mathbf{y};\, t)\mathbf{n}_{S_r}(\mathbf{y}))\, \mathrm{d}\mathbf{y}. \qquad (66)
\end{aligned}$$

Because $\left| \int_0^{+\infty} \int_{S_r} f_{\mathbf{Y}}(\mathbf{y};\, t) \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2\, \mathrm{d}S_r\, \mathrm{d}r \right| = |J(\mathbf{Y})| < +\infty$ and

$$\begin{aligned}
\left| \int_{\mathbb{R}^m} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 \nabla .(f_{\mathbf{Y}}(\mathbf{y};\, t)\mathbf{n}_{S_r}(\mathbf{y}))\, \mathrm{d}\mathbf{y} \right| &\leq \int_{\mathbb{R}^m} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\|^2 |\nabla .(f_{\mathbf{Y}}(\mathbf{y};\, t)\mathbf{n}_{S_r}(\mathbf{y}))|\, \mathrm{d}\mathbf{y} \\
&\leq \mathbf{E}_{\mathbf{Y}} \left(\frac{\|\nabla f_{\mathbf{Y}}(\mathbf{Y};\, t)\|}{f_{\mathbf{Y}}(\mathbf{Y};\, t)} \right)^3,
\end{aligned}$$

the first term in (66) becomes zero and the absolute value of the second term is finite. Thus, since the limit u_2 exists, we have $u_2 = 0$.

In a similar way, we consider the integral from $r = 0$ to $r = +\infty$ of the surface integral in (64):

$$\begin{aligned}
&\left| \int_0^{+\infty} \int_{S_r} \nabla .q(\mathbf{y};\, t) \nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t).\mathbf{n}_{S_r}(\mathbf{y})\, \mathrm{d}S_r\, \mathrm{d}r \right| \\
&\quad \leq \int_0^{+\infty} \int_{S_r} |\nabla .q(\mathbf{y};\, t)| \, \|\nabla \log f_{\mathbf{Y}}(\mathbf{y};\, t)\| \, \|\mathbf{n}_{S_r}(\mathbf{y})\|\, \mathrm{d}S_r\, \mathrm{d}r \\
&\quad \leq \sum_{j=1}^{m} \mathbf{E}_{\mathbf{Y}} \left[\left| \frac{\partial}{\partial Y_j} q_j(\mathbf{Y};\, t) \right| \|\nabla \log f_{\mathbf{Y}}(\mathbf{Y};\, t)\| \right]. \qquad (67)
\end{aligned}$$

https://doi.org/10.1017/jpr.2022.128 Published online by Cambridge University Press

Using (21), we have

$$
\begin{aligned}
&\mathbf{E}(W_{t,j}^2 \mid \mathbf{Y}=\mathbf{y}) = \mathbf{E}_{\mathbf{X}|\mathbf{Y}}[(Y_j - X_j)^2 \mid \mathbf{Y}=\mathbf{y}] \\
&= \int_{\mathbb{R}^m} (y_j - x_j)^2 \frac{f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x})}{f_{\mathbf{Y}}(\mathbf{y}; t)} \, \mathrm{d}\mathbf{x} \\
&= 4\delta^2(\rho, m) t^2 \left(\frac{\frac{\partial^2}{\partial y_j^2} f_{\mathbf{Y}}(\mathbf{y}; t)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \right) + 2\delta(\rho, m) t \\
&\quad - \frac{4t^2 \lambda^2(\rho, m)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \int_{\mathbb{R}^m} \left(\sum_{i=1}^m \Phi^{-1}(F_{X_i}(x_i)) + \frac{1}{\sqrt{t}} \sum_{k \neq j} (y_k - x_k) \right)^2 f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) \, \mathrm{d}\mathbf{x} \\
&\quad - \frac{2t\lambda(\rho, m)}{f_{\mathbf{Y}}(\mathbf{y}; t)} \int_{\mathbb{R}^m} (y_j - x_j) \left(\sum_{i=1}^m \Phi^{-1}(F_{X_i}(x_i)) + \frac{1}{\sqrt{t}} \sum_{k \neq j} (y_k - x_k) \right) f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) \, \mathrm{d}\mathbf{x}.
\end{aligned}
\tag{68}
$$

Also, from (20), we obtain

$$
\begin{aligned}
&\frac{\partial}{\partial y_j} q_j(\mathbf{y}; t) \\
&= \frac{-1}{2\delta(\rho, m) t} \Bigg\{ \int_{\mathbb{R}^m} (y_j - x_j) \left(\sum_{i=1}^m \Phi^{-1}(F_{X_i}(x_i)) + \frac{1}{\sqrt{t}} \sum_{k \neq j} (y_k - x_k) \right) f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) \, \mathrm{d}\mathbf{x} \\
&\quad - 2t\lambda(\rho, m) \int_{\mathbb{R}^m} \left(\sum_{i=1}^m \Phi^{-1}(F_{X_i}(x_i)) + \frac{1}{\sqrt{t}} \sum_{k \neq j} (y_k - x_k) \right)^2 f_{\mathbf{X},\mathbf{W}_t}(\mathbf{x}, \mathbf{y}-\mathbf{x}) \, \mathrm{d}\mathbf{x} \Bigg\}.
\end{aligned}
\tag{69}
$$

Since, for all $j = 1, 2, \ldots, m$, $\mathbf{E}(W_{t,j}^2 \mid \mathbf{Y}=\mathbf{y}) < +\infty$, the first, third, and fourth terms in (68) are finite too and, due to (69), $(\partial/\partial y_j) q_j(\mathbf{y}; t)$ is finite as well. Therefore, from (59), the right-hand side of (67) is finite and, together with the fact that the limit u_3 exists, it follows that $u_3 = 0$.

Similarly, to show that $u_4 = 0$, we find the sequence of relations

$$
\begin{aligned}
&\left| \int_0^{+\infty} \int_{S_r} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\|^2 q(\mathbf{y}; t).\mathbf{n}_{S_r}(\mathbf{y}) \, \mathrm{d}S_r \, \mathrm{d}r \right| \\
&\leq \int_0^{+\infty} \int_{S_r} \|\nabla \log f_{\mathbf{Y}}(\mathbf{y}; t)\|^2 \|q(\mathbf{y}; t)\| \|\mathbf{n}_{S_r}(\mathbf{y})\| \, \mathrm{d}S_r \, \mathrm{d}r = \sqrt{m} \mathbf{E}_{\mathbf{Y}} \big[\|p(\mathbf{Y}; t)\| \|\nabla \log f_{\mathbf{Y}}(\mathbf{Y}; t)\|^2 \big].
\end{aligned}
$$

Using similar steps, we can see that $u_4 = 0$.

Acknowledgements

We express our gratitude to the associate editor and the anonymous reviewers whose comments had a noticeable impact on improving the manuscript.

Funding information

There are no funding bodies to thank relating to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] AMAZIGO, J. AND RUBENFELD, L. (1980). *Advanced Calculus and its Applications to the Engineering and Physical Siences*. Wiley, New York.
[2] ARIAS-NICOLÁS, J. P., FERNÁNDEZ-PONCE, J. M., LUQUE-CALVO, P. AND SUÁREZ-LLORENS, A. (2005). Multivariate dispersion order and the notion of copula applied to the multivariate t-distribution. *Prob. Eng. Inf. Sci.* **19**, 363–375.
[3] ASGARI, F. AND ALAMATSAZ, M. H. (2022). An extension of entropy power inequality for dependent random variables. *Commun. Statist. Theory Meth.* **51**, 4358–4369.
[4] ASGARI, F., ALAMATSAZ, M. H. AND KHOOLENJANI, N. B. (2019). Inequalities for the dependent Gaussian noise channels based on Fisher information and copulas. *Prob. Eng. Inf. Sci.* **33**, 618–657.
[5] BLACHMAN, N. (1965). The convolution inequality for entropy powers. *IEEE Trans. Inf. Theory* **11**, 267–271.
[6] BERGMANS, P. (1974). A simple converse for broadcast channels with additive white Gaussian noise. *IEEE Trans. Inf. Theory* **20**, 279–280.
[7] COSTA, M. (1985). A new entropy power inequality. *IEEE Trans. Inf. Theory* **31**, 751–760.
[8] DEMBO, A. (1989). Simple proof of the concavity of the entropy power with respect to added Gaussian noise. *IEEE Trans. Inf. Theory* **35**, 887–888.
[9] JOE, H. (1997). *Multivariate Models and Dependence Concepts* (Monographs Statist. Appl. Prob. 73). Chapman & Hall, London.
[10] JOHNSON, O. (2004). A conditional entropy power inequality for dependent variables. *IEEE Trans. Inf. Theory* **50**, 1581–1583.
[11] KAY, S. (2009). Waveform design for multistatic radar detection. *IEEE Trans. Aerosp. Electron. Systems* **45**, 1153–1166.
[12] KHOOLENJANI, N. B. AND ALAMATSAZ, M. H. (2016). Extension of de Bruijn's identity to dependent non-Gaussian noise channels. *J. Appl. Prob.* **53**, 360–368.
[13] SHANNON, C. E. (1948). A mathematical of communication. *Bell Systems Tech. J.* **27**, 623–656.
[14] SKLAR, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.
[15] STAM, A. J. (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control* **2**, 101–112.
[16] TAKANO, S., WATANABE, S., FUKUSHIMA, M., PROHOROV, Y. AND SHIRYAEV, A. (1995). The inequalities of Fisher information and entropy power for dependent variables. *Proc. 7th Japan–Russia Symp. Prob. Theory Math. Statist*. pp. 460–470.
[17] VILLANI, C. (2000). A short proof of the concavity of entropy power. *IEEE Trans. Inf. Theory* **46**, 1695–1696.
[18] WEINGARTEN, H., STEINBERG, Y. AND SHAMAI, S. (2006). The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Trans. Inf. Theory* **52**, 3936–3964.