Tight Lower Bound on Differential Entropy for Mixed Gaussian Distributions

Abdelrahman Marconi, Ahmed H. Elghandour, Ashraf D. Elbayoumy, and Amr Abdelaziz

Military Technical College, Cairo, Egypt

https://doi.org/10.26636/jtit.2024.2.1444

Abstract — In this paper, a tight lower bound for the differential entropy of the Gaussian mixture model is presented. First, the probability model of mixed Gaussian distribution that is created by mixing both discrete and continuous random variables is investigated in order to represent symmetric bimodal Gaussian distribution using the hyperbolic cosine function, on which a tighter upper bound is set. Then, this tight upper bound is used to derive a tight lower bound for the differential entropy of the Gaussian mixture model introduced. The proposed lower bound allows to maintain its tightness over the entire range of the model's parameters and shows more tightness when compared with other bounds that lose their tightness over certain parameter ranges. The presented results are then extended to introduce a more general tight lower bound for asymmetric bimodal Gaussian distribution, in which the two modes have a symmetric mean but differ in terms of their weights.

Keywords — differential entropy, lower bound, mixture random variable, multimodal Gaussian

1. Introduction

Mixed Gaussian distribution has been an attractive field of research for many decades, as it represents random variables used in many fields of science, e.g. in signal processing [1], machine learning [2], gas mixture in chemistry, thermodynamics [3], economy, and biology.

In wireless communication systems, mixed Gaussian distribution is used to represent the output of many stochastic information sources, such as AWGN channels and index modulation schemes [5].

Mixed Gaussian distribution is frequently utilized as a noise model in various signal processing applications. This specific noise model is employed to describe different communication problems, such as co-channel interference, in which thermal noise with a Gaussian distribution is combined with artificial "clutter", e.g. signals from communication systems [6].

Mixed Gaussian distribution results from mixing both continuous and discrete random variables. The mixing can be performed using the product or summation of two random variables. The resultant mixture has a number of Gaussian modes that differ in terms of model parameters, such as mean, variance, and weights [7].

The number of modes depends on the number of values the discrete random variable can assume. For example, bimodal Gaussian distribution is generated from a discrete random variable with two values and becomes symmetric when the two modes have the same variance and weight, but their means are symmetric [7], [8]. A special case of mixed Gaussian distribution occurs when the discrete random variable is uniform with two values. Then the resultant mixture will have symmetric bimodal Gaussian distribution in which the two Gaussian components have symmetric means and are identical in terms of weights and variance.

The differential entropy of mixed Gaussian distribution has gained its significance from the importance of Gaussian mixture distribution. In communication systems, it is necessary to calculate the differential entropy in order to find the achievable rate and the rate-distortion for a specific coding scheme [9].

An analytic form of the differential entropy of mixed Gaussian distribution is not available [10], and the integration of $\log(\cosh(z))$ is not possible as shown in Appendix B.

Instead of relying on the analytic form, the differential entropy of a Gaussian mixture may be either estimated [12] or calculated using an approximation based on numerical calculations [10]. Meanwhile, different researchers are using bounds on the differential entropy [13] by setting bounds on $\log(\cosh(z))$. In [4], the authors introduced the bounds on the differential entropy using the Taylor series. However, the introduced lower bound is not tight to the introduced upper bound. In this paper, we combine both series expansion, and a tighter upper bound on the $\log(\cosh(z))$ than the one used in [13] to present a tighter lower bound on the differential entropy for symmetric bimodal Gaussian distribution.

General bounds on the differential entropy (based on different principles) provide basic measures which can be used to compare the results. These bounds are either of the upper variety, such as the maximum entropy upper bound (MEUB) based on the principle of maximum entropy [9] and the separation of upper bound components [10], or of the lower bound variety, based on the concavity of the differential entropy [9]. However, these bounds are very lossy and do not provide an accurate measure of the compared bounds. However, a comparison of the results with the bounds on $\log(\cosh(z))$ provides more accurate results.

In [19], the authors used the principle of entropy concavity deficit to introduce bounds on the differential entropy of a Gaussian mixture. This principle uses the difference between the differential entropy of the mixture and the weighted sum of differential entropies of its constituent components. Meanwhile, the tightness of the resultant bounds is conditioned on variance σ of the Gaussian components of the mixture,

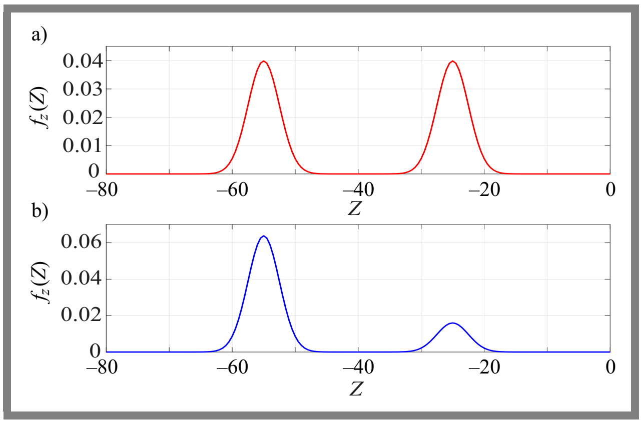

Fig. 1. Symmetric bimodal Gaussian distribution a) and asymmetric bimodal Gaussian distribution b).

such that the bounds lose their tightness for $\sigma^2 < 0.5$, and the error is growing with σ.

In this paper, we introduce a tight lower bound on the differential entropy of bimodal Gaussian distribution and compare the results with [4], [13], by combining the two techniques, using a more tight upper bound on the $\log(\cosh(z))$ than the ones used in [13], as well as using series expansion.

2. Preliminary Requirements and Setup

In this section, we provide the preliminary requirements and the probability setup needed to characterize the proposed model.

2.1. Differential Entropy of a Continuous RV

For a continuous random variable C with a probability density function (PDF) $f(C)$, the differential entropy of C is the expectation of the logarithmic function of its PDF, defined as [9]:

$$\begin{aligned} h(C) &= -E\big[\log(f(C))\big] \\ &= -\int_{-\infty}^{\infty} f(C)\log\left(f(C)\right) dc \ . \end{aligned} \tag{1}$$

2.2. Probability Analysis for Mixture Distribution

For the measure space $(\Omega, \mathcal{F}, \mathbb{P})$, with probability measure $\mathbb{P}$. Let $(\mathbb{R}, B_{\mathbb{R}})$ represents the measurable space on $\mathbb{R}$ with the Boral σ-algebra. The subset $\{d_1, d_2, \ldots\}$ of $\mathbb{R}$ is countable when D is an absolutely discrete random variable (RV) with probability mass function (PMF) $p_i = \mathbb{P}(D = d_i)$, such that $\sum_i p_i = 1$, and an induced probability measure μ_D on $(\mathbb{R}, B_{\mathbb{R}})$.

For the continuous RV C, the probability measure μ_C, induced on $(\mathbb{R}, B_{\mathbb{R}})$, is absolutely continuous with respect to the Lebesgue measure. This probability measure is characterized by the PDF $f(C)$, where $\int_{\mathbb{R}} f(C)\, dc = 1$.

The mixture RV $Z = C \times D$ formed by mixing arbitrary discrete and continuous RVs will have the PDF in the form of:

$$f(Z) = \sum_i p_i f_i(C) \ . \tag{2}$$

For a continuous RV with a Gaussian distribution $f(C) \sim \mathcal{N}(\mu_c, \sigma_c^2)$, the mixture RV is a multimodal Gaussian with PDF:

$$f(Z) = \sum_i p_i \mathcal{N}(\mu_i, \sigma_i^2) \ , \tag{3}$$

where $\mu_i = d_i \mu_c$ and $\sigma_i^2 = d_i^2 \sigma_c^2$.

The number of Gaussian modes of the mixture distribution depends on the number of values that the discrete RV may take, such that a bimodal Gaussian RV results when the discrete RV $D \in \{d_1, d_2\}$.

A special case of the bimodal Gaussian distribution occurs when the discrete random variable is uniform with two values, such that $P(D = d_1) = P(D = d_2) = 0.5$, $d_1 = -d_2$. Then, the resultant mixture distribution is symmetric bimodal Gaussian, such that Eq. (3) becomes:

$$\begin{aligned} f_z(Z) &= 0.5\,\mathcal{N}(\mu, \sigma^2) + 0.5\,\mathcal{N}(-\mu, \sigma^2) \\ &= \mathcal{N}(\mu, \sigma^2) e^{\frac{-\mu z}{\sigma^2}} \cosh(\frac{\mu z}{\sigma^2}) \end{aligned} \ , \tag{4}$$

in which the two Gaussian components have symmetric means and are identical in weights and variance, as shown in Fig. 1a.

When the discrete RV is not uniform, such that $P(D = d_1) \neq P(D = d_2)$, the two Gaussian components differ only in terms of their weights, as shown in Fig. 1b, and the distribution becomes weighted symmetric bimodal Gaussian as:

$$\begin{aligned} f_z(z) &= p_1 \mathcal{N}(\mu, \sigma^2) + p_2 \mathcal{N}(-\mu, \sigma^2) \\ &= \frac{2}{(e^b + e^{-b})} \mathcal{N}(\mu, \sigma^2) e^{\frac{-\mu z}{\sigma^2}} \cosh(\frac{\mu z}{\sigma^2} + b) \\ &= \frac{e^b}{e^b + e^{-b}} \mathcal{N}(\mu, \sigma^2) + \frac{e^{-b}}{e^b + e^{-b}} \mathcal{N}(-\mu, \sigma^2) \end{aligned} \ , \tag{5}$$

where $b = 0.5 \log(\frac{p_1}{p_2})$. A derivation of the second form of Eq. (5) is provided in Appendix A.

3. Tight Lower Bound on the Differential Entropy of Symmetric Bimodal Gaussian Distribution

Theorem 1. A tight lower bound on the differential entropy for the symmetric bimodal Gaussian distribution is:

$$\begin{aligned} h(Z) \geqslant\ & h(C) + \alpha^2 + \frac{2}{3} - \frac{6\alpha}{\sqrt{2\pi}} e^{\frac{-\alpha^2}{2}} - \alpha^2 \operatorname{erf}(\frac{\alpha}{\sqrt{2}}) \\ & - 2\beta \sum_{k=1}^{N} (\frac{M_{k1} V_{k1}}{2} \operatorname{erfc}(\frac{-V_{k1}}{\sigma\sqrt{2}}) + \frac{M_{k2} V_{k2}}{2} \operatorname{erfc}(\frac{-V_{k2}}{\sigma\sqrt{2}})) \ , \end{aligned} \tag{6}$$

where $\alpha = \frac{\mu}{\sigma}$, $\beta = \frac{\alpha}{\sigma}$,

$$M_{k1} = e^{\frac{-(6k\alpha^2 - 9k^2\alpha^2)}{2}} , \quad M_{k2} = e^{\frac{6k\alpha^2 + 9k^2\alpha^2}{2\sigma^2}}$$
$$V_{k1} = \mu - 3k\mu \ , \qquad V_{k2} = -\mu - 3k\mu \ .$$

The proof is outlined in Appendix B. Below, we provide a description of the main axioms on which our proof is based. We shall start by applying the differential entropy definition

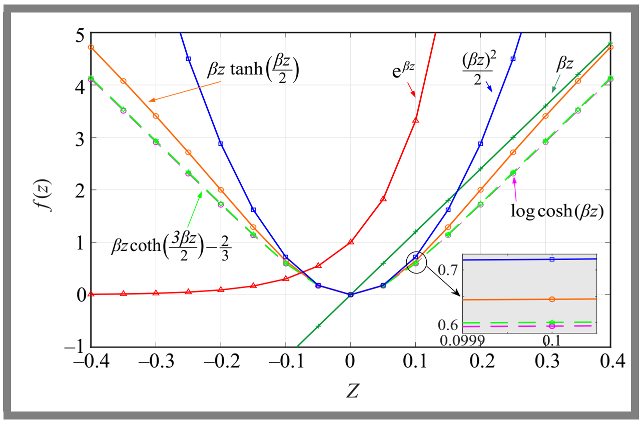

Fig. 2. Comparison of the upper bound on $\log(\cosh(\beta z))$ for different inequalities.

from Eq. (1) to the symmetric bimodal Gaussian distribution given by Eq. (4). The resulting integration of $\log\cosh(\beta z)$ in Eq. (16) and Eq. (18) cannot be solved in a closed form. Instead, many inequalities are used to set lower bounds on $\log(\cosh(\beta z))$, such as $e^{\beta z}$, $\frac{\beta z^2}{2}$ and $z\,\mathrm{tgh}\left(\frac{\beta z}{2}\right)$ [14].

In order to calculate the differential entropy of bimodal Gaussian distribution, the authors in [13] used βz as an upper bound on $\log(\cosh(\beta z))$. Meanwhile, the relationship between $\log\cosh(\beta z)$ and ctgh in [15] shows more tightness as an upper bound such that:

$$\log(\cosh(\beta z)) \leqslant \beta z\,\mathrm{ctgh}\left(\frac{3\beta z}{2}\right) - \frac{2}{3}\,. \tag{7}$$

Therefore, we use Eq. (7) as an upper bound on $\log(\cosh(\beta z))$ due to its tightness, which leads to a tight lower bound on the differential entropy. Moving forward, we use series expansion from Eq. (24) to expand $\mathrm{ctgh}(\frac{3\beta z}{2})$ [16] in order to avoid difficulties in solving the integration in a closed form.

A comparison between different upper bounds on $\log(\cosh(\beta z))$ is shown in Fig. 2, with the zoom-in portion indicating that $\beta z\,\mathrm{ctgh}(\frac{3\beta z}{2}) - \frac{2}{3}$ is the most tight upper bound over the entire range of model parameter Z.

4. Tight Lower Bound on the Differential Entropy of Weighted Symmetric Bimodal Gaussian Distribution

Representing bimodal Gaussian distribution with the hyperbolic cosine, Eq. (5) offers the advantage of controlling the weights of the two Gaussian modes by shifting the hyperbolic cosine on the x-axis which, in turn, changes the weights of the two modes, thus resulting in weighted symmetric bimodal Gaussian distribution. Meanwhile, the use of the symmetric Hyperbolic cosine (shift $= 0$) results in the symmetric bimodal Eq. (4). Therefore, in this section, we use the advantages of the hyperbolic cosine in order to present a tight lower bound on the differential entropy of weighted symmetric bimodal Gaussian distribution.

Theorem 2. For the weighted symmetric bimodal Gaussian distribution, the lower bound on the differential entropy is:

$$\begin{aligned} h(Z) \geqslant (A+B)h(C) + (A+B)\frac{\mu^2}{\sigma^2} + \frac{2(A+B)}{3} \\ + 2(A-B)\mathrm{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) - 2(A\boldsymbol{I}_{12}^{b} + B\boldsymbol{I}_{22}^{b}) \end{aligned}\,, \tag{8}$$

where b is the x-axis shift of the hyperbolic cosine used in modeling weighted symmetric bimodal Gaussian distribution and:

$$A = \frac{e^{b}}{e^{b}+e^{-b}} \text{ and } B = \frac{e^{-b}}{e^{b}+e^{-b}}\,.$$

Both $\boldsymbol{I}_{12}^{b}$ and $\boldsymbol{I}_{22}^{b}$ are defined in Appendix C.

The proof is outlined in Appendix C with an analogy to the proof of Theorem 1.

By calculating the differential entropy based on its basic definition, we substitute the PDF of the weighted bimodal Gaussian Eq. (5) in Eq. (1), and then apply the tight upper bound of the shifted version of Eq. (7):

$$\log\left(\cosh(\beta(z+b))\right) \leqslant \beta(z+b)\,\mathrm{ctgh}\left(\frac{3\beta(z+b)}{2}\right) - \frac{2}{3} \tag{9}$$

on $\log(\cosh(\beta z + b))$, which leads to a tight lower bound on the differential entropy of the proposed mixture model.

Then, we expand $\mathrm{ctgh}(\frac{3\beta(z+b)}{2})$ using the series expansion technique, such that:

$$\mathrm{ctgh}\left(\frac{3\beta z}{2}\right) = 1 + 2\sum_{k=1}^{N} e^{-2k\frac{3\beta z}{2}}\,. \tag{10}$$

In order to get a closed-form solution for the presented lower bound, the exponential term in Eq. (10) inside the integration Eq. (29) is used to form a Gaussian function whose integration can be found using the error function.

5. Simulation Results

In this section, we provide through simulations comparing our lower bound on the differential entropy of mixed Gaussian distribution and different lower bounds. The simulations use different model parameters, such as variance and mean of the Gaussian modes and series approximation order $N = 3$.

Many authors, for instance in [4], compare their results with general bounds that rely on different principles providing basic measures, such as the maximum entropy upper bound (MEUB) based on the principle of maximum entropy [9], separation of components upper bound (SCUB) [10], and concavity of differential entropy [9], as a lower bound. These bounds, however, are very lossy and do not provide an accurate measure of the compared bounds.

Meanwhile, bounds on the differential entropy that are derived from the bounds on $\log(\cosh(z))$ are tighter than general bounds, such as MEUB and SCUB, which provide an accurate and clear comparison. Therefore, we use the upper bound obtained from inequality [13]:

$$\log(\cosh(\beta z)) \geqslant \beta z - \log(2)\,, \tag{11}$$

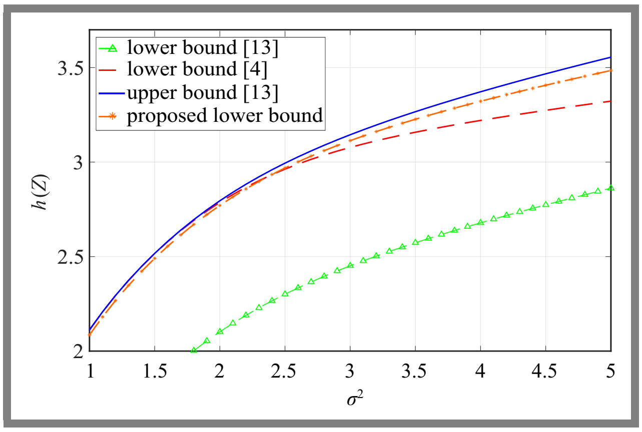

Fig. 3. Plots of different lower bounds on the differential entropy for symmetric bimodal Gaussian distribution in terms of σ^2.

instead of SCUB and MEUB to compare the different lower bounds, because of its tightness.

In Fig. 3, the proposed lower bound on the differential entropy of symmetric bimodal Gaussian distribution is compared with the lower bounds presented in [13] and [4] and the upper bound in [13] is set as an upper rather than SCUB and MEUB limit due to its tightness.

The comparison shows that the proposed lower bounds achieve more tightness over the entire range of variance than the lower bound proposed in [13], as they rely on the inequality given by Eq. (11) instead of βz. Therefore, more tightness is achieved over a high degree of variance than in the case of the lower bound presented in [4].

Figure 4 shows the effect of changing the mean of the Gaussian modes on the differential entropy of symmetric bimodal Gaussian distribution. The proposed lower bound is tighter to the upper bound than the solution proposed in [13], with a large gain over the low range of the model parameter μ compared to the lower bound presented in [4]. In addition to achieving more tightness over the entire range of the same model parameter is used rather than in [13]. Furthermore, a SCUB simulation shows that the use of a general bound is very lossy, which does not facilitate an accurate comparison with the lower bounds. The use of the upper bound of [13], meanwhile, provides an accurate and clear measure for a comparison between different lower bounds.

6. Application

In this section, we apply the bounds proposed in Theorems 1 and 2 on the differential entropy of the channel output to study channel capacity with a peak power constraint. For an AWGN channel, the output Y is:

$$Y = X + Z ,$$

where X is the channel input signal with a finite number of mass points constrained in the interval $[\sqrt{\rho} : \sqrt{\rho}]$ and Z is the additive Gaussian noise, such that $Z \sim \mathcal{N}(0, \sigma_n^2)$. Both X and Z are independent variables.

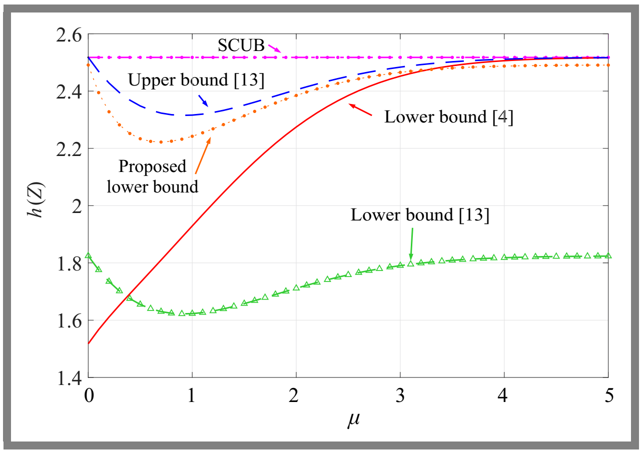

Fig. 4. Plots of different bounds on the differential entropy for symmetric bimodal Gaussian distribution in terms of μ.

In this case, the channel is constrained to peak power ρ and the corresponding constrained channel capacity is:

$$C_\rho(\rho) \triangleq \underset{X\in[\sqrt{\rho}:\sqrt{\rho}]}{SUP} I(X; X + Z) \; . \tag{12}$$

In [17], it has been proven that for an amplitude-constrained signal with a maximum amplitude of $\sqrt{\rho} \leqslant 1.05$, channel capacity is achievable when X is equiprobable with $\{-\sqrt{\rho} : \sqrt{\rho}\}$ values. In this case, channel output Y has a bimodal Gaussian distribution.

We conversely verify the constrained channel capacity presented in [17] by showing that this capacity is achieved when the distribution of channel output Y is of symmetric bimodal Gaussian distribution that results from an equiprobable input signal X with PMF $p(X = -\sqrt{\rho}) = p(X = \sqrt{\rho}) = 0.5$.

The variations in the achievable rate $I(X; X+Z)$ are studied according to the changes in the weights of the two Gaussian modes of the channel output distribution. The weights of the two modes follow the probability of the two values of input signal X.

Then, the lower bound from Theorem 2 is applied on the differential entropy of the channel output and the value of parameter b in Eq. (9) controlling the weights of the Gaussian modes 13, is varied. We show via the simulation presented in Fig. 5, that the maximum achievable rate occurs at $b = 0$, which results in equiprobable modes. Additionally, Theorem 2 converges to Theorem 1.

7. Conclusion

By relying on the advantage of representing bimodal Gaussian distribution using the hyperbolic cosine, we proposed tighter lower bounds on the differential entropy of bimodal Gaussian distribution, rather than the previously used lower bounds in both cases (symmetric/weighted symmetric) using a tight upper bound on the $\log(\cosh(z))$. The proposed lower bound shows more tightness over the entire range of model parameters, with more gain in the lower range of the mean of the Gaussian modes.

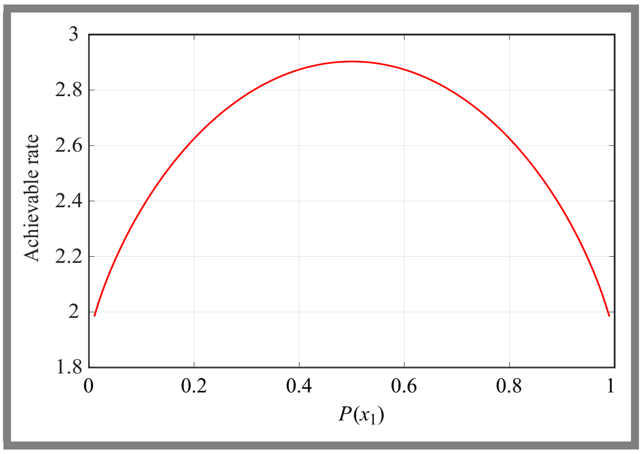

Fig. 5. Simulation of the achievable constrained channel rate with the corresponding input probability.

Appendix A: Weighted symmetric bimodal Gaussian Distribution

$$
\begin{aligned}
f_z(z) &= \frac{2}{e^b+e^{-b}}\mathcal{N}(\mu,\sigma^2)e^{\frac{-\mu z}{\sigma^2}}\cosh\Big(\frac{\mu z}{\sigma^2}+b\Big) \\
&= \frac{2}{e^b+e^{-b}}\mathcal{N}(-\mu,\sigma^2)e^{\frac{\mu z}{\sigma^2}}\cosh\Big(\frac{\mu z}{\sigma^2}+b\Big) \\
&= \frac{e^b}{e^b+e^{-b}}\mathcal{N}(\mu,\sigma^2)2e^{-(\frac{\mu z}{\sigma^2}+b)}\cosh\Big(\frac{\mu z}{\sigma^2}+b\Big) \\
&= \frac{e^b}{e^b+e^{-b}}\mathcal{N}(\mu,\sigma^2)\Big[1+e^{-2(\frac{\mu z}{\sigma^2}+b)}\Big] \\
&= \frac{e^b}{e^b+e^{-b}}\mathcal{N}(\mu,\sigma^2)\Big[1+e^{-2b}e^{\frac{-(z+\mu)^2+(z-\mu)^2}{2\sigma^2}}\Big] \\
&= \frac{1}{e^b+e^{-b}}\mathcal{N}(\mu,\sigma^2)e^{\frac{-\mu z}{\sigma^2}}\Big[e^{\frac{\mu z}{\sigma^2}+b}+e^{\frac{-\mu z}{\sigma^2}-b}\Big] \\
&= \frac{1}{e^b+e^{-b}}\Big[e^b\mathcal{N}(\mu,\sigma^2)+e^{-b}\mathcal{N}(-\mu,\sigma^2)\Big] \\
&= \frac{e^b}{e^b+e^{-b}}\mathcal{N}(\mu,\sigma^2)+\frac{e^{-b}}{e^b+e^{-b}}\mathcal{N}(-\mu,\sigma^2)
\end{aligned}
\tag{13}
$$

Appendix B: Proof of Theorem 1

The differential entropy of the bimodal Gaussian RV Z is:

$$
\begin{aligned}
h(Z) &= -0.5\int_{-\infty}^{\infty}\big[\mathcal{N}(\mu,\sigma^2)+\mathcal{N}(-\mu,\sigma^2)\big]\log\big(f_z(Z)\big)dz \\
&= -0.5\int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\big(f_z(Z)\big)dz \\
&\quad -0.5\int_{-\infty}^{\infty}\mathcal{N}(-\mu,\sigma^2)\log\big(f_z(Z)\big)dz \\
&= -0.5\big(\boldsymbol{I}_1+\boldsymbol{I}_2\big)
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
\boldsymbol{I}_1 &= \int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\Big[\mathcal{N}(\mu,\sigma^2)e^{\frac{-\mu z}{\sigma^2}}\cosh\Big(\frac{\mu z}{\sigma^2}\Big)\Big]dz \\
&= \int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\mathcal{N}(\mu,\sigma^2)dz \\
&\quad -\frac{\mu}{\sigma^2}\int_{-\infty}^{\infty}z\mathcal{N}(\mu,\sigma^2)dz \\
&\quad +\int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\Big(\frac{\mu z}{\sigma^2}\Big)dz \\
&= \log\frac{1}{\sigma\sqrt{2\pi}}-\frac{1}{2\sigma^2}\int_{-\infty}^{\infty}z^2\mathcal{N}(\mu,\sigma^2)dz \\
&\quad -\frac{\mu^2}{2\sigma^2}+\int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\Big(\frac{\mu z}{\sigma^2}\Big)dz \\
&= \log\frac{1}{\sigma\sqrt{2\pi}}-\frac{1}{2\sigma^2}(\sigma^2+\mu^2)-\frac{\mu^2}{2\sigma^2} \\
&\quad +\int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\Big(\frac{\mu z}{\sigma^2}\Big)dz \\
&= \log\frac{1}{\sigma\sqrt{2\pi}}-\frac{1}{2}-\frac{\mu^2}{\sigma^2} \\
&\quad +\int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\Big(\frac{\mu z}{\sigma^2}\Big)dz \\
&= -0.5\log(2\pi\sigma^2)-0.5-\frac{\mu^2}{\sigma^2} \\
&\quad +\int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\Big(\frac{\mu z}{\sigma^2}\Big)dz \\
&= -0.5\log(2\pi e\sigma^2)-\frac{\mu^2}{\sigma^2} \\
&\quad +2\int_{0}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\Big(\frac{\mu z}{\sigma^2}\Big)dz \\
&= -h(C)-\frac{\mu^2}{\sigma^2}+2\boldsymbol{I}_{11}
\end{aligned}
\tag{15}
$$

$$
\boldsymbol{I}_{11} = \int_{0}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\Big(\frac{\mu z}{\sigma^2}\Big)dz \tag{16}
$$

Similarly:

$$
\boldsymbol{I}_2 = -h(C)-\frac{\mu^2}{\sigma^2}+2\boldsymbol{I}_{21} \tag{17}
$$

$$
\boldsymbol{I}_{21} = \int_{0}^{\infty}\mathcal{N}(-\mu,\sigma^2)\log\cosh\Big(\frac{\mu z}{\sigma^2}\Big)dz \tag{18}
$$

Then:

$$
\begin{aligned}
h(Z) &= -0.5\Big(-2h(C)-2\frac{\mu^2}{\sigma^2}+2\boldsymbol{I}_{11}+2\boldsymbol{I}_{21}\Big) \\
&= h(C)+\frac{\mu^2}{\sigma^2}-\big(\boldsymbol{I}_{11}+\boldsymbol{I}_{21}\big)
\end{aligned}
\tag{19}
$$

Applying the inequality Eq. (11):

$$
\begin{aligned}
\boldsymbol{I}_{11} &= \int_{0}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\Big(\cosh\big(\beta z\big)\Big)dz \\
&\leqslant \int_{0}^{\infty}\mathcal{N}(\mu,\sigma^2)\beta z\,\mathrm{ctgh}\Big(\frac{3\beta z}{2}\Big)dz \\
&\quad -\int_{0}^{\infty}\frac{2}{3}\mathcal{N}(\mu,\sigma^2)dz \\
&= \boldsymbol{I}_{12}-\frac{1}{3}\Big(1+\mathrm{erf}\Big(\frac{\mu}{\sigma\sqrt{2}}\Big)\Big)
\end{aligned}
\tag{20}
$$

Where erf is the error function [18],

$$\begin{aligned} \boldsymbol{I}_{21} &\leqslant \int_0^\infty \mathcal{N}(-\mu,\sigma^2)\beta z \,\mathrm{ctgh}\,\big(\frac{3\beta z}{2}\big)dz \\ z &- \int_0^\infty \frac{2}{3}\mathcal{N}(-\mu,\sigma^2)dz \\ &= \boldsymbol{I}_{22} - \frac{1}{3}\big(1-\mathrm{erf}(\frac{\mu}{\sigma\sqrt{2}})\big) \end{aligned} \tag{21}$$

$$\begin{aligned} \boldsymbol{I}_{11}+\boldsymbol{I}_{21} &= \boldsymbol{I}_{12}+\boldsymbol{I}_{22} - \frac{1}{3}\big(1+\mathrm{erf}(\frac{\mu}{\sigma\sqrt{2}})\big) \\ &- \frac{1}{3}\big(1-\mathrm{erf}(\frac{\mu}{\sigma\sqrt{2}})\big) \\ &= \boldsymbol{I}_{12}+\boldsymbol{I}_{22} - \frac{1}{3} - \frac{1}{3}\mathrm{erf}\big(\frac{\mu}{\sigma\sqrt{2}}\big) - \frac{1}{3} \\ &+ \frac{1}{3}\mathrm{erf}\big(\frac{\mu}{\sigma\sqrt{2}}\big) \\ &= \boldsymbol{I}_{12}+\boldsymbol{I}_{22} - \frac{2}{3} \end{aligned} \tag{22}$$

$$h(Z) \geqslant h(C) + \frac{\mu^2}{\sigma^2} + \frac{2}{3} - \big(\boldsymbol{I}_{12}+\boldsymbol{I}_{22}\big) \tag{23}$$

Using series expansion [16]:

$$\mathrm{ctgh}\,\big(\frac{3\beta z}{2}\big) = 1 + 2\sum_{k=1}^{N} e^{-2k\frac{3\beta z}{2}} \tag{24}$$

Where N is the order approximation of the series expansion.

$$\begin{aligned} \boldsymbol{I}_{12} &= \int_0^\infty \mathcal{N}(\mu,\sigma^2)\beta z \,\mathrm{ctgh}\,\big(\frac{3\beta z}{2}\big)dz \\ &= \int_0^\infty \mathcal{N}(\mu,\sigma^2)\beta z\big(1+2\sum_{k=1}^{N} e^{-2k\frac{3\beta z}{2}}\big)dz \\ &= \beta\int_0^\infty z\mathcal{N}(\mu,\sigma^2)dz + 2\beta\sum_{k=1}^{N}\int_0^\infty z\mathcal{N}(\mu,\sigma^2)e^{-3k\beta z}dz \\ &= \beta\frac{\sigma}{\sqrt{2\pi}}e^{\frac{-\mu^2}{2\sigma^2}} + \frac{\beta\mu}{2}\mathrm{erfc}\big(\frac{-\mu}{\sigma\sqrt{2}}\big) + 2\beta\sum_{k=1}^{N}\boldsymbol{I}_{13} \\ &= \frac{\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}} + \frac{\alpha^2}{2}\mathrm{erfc}\big(\frac{-\alpha}{\sqrt{2}}\big) + 2\beta\sum_{k=1}^{N}\boldsymbol{I}_{13} \end{aligned} \tag{25}$$

$$\begin{aligned} \boldsymbol{I}_{22} &= \int_0^\infty \mathcal{N}(-\mu,\sigma^2)\beta z \,\mathrm{ctgh}\,\big(\frac{3\beta z}{2}\big)dz \\ &= \beta\int_0^\infty z\mathcal{N}(-\mu,\sigma^2)dz \\ &+ 2\beta\sum_{k=1}^{N}\int_0^\infty z\mathcal{N}(-\mu,\sigma^2)e^{-3k\beta z}dz \\ &= \frac{\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}} - \frac{\alpha^2}{2}\mathrm{erfc}\big(\frac{\alpha}{\sqrt{2}}\big) + 2\beta\sum_{k=1}^{N}\boldsymbol{I}_{23} \end{aligned} \tag{26}$$

$$\begin{aligned} \boldsymbol{I}_{12}+\boldsymbol{I}_{22} &= \frac{\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}} + \frac{\alpha^2}{2}\mathrm{erfc}\big(\frac{-\alpha}{\sqrt{2}}\big) + 2\beta\sum_{k=1}^{N}\boldsymbol{I}_{13} \\ &+ \frac{\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}} - \frac{\alpha^2}{2}\mathrm{erfc}\big(\frac{\alpha}{\sqrt{2}}\big) + 2\beta\sum_{k=1}^{N}\boldsymbol{I}_{23} \\ &= \frac{2\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}} - \alpha^2\big(1-\mathrm{erfc}(\frac{-\alpha}{\sqrt{2}})\big) \\ &+ 2\beta\sum_{k=1}^{N}\big(\boldsymbol{I}_{13}+\boldsymbol{I}_{23}\big) \\ &= \frac{2\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}} + \alpha^2\mathrm{erf}\big(\frac{\alpha}{\sqrt{2}}\big) + 2\beta\sum_{k=1}^{N}\big(\boldsymbol{I}_{13}+\boldsymbol{I}_{23}\big) \end{aligned} \tag{27}$$

$$\begin{aligned} h(Z) &\geqslant h(C) + \alpha^2 + \frac{2}{3} - \frac{2\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}} - \alpha^2\mathrm{erf}\big(\frac{\alpha}{\sqrt{2}}\big) \\ &- 2\beta\sum_{k=1}^{N}\big(\boldsymbol{I}_{13}+\boldsymbol{I}_{23}\big) \end{aligned} \tag{28}$$

$$\begin{aligned} \boldsymbol{I}_{13} &= \int_0^\infty z\mathcal{N}(\mu,\sigma^2)e^{-3k\beta z}dz \\ &= \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty z e^{\frac{-(z-\mu)^2}{2\sigma^2}}e^{-3k\beta z}dz \\ &= e^{\frac{-(6\mu k\sigma^2\beta - 9k^2\sigma^4\beta^2)}{2\sigma^2}}\int_0^\infty z\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(z-(\mu-3k\sigma^2\beta))^2}{2\sigma^2}}dz \\ &= e^{\frac{-(6\mu k\sigma^2\beta - 9k^2\sigma^4\beta^2)}{2\sigma^2}}\int_0^\infty z\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(z-V_{k1})^2}{2\sigma^2}}dz \\ &= M_{k1}\int_0^\infty z\mathcal{N}(V_{k1},\sigma^2)dz \\ &= M_{k1}\big(\frac{\sigma}{\sqrt{2\pi}}e^{\frac{-V_{k1}^2}{2\sigma^2}} + \frac{V_{k1}}{2}\mathrm{erfc}(\frac{-V_{k1}}{\sigma\sqrt{2}})\big) \\ &= \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\alpha^2}{2}} + \frac{M_{k1}V_{k1}}{2}\mathrm{erfc}\big(\frac{-V_{k1}}{\sigma\sqrt{2}}\big) \end{aligned} \tag{29}$$

$$\begin{aligned} \boldsymbol{I}_{23} &= \int_0^\infty z\mathcal{N}(-\mu,\sigma^2)e^{-3k\beta z}dz \\ &= M_{k2}\int_0^\infty z\mathcal{N}(V_{k2},\sigma^2)dz \\ &= \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\alpha^2}{2}} + \frac{M_{k2}V_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big) \end{aligned} \tag{30}$$

$$\begin{aligned} \boldsymbol{I}_{13}+\boldsymbol{I}_{23} &= \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\alpha^2}{2}} + \frac{M_{k1}V_{k1}}{2}\mathrm{erfc}\big(\frac{-V_{k1}}{\sigma\sqrt{2}}\big) \\ &+ \frac{\sigma}{\sqrt{2\pi}}e^{-\frac{\alpha^2}{2}} + \frac{M_{k2}V_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big) \\ &= \frac{M_{k1}V_{k1}}{2}\mathrm{erfc}\big(\frac{-V_{k1}}{\sigma\sqrt{2}}\big) + \frac{M_{k2}V_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big) \\ &+ \frac{2\sigma}{\sqrt{2\pi}}e^{-\frac{\alpha^2}{2}} \end{aligned} \tag{31}$$

Then, the lower bound on the differential entropy of the symmetric bimodal Gaussian is:

$$\begin{aligned} h(Z) &\geqslant h(C) + \alpha^2 + \frac{2}{3} - \frac{2\alpha}{\sqrt{2\pi}} e^{\frac{-\alpha^2}{2}} - \alpha^2 \mathrm{erf}\big(\frac{\alpha}{\sqrt{2}}\big) \\ &- 2\beta \sum_{k=1}^{N} \big(\frac{M_{k1}V_{k1}}{2}\mathrm{erfc}(\frac{-V_{k1}}{\sigma\sqrt{2}})\big) + \frac{M_{k2}V_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big) \\ &- 2\beta \frac{2\sigma}{\sqrt{2\pi}} e^{-\frac{\alpha^2}{2}} \\ &= h(C) + \alpha^2 + \frac{2}{3} - \frac{2\alpha}{\sqrt{2\pi}} e^{\frac{-\alpha^2}{2}} - \alpha^2 \mathrm{erf}\big(\frac{\alpha}{\sqrt{2}}\big) \\ &- 2\beta \sum_{k=1}^{N} \big(\frac{M_{k1}V_{k1}}{2}\mathrm{erfc}(\frac{-V_{k1}}{\sigma\sqrt{2}})\big) + \frac{M_{k2}V_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big) \\ &- \frac{4\alpha}{\sqrt{2\pi}} e^{-\frac{\alpha^2}{2}} \\ &= h(C) + \alpha^2 + \frac{2}{3} - \frac{6\alpha}{\sqrt{2\pi}} e^{\frac{-\alpha^2}{2}} - \alpha^2 \mathrm{erf}\big(\frac{\alpha}{\sqrt{2}}\big) \\ &- 2\beta \sum_{k=1}^{N} \big(\frac{M_{k1}V_{k1}}{2}\mathrm{erfc}(\frac{-V_{k1}}{\sigma\sqrt{2}})\big) + \frac{M_{k2}V_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big) \end{aligned} \tag{32}$$

Appendix C: Proof of Theorem 2

$$\begin{aligned} h(Z) &= -\int_{-\infty}^{\infty} f_z(Z)\log(f_z(Z))dz \\ &= -\frac{e^b}{e^b+e^{-b}}\int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log(f_z(Z))dz \\ &- \frac{e^{-b}}{e^b+e^{-b}}\int_{-\infty}^{\infty}\mathcal{N}(-\mu,\sigma^2)\log(f_z(Z))dz \\ &= -\big(\frac{e^b}{e^b+e^{-b}}\big)\log\big(\frac{2}{e^b+e^{-b}}\big) - \big(\frac{e^b}{e^b+e^{-b}}\big)\boldsymbol{I}_1^b \\ &- \big(\frac{e^{-b}}{e^b+e^{-b}}\big)\log\big(\frac{2}{e^b+e^{-b}}\big) - \big(\frac{e^{-b}}{e^b+e^{-b}}\big)\boldsymbol{I}_2^b \\ &= -A\log\big(\frac{2}{e^b+e^{-b}}\big) - A\boldsymbol{I}_1^b - B\log\big(\frac{2}{e^b+e^{-b}}\big) - B\boldsymbol{I}_2^b \\ &= -\log\big(\frac{2}{e^b+e^{-b}}\big)(A+B) - A\boldsymbol{I}_1^b - B\boldsymbol{I}_2^b \end{aligned} \tag{33}$$

$$\begin{aligned} \boldsymbol{I}_1^b &= \int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\Big[\mathcal{N}(\mu,\sigma^2)e^{\frac{-\mu z}{\sigma^2}}\cosh\big(\frac{\mu z}{\sigma^2}+b\big)\Big]dz \\ &= \int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\mathcal{N}(\mu,\sigma^2)dz - \frac{\mu}{\sigma^2}\int_{-\infty}^{\infty}z\mathcal{N}(\mu,\sigma^2)dz \\ &+ \int_{-\infty}^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\big(\frac{\mu z}{\sigma^2}+b\big)dz \\ &= -0.5\log(2\pi e\sigma^2) - \frac{\mu^2}{\sigma^2} \\ &+ 2\int_0^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\big(\frac{\mu z}{\sigma^2}+b\big)dz \\ &= -h(C) - \frac{\mu^2}{\sigma^2} + 2\boldsymbol{I}_{11}^b \end{aligned} \tag{34}$$

$$\boldsymbol{I}_{11}^b = \int_0^{\infty}\mathcal{N}(\mu,\sigma^2)\log\cosh\big(\frac{\mu z}{\sigma^2}+b\big)dz \tag{35}$$

$$\begin{aligned} \boldsymbol{I}_2^b &= \int_{-\infty}^{\infty}\mathcal{N}(-\mu,\sigma^2)\log\Big[\mathcal{N}(-\mu,\sigma^2)e^{\frac{\mu z}{\sigma^2}}\cosh\big(\frac{\mu z}{\sigma^2}+b\big)\Big]dz \\ &= \int_{-\infty}^{\infty}\mathcal{N}(-\mu,\sigma^2)\log\mathcal{N}(-\mu,\sigma^2)dz \\ &+ \frac{\mu}{\sigma^2}\int_{-\infty}^{\infty}z\mathcal{N}(-\mu,\sigma^2)dz \\ &+ \int_{-\infty}^{\infty}\mathcal{N}(-\mu,\sigma^2)\log\cosh\big(\frac{\mu z}{\sigma^2}+b\big)dz \\ &= -0.5\log(2\pi e\sigma^2) - \frac{\mu^2}{\sigma^2} \\ &+ 2\int_0^{\infty}\mathcal{N}(-\mu,\sigma^2)\log\cosh\big(\frac{\mu z}{\sigma^2}+b\big)dz \\ &= -h(C) - \frac{\mu^2}{\sigma^2} + 2\boldsymbol{I}_{21}^b \end{aligned} \tag{36}$$

$$\boldsymbol{I}_{21}^b = \int_0^{\infty}\mathcal{N}(-\mu,\sigma^2)\log\cosh\big(\frac{\mu z}{\sigma^2}+b\big)dz \tag{37}$$

$$\begin{aligned} h(Z) &= -\log\big(\frac{2}{e^b+e^{-b}}\big)(A+B) + (A+B)h(C) \\ &+ (A+B)\frac{\mu^2}{\sigma^2} - 2\big(A\boldsymbol{I}_{11}^b + B\boldsymbol{I}_{21}^b\big) \\ &= (A+B)\big[-\log\big(\frac{2}{e^b+e^{-b}}\big) + h(C) + \frac{\mu^2}{\sigma^2}\big] \\ &- 2\big(A\boldsymbol{I}_{11}^b + B\boldsymbol{I}_{21}^b\big) \end{aligned} \tag{38}$$

$$\begin{aligned} \boldsymbol{I}_{11}^b &= \int_0^{\infty}\mathcal{N}(\mu,\sigma^2)\log(\cosh(\beta z+b))dz \\ &\leqslant \int_0^{\infty}\mathcal{N}(\mu,\sigma^2)\Big[(\beta z+b)\,\mathrm{ctgh}\,\big(\frac{3(\beta z+b)}{2}\big) - \frac{2}{3}\Big] \\ &= \int_0^{\infty}\mathcal{N}(\mu,\sigma^2)(\beta z+b)\,\mathrm{ctgh}\,\big(\frac{3(\beta z+b)}{2}\big)dz \\ &- \int_0^{\infty}\frac{2}{3}\mathcal{N}(\mu,\sigma^2)dz \\ &= \boldsymbol{I}_{12}^b - \frac{1}{3}\big(1+\mathrm{erf}(\frac{\mu}{\sigma\sqrt{2}})\big) \end{aligned} \tag{39}$$

$$\boldsymbol{I}_{12}^b = \int_0^{\infty}\mathcal{N}(\mu,\sigma^2)(\beta z+b)\,\mathrm{ctgh}\,\big(\frac{3(\beta z+b)}{2}\big)dz \tag{40}$$

$$\begin{aligned} \boldsymbol{I}_{21}^b &= \int_0^{\infty}\mathcal{N}(-\mu,\sigma^2)\log\big(\cosh(\beta z+b)\big)dz \\ &\leqslant \int_0^{\infty}\mathcal{N}(-\mu,\sigma^2)\Big[(\beta z+b)\,\mathrm{ctgh}\,\big(\frac{3(\beta z+b)}{2}\big) - \frac{2}{3}\Big] \\ &= \int_0^{\infty}\mathcal{N}(-\mu,\sigma^2)(\beta z+b)\,\mathrm{ctgh}\,\big(\frac{3(\beta z+b)}{2}\big)dz \\ &- \int_0^{\infty}\frac{2}{3}\mathcal{N}(-\mu,\sigma^2)dz \\ &= \boldsymbol{I}_{22}^b - \frac{1}{3}\big(1-\mathrm{erf}(\frac{\mu}{\sigma\sqrt{2}})\big) \end{aligned} \tag{41}$$

$$\boldsymbol{I}_{22}^b = \int_0^{\infty}\mathcal{N}(-\mu,\sigma^2)(\beta z+b)\,\mathrm{ctgh}\,\big(\frac{3(\beta z+b)}{2}\big)dz \tag{42}$$

$$\begin{aligned} A\boldsymbol{I}_{11}^b + B\boldsymbol{I}_{21}^b &= A\boldsymbol{I}_{12}^b + B\boldsymbol{I}_{22}^b - \frac{1}{3}(A+B) \\ &- (A-B)\mathrm{erf}\big(\frac{\mu}{\sigma\sqrt{2}}\big) \end{aligned} \tag{43}$$

$$h(Z) \geqslant (A+B)\Big[-\log\big(\frac{2}{e^b+e^{-b}}\big)+h(C)+\frac{\mu^2}{\sigma^2}+\frac{2}{3}\Big] + 2(A-B)\mathrm{erf}\big(\frac{\mu}{\sigma\sqrt{2}}\big)-2\big(A\boldsymbol{I}_{12}^b+B\boldsymbol{I}_{22}^b\big) \tag{44}$$

$$\mathrm{ctgh}\,\big(\frac{3(\beta z+b)}{2}\big)=1+2\sum_{k=1}^{N}e^{-3k(\beta z+b)} \tag{45}$$

$$\begin{aligned}\boldsymbol{I}_{12}^b &= \int_0^\infty \mathcal{N}(\mu,\sigma^2)(\beta z+b)\big(1+2\sum_{k=1}^{N}e^{-2k\frac{3(\beta z+b)}{2}}\big)dz\\ &= \beta\int_0^\infty z\mathcal{N}(\mu,\sigma^2)dz+b\int_0^\infty \mathcal{N}(\mu,\sigma^2)dz\\ &+2\beta\sum_{k=1}^{N}e^{-3kb}\int_0^\infty z\mathcal{N}(\mu,\sigma^2)e^{-3k\beta z}\\ &+2b\sum_{k=1}^{N}e^{-3kb}\int_0^\infty \mathcal{N}(\mu,\sigma^2)e^{-3k\beta z}dz\\ &+2\beta\sum_{k=1}^{N}e^{-3kb}\boldsymbol{I}_{13}^b+2b\sum_{k=1}^{N}e^{-3kb}\boldsymbol{I}_{14}^b\\ &=\beta\frac{\sigma}{\sqrt{2\pi}}e^{\frac{-\mu^2}{2\sigma^2}}+\frac{\beta\mu}{2}\mathrm{erfc}\big(\frac{-\mu}{\sigma\sqrt{2}}\big)+\frac{b}{2}\mathrm{erfc}\big(\frac{-\mu}{\sigma\sqrt{2}}\big)\\ &+2\beta\sum_{k=1}^{N}e^{-3kb}\boldsymbol{I}_{13}^b-2b\sum_{k=1}^{N}e^{-3kb}\boldsymbol{I}_{14}^b\\ &=\frac{\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}}+\big(\frac{\alpha^2}{2}+\frac{b}{2}\big)\mathrm{erfc}\big(\frac{-\alpha}{\sqrt{2}}\big)\\ &+2\beta\sum_{k=1}^{N}e^{-3kb}\boldsymbol{I}_{13}^b-2b\sum_{k=1}^{N}e^{-3kb}\boldsymbol{I}_{14}^b\end{aligned} \tag{46}$$

$$\begin{aligned}\boldsymbol{I}_{22}^b &= \beta\frac{\sigma}{\sqrt{2\pi}}e^{\frac{-\mu^2}{2\sigma^2}}-\frac{\beta\mu}{2}\mathrm{erfc}\big(\frac{\mu}{\sigma\sqrt{2}}\big)+b-\frac{b}{2}\mathrm{erfc}\big(\frac{-\mu}{\sigma\sqrt{2}}\big)\\ &+2\beta\sum_{k=1}^{N}e^{3kb}\boldsymbol{I}_{23}^b-2b\sum_{k=1}^{N}e^{3kb}\boldsymbol{I}_{24}^b\\ &=\beta\frac{\sigma}{\sqrt{2\pi}}e^{\frac{-\mu^2}{2\sigma^2}}-\beta\mu+\frac{\beta\mu}{2}\mathrm{erfc}\big(\frac{-\mu}{\sigma\sqrt{2}}\big)\\ &+b-\frac{b}{2}\mathrm{erfc}\big(\frac{-\mu}{\sigma\sqrt{2}}\big)2\beta\sum_{k=1}^{N}e^{3kb}\boldsymbol{I}_{23}^b\\ &-2b\sum_{k=1}^{N}e^{3kb}\boldsymbol{I}_{24}^b\\ &=\frac{\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}}-\big(\frac{\alpha^2}{2}+\frac{b}{2}\big)\mathrm{erfc}\big(\frac{-\alpha}{\sqrt{2}}\big)+b\\ &+2\beta\sum_{k=1}^{N}e^{-3kb}\boldsymbol{I}_{23}^b-2b\sum_{k=1}^{N}e^{-3kb}\boldsymbol{I}_{24}^b\end{aligned} \tag{47}$$

$$\begin{aligned}A\boldsymbol{I}_{12}^b+B\boldsymbol{I}_{22}^b &= \frac{(A+B)\alpha}{\sqrt{2\pi}}e^{\frac{-\alpha^2}{2}}\\ &+(A-B)\big(\frac{\alpha^2}{2}+\frac{b}{2}\big)\mathrm{erfc}\big(\frac{-\alpha}{\sqrt{2}}\big)\\ &+bB+2\beta\sum_{k=1}^{N}e^{-3kb}\big(A\boldsymbol{I}_{13}^b+B\boldsymbol{I}_{23}^b\big)\\ &-2b\sum_{k=1}^{N}e^{-3kb}\big(A\boldsymbol{I}_{14}^b+B\boldsymbol{I}_{24}^b\big)\end{aligned} \tag{48}$$

Using Eq. (31)

$$\begin{aligned}A\boldsymbol{I}_{13}^b+B\boldsymbol{I}_{23}^b &= A\frac{M_{k1}V_{k1}}{2}\mathrm{erfc}\big(\frac{-V_{k1}}{\sigma\sqrt{2}}\big)\\ &+B\frac{M_{k2}V_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big)\\ &+\frac{(A+B)\sigma}{\sqrt{2\pi}}e^{-\frac{\alpha^2}{2}}\end{aligned} \tag{49}$$

$$\begin{aligned}\boldsymbol{I}_{14}^b &= \int_0^\infty \mathcal{N}(\mu,\sigma^2)e^{-3k\beta z}dz\\ &=\frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty e^{\frac{-(z-\mu)^2}{2\sigma^2}}e^{-3k\beta z}dz\\ &=e^{\frac{-(6\mu k\sigma^2\beta-9k^2\sigma^4\beta^2)}{2\sigma^2}}\int_0^\infty\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(z-(\mu-3k\sigma^2\beta))^2}{2\sigma^2}}dz\\ &=e^{\frac{-(6\mu k\sigma^2\beta-9k^2\sigma^4\beta^2)}{2\sigma^2}}\int_0^\infty\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(z-V_{k1})^2}{2\sigma^2}}dz\\ &=M_{k1}\int_0^\infty\mathcal{N}(V_{k1},\sigma^2)dz\\ &=\frac{M_{k1}}{2}\mathrm{erfc}\big(\frac{-V_{k1}}{\sigma\sqrt{2}}\big)\end{aligned} \tag{50}$$

$$\begin{aligned}\boldsymbol{I}_{14}^b &= \int_0^\infty\mathcal{N}(-\mu,\sigma^2)e^{-3k\beta z}dz\\ &=\frac{M_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big)\end{aligned} \tag{51}$$

$$A\boldsymbol{I}_{14}^b+B\boldsymbol{I}_{24}^b=A\frac{M_{k1}}{2}\mathrm{erfc}\big(\frac{-V_{k1}}{\sigma\sqrt{2}}\big)+B\frac{M_{k2}}{2}\mathrm{erfc}\big(\frac{-V_{k2}}{\sigma\sqrt{2}}\big) \tag{52}$$

References

[1] J. Jose *et al.*, "An Image Quality Enhancement Scheme Employing Adolescent Identity Search Algorithm in the NSST Domain for Multimodal Medical Image Fusion", *Biomedical Signal Processing and Control*, vol. 66, art. no. 102480, 2021 (`https://doi.org/10.1016/j.bspc.2021.102480`).

[2] J. Xu and Z. Cai, "Gaussian Mixture Deep Dynamic Latent Variable Model with Application to Soft Sensing for Multi-mode Industrial Processes", *Applied Soft Computing*, vol. 114, art. no. 108092, 2022 (`https://doi.org/10.1016/j.asoc.2021.108092`).

[3] J. Skvara and I. Nezbeda, "Thermodynamics and Structure of Supercooled Water. II.", *Journal of Molecular Liquids*, vol. 367, art. no. 120508, 2022 (`https://doi.org/10.1016/j.molliq.2022.120508`).

[4] K. Moshksar and A.K. Khandani, "Arbitrarily Tight Bounds on Differential Entropy of Gaussian Mixtures", *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3340–3354, 2016 (`https://doi.org/10.1109/tit.2016.2553147`).

[5] M. Wen, X. Cheng, and L. Yang, *Index Modulation for 5G Wireless Communications*, Springer, 154 p., 2017 (https://doi.org/10.1007/978-3-319-51355-3).

[6] V. Bhatia and B. Mulgrew, "Non-parametric Likelihood Based Channel Estimator for Gaussian Mixture Noise", *Signal Processing*, vol. 87, no. 11, pp. 2569–2586, 2007 (https://doi.org/10.1016/j.sigpro.2007.04.006).

[7] D. Peel and G.J. McLachlan, *Finite Mixture Models*, John Wiley & Sons, 427 p., 2000 (https://doi.org/10.1002/0471721182).

[8] G.J. McLachlan, S.X. Lee, and S.I. Rathnayake, "Finite Mixture Models", *Annual Review of Statistics and its Application*, vol. 6, pp. 335–378, 2019 (https://doi.org/10.1146/annurev-statistics-031017-100325).

[9] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 748 p., 2005 (https://doi.org/10.1002/047174882X).

[10] M.F. Huber, T. Bailey, H. Durrant-Whyte, and U.D. Hanebeck, "On Entropy Approximation for Gaussian Mixture Random Vectors", *2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Seoul, South Korea, 2008 (https://doi.org/10.1109/mfi.2008.4648062).

[11] J. Melbourne, S. Talukdar, S. Bhaban, and M.V. Salapaka, "Error Bounds on a Mixed Entropy Inequality", *2018 IEEE International Symposium on Information Theory (ISIT)*, Vail, USA, 2018 (https://doi.org/10.1109/isit.2018.8437601).

[12] A. Kolchinsky and B.D. Tracey, "Estimating Mixture Entropy with Pairwise Distances", *Entropy*, vol. 19, no. 7, art. no. 361, 2017 (https://doi.org/10.3390/e19070361).

[13] L.V. Michalowicz, J.M. Nichols, and F. Bucholtz, "Calculation of Differential Entropy for a Mixed Gaussian Distribution", *Entropy*, vol. 10, no. 3, pp. 200–206, 2008 (https://doi.org/10.3390/entropy-e10030200).

[14] J. Sandor, "On Some Inequalities for the Identric, Logarithmic and Related Means", *Aequationes Mathematicae*, vol. 40, pp. 261–270, 1990 (https://doi.org/10.1007/BF02112299).

[15] K.B. Stolarsky, "Holder Means, Lehmer Means, andx^{-1} log cosh$x$", *Journal of Mathematical Analysis and Applications*, vol. 202, no. 3, pp. 810–818, 1996 (https://doi.org/10.1006/jmaa.1996.0349).

[16] W.J. Weltner, P. Schuster, and K. Weltner, *Mathematics for Engineers and Scientists*, Stanley Thornes, 520 p., 1986 (ISBN: 9780859501200).

[17] M. Raginsky, "On the Information Capacity of Gaussian Channels under Small Peak Power Constraints", *46th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, USA, 2008 (https://doi.org/10.1109/allerton.2008.4797569).

[18] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Elsevier, 1160 p., 1980 (https://doi.org/10.1016/C2013-0-10754-4).

[19] J. Melbourne, S. Talukdar, S. Bhaban, M. Madiman, and M.V. Salapaka, "The Differential Entropy of Mixtures: New Bounds and Applications", Entropy, vol. 68, no. 4, pp. 2123–2146, 2022 (https://doi.org/10.48550/arXiv.1805.11257).

Abdelrahman Marconi, M.Sc.
Communications Department
https://orcid.org/0009-0006-5181-4946
E-mail: marconi@mtc.edu.eg
Military Technical College, Cairo, Egypt
https://www.mtc.edu.eg

Ahmed H. Elghandour, Ph.D.
Communications Department
https://orcid.org/0000-0002-5250-4847
E-mail: elghandour@mtc.edu.eg
Military Technical College, Cairo, Egypt
https://www.mtc.edu.eg

Ashraf D. Elbayoumy, Professor
Communications Department
https://orcid.org/0000-0001-6356-5285
E-mail: adiaa@mtc.edu.eg
Military Technical College, Cairo, Egypt
https://www.mtc.edu.eg

Amr Abdelaziz, Ph.D.
Communications Department
https://orcid.org/0000-0003-3709-3653
E-mail: amrashry@mtc.edu.eg
Military Technical College, Cairo, Egypt
https://www.mtc.edu.eg