Weighted Performance Function for (r, s)-Entropy of Discrete Memoryless Communication Channel under Single Constraint

H. D. Arora, Vijay Kumar, and Anjali Dhiman

Abstract—**In almost every engineering and technology fields, information channel capacity is the tightest upper bound on the amount of information that can be reliably transmitted over a communications channel. Brijpaul and Sharma [1] gave a direct method of computing the performance function of a discrete memoryless communication channels. They obtained the performance function for discrete channels by using Shannon entropy which is additive in nature, but in some situations additivity does not hold well. In such situations, non-additivity prevails. Arimoto [2] and Blahut [3] proposed an iterative method to compute the channel capacity of a discrete memoryless channel. In this paper we present an algorithm for computing performance function for useful (r, s) – entropy under single and multiple constraints by defining mutual information in terms of Sharma and Mittal [4] entropy of order r and degree s which is non-additive in nature.**

Index Terms—**(r, s) Entropy, performance function, channel capacity, communication channel, additivity.**

I. Introduction

Channel capacity is a fundamental concept in information theory and was introduced by Shannon [5]. The channel capacity is the maximum rate at which information can be transmitted with a single use of the channel with arbitrarily low probability of error. It is usually expressed as bits per second. The capacity of the channel is a useful measure as it tells us the highest rate at which information can be reliably transmitted. A general method for determining the capacity of discrete memoryless channel has been suggested by Cheng [6] and Takano [7]; Zhao and Bose [8]. While Meister and Oettli [9] proposed an iterative procedure based on the method of concave programming and showed that it converges to capacity. An insight into the concept of Channel Capacity was presented by Costello and Forney [10]. Arimoto [2] and Blahut [3] also proposed another iteration method to compute the capacity, which is very simple and systematic.

Let $X = (x_1, x_2, x_{3\ldots}\ x_n)$ and $Y = (y_1, y_2, y_3, \ldots\ y_m)$ represent the set of input alphabet with n letters and the set of output alphabet with m letters respectively. Let $p\ (x_i)$ and $p\ (y_j)$ where $i = 1, 2, 3, \ldots\ n$ and $j = 1, 2, 3, \ldots\ m$ be the probability distribution functions defined on X and Y respectively. Shannon's [5] measure of information is given by

$$H\ (X) = -\sum_{i=1}^{n} p(\chi_i) \log p(\chi_i)\ ,\ p\ (x_i) \geq 0,\ \sum_{i=1}^{n} p(x_i) = 1 \tag{1.1}$$

The conditional entropy is defined as

$$H\ (X/Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j) \log p\ (x_i / y_j) \tag{1.2}$$

where $p\ (x_i\ /\ y_j)$ and $p\ (x_i,\ y_j)$ are the conditional and joint probabilities respectively. The average mutual information is given by

$$I\ (X\ /\ Y) = H\ (X) - H\ (X\ /\ Y) \tag{1.3}$$

The channel capacity are defined by Shannon is given by

$$C = \text{Max.}\ \{I(X/Y):\ \sum_{i=1}^{n} p(x_i) = 1\ \} \tag{1.4}$$

Brijpaul and Sharma [11] formulated the effects of the restrictions as constraints and found the maximum rate under these constraints. They call it as channel performance. They have defined the channel – performance as

$$C = \text{Max.}\{\text{I}(X/Y): \sum_{i=1}^{n} p(x_i) = 1\ , \sum_{i=1}^{n} C_{ki}\ p(x_i) \leq \delta_{\ k}\};\quad k = 1, 2, 3, \ldots\ \ell \tag{1.5}$$

where C_{ki} is the k^{th} type of cost associated with the symbol x_i and $\delta_{\ k}$, $k = 1, 2, 3, \ldots\ldots\ \ell$ are non – negative previously fixed constants arising from practical considerations.

Brijpaul and Sharma [11] obtained the performance function for discrete channels by using Shannon entropy which is additive in nature, but in some situations additivity does not hold good. In such situations non-additivity prevails. Belis and Guiasu [12] attached with the probability scheme, a utility (weighted) distribution. Shimar and Taneja [13] have computed the performance function of γ entropy using Belis and Guiasu [12] measure under single and multiple constraints. In this paper, we are giving an algorithm for computing the performance function for

Manuscript received August 12, 2013; revised October 23, 2013

H. D. Arora is with the Department of Mathematics, AIAS, Amity University, Noida, India (e-mail: hdarora@amity.ed).

Vijay Kumar is with the Department of Mathematics, Amity School of Engineering & Technology, New Delhi, India (e-mail: vijay_parashar@yahoo.com).

Anjali Dhiman is with the Department of Mathematics, Ideal Institute of Mgmt and Tech, G. G. S. I. P. Universi0ty, Delhi, India (e-mail: anjalidhiman2006@gmail.com).

DOI: 10.7763/IJMO.2014.V4.340

generalized weighted β entropy of discrete memoryless communication channel under single and multiple constraints in which the mutual information has been defined in terms of Havrda and Charvat [14] entropy of degree β which is non-additive in nature which depends upon the utility distribution.

$$U = \left(u(x_1), u(x_1), u(x_1), \ldots u(x_n) : u(x_n) \geq 0 \right)$$

II. Algorithm

Let us denote a discrete memoryless channel with n input and m output symbols by $m \times n$ matrix A i.e. where

$$H_\beta(P\ ;\ U) = (2^{1-\beta} - 1)^{-1} \sum_{i=1}^{n} u(x_i)\, p^{\beta}(x_i) - 1\} \ ; \beta \neq 1,\ \beta > 0 \tag{2.1}$$

$$H_\beta(A{:}\ P;\ U) = (2^{1-\beta} - 1)^{-1}\{\sum_{j=1}^{m}\sum_{i=1}^{n}[\,u(x_i)\, p(x_i)\, p(y_j/x_i)\,]^{\beta} - \sum_{j=1}^{m}[\sum_{i=1}^{n} u(x_i)\, p(x_i)\, p\,(y_j/x_i)\,]^{\beta}\} \tag{2.2}$$

where $H_\beta\,(A : P\ ;\ U)$ is the conditional weighted entropy of degree β.

We define the weighted capacity of degree β of a discrete memoryless channel A as

$$C_\beta\,(P\ ;\ U) = \max\ I_\beta\,(A : P\ ;\ U) \tag{2.3}$$

In view of various cost factors, we define the generalized weighted β – performance function as

$$C_\beta\,(\delta\ ;\ U) = \max\ \{\, I_\beta\,(A : P\ ;\ U)\ :\ \sum_{i=1}^{n} p(x_i) = 1\ ,\ \sum_{i=1}^{n} C_{ki}\, p(x_i) \leq \delta_k \}\ ;\ k = 1, 2, 3, \ldots\ \ell \tag{2.4}$$

where C_{ki} is the k^{th} type of cost associated with the symbol x_i.

Introduce a stochastic matrix $n \times m$ as

$$q = \{\ q\,(x_i / y_j)\ ,\ i = 1, 2, 3, \ldots n\ \&\ j = 1, 2, 3, \ldots m\ \}$$

where

$$q\,(x_i / y_j) \geq 0\ ,\ \sum_{i=1}^{n} q\,(x_i / y_j) = 1 \tag{2.5}$$

We define generalize conditional entropy of degree β as $J_\beta\,(A{:}\ P;\ Q;\ U) = (2^{1-\beta} - 1)^{-1}$

$$\{\sum_{j=1}^{m}\sum_{i=1}^{n} u^{\beta}(x_i)\, p^{\beta}(x_i)\, p^{\beta}(y_j/x_i)\,(\,1 - q^{1-\beta}(x_i/y_j)\ )\ \} ; \beta \neq 1,\ \beta > 0 \tag{2.6}$$

Then, if q is defined by

$$q\,(x_i / y_j) = \frac{u\,(x_i)\, p(x_i)\, p(y_j/x_i)}{\sum_{i=1}^{n} u(x_i)\, p(x_i)\, p(y_j/x_i)} = P\times (x_i / y_j) \tag{2.7}$$

Using (2.7) in (2.6), the expression reduces to (2.2) i.e. to say

$$J_\beta\,(A : P\ ;\ q\ ;\ U) = H_\beta\,(A : P\ ;\ U)$$

Furthermore, we can easily prove the inequality

$$J_\beta\,(A : P\ ;\ q\ ;\ U) \geq J_\beta\,(A : P\ ;\ P^*\ ;\ U\,) \tag{2.8}$$

where P^* is the stochastic matrix whose $(i\ ,\ j)^{\text{th}}$ entry is $A\,(x_i / y_j)$ as defined in (2.9).

Theorem 2. 1 Performance function for generalized weighted β entropy under single constraint

For any fixed β and q, $J_\beta\,(A;\ P;\ q;\ U)$ is maximized by

$$p\,(x_i) = \left[\frac{(\lambda + \theta\, C_i)\,(2^{1-\beta} - 1)}{\beta\ \ u(x_i)\, a_i}\right]^{\frac{1}{\beta - 1}} \tag{2.9}$$

where

$$a_i = 1 - \sum_{j=1}^{m} p^{\beta}\,(y_j/x_i)\ \ (1 - q^{1-\beta}\,(x_i/y_j)\,) \tag{2.10}$$

λ and θ are lagrange's multipliers, subject to

$$\sum_{i=1}^{n} p(x_i) = 1 \ \text{ and } \sum_{i=1}^{n} C_i\, p(x_i) = \delta \tag{2.11}$$

And

$$\delta = \sum_{i=1}^{n} C_i \left[\frac{(\lambda + \theta\, C_i)\,(2^{1-\beta} - 1)}{\beta\ \ u(x_i)\, a_i}\right]^{\frac{1}{\beta - 1}} \tag{2.12}$$

And then

$$C_\beta\,(\delta\ ;\ q\ ;\ U) = (2^{1-\beta} - 1)^{-1} \sum_{i=1}^{n} \left[\frac{(\lambda + \theta C_i)\,(2^{1-\beta} - 1)}{\beta}\right]^{\frac{\beta}{\beta - 1}} (a_i)^{\frac{1}{1-\beta}} - 1 \leq C_\beta\,(\delta\ ;U) \tag{2.13}$$

Proof:

The function which we want to maximize is of the following form:

$$H_\beta\,(P\ ;\ U) - J_\beta\,(A{:}\ P\ ;\ q\ ;\ U)$$

Using the lagrange's method of multipliers, we have

$$\phi = H_\beta\,(\mathrm{P}\ ;\ \mathrm{U}) - J_\beta\,(\mathrm{A\ I\ P}\ ;\ q\ ;\ \mathrm{U}) + \lambda\,(1 - \sum_{i=1}^{n} p(x_i)\,) + \theta\ (\,\delta - \sum_{i=1}^{n} C_i\, p(x_i)\,) \tag{2.14}$$

We will obtain a maximum of (2.14) by differentiating ϕ with respect to $p\,(x_i)$ and setting this partial derivative equal to zero. i.e.

$$\phi = (2^{1-\beta}-1)^{-1}\{\sum_{i=1}^{n} u(x_i)\,p^{\beta}(x_i) - 1 - \sum_{j=1}^{m}\sum_{i=1}^{n} p^{\beta}(y_j/x_i)\,u^{\beta}(x_i)\,p^{\beta}(x_i) + \sum_{j=1}^{m}\sum_{i=1}^{n} p^{\beta}(y_j/x_i)\,u^{\beta}(x_i)\,p^{\beta}(x_i)\,q^{1-\beta}(x_i/y_j)\}$$

$$+ \lambda(1 - \sum_{i=1}^{n} p(x_i)) + \theta(\delta - \sum_{i=1}^{n} C_i\,p(x_i) \qquad (2.15)$$

Now $$\frac{\partial \varphi}{\partial p(x_i)} = (2^{1-\beta}-1)^{-1}\{\beta\,u(x_i)p^{\beta-1}(x_i) - \beta\sum_{j=1}^{m} p^{\beta}(y_j/x_i)\,u(x_i)\,p^{\beta-1}(x_i) + \sum_{j=1}^{m} p^{\beta}(y_j/x_i)\,u(x_i)\,p^{\beta-1}(x_i)\,q^{1-\beta}(x_i/y_j)\}$$

$$-\lambda - \theta C_i = 0$$

$$\Rightarrow (\lambda + \theta C_i)(2^{1-\beta}-1) = \beta\{u(x_i)p^{\beta-1}(x_i) - \sum_{j=1}^{m} p^{\beta}(y_j/x_i)\,u(x_i)\,p^{\beta-1}(x_i) + \sum_{j=1}^{m} p^{\beta}(y_j/x_i)\,u(x_i)\,p^{\beta-1}(x_i)\,q^{1-\beta}(x_i/y_j)\}$$

$$\Rightarrow (\lambda + \theta C_i)(2^{1-\beta}-1) = \beta\,u(x_i)\,p^{1-\beta}(x_i)\,\{1 - \sum_{j=1}^{m} p^{\beta}(y_j/x_i) + \sum_{j=1}^{m} p^{\beta}(y_j/x_i)\,q^{1-\beta}(x_i/y_j)\}$$

$$\Rightarrow (\lambda + \theta C_i)(2^{1-\beta}-1) = \beta\,u(x_i)\,p^{1-\beta}(x_i)\,\{1 - \sum_{j=1}^{m} p^{\beta}(y_j/x_i)\,[1 - \sum_{j=1}^{m} q^{1-\beta}(x_i/y_j)]\}$$

$$\Rightarrow (\lambda + \theta C_i)(2^{1-\beta}-1) = \beta\,u(x_i)\,p^{\beta-1}(x_i)\,a_i,$$

where

$$a_i = 1 - \sum_{j=1}^{m} p^{\beta}(y_j/x_i)\,(1 - q^{1-\beta}(x_i/y_j))$$

$$\Rightarrow p^{\beta-1}(x_i) = \frac{(\lambda+\theta C_i)(2^{1-\beta}-1)}{\beta\,u(x_i)\,a_i}$$

$$\Rightarrow p\times(x_i) = \left[\frac{(\lambda+\theta C_i)(2^{1-\beta}-1)}{\beta\,u(x_i)\,a_i}\right]^{\frac{1}{\beta-1}} \quad ; i = 1, 2, 3, n \qquad (2.16)$$

Which is the required probability distribution? Now multiplying the above equation by C_i and summing over i and using the relation $\sum_{i=1}^{n} C_i\,p(x_i) = \delta$, we have

$$\delta = \sum_{i=1}^{n} C_i\left[\frac{(\lambda+\theta C_i)(2^{1-\beta}-1)}{\beta\,u(x_i)\,a_i}\right]^{\frac{1}{\beta-1}}$$

Now, $C_\beta(\delta\,;q\,;U) = \underset{P\in\Delta_n}{\mathrm{Max}}\{H_\beta(P\,;U) - J_\beta(A\,;P\,;q\,;U)\,:\,\sum_{i=1}^{n} C_i\,p(x_i) = \delta\}$

$$= (2^{1-\beta}-1)^{-1}\{\sum_{i=1}^{n} u^{\beta}(x_i)p^{\beta}(x_i) - 1 - \sum_{j=1}^{m}\sum_{i=1}^{n} u^{\beta}(x_i)\,p^{\beta}(y_j/x_i)\,p^{\beta}(x_i) + \sum_{j=1}^{m}\sum_{i=1}^{n} u^{\beta}(x_i)\,p^{\beta}(y_j/x_i)\,p^{\beta}(x_i)\,q^{1-\beta}(x_i/y_j)\}$$

$$= (2^{1-\beta}-1)^{-1}\{\sum_{i=1}^{n} u^{\beta}(x_i)p^{\beta}(X_i) - 1 - \sum_{j=1}^{m}\sum_{i=1}^{n} u^{\beta}(x_i)\,p^{\beta}(y_j/x_i)\,p^{\beta}(x_i)\,[1 - q^{1-\beta}(\mathrm{x}\,X_i/Y_j)]\}$$

$$= (2^{1-\beta}-1)^{-1}\{\sum_{i=1}^{n} u^{\beta}(x_i)\left[\frac{(\lambda+\theta C_i)(2^{1-\beta}-1)}{\beta\,u(x_i)\,a_i}\right]^{\frac{\beta}{\beta-1}}[1 - \sum_{j=1}^{m} p^{\beta}(y_j/x_i)\,(1 - q^{1-\beta}(x_i/y_j))] - 1\}$$

$$= (2^{1-\beta}-1)^{-1}\sum_{i=1}^{n}\left[\frac{(\lambda+\theta C_i)(2^{1-\beta}-1)}{\beta\,u^{\beta}(x_i)\,a_i}\right]^{\frac{\beta}{\beta-1}} a_i - 1$$

$$C_\beta(\delta\,;q;U) = (2^{1-\beta}-1)^{-1}$$

$$\sum_{i=1}^{n}\left[\frac{(\lambda+\theta C_i)(2^{1-\beta}-1)}{\beta}\right]^{\frac{\beta}{\beta-1}}(a_i)^{\frac{1}{1-\beta}} - 1 \le C_\beta(\delta\,;U)$$

which is the required result.

Particular case: If by putting $u_i = 1$ in equation (1.12) and (1.16), the above equation reduces to the results obtained by Gupta and Arora [15].

III. Conclusion

Information theory, developed by Claude E. Shannon during World War II, defines the notion of channel capacity and provides a mathematical model by which one can compute it. The key result states that the Performance function of the DMC channel, as defined above, is given by the maximum of the mutual information between the input and output of the channel, where the maximization is with respect to the input distribution This paper focuses on computing the performance function for (r, s) entropy of a discrete memoryless communication channel using the utility function under single constraint. This result can also be extended to find the performance function of a discrete memoryless communication channel under multiple constraints.

REFERENCES

[1] K. Brijpaul and B. D. Sharma, "Direct methods for computing the performance of Discrete Memoryless communication channels," *Journal of Combinatorics, information and System sciences*, vol. 13, no. 3-4, pp. 86-105, 1988.

[2] S. Arimoto, "An Algorithm for computing the capacity of Arbitrary discrete memoryless channels," *IEEE Trans. on Inform. Theory*, vol. 18, pp. 14-20, 1972.

[3] R. E. Blahut, "Computation of channel capacity and rate- distortion functions," *IEEE Trans. Inform. Theory*, vol. 18, no. 4, pp. 460-473, 1972.

[4] B. D. Sharma and D. P. Mittal, "New non-additive measures of inaccuracy," *J. Math. Sci.*, vol. 10, pp. 122-133, 1975.

[5] C. E. Shannon, "The mathematical theory of Communication," *Syst. Tech. Jr.*, vol. 27, pp. 279-423, 1948.

[6] M. C. Cheng, "On the computation of capacity of DMC," *Information and Control*, vol. 24, pp. 292-298, 1974.

[7] S. Takano, "On a method of calculating the capacity of a discrete Memoryless channel," *Information and Control*, vol. 29, pp. 324-336, 1975.

[8] Y. Q. Zhao and J. A. Bose, "Rayleigh flat fading channel capacity," in *Proc. the 3rd Annual Communication Networks and Services Research Conference*, 2005, pp. 214-217.

[9] B. Meister and W. Oettli, "On the capacity of a discrete constant channel," *Information and Control*, vol. 11, pp. 341 – 351, 1967

[10] D. J. Costello and G. D. Forney, "Channel coding: the road to channel capacity," in *Proc. the IEEE*, June 2007, vol. 95, no. 6, pp. 1150-1177.

[11] K. Brijpaul and B. D. Sharma, "The performance function of a discrete Memoryless communication channel," *Journal of information and Optimization sciences*, vol. 11, pp. 327-336, 1990.

[12] M. Belis and S. A. Guiasu, "Qualitiative-Quantitative Measure of Information in Cubernetics Systems," *IEEE Trans. Inf. Theory*, pp. 593-594, 1968.

[13] N. S. Shimar and H. C. Taneja, "The weighted performance of a discrete memoryless channel," *Soochow Journal of mathematics*, vol. 21, no. 1, pp. 13-20, 1995.

[14] J. Havrda and F. Charvat, "Quantification Method of Classification Processes Concepts of structural β – Entropy," *Kybernetika*, vol. 3, pp. 30 – 35, 1967.

[15] P. Gupta and H. Arora, "Computation of (r, s)entropy of a DMCC under single and Multiple Constraints," *Indian Journal of Mathematics*, vol. 48, pp. 153-164, 2006.

H. D. Arora is serving as an associate professor and former head of Mathematics Department, Amity Institute of Applied Sciences, Amity University, Noida. An M.Phil in operations research, M.B.A. in Marketing, PhD in information theory, Dr. Arora is a gold medalist in M.Sc Operations Research. He has a vast experience of more than 13 years in the education industry. He has published more than 7 books in the field of engineering mathematics, business statistics and allied areas. Dr. Arora has made valuable contribution to research in Information theory and has a number of publications to his credit in National and International Journals of high repute.

Vijay Kumar received his M.Sc. degree in applied mathematics and M.Phil. degree in mathematics from the Indian Institute of Technology, Roorkee, in 1998 and 2000, respectively. He received his Ph.D. degree in software reliability and optimization (operational research) from the Department of operational research, University of Delhi. He is also working as an assistant professor in the Department of Mathematics at Amity School of Engineering & Technology, New Delhi, India. He has published several research papers in the area of software reliability, marketing models and optimization in international journals and conferences. His current research interests include software reliability growth modeling, optimal control theory and marketing models in the context of innovation diffusion theory. He is a Life Member of Society for Reliability Engineering, Quality and Operations Management (SREQOM).

Anjali Dhiman was born in Hoshiarpur, Punjab, on August 10, 1983. She obtained her M. Phil in mathematics at Global Open University, India in 2009. She did her M. Sc (Maths) at Hansraj College, University of Delhi in 2006 and B. Sc (Maths Honours) at Hansraj College, University of Delhi in 2004. Her areas of interest are Information theory, fuzzy information measures.

She is currently an assistant professor at the Ideal Institute of Mgmt & Tech, G.G.S. Indraprastha University, and Delhi. Prior to her appointment at IIMT, she was a lecturer at Chanderprabhu jain Institute of Mgmt & Tech, Delhi. She has presented research paper at International Conferences including ICQRITITM-2012, Delhi and ICRITO-2013.