ON OPTIMAL CHANNEL CAPACITY THEOREMS VIA VERMA INFORMATION MEASURE WITH TWO-SIDED INPUT IN NOISY STATE

Abstract

The capacity for which the value R_G or $R(a^*)$ is its lower bound is referred to as the optimal channel capacity. Through this communication, we attempt to study the additive Gaussian interference (S_1, S_2) that arises from the Gaussian channel's two-sided state information, which is non-causally known at both the transmitter and receiver and reliant on the channel's noise Z and input X. The author has highlighted his own entropy maximized using normal distribution in this investigation, which can be significant in the area of communication technology.

Keywords: Optimal channel capacity, Verma information measures, Gaussian channel etc.

1. INTRODUCTION

Shannon [5] and Kosnetsov-Tsybakov [4] conducted substantial research on side information channels. As an extension of the GP [3] theory, Cover-Chiang [2] obtained the capacity theorem in 2002 by analyzing the channel with two-sided and correlated state information that was not causally known at the transmitter and receiver. Open issues in information theory include the Cover-Chiang and Gaussian versions of the GP. Numerous more significant scholars who researched the issue in unique circumstances may be found in the literature and its references, in addition to Costa's **Dirty Paper Coding (DPC)**. The channel with one-sided additive interference, which is known as the transmitter, is examined. In this channel, noise and interference have a random joint distribution, and noise is influenced by both interference and channel input. Costa [1] looked at the channel's Gaussian variant with transmitter-accessible side information. At the receiver end, the side information is seen as additive interference. Costa demonstrated the channel's capability $\frac{1}{2}\log\left(1+\frac{P}{N}\right)$, which is the same for channels with no interference. According to Costa's theorem, the maximum rate corresponds to independent X and S_1, and U in form of linear combinatation of X and S_1. Now, we define F_C' as a subset of F_C with elements $f'(y, x, u, s_1)$ having the following properties as well as the properties mentioned before:

(i) X is a zero mean Gaussian random variable with the maximum variance P and independent of S_1.

(ii) The auxiliary random variable U takes the linear form $U = aS_1 + X$.

It is clear that the set F'_C and their marginal and conditional distributions are subset of corresponding F_C's.

Now, suppose $A_{opt}$is the covariance matrix, for the random variables (X, S_1, S_2, Z)in optimum situation of the channel, which is defined such that

$$A_{opt} = \begin{pmatrix} P & E_{0,1} & E_{0,2} & E_{0,3} \\ E_{0,1} & Q_1 & E_{1,2} & E_{1,3} \\ E_{0,2} & E_{1,2} & Q_2 & E_{2,3} \\ E_{0,3} & E_{1,3} & E_{2,3} & N \end{pmatrix}$$ and $D = \det(A_{opt})$. Here the minors of the determinant D

$i.e. D_1, D_3, D_5, D_7, D_9, D_{11}$ are having 3×3 determinants and the minors $D_2, D_4, D_6, D_8, D_{10}, D_{12}, D_{13}$ are having 2×2 determinants which, are as follows:

$$D_1 = \begin{vmatrix} Q_1 & E_{1,2} & E_{1,3} \\ E_{1,2} & Q_2 & E_{2,3} \\ E_{1,3} & E_{2,3} & N \end{vmatrix}$$ etc. and $$D_2 = \begin{vmatrix} Q_1 & E_{1,2} \\ E_{1,2} & Q_2 \end{vmatrix}$$ etc. Also, $$D_1^N = \begin{vmatrix} 1 & \rho_{s_1 s_2} & \rho_{s_1 z} \\ \rho_{s_1 s_2} & 1 & \rho_{s_2 z} \\ \rho_{s_1 z} & \rho_{s_2 z} & 1 \end{vmatrix}$$

Later in 2013, Verma [6, 7] presents the modified version of the new parametric measure of information which, is given by

$$V_a(P) = \sum_{i=1}^{n} \ln(1 + ap_i) - \sum_{i=1}^{n} p_i \ln p_i - \ln(1 + a), \qquad a > 0 \tag{1.1}$$

He also discuss the maximization of this measure (1.1), for the set of random variables $X_1, X_2, \dots \dots \dots, X_n$, by multivariate normal distribution $i.e.$

$$H(f) = \int_b^d \ln\big(1 + af(x)\big)\,dx - \int_b^d f(x) \ln f(x) dx$$

$$= \int_b^d \ln\left\{\frac{((2\pi)^n |K|)^{\frac{1}{2}} + ae^{-\frac{1}{2}(X-\mu)^T K^{-1}(X-\mu)}}{((2\pi)^n |K|)^{\frac{1}{2}}}\right\} dx$$

$$+ \frac{1}{2}\int_b^d f(x)(X-\mu)^T K^{-1}(X-\mu)\,dx$$

$$+ \int_b^d f(x) \ln((2\pi)^n |K|)^{\frac{1}{2}}\,dx$$

$$= \log\left\{((2\pi)^n|K|)^{\frac{1}{2}} + ae^{-\frac{n}{2}}\right\} - \frac{n}{2}\log((2\pi)^n e^{-1}|K|) + \frac{1}{2}\log((2\pi)^n|K|) \quad \text{bits.}$$

and provided further information, if the arithmetic average of the probability of error $i.e.$ $P_e^{(n)} \to 0$ for a sequence of $(2^{nR}, n)$ codes for a Gaussian channel with power constraint P, achievable rate R, channel capacity C with noise N then

$$R \leq C = \log\left(1 + \frac{P}{N}\right) \qquad \text{bits per transmission.}$$

2. OUR RESULTS

2.1 ACHIEVABLE RATE FOR MODIFIED VERMA [6, 7] CHANNEL

Theorem 2.1.1 For every, memoryless channels with discrete time and continuous alphabets when modified Verma [6, 7] entropy is maximized through Gaussian distribution, the achievable rate $R(a)$ is given by

$$R(a) = \log\left(\frac{P(P+Q_1+N)}{PQ_1(1-a)^2 + N(P+a^2Q_1)}\right)$$

and $$\max_a R(a) = R(a^*) = \log\left(1 + \frac{P}{N}\right) \text{ where } a^* = \frac{P}{P+N}.$$

Proof. The above result shows the achievable rate for Verma channel

For this purpose, first of all the Verma channel capacity for memoryless channel with time discrete and continuous alphabets is given by

$$C_{verma} = \max_{f(u,x|s_1)} [I(U;Y) - I(U;S_1)]$$

Where the maximum is over all $f(y,x,u,s_1)$'s in F_C. Since $F_C' \subseteq F_C$ we have

$$C_{verma} \geq \max_{f(u,x|s_1)} [I(U;Y) - I(U;S_1)]$$

$$= \max_{f'(u,x|s_1)f'(x|s_1)} [I(U;Y) - I(U;S_1)]$$

$$= \max_a [I(U;Y) - I(U;S_1)] \qquad (2.2.1)$$

$$R(a) = \log\left(\frac{P(P+Q_1+N)}{PQ_1(1-a)^2 + N(P+a^2Q_1)}\right)$$

and $$\max_a R(a) = R(a^*) = \log\left(1 + \frac{P}{N}\right)$$

where $a^* = \frac{P}{P+N}$. This completes the proof.

2.2 UPPER BOUND OF THE MODIFIED VERMA [6, 7] CHANNEL CAPACITY

Theorem 2.2.1 For every, memoryless channels with discrete time and continuous alphabets having the side information S_1 the upper bound for the modified Verma channel capacity is given by

$$C_{verma} = \log\left(1 + \frac{P}{N}\right)$$

Proof. The proof is completed as follows:

Since, from the given information, we have

$$\begin{aligned} I(U;Y) - I(U;S_1) &= -H(U|Y) + H(U|S_1) \\ &\leq -H(U|Y,S_1) + H(U|S_1) \\ &= I(U;Y,S_1) \\ &\leq I(X;Y,S_1) \end{aligned}$$

To solve the problem, we use the Markov chain $U \rightarrow XS_1 \rightarrow Y \; f(y,x,u,s_1)$

$$\begin{aligned} C_{costa} &= \max_{f(u,x|s_1)}[I(U;Y) - I(U;S_1)] \\ &\leq \max_{f(x|s_1)}[I(X;Y|S_1)] \\ &= \max_{f(x|s_1)}[H(Y|S_1) - H(Y|X,S_1)] \\ &= \max_{f(x|s_1)}[H(X+Z|S_1) - H(Z|X,S_1)] \\ &\leq \max_{f(x|s_1)}[H(X+Z) - H(Z)] \\ &= \log\left(1 + \frac{P}{N}\right) \end{aligned}$$

Now, transmission in the absence of interference we can apply the condition $U = X$ and

$$C = \log\left(1 + \frac{P}{N+Q_1+Q_1'}\right)$$

Now, transmission when S_1 is known, then $U = X + aS_1$ and $C = \log\left(1 + \frac{P}{N+Q_1'}\right)$

Also, transmission when S_1 and S_1' are known, then $U = X + aS_1 + bS_1'$ and $C = \log\left(1 + \frac{P}{N}\right)$.

2.3 LOWER BOUND OF THE MODIFIED VERMA [6, 7] CHANNEL CAPACITY

Theorem 2.3.1 For every, time discrete memoryless channels with two side informations S_1 and S_2 then the minimal capacity for the modified Verma channel is given by

$$R_G = \log\left(1 + \frac{\left[\sigma_x\left(1-\rho_{xs_1}^2\right)-\sigma_z\left(\rho_{xs_1}\rho_{s_1z}-\rho_{xz}\right)\right]^2\left(1-\rho_{s_1s_2}^2\right)}{\sigma_Z^2\left(\left(1-\rho_{xs_1}^2\right)D_1^N-\left(\rho_{xs_1}\rho_{s_1z}-\rho_{xz}\right)^2\left(1-\rho_{s_1s_2}^2\right)\right)}\right)$$

and $$R_G = \log\left(1 + \frac{\left(1+\rho_{xz}\frac{\rho_Z}{\sigma_X}\right)^2\sigma_X^2}{1-\rho_{Xz}^2-\rho_{s_1z}^2\sigma_Z^2}\right) \text{ when } \rho_{xs_1} = \rho_{s_1s_2} = 0.$$

Proof. On using the extension of Cover-Chiang capacity [2] theorem for random variables with continuous alphabets, the capacity of modified Verma channel may be expressed as

$$C = \max_{f(u,x|s_1)}[I(U;Y,S_2) - I(U;S_1)].$$

Where the maximum is over all distributions $f(y,x,u,s_1,s_2) \in F$. But $F \subseteq F^*$ then, we have

$$C \geq \max_{f^*(u,x|s_1)f'(x|s_1)}[I(U;Y,S_2) - I(U;S_1)]$$

$$= \max_a[I(U;Y,S_2) - I(U;S_1)].$$

Now, define $$R(a) = I(U;Y,S_2) - I(U;S_1).$$

Then we have $$C \geq \max_a R(a) = R(a^*).$$

To compute $$I(U;Y,S_2) = H(U) + H(Y,S_2) - H(U,Y,S_2)$$

and $$I(U;S_1) = H(U) + H(S_1) - H(U,S_1)$$

For $$H(Y,S_2) = \log\left((2\pi e)^2 \det\left(cov\,(Y,S_2)\right)\right)$$

Where $cov(Y, S_2) = [e_{ij}]_{2\times 2}$

and $e_{11} = P + Q_1 + Q_2 + N + 2E_{0,1} + 2E_{0,2} + 2E_{1,2} + 2E_{0,3} + 2E_{1,3} + 2E_{2,3}$,

$e_{11} = e_{21} = Q_2 + E_{0,2} + E_{1,2} + E_{2,3}$, $e_{22} = Q_2$. After manipulations we get the following result

$\det(cov(Y, S_2)) = D_2 + D_6 + D_8 + 2D_{10} - 2D_{12} - 2D_{13}$.

Hence, $H(U, Y, S_2) = \log((2\pi e)^3 \det(cov(U, Y, S_2)))$ and $cov(U, Y, S_2) = [e_{ij}]_{3\times 3}$.

Also, $e_{11} = P + a^2 Q_1 + 2aE_{0,1}$,

$e_{12} = e_{21} = P + aQ_1 + (a + 1)E_{0,1} + aE_{1,2} + aE_{1,3} + E_{0,2} + E_{0,3}$,

$e_{13} = e_{31} = aE_{1,2} + E_{0,2}$, $e_{23} = e_{32} = Q_2 + E_{0,2} + E_{1,2} + E_{2,3}$,

$e_{22} = P + Q_1 + Q_2 + N + 2E_{0,1} + 2E_{0,2} + 2E_{1,2} + 2E_{0,3} + 2E_{1,3} + 2E_{2,3}$ and $e_{33} = Q_2$.

After manipulations we get the following result

$\det(cov(U, Y, S_2)) = a^2 D_1 + 2a(a-1)D_3 + 2(a-1)D_7 + (a-1)^2 D_9 - 2aD_{11} - 2D_5$.

Now, for $H(S_1)$ and $H(U, S_1)$ we have $H(S_1) = \log((2\pi e)Q_1)$ and

$H(U, S_1) = \log((2\pi e)^3 \det(cov(U, S_1)))$ where

$cov(U, S_1) = \begin{pmatrix} a^2 Q_1 + P + 2aE_{0,1} & aQ_1 + E_{0,1} \\ aQ_1 + E_{0,1} & Q_1 \end{pmatrix}$ and $\det(cov(U, S_1)) = D_4$. Now, from

$$I(U; Y, S_2) = H(U) + H(Y, S_2) - H(U, Y, S_2)$$

and
$$I(U; S_1) = H(U) + H(S_1) - H(U, S_1)$$

we acquire the following result

$$R(a) = \log\left[\frac{D_4\{D_8 + D_2 + D_6 + 2D_{10} - 2D_{12} - 2D_{13}\}}{Q_1\{(a-1)^2 D_9 + a^2 D_1 + 2a(a-1)D_3 + 2aD_{11} + 2(a-1)D_7 + D_5\}}\right].$$

The optimum value of a corresponding to maximum value of $R(a)$ can be evaluated

$$a^* = \frac{(D_3+D_9)-(D_7+D_{11})}{2D_3+D_1+D_9}.$$

Finally, we conclude that in brevity

$$R_G = \log\left(1 + \frac{\left[\sigma_x(1-\rho_{xs_1}^2)-\sigma_z(\rho_{xs_1}\rho_{s_1z}-\rho_{xz})\right]^2(1-\rho_{s_1s_2}^2)}{\sigma_z^2\left((1-\rho_{xs_1}^2)D_1^N-(\rho_{xs_1}\rho_{s_1z}-\rho_{xz})^2(1-\rho_{s_1s_2}^2)\right)}\right).$$

Setting $\rho_{xs_1} = \rho_{s_1s_2} = 0$ in $C_{verma} = \max_a[I(U;Y) - I(U;S_1)]$ then $D_1^N = 1 - \rho_{xz}^2$. Hence,

$$R_G = \log\left(1 + \frac{\left(1+\rho_{xz}\frac{\rho_z}{\sigma_x}\right)^2\sigma_X^2}{1-\rho_{Xz}^2-\rho_{S_1z}^2\sigma_z^2}\right).$$

Which is the lower bound of modified Verma channel capacity and this completes the proof.

CONCLUSION

The following results emerge from the above investigation:

(i) Optimal channel capacity is the capacity for which the value of channel capacity equalize with its lower bounds.

(ii) It is observe that the lower and upper bounds of the Verma channel capacity coincide and therefore the modified Verma channel capacity is equal to

$$\log\left(1 + \frac{P}{N}\right)$$

References

1. Costa, M. (1983): Writing on dirty paper (corresp.), Information Theory, IEEE Transactions on, vol. 29, no. 3, pp.439-441.

2. Cover, T. M. and Chiang, M. (2002): Duality between channel capacity and rate distortion with two-sided state information, Information Theory, IEEE Transactions on, vol. 48, no. 6, pp. 1629-1638.

3. Gel'fand, S. I. and Pinsker, M. S. (1980): Coding for channel with random parameters, Probl. Contr. Inform. Theory, vol. 9, no. 1, pp. 19-31.

4. Kosnetsov, A. V. and Tsybakov, B. S. (1974): Coding in a memory with defective cells, Probl. Pered. Inform. vol. 10, no. 2, pp. 52-60.

5. Shannon, C. E. (1958): Channels with side information at the transmitter, IBM Journal of Research and Development, vol. 2, no. 4, pp. 289-293.

6. Verma, R. K., Dewangan, C. L. and Jha, P. (2012): An Unorthodox Parametric Measures of Information and Corresponding Measures of Fuzzy Information, Int. Jour. of Pure and Appl. Mathematics, Bulgaria, Vol. 76, No. 4, pp. 599-614.

7. Verma, R. K. and Verma, Babita (2013): A New Approach in Mathematical Theory of Communication (A New Entropy with its Application), Lambert Academic Publishing, Germany.