

Article

# Gaussian Optimality for Derivatives of Differential Entropy Using Linear Matrix Inequalities <sup>†</sup>

Xiaobing Zhang <sup>1,2,3</sup>, Venkat Anantharam <sup>4</sup> and Yanlin Geng <sup>2,5,\*</sup>

<sup>1</sup> Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China; xiaobingzhangupc@gmail.com

<sup>2</sup> School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China

<sup>3</sup> University of Chinese Academy of Sciences, Beijing 100049, China

<sup>4</sup> Department of Electrical Engineering and Computer Sciences (EECS), University of California, Berkeley, CA 94720-1234, USA; ananth@eecs.berkeley.edu

<sup>5</sup> Shanghai Institute of Fog Computing Technology, Shanghai Tech University, Shanghai 201210, China

\* Correspondence: gengyanlin@gmail.com; Tel.: +86-21-2068-5381

<sup>†</sup> This paper is an extended version of our paper submitted to 2018 IEEE International Symposium on Information Theory (ISIT), Vail, CO, USA, 17–22 June 2018.

Received: 23 January 2018; Accepted: 5 March 2018; Published: 9 March 2018

**Abstract:** Let  $Z$  be a standard Gaussian random variable,  $X$  be independent of  $Z$ , and  $t$  be a strictly positive scalar. For the derivatives in  $t$  of the differential entropy of  $X + \sqrt{t}Z$ , McKean noticed that Gaussian  $X$  achieves the extreme for the first and second derivatives, among distributions with a fixed variance, and he conjectured that this holds for general orders of derivatives. This conjecture implies that the signs of the derivatives alternate. Recently, Cheng and Geng proved that this alternation holds for the first four orders. In this work, we employ the technique of linear matrix inequalities to show that: firstly, Cheng and Geng’s method may not generalize to higher orders; secondly, when the probability density function of  $X + \sqrt{t}Z$  is log-concave, McKean’s conjecture holds for orders up to at least five. As a corollary, we also recover Toscani’s result on the sign of the third derivative of the entropy power of  $X + \sqrt{t}Z$ , using a much simpler argument.

**Keywords:** differential entropy; entropy power; log-concavity; linear matrix inequality; Gaussian optimality

## 1. Introduction

For a general continuous random variable  $X$  with probability density function  $f(x)$ , its differential entropy is defined as

$$h(X) = - \int_{-\infty}^{+\infty} f(x) \ln f(x) dx,$$

given that the above integral exists. In [1], Shannon considered the entropy power  $N(X) = \frac{1}{2\pi e} e^{2h(X)}$ , and introduced the celebrated Entropy Power Inequality (EPI):

$$e^{2h(X+Y)} \geq e^{2h(X)} + e^{2h(Y)},$$

where  $X$  and  $Y$  are independent, and equality holds if and only if  $X$  and  $Y$  are Gaussian. This inequality is nontrivial and was rigorously proved later by Stam [2].

**Remark 1.** This is the full version of a conference paper submitted to ISIT 2018 [3].

There have been numerous generalizations of the EPI. In [4], Costa considered the case where  $X$  is perturbed by an independent standard Gaussian  $Z$ , and showed that  $N(X + \sqrt{t}Z)$  is concave in  $t$  for  $t > 0$ :

$$\frac{d^2}{dt^2} N(X + \sqrt{t}Z) \leq 0, \quad t > 0.$$

Toscani [5] further showed that  $\frac{d^3}{dt^3} N(X + \sqrt{t}Z) \geq 0$ , under the condition that the probability density function of  $X + \sqrt{t}Z$  is log-concave.

Later, Villani [6] simplified Costa's proof by directly studying the second derivative in  $t$  of the differential entropy instead of the second derivative of the entropy power. In the proof, it was noticed [6–8] that the signs of the first two derivatives of  $h(X + \sqrt{t}Z)$  alternate. Along this line, Cheng and Geng [9] showed that this alternation holds for the first four derivatives, and they made the following conjecture that the alternation is true for general orders of derivatives.

**Conjecture 1** ([9]). *The derivatives of the differential entropy  $h(X + \sqrt{t}Z)$  satisfy  $(-1)^{n-1} \times \frac{d^n}{dt^n} h(X + \sqrt{t}Z) \geq 0$  for  $t > 0$  and  $n \geq 1$ .*

According to Equation (3) in Lemma 2 and the comments,  $2 \times \frac{d}{dt} h(X + \sqrt{t}Z)$  is the Fisher information  $J(X + \sqrt{t}Z)$ . The above conjecture is equivalent to hypothesizing that the Fisher information of  $X + \sqrt{t}Z$  is completely monotone, thus admitting a very simple characterization using the Laplace Transform [10]: there exists a finite Borel measure  $\mu(\cdot)$  such that

$$J(X + \sqrt{t}Z) = \int_0^{+\infty} e^{-\lambda t} \mu(d\lambda). \quad (1)$$

Back in 1966, McKean [7] also studied the derivatives in  $t$  of  $h(X + \sqrt{t}Z)$ , and noticed that Gaussian  $X$  achieves the minimum of  $\frac{d}{dt} h(X + \sqrt{t}Z)$  and  $-\frac{d^2}{dt^2} h(X + \sqrt{t}Z)$ , subject to  $\text{Var}(X) = \sigma^2$ . Then, McKean implicitly made the following conjecture that Gaussian optimality holds generally:

**Conjecture 2** ([7]). *Subject to  $\text{Var}(X) = \sigma^2$ , Gaussian  $X$  with variance  $\sigma^2$  achieves the minimum of  $(-1)^{n-1} \times \frac{d^n}{dt^n} h(X + \sqrt{t}Z)$  for  $t > 0$  and  $n \geq 1$ .*

Notice that, if  $X_G$  is Gaussian with variance  $\sigma^2$ , by routine calculation,

$$\begin{aligned} 2h(X_G + \sqrt{t}Z) &= \ln 2\pi e(\sigma^2 + t), \\ (-1)^{n-1} \times 2 \frac{d^n}{dt^n} h(X_G + \sqrt{t}Z) &= (n-1)! \times (\sigma^2 + t)^{-n} > 0. \end{aligned} \quad (2)$$

Hence, McKean's conjecture implies the one by Cheng and Geng.

Compared with the progress made by Cheng and Geng [9] on Conjecture 1, there has been little progress on Conjecture 2. Most of the existing results are on the second derivative of the differential entropy (or the mutual information), and on generalizing the EPI to other settings. For example: Guo et al. [11] represents the derivatives in the signal-to-noise ratio of the mutual information in terms of the minimum mean-square estimation error, building on de Bruijn's identity [2]; Wibisono and Jog [12] study the mutual information along the density flow defined by the heat equation and show that it is a convex function of time if the initial distribution is log-concave; Wang and Madiman [13] recover the proof of the EPI via rearrangements; Courtade [14] generalizes Costa's EPI to non-Gaussian additive perturbations; and König and Smith [15] propose a quantum version of the EPI.

In this paper, we work on Conjecture 2. The main results are to show that Conjecture 2 holds for higher orders up to at least five under the log-concavity condition, and the introduction of the technique of linear matrix inequalities.

The paper is organized as follows: in Section 2, we obtain the formulae for the derivatives of the differential entropy  $h(X + \sqrt{t}Z)$  (Theorem 1) and show that McKean's conjecture holds for

higher orders up to at least five under the log-concavity condition (Corollary 1). As a corollary, we recover Toscani's result [5] on the third derivative of the entropy power, using the Cauchy–Schwarz inequality, which is much simpler. In Section 3, we introduce the linear matrix inequality approach, and transform the above two conjectures to the feasibility check of semidefinite programming problems. With this approach, we can easily obtain the coefficients in Theorem 1. Then, we show that the direct generalization of the method by Cheng and Geng might not work for orders higher than four for proving Conjecture 1. In Section 4, we prove the main theorem of Section 2.

## 2. Main Results

We first introduce the notation that is used throughout this paper. When the functions are single-variate, we use  $\frac{d}{dy}$  for its derivative. For the multi-variate cases, we use  $\frac{\partial}{\partial y}$  for the partial derivative. To simplify the notation, for the derivatives of a general single-variate function  $g(y)$ , we also use  $g'(y)$ ,  $g''(y)$  and  $g'''(y)$  to represent the first, second and third derivatives, respectively; and  $g^{(n)}(y)$  denotes the  $n$ -th derivative for  $n \geq 1$ .

In the rest of the paper, let  $Z$  be a standard Gaussian random variable, and  $X$  be independent of  $Z$ . Denote

$$Y_t := X + \sqrt{t}Z, \quad t > 0.$$

According to [4,16],  $Y_t$  has nice properties: The probability density function  $f(y, t)$  of  $Y_t$  exists, is strictly positive and infinitely differentiable; The differential entropy  $h(Y_t)$  exists. Denote

$$f_n := \frac{\partial^n}{\partial y^n} f(y, t),$$

$$T_n := \frac{\partial^n}{\partial y^n} \ln f(y, t), \quad n = 0, 1, 2, \dots,$$

where it is understood that  $f_n$  and  $T_n$  are functions of  $(y, t)$ . We also present some properties of  $f(y, t)$  in the following lemma. The proof can be found in, say, [2,16] and Propositions 1 and 2 in [9].

**Lemma 1.** For  $t > 0$ , the probability density function  $f(y, t)$  satisfies the following properties:

- (1) The heat equation holds:  $\frac{\partial}{\partial t} f = \frac{1}{2} \frac{\partial^2}{\partial y^2} f$ .
- (2)  $\lim_{|y| \rightarrow \infty} f_n = 0, \forall n \geq 0, \forall t > 0$ .
- (3) The expectation of the product of the  $T_i$ ,  $E[\prod_i T_i]$  exists, and  $\lim_{|y| \rightarrow \infty} f \prod_i T_i = 0, \forall t > 0$ .

In Lemma 1, part (3), in writing  $E[\prod_i T_i]$ , we think of each  $T_i$  as a function of  $(Y_t, t)$ .

Notice that, given  $X$  and  $Z$ , the differential entropy  $h(X + \sqrt{t}Z)$  is a function of  $t$ . The formulae for the first and second derivatives of  $h(X + \sqrt{t}Z)$  are presented in the following lemma. According to Stam [2], the first equality is due to de Bruijn, and the right-hand side is actually the Fisher information (page 671 of [17]); the second one is due to McKean [7], Toscani [8] and Villani [6]; the Gaussian optimality is due to McKean [7].

**Lemma 2.** For the first and second derivatives of the differential entropy  $h(X + \sqrt{t}Z)$ , the following expressions hold for  $t > 0$ :

$$2h'(X + \sqrt{t}Z) = E \left[ \left( \frac{f_1}{f} \right)^2 \right], \quad (3)$$

$$-2h''(X + \sqrt{t}Z) = E \left[ \left( \frac{f_2}{f} - \frac{f_1^2}{f^2} \right)^2 \right]. \quad (4)$$

Subject to  $\text{Var}X = \sigma^2$ , Gaussian  $X$  with variance  $\sigma^2$  minimizes  $h'(X + \sqrt{t}Z)$  and  $-h''(X + \sqrt{t}Z)$ .

By standard manipulations, one has

$$T_1 = \frac{f_1}{f}, \quad T_2 = \frac{f_2}{f} - \frac{f_1^2}{f^2}. \quad (5)$$

Thus, it is straightforward to rewrite the derivatives as

$$2h'(X + \sqrt{t}Z) = \mathbb{E} \left[ T_1^2 \right], \quad (6)$$

$$-2h''(X + \sqrt{t}Z) = \mathbb{E} \left[ T_2^2 \right]. \quad (7)$$

For the third and fourth derivatives, one can refer to Theorems 1 and 2 in [9], where they were represented by the  $f_i$ . Notice that these representations are not unique, and the ones in [9] are sufficient for identifying the signs. Instead, in Theorem 1, we use the  $T_i$ , and this will facilitate our proof of the Gaussian optimality in Corollary 1.

**Theorem 1.** For  $t > 0$ , the derivatives of the differential entropy  $h(X + \sqrt{t}Z)$  can be expressed as:

$$2h^{(3)}(X + \sqrt{t}Z) = \mathbb{E} \left[ T_3^2 - 2T_2^3 \right], \quad (8)$$

$$-2h^{(4)}(X + \sqrt{t}Z) = \mathbb{E} \left[ T_4^2 + 6T_2^4 - 12T_3^2T_2 \right], \quad (9)$$

$$2h^{(5)}(X + \sqrt{t}Z) = \mathbb{E} \left[ T_5^2 - 24T_2^5 - 8T_4^2T_2 - 6T_3^2T_2T_1^2 + 12T_5T_3T_2 + 114T_3^2T_2^2 \right]. \quad (10)$$

The proof to this theorem is left to Section 4. The existence of such expressions and how to obtain the coefficients are left to Section 3, where the method of linear matrix inequalities is introduced.

#### Log-Concave Case

Lemma 2 already ensures the optimality of Gaussians, subject to  $\text{Var}(X) = \sigma^2$ , for the first and second derivatives. For higher ones, we do not know if we can show the optimality based on the expressions in Theorem 1. Here, we impose the constraint of log-concavity on  $f(y, t)$  and summarize the results in Corollaries 1–3.

A nonnegative function  $f(\cdot)$  is logarithmically concave (or log-concave for short) if its domain is convex and it satisfies the inequality

$$f(\theta x + (1 - \theta)y) \geq f(x)^\theta f(y)^{1-\theta}$$

for all  $x, y$  in the domain and  $0 < \theta < 1$ . If  $f$  is strictly positive, this is equivalent to saying that the logarithm of the function is concave (Section 2.5 of [18]). In our case, assuming that  $f(y, t)$  is log-concave in  $y$  is equivalent to  $T_2 \leq 0$ .

Examples of log-concave distributions include the Gaussian, exponential, Laplace, and the Gamma with parameter larger than one. Notice that, if the probability density function of  $X$  is log-concave, then so is that of  $X + \sqrt{t}Z$  (Section 3.5.2 of [18]).

**Corollary 1.** If the probability density function of  $X + \sqrt{t}Z$  is log-concave, then, subject to  $\text{Var}(X) = \sigma^2$ , Gaussian  $X$  with variance  $\sigma^2$  achieves the minimum of  $(-1)^{n-1}h^{(n)}(X + \sqrt{t}Z)$  for  $t > 0$  and  $3 \leq n \leq 5$ .

**Proof.** Let  $X_G$  be Gaussian with mean  $\mu$  and variance  $\sigma^2$ . The probability density function of  $Y_G := X_G + \sqrt{t}Z$  is

$$f(y_G, t) = \frac{1}{\sqrt{2\pi(\sigma^2 + t)}} \times \exp\left\{-\frac{(y_G - \mu)^2}{2(\sigma^2 + t)}\right\}.$$

The key observation is that the second derivative of the logarithm in the Gaussian case is

$$T_{2,G} := \frac{\partial^2}{\partial y_G^2} \ln f(y_G, t) = -(\sigma^2 + t)^{-1}.$$

Hence, from Equation (2), the derivatives of the differential entropy in the Gaussian case are

$$(-1)^{n-1} \times 2h^{(n)}(X_G + \sqrt{t}Z) = (n-1)! \times (\sigma^2 + t)^{-n} = (n-1)! \times E[-T_{2,G}]^n.$$

Now, if one can show the following chain of inequalities:

$$(-1)^{n-1} \times 2h^{(n)}(X + \sqrt{t}Z) \stackrel{(a)}{\geq} (n-1)! \times E[(-T_2)^n] \stackrel{(b)}{\geq} (n-1)! \times E[-T_2]^n \stackrel{(c)}{\geq} (n-1)! \times E[-T_{2,G}]^n, \quad (11)$$

then one is done.

For inequality (b), the log-concavity condition, namely  $T_2 \leq 0$ , suffices.

For inequality (c), it suffices to show that  $E[-T_2] \geq E[-T_{2,G}] \geq 0$ . This can be proved using Lemma 2: Notice that

$$E\left[\frac{f_2}{f}\right] = \int_{-\infty}^{+\infty} f_2(y, t) dy = \int_{-\infty}^{+\infty} df_1(y, t) = f_1(y, t) \Big|_{-\infty}^{+\infty} = 0 - 0,$$

where the last equality is due to Lemma 1.

Now, from Equation (5),

$$E[-T_2] = E\left[-\frac{f_2}{f} + \frac{f_1^2}{f^2}\right] = E\left[\frac{f_1^2}{f^2}\right]. \quad (12)$$

Combining this with Lemma 2, one has

$$E[-T_2] = 2h'(X + \sqrt{t}Z) \geq 2h'(X_G + \sqrt{t}Z) = E[-T_{2,G}].$$

This part is finished by noticing that  $E[-T_{2,G}] > 0$  from Equation (2).

For inequality (a), we show each case of  $n$  using Theorem 1 and the condition  $T_2 \leq 0$ . For  $n = 3$ ,

$$2h^{(3)}(X + \sqrt{t}Z) = E[T_3^2 - 2T_2^3] \geq E[-2T_2^3] = (n-1)! \times E[(-T_2)^n], \quad n = 3.$$

For  $n = 4$ ,

$$-2h^{(4)}(X + \sqrt{t}Z) = E[T_4^2 + 6T_2^4 - 12T_3^2T_2] \geq E[6T_2^4] = (n-1)! \times E[(-T_2)^n], \quad n = 4,$$

where the inequality is due to  $T_2 \leq 0$ , thus  $E[-12T_3^2T_2] \geq 0$ . For  $n = 5$ ,

$$\begin{aligned} 2h^{(5)}(X + \sqrt{t}Z) &= E\left[T_5^2 - 24T_2^5 - 8T_4^2T_2 - 6T_3^2T_2T_1^2 + 12T_5T_3T_2 + 114T_3^2T_2^2\right] \\ &= E\left[(T_5 + 6T_3T_2)^2 - 24T_2^5 - 8T_4^2T_2 - 6T_3^2T_2T_1^2 + 78T_3^2T_2^2\right] \\ &\geq E\left[-24T_2^5\right] \\ &= (n-1)! \times E[(-T_2)^n], \quad n = 5. \end{aligned}$$

Now, the proof is finished.  $\square$

The following corollary deals with the fifth-order case in [9], under the log-concavity assumption. The proof follows directly from Corollary 1 and Equation (2).

**Corollary 2.** *If the probability density function of  $X + \sqrt{t}Z$  is log-concave, then the fifth derivative of the differential entropy is strictly positive:  $h^{(5)}(X + \sqrt{t}Z) > 0$  for  $t > 0$ .*

Regarding the entropy power, it is already known that  $N'(X + \sqrt{t}Z) \geq 0$  from the connection with Fisher information, and  $N''(X + \sqrt{t}Z) \leq 0$  according to [4]. For the third derivative, Toscani showed that  $N^{(3)}(X + \sqrt{t}Z) \geq 0$ , under the log-concavity assumption. Here, we simplify Toscani's proof, using a Cauchy–Schwartz argument.

**Corollary 3.** *If the probability density function of  $X + \sqrt{t}Z$  is log-concave, then the third derivative of the entropy power is nonnegative:  $N^{(3)}(X + \sqrt{t}Z) \geq 0$  for  $t > 0$ .*

**Proof.** For brevity, let  $h' := h'(X + \sqrt{t}Z)$ , and, similarly, we omit the arguments for higher orders. Routine manipulations yield that

$$N^{(3)}(X + \sqrt{t}Z) = \frac{d^3}{dt^3} \frac{1}{2\pi e} e^{2h(X + \sqrt{t}Z)} = \frac{1}{2\pi e} e^{2h(X + \sqrt{t}Z)} \left[ (2h')^3 + 3 \times 2h' \times 2h'' + 2h''' \right].$$

Thus, it suffices to show  $(2h')^3 + 3 \times 2h' \times 2h'' + 2h''' \geq 0$ . First, we express  $h'$ ,  $h''$  and  $h'''$  in the form of the  $T_i$ : according to Lemma 2 and Equation (12),  $2h' = E[-T_2]$ ; from Equation (7),  $2h'' = -E[T_2^2]$ ; copying from Equation (8),  $2h''' = E[T_3^2 - 2T_2^3]$ .

Also notice that, from Lemma 2,  $2h'(X + \sqrt{t}Z) \geq 2h'(X_G + \sqrt{t}Z) = (\sigma_X^2 + t)^{-1} > 0$ . Hence,  $E[-T_2] > 0$  (it cannot be zero). Now, under the log-concavity condition, namely  $T_2 \leq 0$ , from the Cauchy–Schwartz inequality for random variables, we have:

$$E[-T_2]E[-T_2^3] = E\left[\sqrt{-T_2}^2\right]E\left[\sqrt{-T_2^3}^2\right] \geq E\left[\sqrt{-T_2}\sqrt{-T_2^3}\right]^2 = E\left[T_2^2\right]^2.$$

Thus, we have

$$\begin{aligned} (2h')^3 + 3 \times 2h' \times 2h'' + 2h''' &= E[-T_2]^3 - 3 \times E[-T_2] \times E[T_2^2] + E[T_3^2 - 2T_2^3] \\ &\geq E[-T_2]^3 - 3 \times E[-T_2] \times E[T_2^2] + E[-2T_2^3] \\ &= (E[-T_2])^{-1} \left( E[-T_2]^4 - 3 \times E[-T_2]^2 \times E[T_2^2] + 2E[-T_2]E[-T_2^3] \right) \\ &\geq (E[-T_2])^{-1} \left( E[-T_2]^4 - 3 \times E[-T_2]^2 \times E[T_2^2] + 2E[T_2^2]^2 \right) \\ &= (E[-T_2])^{-1} \left( E[T_2^2] - E[-T_2]^2 \right) \left( 2E[T_2^2] - E[-T_2]^2 \right). \end{aligned}$$

The proof is finished by noticing that  $E[T_2^2] \geq E[-T_2]^2 \geq 0$ , which implies that the right-hand side is nonnegative.  $\square$

### 3. Linear Matrix Inequalities

In this section, we introduce the method of linear matrix inequalities (LMI), and transform the proof of Conjectures 1 and 2 to the feasibility problem of LMI. This transformation also enables us to find the right coefficients in Theorem 1.

Recall that, in [9], the authors first obtained the fourth derivative as the following (Equation (27) in [9])

$$2h^{(4)}(X + \sqrt{t}Z) = E \left[ -\frac{f_4^2}{f^2} - \frac{4f_2f_3^2}{f^3} + \frac{4f_1^2f_3^2}{f^4} - \frac{3f_2^4}{f^4} + \frac{24f_1^2f_2^3}{f^5} - \frac{36f_1^4f_2^2}{f^6} + \frac{90f_1^8}{7f^8} \right]. \quad (13)$$

Then, with some equalities (from integration by parts), they showed this derivative can be expressed as the negative of a sum of squares (Theorem 2 in [9]):

$$2h^{(4)}(X + \sqrt{t}Z) = -E \left[ \left( \frac{f_4}{f} - \frac{6}{5} \frac{f_1 f_3}{f^2} - \frac{7}{10} \frac{f_2^2}{f^2} + \frac{8}{5} \frac{f_1^2 f_2}{f^3} - \frac{1}{2} \frac{f_1^4}{f^4} \right)^2 + \left( \frac{2}{5} \frac{f_1 f_3}{f^2} - \frac{1}{3} \frac{f_1^2 f_2}{f^3} + \frac{9}{100} \frac{f_1^4}{f^4} \right)^2 \right. \\ \left. + \left( -\frac{4}{100} \frac{f_1^2 f_2}{f^3} + \frac{4}{100} \frac{f_1^4}{f^4} \right)^2 + \frac{1}{300} \frac{f_2^4}{f^4} + \frac{56}{90,000} \frac{f_1^4 f_2^2}{f^6} + \frac{13}{70,000} \frac{f_1^8}{f^8} \right]. \quad (14)$$

Hence, the fourth derivative is nonpositive. The sum of squares has a natural connection with positive semidefinite matrices. The right-hand side of Equation (14) can be written as  $-E[u^T F u]$ , where  $u$  is the column vector with coordinates  $u = (f_4/f, f_1 f_3/f^2, f_2^2/f^2, f_1^2 f_2/f^3, f_1^4/f^4)$  and  $F$  is a positive semidefinite matrix. Thus, the method in [9] is actually to verify the existence of a suitable positive semidefinite matrix  $F$ . This can be cast as the feasibility of a linear matrix inequality.

A linear matrix inequality (Chapter 2 of [18]) has the form

$$F(x, y) := F_0 + \sum_{i=1}^I x_i F_i + \sum_{j=1}^J y_j G_j \succeq 0, \quad (15)$$

where the  $m \times m$  symmetric matrices  $F_0, F_i, G_j, i = 1, \dots, I, j = 1, \dots, J$  are given, variables  $x_i$  are real and  $y_j$  are nonnegative, and the notation  $F(x, y) \succeq 0$  means  $F(x, y)$  is positive semidefinite. The feasibility problem refers to identifying if there exists a set of  $x_i$  and  $y_j$  such that  $F(x, y)$  is positive semidefinite.

To reformulate the method used by Cheng and Geng [9] as an LMI feasibility problem, using the fourth derivative as an illustrative example, the main idea is: first, transform the original expression of the derivative to the form

$$-2h^{(4)}(X + \sqrt{t}Z) = E[u^T F_0 u].$$

Then, transform the equalities resulting from integration by parts to the form

$$0 = E[u^T F_i u], \quad i = 1, 2, \dots, I.$$

Finally, try to find a set of variables  $x_i$  such that  $F_0 + \sum_i x_i F_i \succeq 0$ , which is sufficient to show that

$$-2h^{(4)}(X + \sqrt{t}Z) = E[u^T F_0 u] = E[u^T (F_0 + \sum_i x_i F_i) u] \geq 0.$$

One can notice that there is no matrix  $G_j$  in the above statement. This is mainly because only equalities were available in [9]. When one imposes inequality constraints, for example  $T_2 \leq 0$ , as in this paper, then one will be able to construct matrices  $G_j$ .

Before we proceed to introduce the details on constructing those matrices, the following observations are clear regarding  $u = (f_4/f, f_1 f_3/f^2, f_2^2/f^2, f_1^2 f_2/f^3, f_1^4/f^4)$  and the fourth derivative  $2h^{(4)}(X + \sqrt{t}Z)$  (see Equation (13)): (a) the sum-order of derivatives for each entry of  $u$  is four, for example, the sum-order of  $f_1^2 f_2/f^3$  is  $1 \times 2 + 2 = 4$ ; (b) the highest order of a single term in the entries of  $u$  is four, namely  $f_4/f$ ; (c) the sum-order of each entry in the fourth derivative is eight, which is twice that of  $u$ .

In the following, we take the fourth derivative as an example, and show how to construct these matrices  $F_0$  (Section 3.3),  $F_i$  (Sections 3.1 and 3.2), and  $G_j$  (Section 3.4). We decide to use the  $T_k$  as the entries of  $u$ , instead of the  $f_k$ , the motivation for which is clear from the proof of Corollary 1 and the desire to exploit the assumption  $T_2 \leq 0$ . Based on the above observation and the expressions in Equation (5), our vector  $u$  is



$$u = (T_4, T_3T_1, T_2^2, T_2T_1^2, T_1^4). \quad (16)$$

Thus,  $F_0, F_i, G_j$  are  $5 \times 5$  symmetric matrices. Here, we mention that the expressions appearing as coordinates in  $u$  correspond to the integer partitions of four.

The organization of this section is as follows: Sections 3.1–3.3 deal with the sign of the fourth derivative with only equality constraints (see Conjecture 1); Section 3.4 further incorporates the inequality constraints, namely  $T_2 \leq 0$ ; Section 3.5 shows the manipulation for the optimality of Gaussian inputs (see Conjecture 2). In Section 3.6, we consider the sign and the Gaussian optimality for the fifth derivative.

### 3.1. Matrices $F_i$ from Multiple Representations

The matrices  $F_i$  are such that  $E[u^T F_i u] = 0$ . A trivial case is to notice that different products of the form  $u(i)u(j)$  may map to the same term, for example

$$T_2^2 T_1^4 = (T_2^2)(T_1^4) = (T_2 T_1^2)(T_2 T_1^2), \implies u(3)u(5) = u(4)u(4).$$

That is,  $T_2^2 T_1^4$  admits multiple representations as  $u(i)u(j)$ . It is easy to construct the corresponding matrix  $F_1$  such that  $u^T F_1 u = 0$ :

$$F_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}.$$

For the fourth derivative, only one term has multiple representations. There is none for the third derivative, and three for the fifth ( $F_1, F_2$  and  $F_3$  in Section 3.6).

### 3.2. Matrices $F_i$ from Integration by Parts

The equalities of the type  $E[u^T F_i u] = 0$  used in [9] are from integration by parts. Here, we list them one by one.

Notice that all the possible terms with sum-order eight and highest-order four are the following (the numbers in the left column are indices):

$$\begin{array}{ll} 1-5: & T_4^2, \quad T_4 T_3 T_1, \quad T_4 T_2^2, \quad T_4 T_2 T_1^2, \quad T_4 T_1^4, \\ 6-9: & T_3^2 T_1^2, \quad T_3 T_2^2 T_1, \quad T_3 T_2 T_1^3, \quad T_3 T_1^5, \\ 10-13: & T_2^4, \quad T_2^3 T_1^2, \quad T_2^2 T_1^4, \quad T_2 T_1^6, \\ 14: & T_1^8, \\ 15: & T_3^2 T_2. \end{array}$$

Denote this vector as  $w$ .

These terms are arranged in the order such that the first (fourteen) terms can be expressed as  $u(i)u(j)$  for some  $i$  and  $j$ , while the last term(s) cannot be. We call the first terms the quadratic part  $w^{qua}$ , and the last term(s) the non-quadratic part  $w^{non}$ . Thus,  $w = (w^{qua}, w^{non})$ .

It is not difficult to conclude that, for non-repetition, one only needs to perform integration by parts on the entries whose highest-order term is of power one. All of these entries are (eight in total):

$$T_4 T_3 T_1, \quad T_4 T_2^2, \quad T_4 T_2 T_1^2, \quad T_4 T_1^4, \quad T_3 T_2^2 T_1, \quad T_3 T_2 T_1^3, \quad T_3 T_1^5, \quad T_2 T_1^6.$$



Taking  $T_4T_3T_1$  as an example, one can show that (Equation (18), see the end of this subsection)

$$E \left[ 2T_4T_3T_1 + T_3^2T_1^2 + T_3^2T_2 \right] = 0.$$

In addition, this can be written as  $E[c_1^T w] = 0$ , where

$$c_1 \in \mathbb{R}^{15}, \quad c_1([2, 6, 15]) = [2, 1, 1].$$

There are eight equalities in total and hence there are vectors  $c_1, \dots, c_8$ . We put each  $c_i$  as the  $i$ -th row of  $C \in \mathbb{R}^{8 \times 15}$ , and write those equalities as

$$E[Cw] = \mathbf{0},$$

where

$$C = \begin{bmatrix} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ & 2 & & & & & 1 & & & & & & & & & \\ & & 1 & & & & & 1 & & & & & & & & \\ & & & 1 & & & 1 & 2 & 1 & & & & & & & \\ & & & & 1 & & & 4 & 1 & & & & & & & \\ & & & & & 3 & & & & 1 & 1 & & & & & \\ & & & & & & 2 & & & & 3 & 1 & & & & \\ & & & & & & & 1 & & & & 5 & 1 & & & \\ & & & & & & & & & & & & 7 & 1 & & \end{bmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix}. \quad (17)$$

The entries can be found in Equations (18)–(25).

We need to extract matrices  $F$  from these eight equalities  $E[Cw] = \mathbf{0}$ , such that  $E[u^T F u] \equiv 0$ . The main problem is that  $c_k^T w$  may contain entries that are not expressible as  $u(i)u(j)$ . In particular, for the fourth derivative, this happens when  $c_k(15) \neq 0$ . One needs to do some work to cancel these entries. The general method, which can also be used in higher-order cases, is stated below:

1. Firstly, since  $w = (w^{qua}, w^{non})$ , we separate the blocks of  $C$  accordingly,

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & \mathbf{0} \end{bmatrix}, \quad E \left[ \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & \mathbf{0} \end{bmatrix} \begin{bmatrix} w^{qua} \\ w^{non} \end{bmatrix} \right] = \mathbf{0}.$$

In the above,  $C_{11} \in \mathbb{R}^{2 \times 14}$ ,  $C_{12} \in \mathbb{R}^{2 \times 1}$ ,  $C_{21} \in \mathbb{R}^{6 \times 14}$ .

2. Secondly, each row of  $C_{21}$  corresponds to a symmetric matrix  $F_i$  such that  $E[u^T F_i u] \equiv 0$ . In particular, for the first row of  $C_{21}$ , the matrix is

$$F_2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 2 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

such that  $\frac{1}{2}u^T F_2 u = T_4T_2T_1^2 + T_3^2T_1^2 + 2T_3T_2^2T_1 + T_3T_2T_1^3$ . Notice a scaling of a factor of two is added here just for conciseness, and this does not affect the feasibility of (15). Similarly, the other five matrices, corresponding to the remaining rows of  $C_{21}$ , are

$$F_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad F_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 3 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad F_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

$$F_6 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 5 & 1 & 0 \end{bmatrix}, \quad F_7 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{bmatrix}.$$

3. Thirdly, for  $C_{11}$  and  $C_{12}$ , the equalities are  $E[C_{11}w^{qua} + C_{12}w^{non}] = \mathbf{0}$ . Notice  $w^{non}$  cannot be expressed in a quadratic form. Supposing that we can find a column vector  $z$  such that  $z^T C_{12} = 0$ , then  $E[z^T C_{11}w^{qua}] = E[z^T (C_{11}w^{qua} + C_{12}w^{non})] = 0$ . The vector  $z$  actually lies in the null space of  $C_{12}^T$ , and it suffices to find the basis. One way is to do the QR decomposition:

$$C_{12} = Q \begin{bmatrix} U \\ \mathbf{0} \end{bmatrix},$$

where  $U$  is upper-triangular. The null-space of  $C_{12}^T$  has the same dimensions as the number of rows of  $\mathbf{0}$  above, and a basis as the last several columns of  $Q$ —in particular, for the fourth derivative

$$C_{12} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = QR = \begin{bmatrix} -\frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} -\sqrt{5} \\ 0 \end{bmatrix}.$$

Hence, one takes  $z$  as the second column of  $Q$ , which is (after scaling for conciseness)  $z^T = [-2, 1]$ . Then, one calculates  $z^T C_{11}w^{qua} = -4T_4T_3T_1 + T_4T_2^2 - 2T_3^2T_1^2 + T_3T_2^2T_1$ , and the corresponding matrix  $F_8$  (scaled by a factor of two) is

$$F_8 = \begin{bmatrix} 0 & -4 & 1 & 0 & 0 \\ -4 & -4 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The rest of this subsection is devoted to calculating the equalities obtained from integration by parts. This is similar to that in [9], except in the form of the  $T_i$ . To begin, we need the following lemma.

**Lemma 3.** Let  $A$  be a linear combination of terms of products of the  $T_i$ , then, for  $n \geq 2$ ,

$$E \left[ T_n A + T_{n-1} \frac{\partial}{\partial y} A + T_{n-1} T_1 A \right] = 0.$$

**Proof.** From calculus,

$$\begin{aligned}
 E[T_n A] &= \int f T_n A dy = \int f A dT_{n-1} \\
 &\stackrel{(a)}{=} 0 - \int T_{n-1} d(f A) \\
 &= - \int T_{n-1} \left( f_1 A + f \frac{\partial}{\partial y} A \right) dy \\
 &\stackrel{(b)}{=} - \int T_{n-1} \left( f T_1 A + f \frac{\partial}{\partial y} A \right) dy \\
 &= -E \left[ T_{n-1} \frac{\partial}{\partial y} A + T_{n-1} T_1 A \right],
 \end{aligned}$$

where (a) is due to Lemma 1, and (b) is due to Equation (5).  $\square$

Now, using Lemma 3, one obtains the following equalities:

$$T_n = T_4, A = T_3 T_1, \implies E \left[ 2T_4 T_3 T_1 + T_3^2 T_2 + T_3^2 T_1^2 \right] = 0, \quad (18)$$

$$T_n = T_4, A = T_2^2, \implies E \left[ T_4 T_2^2 + 2T_2^2 T_2 + T_3 T_2^2 T_1 \right] = 0, \quad (19)$$

$$T_n = T_4, A = T_2 T_1^2, \implies E \left[ T_4 T_2 T_1^2 + T_3^2 T_1^2 + 2T_3 T_2^2 T_1 + T_3 T_2 T_1^3 \right] = 0, \quad (20)$$

$$T_n = T_4, A = T_1^4, \implies E \left[ T_4 T_1^4 + 4T_3 T_2 T_1^3 + T_3 T_1^5 \right] = 0, \quad (21)$$

$$T_n = T_3, A = T_2^2 T_1, \implies E \left[ 3T_3 T_2^2 T_1 + T_2^4 + T_2^2 T_1^2 \right] = 0, \quad (22)$$

$$T_n = T_3, A = T_2 T_1^3, \implies E \left[ 2T_3 T_2 T_1^3 + 3T_2^3 T_1^2 + T_2^2 T_1^4 \right] = 0, \quad (23)$$

$$T_n = T_3, A = T_1^5, \implies E \left[ T_3 T_1^5 + 5T_2^2 T_1^4 + T_2 T_1^6 \right] = 0, \quad (24)$$

$$T_n = T_2, A = T_1^6, \implies E \left[ 7T_2 T_1^6 + T_1^8 \right] = 0. \quad (25)$$

With these equalities, matrix (17) can be constructed.

### 3.3. Matrix $F_0$ from the Derivative

Suppose we have already obtained the fourth derivative in the form (see Equation (30) later)

$$-2h^{(4)}(X + \sqrt{t}Z) = E[d^T w] = E[d_1^T w^{qua} + d_2^T w^{non}],$$

where  $d_1 \in \mathbb{R}^{14}$ ,  $d_2 \in \mathbb{R}^1$ . Then, similar to  $F_8$ , we can find the matrix  $F_0$  such that  $-2h^{(4)}(X + \sqrt{t}Z) = E[u^T F_0 u]$ .

To cancel the non-quadratic term  $d_2^T w^{non}$ , we solve for  $z_2^T C_{12} = d_2^T$  (the solution  $z_2$  should exist, otherwise it is not possible to find a quadratic form and the LMI approach fails). Then, since  $E[C_{11} w^{qua} + C_{12} w^{non}] = \mathbf{0}$ , we have

$$\begin{aligned}
 -2h^{(4)}(X + \sqrt{t}Z) &= E \left[ d_1^T w^{qua} + d_2^T w^{non} \right] \\
 &= E \left[ d_1^T w^{qua} + d_2^T w^{non} - z_2^T (C_{11} w^{qua} + C_{12} w^{non}) \right] \\
 &= E \left[ (d_1^T - z_2^T C_{11}) w^{qua} \right].
 \end{aligned}$$

Now,  $F_0$  can be constructed from  $d_1^T - z_2^T C_{11}$ .

The details are as follows. First, we need to express the derivative using the entries of  $w$ . This can be done recursively using the following lemma.

**Lemma 4.** Let  $A$  be a linear combination of terms of products of the  $T_i$ . The following equalities hold:

$$2 \frac{\partial}{\partial t} T_n = \frac{\partial^n}{\partial y^n} (T_2 + T_1^2) = T_{n+2} + \sum_{k=0}^n C_n^k T_{k+1} T_{n-k+1}, \quad n \geq 0, \quad (26)$$

$$\frac{d}{dt} E[A] = E \left[ \frac{1}{2} (T_2 + T_1^2) A + \frac{\partial}{\partial t} A \right], \quad (27)$$

$$\frac{d}{dt} E[T_n^2] = E \left[ -T_{n+1}^2 + T_n \sum_{k=1}^{n-1} C_n^k T_{k+1} T_{n-k+1} \right], \quad n \geq 1. \quad (28)$$

The proof is left to Appendix A. Now, with Equation (7):

$$-2h''(X + \sqrt{t}Z) = E[T_2^2],$$

and Equation (28), one can easily obtain that

$$2h^{(3)}(X + \sqrt{t}Z) = -\frac{d}{dt} E[T_2^2] = E[T_3^2 - 2T_2^3]. \quad (29)$$

For the fourth derivative,

$$\begin{aligned} -2h^{(4)}(X + \sqrt{t}Z) &= -\frac{d}{dt} E[T_3^2 - 2T_2^3] \\ &\stackrel{(a)}{=} -E[-T_4^2 + T_3(3T_2T_3 + 3T_3T_2)] - \frac{d}{dt} E[-2T_2^3] \\ &\stackrel{(b)}{=} -E[-T_4^2 + T_3(3T_2T_3 + 3T_3T_2)] - E\left[-(T_2 + T_1^2)T_2^3 - 3T_2^2 \frac{\partial}{\partial t} T_2\right] \\ &\stackrel{(c)}{=} -E[-T_4^2 + 6T_3^2T_2 - T_2^4 - T_2^3T_1^2 - 3T_2^2(T_4 + 2T_3T_1 + 2T_2T_2)] \\ &= E[T_4^2 + 3T_4T_2^2 - 6T_3^2T_2 + 6T_3T_2^2T_1 + 7T_2^4 + T_2^3T_1^2], \end{aligned} \quad (30)$$

where (a), (b), (c) are due to Equations (28), (27), (26), respectively.

Hence, we have the vector  $d = (d_1, d_2) \in \mathbb{R}^{15}$  and its blocks  $d_1 \in \mathbb{R}^{14}$ ,  $d_2 \in \mathbb{R}^1$ :

$$\begin{aligned} d[1, 3, 7, 10, 11] &= [1, 3, 6, 7, 1], \\ d[15] &= -6. \end{aligned}$$

One solves for  $z_2$  such that  $z_2^T C_{12} = d_2^T$  and obtains

$$z_2^T = [0, -3].$$

Now,  $d_1^T - z_2^T C_{11}$  has nonzero entries at locations  $[1, 3, 7, 10, 11]$ , with values  $[1, 6, 9, 7, 1]$ , respectively. Furthermore,  $F_0$  (scaled by a factor of two) is found as

$$F_0 = \begin{bmatrix} 2 & 0 & 6 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 6 & 9 & 14 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

By the end of this subsection, it is easy to see that Cheng and Geng's method can be reformulated as identifying if there exist  $x_1, \dots, x_8 \in \mathbb{R}$  such that

$$F_0 + \sum_{i=1}^8 x_i F_i \succeq 0.$$

We use the convex optimization package [19] to identify the feasibility of the above LMI problem, and it turns out to be feasible as it should be according to Equation (14).

### 3.4. Matrices $G_j$ from Log-Concavity

Recall that, in [9], there is no matrix  $G_j$ , since there is no inequality constraint. In this paper, we consider the log-concave case  $T_2 \leq 0$ , thus introducing inequality constraints.

For the fourth order,  $T_2 \leq 0$  actually implies that the following entries in  $w$  are nonpositive:

$$T_2^3 T_1^2, \quad T_2 T_1^6, \quad T_2 T_3^2.$$

It is clear that the powers of  $T_2$  are odd, and the others are even.

To transform these nonpositive terms into matrices  $G_j$ , the first two terms,  $T_2^3 T_1^2$  and  $T_2 T_1^6$  are trivial, since they can be expressed by  $u(i)u(j)$  directly:

$$\begin{aligned} 0 &\geq 2E[T_2^3 T_1^2] = E[u^T G_1 u], \\ 0 &\geq 2E[T_2 T_1^6] = E[u^T G_2 u], \end{aligned}$$

where

$$G_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad G_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

For the term  $T_2 T_3^2$ , the idea is similar to the third part in Section 3.2. One first finds  $z_3 \in \mathbb{R}^2$  such that  $z_3^T C_{12} w^{non} = T_2 T_3^2$ , namely  $z_3^T C_{12} = 1$ . The solution is  $z_3^T = [0, \quad 1/2]$ . Then,

$$E[T_2 T_3^2] = E[T_2 T_3^2 - z_3^T (C_{11} w^{qua} + C_{12} w^{non})] = E[-z_3^T C_{11} w^{qua}] = E[-\frac{1}{2} T_4 T_2^2 - \frac{1}{2} T_3 T_2^2 T_1].$$

Now, it is routine to obtain

$$0 \geq 4E[T_2 T_3^2] = E[u^T G_3 u],$$

where

$$G_3 = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

At this point, we are done with the procedure for calculating all these matrices  $F_0$ , the  $F_i$  and the  $G_j$ . To show the negativity of the fourth derivative, it suffices to find a set of variables  $x_i \in \mathbb{R}$  and  $y_j \geq 0$  such that

$$F_0 + \sum_{i=1}^8 x_i F_i + \sum_{j=1}^3 y_j G_j \succeq 0.$$

**Remark 2.** The matrix  $G_2$  is actually redundant, since we know that  $E[T_2 T_1^6] \equiv -\frac{1}{7} E[T_1^8] \leq 0$ , which is already included in the matrices  $F_i$  (in particular, matrix  $F_7$  in Section 3.2). Including  $G_2$  will not affect the feasibility check.

### 3.5. Matrix $\tilde{F}_0$ for Gaussian Optimality

However, to show the optimality of the Gaussian, the above formulation is not enough. According to inequality (a) in Equation (11), it would suffice to show that

$$(-1)^{n-1} \times 2h^{(n)}(X + \sqrt{t}Z) - (n-1)! \times E[(-T_2)^n] \geq 0$$

instead of  $(-1)^{n-1} \times 2h^{(n)}(X + \sqrt{t}Z) \geq 0$ . Thus, one needs to calculate the matrix  $F_0$  such that

$$(-1)^{n-1} \times 2h^{(n)}(X + \sqrt{t}Z) - (n-1)! \times E[(-T_2)^n] = E[u^T F_0 u].$$

The procedure is the same as that in Section 3.3.

In particular, for the fourth derivative, since  $n = 4$  is even, we directly have the quadratic form  $E[(-T_2)^n] = u(3)u(3)$ . It is straightforward to construct the matrix  $\tilde{F}_0$  (scaled by a factor of two) here

$$\tilde{F}_0 = F_0 - \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 \times 3! & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 6 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 6 & 9 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

such that  $E[u^T \tilde{F}_0 u] = -4h^{(4)}(X + \sqrt{t}Z) - 12E[T_2^4]$ .

Now, the LMI is updated as  $\tilde{F}_0 + \sum_{i=1}^8 x_i F_i + \sum_{j=1}^3 y_j G_j \succeq 0$ . Again, we use the convex optimization package [19] to check the feasibility. It turns out to be feasible and the solution helps us to identify the coefficients in Equation (9).

### 3.6. Fifth Derivative

For the fifth derivative, we omit the details of the manipulations since they are routine, and just provide the matrices here. For brevity, we only list out the nonzero entries of the upper-triangular part of a symmetric matrix. These matrices (with scaling) are

$$\begin{aligned} F_0 : & F[(1,1), (1,3), (1,5), (2,3), (2,5), (3,3), (3,4), (3,5), (3,6), (5,5), (5,6)] \\ & = [2, 20, 29, \frac{214}{3}, \frac{49}{2}, \frac{178}{3}, -\frac{37}{3}, 58, -6, 45, -\frac{1}{2}], \\ F_1 : & F[(3,6), (4,5)] = [-1, 1], \\ F_2 : & F[(3,7), (4,6)] = [-1, 1], \\ F_3 : & F[(5,7), (6,6)] = [-1, 2], \\ F_4 : & F[(1,4), (2,2), (2,3), (2,4)] = [1, 2, 2, 1], \\ F_5 : & F[(1,6), (2,4), (2,5), (2,6)] = [1, 1, 3, 1], \\ F_6 : & F[(1,7), (2,6), (2,7)] = [1, 5, 1], \\ F_7 : & F[(2,4), (3,4), (4,4)] = [2, 3, 2], \\ F_8 : & F[(2,5), (3,4), (3,5), (3,6)] = [1, 2, 2, 1], \\ F_9 : & F[(2,6), (3,6), (3,7), (4,4)] = [1, 4, 1, 2], \\ F_{10} : & F[(2,7), (3,7), (4,7)] = [1, 6, 1], \\ F_{11} : & F[(3,6), (5,5), (5,6)] = [3, 6, 1], \\ F_{12} : & F[(3,7), (5,6), (5,7)] = [2, 5, 1], \\ F_{13} : & F[(4,7), (5,7), (6,7)] = [1, 7, 1], \\ F_{14} : & F[(6,7), (7,7)] = [9, 2], \\ F_{15} : & F[(1,2), (1,3), (2,2), (2,3), (3,3), (3,4)] = [6, -3, 6, -5, -2, -1], \\ F_{16} : & F[(1,5), (2,3), (2,5), (3,3), (3,5)] = [-1, -2, -1, 6, 1]. \end{aligned}$$

For the sign of the fifth derivative, we used the convex optimization package [19] to solve the following LMI problem,

$$F_0 + \sum_{i=1}^{16} x_i F_i \succeq 0,$$

but could not find a feasible solution  $x_1, \dots, x_{16} \in \mathbb{R}$ . This suggests to us that a direct generalization of Cheng and Geng's method may not work for the fifth derivative.

Instead, if we consider the log-concavity constraint  $T_2 \leq 0$  and check the optimality of Gaussian inputs, then we have a new matrix  $\tilde{F}_0$  (similar to Section 3.5) and several matrices  $G_j$  as the following:

$$\begin{aligned} \tilde{F}_0 : & F[(1,1), (1,3), (1,5), (2,3), (2,5), (3,3), (3,4), (3,5), (3,6), (5,5), (5,6)] \\ & = [2, 20, 29, \frac{214}{3}, \frac{49}{2}, \frac{178}{3}, -\frac{37}{3}, -38, -6, -3, -\frac{1}{2}], \\ G_1 : & G[(3,4)] = [1], \\ G_2 : & G[(5,6)] = [1], \\ G_3 : & G[(6,7)] = [1], \\ G_4 : & G[(1,3), (2,3), (3,3), (3,4)] = [-3, -5, -2, -1], \\ G_5 : & G[(3,5), (5,5)] = [-2, -1]. \end{aligned}$$

Now, one would like to find  $x_1, \dots, x_{16} \in \mathbb{R}$  and  $y_1, \dots, y_5 \in \mathbb{R}^+$  such that

$$\tilde{F}_0 + \sum_{i=1}^{16} x_i F_i + \sum_{i=1}^5 y_i G_i \succeq 0.$$

This can be solved by the convex optimization package [19]. Again, the solution helps us to arrive at Equation (10).

#### 4. Proof of Theorem 1

**Proof.** For the third derivative, according to Equation (29), we have

$$2h^{(3)}(X + \sqrt{t}Z) = \mathbb{E} [T_3^2 - 2T_2^3].$$

For the fourth derivative, according to Equation (30):

$$-2h^{(4)}(X + \sqrt{t}Z) = \mathbb{E} [T_4^2 + 3T_4T_2^2 - 6T_3^2T_2 + 6T_3T_2^2T_1 + 7T_2^4 + T_2^3T_1^2].$$

Adding multiples of the left-hand sides of the equations:  $-3 \times (19) - (22)$ , we obtain

$$\begin{aligned} -2h^{(4)}(X + \sqrt{t}Z) &= \mathbb{E} [T_4^2 + 3T_4T_2^2 - 6T_3^2T_2 + 6T_3T_2^2T_1 + 7T_2^4 + T_2^3T_1^2] \\ &\stackrel{(a)}{=} \mathbb{E} [T_4^2 + (-6T_3^2T_2 - 3T_3T_2^2T_1) - 6T_3^2T_2 + 6T_3T_2^2T_1 + 7T_2^4 + T_2^3T_1^2] \\ &= \mathbb{E} [T_4^2 - 12T_3^2T_2 + 3T_3T_2^2T_1 + 7T_2^4 + T_2^3T_1^2] \\ &\stackrel{(b)}{=} \mathbb{E} [T_4^2 - 12T_3^2T_2 + (-T_2^4 - T_2^3T_1^2) + 7T_2^4 + T_2^3T_1^2] \\ &= \mathbb{E} [T_4^2 - 12T_3^2T_2 + 6T_2^4], \end{aligned}$$

where (a) is due to Equation (19), and (b) is due to Equation (22).



For the fifth derivative,

$$2h^{(5)}(X + \sqrt{t}Z) = \frac{d}{dt}E \left[ -T_4^2 - 6T_2^4 + 12T_3^2T_2 \right] = \frac{d}{dt}E \left[ -T_4^2 \right] + \frac{d}{dt}E \left[ -6T_2^4 \right] + \frac{d}{dt}E \left[ 12T_3^2T_2 \right].$$

For each term above on the right-hand side: According to Equation (28),

$$\frac{d}{dt}E \left[ -T_4^2 \right] = E \left[ T_5^2 - 8T_4^2T_2 - 6T_4T_3^2 \right]. \quad (31)$$

For the second term,

$$\begin{aligned} \frac{d}{dt}E \left[ -6T_2^4 \right] &\stackrel{(c)}{=} E \left[ -3(T_2 + T_1^2)T_2^4 - 12T_2^3 \times \frac{\partial}{\partial t}(2T_2) \right] \\ &\stackrel{(d)}{=} E \left[ -3T_2^5 - 3T_2^4T_1^2 - 12T_2^3 \left( T_4 + 2T_3T_1 + 2T_2^2 \right) \right] \\ &= E \left[ -3T_2^4T_1^2 - 12T_4T_2^3 - 24T_3T_2^3T_1 - 27T_2^5 \right], \end{aligned} \quad (32)$$

where (c) is due to Equation (27), and (d) is due to Equation (26). For the third term, according to Equation (27),

$$\frac{d}{dt}E \left[ 12T_3^2T_2 \right] = E \left[ 6 \left( T_2 + T_1^2 \right) T_3^2T_2 + \frac{\partial}{\partial t} \left( 12T_3^2T_2 \right) \right],$$

where the last term is

$$\begin{aligned} \frac{\partial}{\partial t} \left( 12T_3^2T_2 \right) &= 12T_3^2 \frac{\partial}{\partial t} T_2 + 24T_3T_2 \times \frac{\partial}{\partial t} T_3 \\ &\stackrel{(e)}{=} 6T_3^2 \left( T_4 + 2T_3T_1 + 2T_2^2 \right) + 12T_3T_2 \left( T_5 + 2T_4T_1 + 6T_3T_2 \right) \\ &= 12T_5T_3T_2 + 6T_4T_3^2 + 24T_4T_3T_2T_1 + 12T_3^3T_1 + 84T_3^2T_2^2, \end{aligned}$$

where (e) is due to Equation (26). Hence

$$\frac{d}{dt}E \left[ 12T_3^2T_2 \right] = E \left[ 12T_5T_3T_2 + 6T_4T_3^2 + 24T_4T_3T_2T_1 + 6T_3^2T_2T_1^2 + 12T_3^3T_1 + 90T_3^2T_2^2 \right]. \quad (33)$$

Finally, combining Equations (31)–(33), we get

$$\begin{aligned} 2h^{(5)}(X + \sqrt{t}Z) &= \frac{d}{dt}E \left[ -T_4^2 \right] + \frac{d}{dt}E \left[ -6T_2^4 \right] + \frac{d}{dt}E \left[ 12T_3^2T_2 \right] \\ &= E \left[ T_5^2 - 8T_4^2T_2 - 3T_2^4T_1^2 - 12T_4T_2^3 - 24T_3T_2^3T_1 \right. \\ &\quad \left. - 27T_2^5 + 12T_5T_3T_2 + 24T_4T_3T_2T_1 + 6T_3^2T_2T_1^2 + 12T_3^3T_1 + 90T_3^2T_2^2 \right]. \end{aligned} \quad (34)$$

To simplify Equation (34), using Lemma 3, one first obtains the following equalities:

$$T_n = T_4, A = T_3T_2T_1, \implies E \left[ 2T_4T_3T_2T_1 + T_3^3T_1 + T_3^2T_2^2 + T_3^2T_2T_1^2 \right] = 0, \quad (35)$$

$$T_n = T_3, A = T_2^3T_1, \implies E \left[ 4T_3T_2^3T_1 + T_2^5 + T_2^4T_1^2 \right] = 0, \quad (36)$$

$$T_n = T_4, A = T_2^3, \implies E \left[ T_4T_2^3 + 3T_3^2T_2^2 + T_3T_2^3T_1 \right] = 0. \quad (37)$$

Then, adding multiples of the left-hand sides of Equations (35)–(37), we have

$$\begin{aligned} 2h^{(5)}(X + \sqrt{t}Z) &= 2h^{(5)}(X + \sqrt{t}Z) - 12 \times (35) + 3 \times (36) + 12 \times (37) \\ &= E \left[ T_5^2 - 24T_2^5 - 8T_4^2T_2 - 6T_3^2T_2T_1^2 + 12T_5T_3T_2 + 114T_3^2T_2^2 \right]. \end{aligned}$$

□

## 5. Discussion

### 5.1. On the Derivatives

We are not able to say anything conclusive about the sign of the fifth derivative of the differential entropy  $h(X + \sqrt{t}Z)$ . If we impose the log-concavity condition, namely  $T_2 \leq 0$ , then the fifth derivative is at least  $4! \times E[(-T_2)^5]$ . This motivates us to consider the following problem: Without additional constraints, what are the values  $c_5 > 0$  such that

$$2h^{(5)}(X + \sqrt{t}Z) \geq c_5 \times E[(-T_2)^5].$$

If one finds such a value  $c_5$ , then so long as  $E[(-T_2)^5] \geq 0$ , the sign of the fifth derivative is determined. This condition is much weaker than  $T_2 \leq 0$ .

For the computational part, one only needs to construct the matrix  $\tilde{F}_0$  such that  $2h^{(5)}(X + \sqrt{t}Z) - c_5 \times E[(-T_2)^5] = E[u^T \tilde{F}_0 u]$ , and then solve the problem (see Section 3.6 for the matrices  $F_i$ )

$$\tilde{F}_0 + \sum_{i=1}^{16} x_i F_i \succeq 0.$$

It turns out that  $c_5 = 0.13$  works, while  $c_5 = 0.125$  fails. The authors guess that  $c_5 \in [0.13, 24]$  works, but, at the moment, can just partly confirm this with limited simulation.

Notice that the third derivative of the entropy power  $N(X + \sqrt{t}Z)$  was shown to be nonnegative under the log-concavity condition [5], and we recover this in Corollary 3. We also considered the fourth derivative, but failed to obtain the sign because we were unable to apply the Cauchy–Schwartz inequality as we did for the third derivative.

### 5.2. Possible Proofs

To prove Conjecture 1, besides the method proposed in [9], we are also considering the following ways: the first one is constructive and inspired by Equation (1). Given a random variable  $X$ , if we can construct a proper measure  $\mu(\cdot)$  such that Equation (1) holds, then one proves Conjecture 1. However, this is difficult even when  $X$  is binary symmetric, which is a very simple random variable.

The second one is recursive. Suppose one can find a formula for the  $n$ -th derivative such that

$$\begin{aligned} (-1)^{n-1} \times h^{(n)}(X + \sqrt{t}Z) &= E \left[ \sum_{i=1}^{k_n} A_i^2 \right], \\ \frac{d}{dt} E[A_i^2] &= E[-B_i^2], \end{aligned}$$

then it is clear that

$$(-1)^n \times h^{(n+1)}(X + \sqrt{t}Z) = E \left[ \sum_{i=1}^{k_n} B_i^2 \right] \geq 0.$$

However, this fails for  $n = 2$  (see Equation (7) and Theorem 1). Instead, one may expect that

$$\frac{d}{dt} E[A_i^2] = E[-B_i^2 - C_i^2 + B_{i+1}^2],$$

and then

$$(-1)^n \times h^{(n+1)}(X + \sqrt{t}Z) = E[-B_1^2 + B_{k_n+1}^2 - \sum_{i=1}^{k_n} C_i^2].$$

If further one can show that  $E[-B_1^2 + B_{k_n+1}^2] = E[-C_{k_n+1}^2]$  for some  $C_{k_n+1}$ , then one finishes the proof. Notice here that a clever observation is needed for this way to work.

### 5.3. Applications

The topic of Gaussian optimality has wide applications, for example in [20,21]. In this work, besides the Gaussian optimality, we also have some new observations. In [11], the derivatives in the signal-noise ratio (snr) of  $I(X; \sqrt{\text{snr}}X + Z)$  are studied. In particular, the first four derivatives are obtained in the language of the minimum mean-square error (Equations (69)–(72) in Corollary 1 of [11]). However, it is not clear whether some of these derivatives are signed or not.

With some standard manipulations, it is not difficult to show that

$$I(X; \sqrt{\text{snr}}X + Z) = h(\sqrt{\text{snr}}X + Z) - h(Z) = h(X + \frac{1}{\sqrt{\text{snr}}}Z) + \log \sqrt{\text{snr}} - \frac{1}{2} \log 2\pi e.$$

By letting  $t = 1/\sqrt{\text{snr}}$ , one can easily connect the minimum mean-square error formulae in [11] with the signs of the derivatives of  $h(X + \sqrt{t}Z)$  in  $t$ . The verification of Conjectures 1 and 2 would imply the bounding and extremal properties of Equations (69)–(72) in [11], and thus deepen our understanding of the minimum mean-square error estimation under the additive-Gaussian setting.

In addition, notice that the probability density function  $f(y, t)$  of  $Y = X + \sqrt{t}Z$  is the solution of the heat equation  $\frac{\partial}{\partial t} f(y, t) = \frac{1}{2} \frac{\partial^2}{\partial y^2} f(y, t)$  with the initial condition that  $f(y, 0) = f_X(y)$ . Hence, Conjectures 1 and 2, if true, reveal the properties of the differential entropy of functions that satisfy the heat equation. For more results related to diffusion equations, one may refer to [22].

## 6. Conclusions

In this paper, we studied two conjectures on the derivatives of the differential entropy of a general random variable with added Gaussian noise. Regarding the conjecture on the signs of the derivatives made by Cheng and Geng, we introduced the linear matrix inequality approach to provide evidence that their original method might not generalize to orders higher than four. Instead, we considered imposing an additional constraint, namely the log-concavity assumption, and showed the optimality of Gaussian random variables for orders three, four and five. Thus, we made progress on McKean's conjecture, under a mild condition.

**Acknowledgments:** The authors would like to thank Professor Chandra Nair for his valuable suggestions. The work of Venkat Anantharam was supported by the National Science Foundation (NSF) grants ECCS-1343398, CNS-1527846, CCF-1618145, the NSF Science and Technology Center grant CCF-0939370 (Science of Information), and the William and Flora Hewlett Foundation supported Center for Long Term Cybersecurity at Berkeley. The work of Yanlin Geng was supported in part by the National Natural Science Foundation of China under Grant 61601288, and the Science and Technology Commission of Shanghai Municipality under Grant 15YF1407900.

**Author Contributions:** Venkat Anantharam proposed the linear matrix inequality approach. Xiaobing Zhang performed the experiments to find the coefficients in the theorem and wrote the paper. Yanlin Geng proved the main results. Venkat Anantharam and Yanlin Geng reviewed and edited the manuscript. All authors read and approved the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Proof of Lemma 4

**Proof.** For Equation (26), according to Lemma 1,  $f(y, t)$  satisfies the following (heat) equation:

$$f_t := \frac{\partial}{\partial t} f = \frac{1}{2} f_2.$$

In addition, according to Equation (5),

$$T_1 = \frac{f_1}{f}, \quad T_2 = \frac{f_2}{f} - \frac{f_1^2}{f^2}.$$

Hence,

$$2 \frac{\partial}{\partial t} T_0 = 2 \frac{\partial}{\partial t} \ln f(y, t) = \frac{2f_t}{f} = \frac{f_2}{f} = T_2 + T_1^2.$$

Now, it follows that, for  $n \geq 0$ ,

$$2 \frac{\partial}{\partial t} T_n = 2 \frac{\partial}{\partial t} \left( \frac{\partial^n}{\partial y^n} T_0 \right) = \frac{\partial^n}{\partial y^n} \left( \frac{\partial}{\partial t} 2T_0 \right) = \frac{\partial^n}{\partial y^n} (T_2 + T_1^2) = T_{n+2} + \sum_{k=0}^n C_n^k T_{k+1} T_{n-k+1}.$$

For Equation (27),

$$\frac{d}{dt} E[A] = \frac{d}{dt} \int f A dy = \int \left( f_t A + f \frac{\partial}{\partial t} A \right) dy = \int \left( f \frac{1}{2} \frac{f_2}{f} A + f \frac{\partial}{\partial t} A \right) dy = E \left[ \frac{1}{2} (T_2 + T_1^2) A + \frac{\partial}{\partial t} A \right].$$

For Equation (28), the derivative is

$$\frac{d}{dt} E[T_n^2] \stackrel{(a)}{=} \int \left( f \frac{1}{2} (T_2 + T_1^2) T_n^2 + f \frac{\partial}{\partial t} T_n^2 \right) dy = \int \left( f \frac{1}{2} (T_2 + T_1^2) T_n^2 + f T_n \times 2 \frac{\partial}{\partial t} T_n \right) dy,$$

where (a) is due to Equation (27).

For the first term of the right-hand side, from Lemma 1 and integration by parts,

$$\begin{aligned} \int f T_{n+1} T_n T_1 dy &= \int f T_n T_1 dT_n = 0 - \int T_n (f_1 T_n T_1 + f T_{n+1} T_1 + f T_n T_2) dy, \\ \Rightarrow \int f T_{n+1} T_n T_1 dy &= -\frac{1}{2} \int f T_n^2 (T_2 + T_1^2) dy. \end{aligned}$$

For the second term, we have

$$\begin{aligned} \int f T_n \times 2 \frac{\partial}{\partial t} T_n dy &\stackrel{(b)}{=} \int f T_n \left( T_{n+2} + \sum_{k=0}^n C_n^k T_{k+1} T_{n-k+1} \right) dy \\ &= \int f T_n \left( 2T_{n+1} T_1 + \sum_{k=1}^{n-1} C_n^k T_{k+1} T_{n-k+1} \right) dy + \int f T_n dT_{n+1} \\ &= \int \left( f T_n (2T_{n+1} T_1 + \sum_{k=1}^{n-1} C_n^k T_{k+1} T_{n-k+1}) - T_{n+1} (f T_1 T_n + f T_{n+1}) \right) dy \\ &= \int f \left( -T_{n+1}^2 + T_{n+1} T_n T_1 + T_n \sum_{k=1}^{n-1} C_n^k T_{k+1} T_{n-k+1} \right) dy, \end{aligned}$$

where (b) is due to Equation (26).

Combining these two terms together, the third equality is proved.  $\square$

## References

- Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
- Stam, A.J. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control* **1959**, *2*, 101–112.
- Zhang, X.; Anantharam, V.; Geng, Y. Gaussian Extremality for Derivatives of Differential Entropy under the Additive Gaussian Noise Flow. *IEEE Int. Symp. Inf. Theory* **2018**, submitted.
- Costa, M. A new entropy power inequality. *IEEE Trans. Inf. Theory* **1985**, *31*, 751–760.
- Toscani, G. A concavity property for the reciprocal of Fisher information and its consequences on Costa's EPI. *Phys. A Stat. Mech. Appl.* **2015**, *432*, 35–42.
- Villani, C. A short proof of the "concavity of entropy power". *IEEE Trans. Inf. Theory* **2000**, *46*, 1695–1696.
- McKean, H.P. Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas. *Arch. Ration. Mech. Anal.* **1966**, *21*, 343–367.
- Toscani, G. Entropy production and the rate of convergence to equilibrium for the Fokker-Planck equation. *Q. Appl. Math.* **1999**, *57*, 521–541.
- Cheng, F.; Geng, Y. Higher order derivatives in Costa's entropy power inequality. *IEEE Trans. Inf. Theory* **2015**, *61*, 5892–5905.
- Bernstein, S. Sur les fonctions absolument monotones. *Acta Math.* **1929**, *52*, 1–66.
- Guo, D.; Wu, Y.; Shitz, S.S.; Verdú, S. Estimation in Gaussian noise: Properties of the minimum mean-square error. *IEEE Trans. Inf. Theory* **2011**, *57*, 2371–2385.
- Wibisono, A.; Jog, V. Convexity of mutual information along the heat flow. *arXiv* **2018**, arXiv:1801.06968.
- Wang, L.; Madiman, M. Beyond the entropy power inequality, via rearrangements. *IEEE Trans. Inf. Theory* **2014**, *60*, 5116–5137.
- Courtade, T.A. Concavity of entropy power: Equivalent formulations and generalizations. In Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, 25–30 June 2017; pp. 56–60.
- König, R.; Smith, G. The entropy power inequality for quantum systems. *IEEE Trans. Inf. Theory* **2014**, *60*, 1536–1548.
- Rioul, O. Information theoretic proofs of entropy power inequalities. *IEEE Trans. Inf. Theory* **2011**, *57*, 33–55.
- Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: New York, NY, USA, 2006.
- Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: Cambridge, UK, 2004.
- Grant, M.; Boyd, S.; Ye, Y. CVX: Matlab Software for Disciplined Convex Programming. 2008. Available online: <http://cvxr.com/cvx/> (accessed on 5 March 2018).
- Weingarten, H.; Steinberg, Y.; Shamai, S.S. The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Trans. Inf. Theory* **2006**, *52*, 3936–3964.
- Geng, Y.; Nair, C. The capacity region of the two-receiver Gaussian vector broadcast channel with private and common messages. *IEEE Trans. Inf. Theory* **2014**, *60*, 2087–2104.
- Toscani, G. Diffusion Equations and Entropy Inequalities. *preprint* **2016**. Available online: <http://mate.unipv.it/toscani/publi/Note-Ravello-2016.pdf> (accessed on 5 March 2018).



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).