Constellation Constrained Capacity of Additive Gaussian Mixture Noise Channels

J. Harshan and Emanuele Viterbo

Department of Electrical and Computer Systems Engineering,
Monash University, Australia
{harshan.jagadeesh and emanuele.viterbo}@monash.edu

Abstract— **Communication channels that are characterized by additive Gaussian noise have been well studied. However, many practical systems are also known to experience non-Gaussian noise. A convenient method to analyse such systems is by modelling the non-Gaussian noise using Gaussian mixture densities. In this paper we compute the constellation constrained (CC) capacity of additive Gaussian mixture (GM) noise channels with finite input alphabets. We study a wide spectrum of GM densities covering single lobe, multi-lobe, symmetric tapering and asymmetric tapering densities. We show that the CC capacity of GM densities is larger than that of the Gaussian density of the same variance at low SNR values. This observation points at the drawback of the existing capacity achieving codes matched to Gaussian channels, and highlights the need for constructing new codes for such channels. We also study GM noise models with data dependent density parameters, that have been recently shown to approximate the NAND flash memory channels.**

I. INTRODUCTION AND PRELIMINARIES

Communication in the presence of additive Gaussian noise is well understood and has been extensively studied in the past. In such systems, sufficient number of noise components from individual sources combine together through the central limit theorem to contribute Gaussian statistics [1]. However, many practical systems are characterized by non-Gaussian noise due to inadequate number of identical noise sources. Some well known examples for non-Gaussian channels include electromagnetic interference [2], underwater acoustic channels [3], power line communication [4], and NAND flash memory channels [5], [6], [7]. In these physical channels, the ambient noise has been confirmed non-Gaussian through experimental measurements.

Non-Gaussian noise sources have posed significant challenges to fit a suitable model to approximate their characteristics. Modelling such sources is paramount from a theoretical stand point to understand the limits of communication over non-Gaussian channels. Among several approximation techniques, Gaussian mixture (GM) densities have been used to model a variety of non-Gaussian noise environments [1], [8].

Many works have addressed the problem of designing signal processing and detection algorithms in Gaussian mixture noise channels [9], [10], [1], [11], [12]. However, there are no works which study the mutual information of additive GM noise with finite input alphabet. It is important from a practical view point to study the maximum achievable rate of communication with finite input alphabet.

In this paper we study the mutual information of the additive GM noise channel with uniform distribution on the finite input alphabet. Similar studies have been reported in the past for Gaussian distributions in multiple access channels [13], interference channels [14], wiretap channels [15] and cognitive radio networks [16] (also see the references within). Such capacity expressions are referred to as *constellation constrained* (CC) capacity due to *finite* size on the input alphabet. Throughout the paper, we refer the mutual information of the additive GM noise with uniform input distribution as the CC capacity of the additive GM noise channel. The contributions of this paper are:

- We compute the CC capacity of the additive Gaussian mixture noise with uniform distribution on the finite input alphabets.
- We study a wide range of densities ranging from symmetric tapering, skewed tapering and multi-lobe cases. For these cases, we numerically show that the CC capacity of the GM is indeed larger than that of the Gaussian distribution of same variance at low SNR values. Interestingly, for the case of symmetric tapering and skewed tapering, we observe a cross-over behaviour between the CC capacity curves of the GM and the equivalent Gaussian noise.
- We also study the case when the non-Gaussian densities are a function of the input information symbols which are known to arise in the context of NAND flash memory channels as shown in [6].

Our results point at the drawback of the existing capacity achieving codes for Gaussian channels, and generate interest in the construction of new codes for such channels.

II. SYSTEM MODEL

We consider a discrete time communication channel that is characterized by the non-Gaussian additive noise Z. The input to the channel is denoted by X which takes value from a *finite real* alphabet $\mathcal{S} = \{x_1, x_2, \ldots, x_N\}$ of size N. The set $\mathcal{S}$ is normalized such that $\sum_{k=1}^{N} x_k^2 = N$. We also assume uniform input distribution on X. The signal model for this channel is given by

$$Y = \sqrt{P}X + Z, \qquad (1)$$

where the scalar $\sqrt{P}$ is used to vary the average transmit power. We use the two-term Gaussian mixture (GM) model

for the noise samples of Z. The probability density function (pdf) of Z is given by

$$p_Z(z) = p\mathcal{N}(\mu_1, \sigma_1^2) + (1-p)\mathcal{N}(\mu_2, \sigma_2^2),$$

where $0 \leq p \leq 1$ and $\mathcal{N}(\cdot, \cdot)$ denotes a Gaussian distribution. This model is an approximation to Middleton Class A noise model [8] and has been used to model physical noise in radio and acoustic channels. It is easy to verify that the mean μ and the variance σ^2 of the GM are respectively given by

$$\mu = p_1\mu_1 + (1-p)\mu_2, \tag{2}$$

and

$$\sigma^2 = p(\mu_1^2 + \sigma_1^2) + (1-p)(\mu_2^2 + \sigma_2^2) - (p\mu_1 + (1-p)\mu_2)^2. \tag{3}$$

Using the scale factor $\sqrt{P}$, we define SNR as

$$\text{SNR} \triangleq \frac{PE[X^2]}{\sigma^2} = \frac{P}{\sigma^2}.$$

With finite input alphabet, the CC capacity of the GM noise with uniform input distribution is given by $I(X;Y) = h(Y) - h(Y|X)$, where $h(\cdot)$ denotes the differential entropy function. Due to the *finite* size constraint, it is not possible to have closed form expressions for the CC capacity. However, these numbers can be numerically computed either using Gauss-Hermite quadrature integration [17] or Monte Carlo simulations. Note that closed form expressions for the CC capacity are also not available in the literature for Gaussian noise distributions. As a result, even if one of the components in the mixture has negligible contribution, it is not possible to evaluate the CC capacity in closed form. Despite this shortcoming, many groups have studied the CC capacity of Gaussian noise channels through computer simulations, and have proposed efficient precoding and signal processing techniques for such channels [13], [14], [15], [16]. This work is one along that direction where we rely on extensive numerical results to understand the limits of the GM noise channel.

III. CAPACITY FOR GENERAL GAUSSIAN MIXTURE NOISE CHANNELS

In this section we compute the CC capacity of GM noise for various combinations of the GM parameters $\{p, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2\}$. We encompass different cases of densities covering symmetric tapering, asymmetric tapering and multi-lobe densities. For the numerical results, we use $\mathcal{S} = \sqrt{\frac{4}{14}}\{0, 1, 2, 3\}$ as the input alphabet. For a given combination of the GM parameters, we first compute the variance given in (3). Then, we compute the CC capacity values as a function of SNR $= \frac{P}{\sigma^2}$ by varying the scalar P.[1] In addition, for comparison purposes, we present the CC capacity values for a Gaussian distribution with the same variance as that of the GM. In the following subsections, we discuss different cases of GM densities and their CC capacity behaviour.

Fig. 1. pdf of the GM (in dashed lines) with different degrees of tapering. pdf of the Gaussian distribution (in solid lines) with same variance.
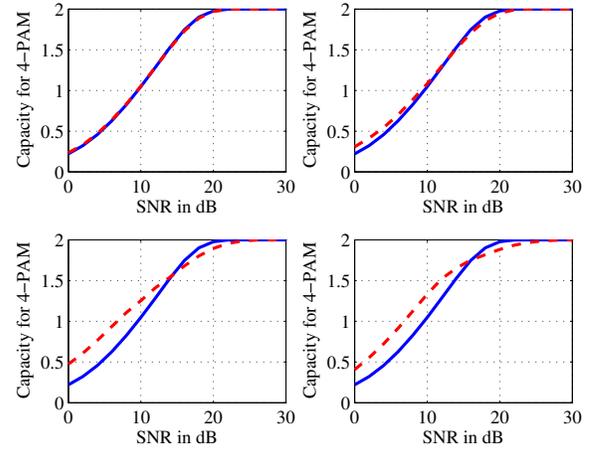
Fig. 2. Capacity of 4-PAM constellation for noise densities presented in Fig. 1.

A. Sharp Tapering with Symmetry

We compute the CC capacity for different levels of steepness in its tail distribution. We use the distribution

$$p_Z(z) = p\mathcal{N}(0, 1) + (1-p)\mathcal{N}(0, 20),$$

and then vary the value of p to obtain several degrees of steepness. In Fig. 1, we present the pdf of the GM when $p \in \{0.1, 0.3, 0.7, 0.9\}$. From Fig. 1, we see that the tapering of the GM density is sharper as p increases. Note that the Gaussian distribution with the same variance is also depicted in the same figure. The corresponding CC capacity curves are given in Fig. 2 which shows that the CC capacity of the GM is larger than that of the Gaussian at low SNR values, while there is a cross-over in the CC capacity after a certain *threshold* SNR. An interesting observation is that the gains in the CC capacity increases as p increases.

[1]The presented results are obtained using Gauss-Hermite quadrature integration with 40 samples to obtain the Hermite approximation. Due to space considerations, we do not showcase the algorithm to calculate the CC capacity.
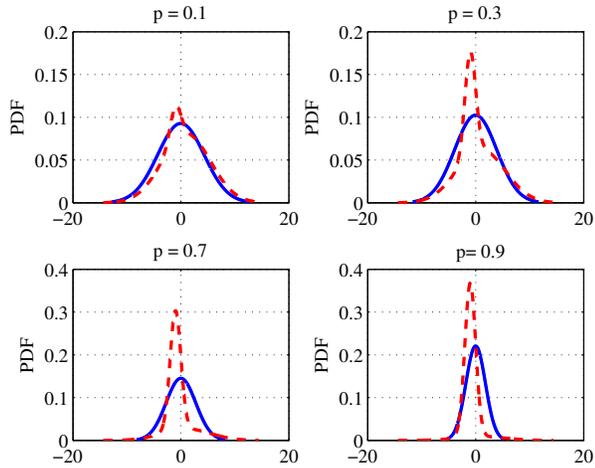
Fig. 3. pdf of the GM (in dashed lines) with different degrees of skewed tapering. pdf of the Gaussian distribution (in solid lines) with same variance.
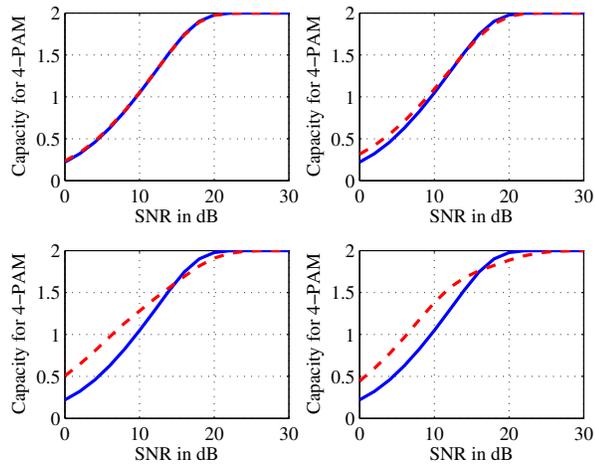
Fig. 5. pdf of the GM (in dashed lines) with different degrees of multiple lobes. pdf of the Gaussian distribution (in solid lines) with same variance. Clockwise direction from the top left figure,, μ_1 values of $1, 2, 6,$ and 4 are used, respectively.

Fig. 4. Capacity of 4-PAM constellation for noise densities presented in Fig. 3.

Fig. 6. Capacity of 4-PAM constellation for noise densities presented in Fig. 5.

B. Sharp Tapering with Asymmetry

In this part we compute the CC capacity for different degrees of asymmetric tapering. To bring in asymmetry, we use

$$p_Z(z) = p\mathcal{N}(-1, 1) + (1-p)\mathcal{N}(1, 20).$$

Similar to the previous subsection, we vary the value of p to obtain several degrees of asymmetric tapering. In Fig. 3, we present the pdf of noise when $p \in \{0.1, 0.3, 0.7, 0.9\}$. The corresponding CC capacity values are given in Fig. 4. Note that the CC capacity behaviour is similar to that of the symmetric tapering case.

C. GM with Multi-lobe Distributions

We compute the CC capacity with multiple lobes in the GM distribution. In order to create multi-lobe structure, we vary the means μ_1 and μ_2 such that $\mu_1 = -\mu_2$. We use the

distribution

$$p_Z(z) = 0.5\mathcal{N}(\mu_1, 1) + 0.5\mathcal{N}(\mu_2, 20).$$

In Fig. 5, we present the pdf of GM when $\mu_1 = 1, 2, 4,$ and 6. Note that the Gaussian distribution with the same variance as that of the GM is also depicted. The corresponding CC capacity values are given in Fig. 6 which shows that the CC capacity of the GM is larger than that of the corresponding Gaussian at low SNR values. However, note that the CC capacity curve does not display the cross-over behaviour with that of the Gaussian distribution. Finally, it is observed that the advantage in the CC capacity increases as the multiple lobes become prominent in the distribution.

D. Explanation for the Capacity Behaviour of Gaussian Mixture

We give an explanation to the low SNR improvement in the CC capacity of the GM. To facilitate the argument, we use the definition $I(X;Y) = h(Y) - h(Y|X)$. Since the scale factor $\sqrt{P}$ is associated with the symbol X, the term $h(Y|X)$ is independent of SNR. As a result, our argument is based on the behaviour of $h(Y)$, which in turn is based on the pdf of Y (denoted by $p_Y(y)$).

At low SNR, the symbols $\sqrt{P}\{x_1, x_2, \ldots, x_N\}$ lie close to each other. As a result, $p_Y(y)$ which is defined as

$$p_Y(y) = \sum_{k=1}^{N} \frac{1}{N} p_{Y/X}(y|X = x_k),$$

looks similar for both the Gaussian and the GM, i.e., the effect of sharp tapering of the GM is not prominent as their tails overlap. Hence, the $h(Y)$ values of the Gaussian and the GM are close at low SNR. However, at moderate and high SNR, the symbols $\sqrt{P}\{x_1, x_2, \ldots, x_N\}$ move farther apart, and hence, the sharp tapering of the GM is prominent. Thus, $p_Y(y)$ of Gaussian and the GM are clearly *distinct*. Hence, the corresponding values of $h(Y)$ are not close at moderate to high SNR. To exemplify this point, in Fig. 7 we plot $p_Y(y)$ for both the Gaussian and the GM at SNR = 3 dB (a low SNR case) and SNR = 10 dB (a high SNR case). To obtain Fig. 7, we use

$$p_Z(z) = 0.5\mathcal{N}(0, 1) + 0.5\mathcal{N}(0, 20).$$

The top figure of Fig. 7 shows the closeness of the two distributions at low SNR.

Finally, since the Gaussian distribution maximises the entropy for a given variance, $h(Y|X)$ for the GM is smaller than that of the Gaussian distribution. Therefore, when $h(Y)$ values are close, the formula $I(X;Y)$ makes sure that the GM gets larger CC capacity than the Gaussian. Note that this is not a formal proof for the low SNR dominance of the GM. This is an intuitive explanation that we could gather after a thorough analysis of the numerical results.

IV. DATA DEPENDENT NOISE DISTRIBUTION FROM GAUSSIAN MIXTURES

In this section we study a communication channel where the additive noise distribution depends on the input X. The signal model is similar to that in Section II except for the distribution of the additive noise Z. The conditional pdf of Z given $X = x_k$ for $k \in \{1, 2, \ldots, N\}$ is given by

$$p_{Z/X}(z|X = x_k) = p_k\mathcal{N}(\mu_{1,k}, \sigma_{1,k}^2) + (1 - p_k)\mathcal{N}(\mu_{2,k}, \sigma_{2,k}^2),$$

where $0 \leq p_k \leq 1$ and $\mathcal{N}(\cdot, \cdot)$ denotes a Gaussian distribution. The subscript k in the GM parameters highlights their dependence on the channel input x_k. With uniform distribution on the input symbols, the pdf of Z is given by

$$p_Z(z) = \frac{1}{N} \sum_{k=1}^{N} \{p_k\mathcal{N}(\mu_{1,k}, \sigma_{1,k}^2) + (1 - p_k)\mathcal{N}(\mu_{2,k}, \sigma_{2,k}^2)\},$$
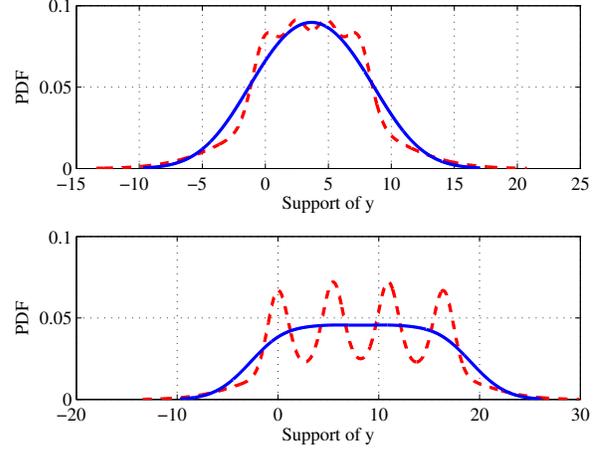
Fig. 7. pdf of y for Gaussian (in solid lines) and GM (in dashed lines) at SNR of 3 dB (top figure) and 10 dB (bottom figure).

It is straightforward to verify that the variance σ^2 of Z is given by

$$\sigma^2 = \frac{1}{N} \sum_{k=1}^{N} \{\{p_k(\mu_{1,k}^2 + \sigma_{1,k}^2) + (1 - p_k)(\mu_{2,k}^2 + \sigma_{2,k}^2)\} - \mu^2,$$

where μ is the mean of the GM. Similar to Section II, we define SNR as P/σ^2. A known example for input dependent additive noise channel is the NAND flash memory channel [5], [6].

A. Numerical Results

For illustrative purposes, we consider a channel with an input dependent GM distribution given by

$$p_{Z/X}(z|X = x_1) = \mathcal{N}(0, \kappa 10.5), \tag{4}$$

where $\kappa > 0$ and

$$P_{Z/X}(z|X = x_k) = 0.5\mathcal{N}(0, 1) + 0.5\mathcal{N}(0, 20), \tag{5}$$

for $k = 2, 3, 4$. The input alphabet is $\mathcal{S} = \sqrt{\frac{4}{14}}\{0, 1, 2, 3\}$. We present the CC capacity values as a function of SNR for $\kappa = 1, 2, 4, 8$. The different values of κ leads to cases when the conditional distribution around x_1 is wider than that around x_2, x_3, x_4. In Fig. 8, we plot the conditional pdfs in (4) and (5) to capture the difference in the conditional distributions. We present the corresponding CC capacity values as a function of SNR in Fig. 9. In the same figure, we plot the CC capacity with the Gaussian distribution of the same variance as that of the GM. Similar to the observations in the previous section, we see low SNR gains in the CC capacity of the GM over the Gaussian distribution. However, in this case, an interesting observation is that the cross-over point between the GM and the Gaussian distribution occurs at lower SNR values as κ increases.

In summary, our numerical results indicate that additive noise channels with input *dependent* GM noise distributions have similar CC capacity characteristics as that of the input
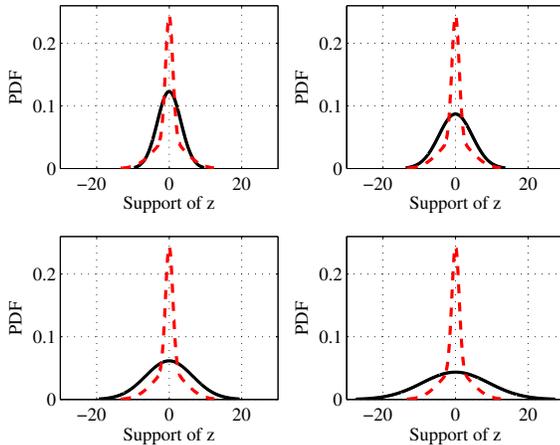
Fig. 8. Different degrees of the conditional pdf around x_1 (in solid lines) and conditional pdf around x_k for $k = 2, 3, 4$ (in dashed lines). Clockwise direction from the top left figure, κ values of $1, 2, 8,$ and 4 are used, respectively.

Fig. 9. In dashed lines: Capacity of 4-PAM constellation for input dependent noise densities in Fig. 8. In solid lines: Capacity results for the equivalent Gaussian distribution of same variance.

independent case. However, the SNR value at which the crossover in the CC capacity occurs depends on the difference between the variance of the conditional distributions. In general, CC capacity of input dependent GM models can be computed for various combination of the GM parameters. We have chosen a specific combination that is inspired by the conditional distributions that were reported through measurements for NAND flash channels (see [6, Sec. II.D]).

V. CONCLUSION

We have presented the CC capacity of additive Gaussian mixture noise channels with finite input alphabets. Though there have been many works that have addressed detection and signal processing problems in the presence of the GM noise, there were no prior works studying the mutual information in such environments. Our research gives an idea of the differences in the achievable rates for the GM noise in comparison with the Gaussian noise. The numerical results also highlight the low SNR gains and the cross-over behavior of the GM density with the Gaussian density. We have also studied the CC capacity of channel models with input dependent noise distributions. An interesting direction for future research is to design coding schemes for GM models.

ACKNOWLEDGMENT

REFERENCES

[1] D. W. J. Stein, "Detection of random signals in Gaussian mixture noise," *IEEE Transactions on Information Theory*, Vol. 41, No. 06, pp. 1788–1801, Nov. 1995.

[2] D. Middleton and A. Spaulding, "Optimum reception in non-Gaussian electromagnetic interference environments. Part 2: Optimum and suboptimum threshold signal detection in class A and B noise," NTIA Rep. 83–120, May 1983.

[3] F. W. Machell and C. S. Penrod, *Probability density functions of ocean acoustic noise processes*, in Statistical Signal Processing, E. J. Wegman and J. G. Smith, Eds. New York: Marcel Dekker, 1984, pp. 211–221.

[4] M. Zimmermann and K. Dostert, "Analysis and modelling of impulsive noise in broad-band powerline communications," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 44, No. 01, pp. 249–258, Feb. 2002.

[5] J. Kim, D. Lee, and W. Sung, "Performance of rate 0.96 (68254, 65536) EG-LDPC Ccde for NAND flash memory error correction," in the Proc. of *IEEE ICC 2012*, Ottawa, Canada, pp. 7029–7033.

[6] G. Dong, Y. Pan, N. Xie, C. Varanasi, and T. Zhang, "Estimating information-theoretical NAND flash memory storage capacity and its implication to memory system design space exploration," *IEEE Transactions on VLSI Systems*, Vol. 20, No. 9, pp. 1705–1714, Sept. 2012.

[7] Y. Cai, O. Mutlu, E. F. Haratsch, and K. Mai, "Program interference in MLC NAND flash memory: characterization, modelling, and mitigation," in the Proc. of *IEEE International Conference on Computer Design (ICCD)*, Asheville, NC, October 2013, pp. 123–130.

[8] D. Middleton, "Non-Gaussian noise models in signal processing for telecommunications: new methods an results for class A and class B noise models," *IEEE Transactions on Information Theory*, Vol. 45, No. 04, pp. 1129–1149, May 1999.

[9] A. Kenarsari-Anhari and L. Lampe, "Performance analysis for BICM transmission over Gaussian mixture noise fading channels", *IEEE Transactions on Communications*, Vol. 58, No. 07, pp. 1962–1972, July 2010.

[10] X. Wang and V. Poor, "Robust multiuser detection in non-Gaussian channels," *IEEE Transactions on Signal Processing*, Vol. 47, No. 02, pp. 289–305, Feb. 1999.

[11] T. Li, Wai Ho Mow, and M. H. Siu, "Bit-interleaved coded modulation in the presence of unknown impulsive noise," in the Proc. of *IEEE ICC 2006*, Instanbul, Turkey, pp. 3014–3019.

[12] A. Nasri and R. Schober, "Performance of BICM-SC and BICM-OFDM systems with diversity reception in non-Gaussian noise and interference," *IEEE Transactions on Communications*, vol. 57, no. 11, pp. 3316-3327, Nov. 2009.

[13] J. Harshan and B. S. Rajan, "On two-user Gaussian multiple access channels with finite input constellations," *IEEE Transactions on Information Theory*, Vol. 57, No. 3, pp. 1299–1327, Mar. 2011.

[14] K. Moshksar, A. Ghasemi, and A. Khandani, "On Gaussian interference channels with constellation-based transmitters," *IEEE Communication Letters*, Vol. 16, No. 12, pp. 1941–1943, Dec. 2012.

[15] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "Linear precoding for finite-alphabet signaling over MIMOME wiretap channels," *IEEE Transactions On Vehicular Technology*, Vol. 61, No. 6, pp. 2599–2612, Jul. 2012.

[16] W. Zeng, C. Xiao, J. Lu, and K. B Letaief, "Globally optimal precoder design with finite-alphabet inputs for cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, Vol. 30, No. 10, pp. 1861–1874, Nov. 2012.

[17] A. H. Stroud and D. Secrest, *Gaussian Quadrature Formulas*, Englewood Cliffs, NJ: Prentice-Hall, 1966.