Total Sum Capacity Computation For Discrete Memoryless Multi-Access Channels

Archit Kumar(B15405), Rajanish Kumar Upadhyay(B15126)

SCEE, IIT MANDI

November 30, 2018

Overview

Problem Statement

To understand, implement and numerically compare novel algorithms to compute inner and outer bounds for the MAC capacity region.

Solution Approach

1. Implementation of BA Algorithm for
   - Point-to-point channels, binary input alphabet(solved using the Lagrangian method).
   - Point-to-point channels, n-nary input alphabet.
2. Generalisation to DMC-MAC.
3. Implementation of algorithms to compute inner and outer bounds for the MAC capacity region.
4. Implementation of a randomized algorithm for sum-capacity computation of MAC Channels.
5. Comparision of performance of existing vs new algorithm.

Applications

- This project provide an estimation of maximum and minimum capacity of those real-world channels, which can modelled as DMC MACs.
- Any breakthrough in calculation of capacity region for DMC-MACs will mean efficient evaluation of multi access schemes.
- Multi access channel model has applications in the uplink of the cellular networks.
- Theoretical advances in solving non-convex optimization problems having similar structure.
- Generation of better encoding and decoding schemes.

- A **communication channel** is a medium to convey information.
- A **discrete** channel is a system consisting of
  1. an input alphabet X
  2. an output alphabet Y
  3. a probability transition matrix $p(y|x)$

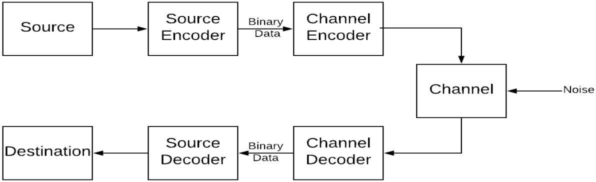


Figure: Shannon's Communication Model.

Channel Definitions

- The channel is said to be **memoryless** if:

$$P_{(Y^n|X^n)}(y^n|x^n) = \prod_{i=1}^{n}(P_{(Y|X)}(y[i]|x[i])) \tag{1}$$

- **Mutual information**: Information provided about the occurrence of an event by some other event.
- **Multi-terminal Channel**: A channel with multiple terminals.

- **Multi-access Channel**: A type of multi-terminal channel, multiple senders transmit multiple messages over a shared medium to one destination.

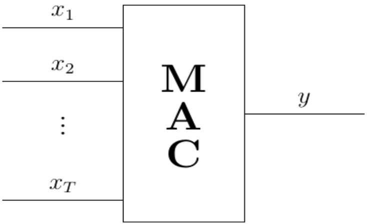


Figure: Schematic representation of a simple MAC Channel.

Shannon coding theorem

- **Noisy-channel coding theorem**: It establishes that for any given degree of noise in a channel, we can communicate discrete data nearly error-free up to a computable maximum rate.
- **Capacity**[1]: The highest rate in bits per channel use at which information can be sent, error-free.
- **Sum Capacity**: The sum capacity in a multi-user channel is the maximum aggregation of all the users data rates.

Example

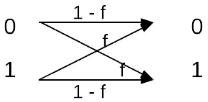


Figure: An elementary channel.

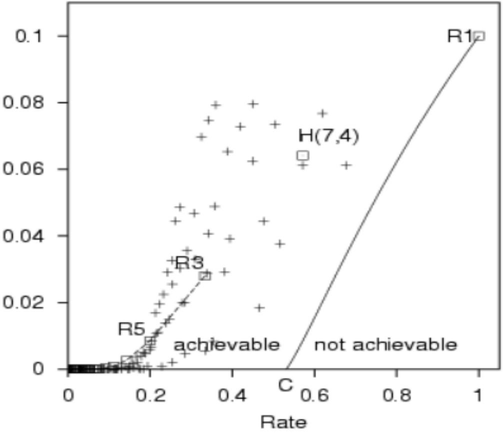


Figure: Shannon limit.

Optimisation Problem

- Capacity region determination for multi-terminal channels has attracted much attention in information theory.
  1. In many cases, single-letter representations are not known.
  2. In cases where such representations have been found, such as the discrete memoryless multiple-access channel (DMC-MAC), evaluation is difficult.
- Specifically, computation of the boundary of the capacity region for DMC-MACs is a non-convex constraint optimization problem.

Single User Case

- For single-user channel, capacity can be numerically approximated to arbitrary precision using Blahut Arimoto(BA) algorithm.
  1. Most common implementation: Expectation-Maximization(EM) paradigm
  2. Interior Point Methods
  3. Cutting plane Algorithms

Blahut-Arimoto Algorithm

For a DMC with transition matrix $p(y|x)$, input distribution $r(x)$ and reverse transition matrix $q(x|y)$, the capacity is given by:

$$C = \sup_{r>0} \max_{q} \sum_{x} \sum_{y} r(x)p(y|x) \log q(x|y)/r(x) \tag{2}$$

where maximisation is taken over all q such that
$q(x|y) = 0$ iff $p(y|x) = 0$

- <u>E-step</u>
  In the expectation step, we calculate $q(x|y)$ from the value of $r(x)$ computed in the last step; this is the inference we draw after observing the current input distribution and include in our prediction of what might have caused the current output.

$$q^k(x|y) = \frac{r^{(k-1)}(x)p(y|x)}{\sum_{x'} r^{(k-1)}(x')p(y|x')} \tag{3}$$

- <u>M-step</u>
  In the maximization step, $r(x)$ is calculated using the updated value of $q(x|y)$. This is the change we incorporate in our input distribution, after observing the behavior of the channel.

$$r^k(x) = \frac{\prod_y q^k(x|y)^{P(y|x)}}{\sum_{x'} \prod_y q^k(x'|y)^{p(y|x')}} \tag{4}$$

EM - continued...

After having adjusted the input distribution, C is calculated using (2).

- if increment is less than certain threshold we terminate the algorithm.
- else we continue with our E and M steps.

Generalized BA

- For a MAC, this convexity is missing
- Recent results [2] have shown that the Kuhn–Tucker condition is:
  - either sufficient for optimality in a MAC
  - or the channel can be decomposed into subchannels for which the Kuhn–Tucker condition is sufficient for optimality. In the this case, at least one subchannel has an optimal distribution that achieves the capacity of the original channel.
- The BA Algorithm is generalized for a MAC channel. We use ideas similar to that used for the single-user case. The optimization is taken over all input distributions (or users).

Total capacity (2) can be computed[3] as:

$$C_{total} = \max_{P_1(X_1)P_2(X_2)...P_m(X_m)} I(X_{\mathcal{M}}; Y) \tag{5}$$

where $X_{\mathcal{M}} = (X_1, X_2, ..., X_m)$

Mutual Information is defined as:

$$I(x;Y) = D(P(Y|x)||P(Y)) \tag{6}$$

where $D(\cdot||\cdot)$ is the Kullback-Leibler distance

$$D(p(Y)||p'(Y)) = \sum p(Y) \log \frac{p(y)}{p'(y)} \tag{7}$$

So (6) simplies into :

$$I(x;Y) = \sum_y p(y|x) \log \frac{p(y|x)}{p(y)} \tag{8}$$

The information function can be extended to a set of variables $X_i \in \mathcal{X}_\rangle$, namely

$$I(x_1, x_2, ..., x_M; Y) = D(P(Y|x1, x2, ..., x_M)||P(Y)). \tag{9}$$

Generalized BA – continued...

For $S \subset \mathcal{M} \doteq 1, 2, ..., M$, let $x_S = x_i : i \in S$ and denote the marginalisation of $I(x_1, x_2, ..., x_M; Y)$ to S by

$$I_S(x_s; Y) = \sum_{x_{\tilde{S}}} P_{\tilde{S}}(x_{\tilde{S}}) I(x_{\mathcal{M}}; Y) \tag{10}$$

where $\tilde{S} \doteq \mathcal{M} \setminus S$

The marginalized mutual information for each user becomes:

$$\begin{aligned} I_m(x_m; Y) &= \sum_{x_{\tilde{m}}} P_{\tilde{m}}(x_{\tilde{m}}) I(x_{\mathcal{M}}; Y) \\ &= \sum_{x_{\tilde{m}}} P_{\tilde{m}}(x_{\tilde{m}}) \sum_{y} p(y|x_{\mathcal{M}}) \log \frac{p(y|x_{\mathcal{M}})}{p(y)} \end{aligned} \tag{11}$$

This marginalized mutual information is then used to adjust the input distributions of each user so that the capacity is maximized.

$$P_m^{r+1}(x_m) = P_m^r(x_m) \frac{f(I_m(x_m; Y))}{\sum_{x_m} P_m^r(x_m) f(I_m(x_m; Y))} \tag{12}$$

where f[4] is the exponential function.

Expectation-Maximisation

- E-step: In the expectation step, we compute the marginal mutual information for each user. This is done by using the current overall mutual information as input in the equation (11).
- M-step: In the maximization step, we compute the adjusted input distributions for each user using the results of the E-step as input to the the equation (12).
- After these two steps we compute the current capacity by using the mutual information equation (9).
  - if the increment is less than a certain threshold we terminate the algorithm
  - else we continue with our E and M steps.

Inner and Outer Bound

- The capacity region for DMC-MACs is non-convex and solving it is an optimisation problem.

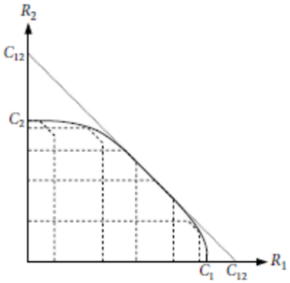


Figure: Capacity Region

In the figure C_ij is the sum capacity.

- It cannot be solved in polynomial time and hence approximations are found by solving a approximate convex problem and projecting the solution of the relaxed problem to the original.

Inner and Outer Bound

- The optimisation problem can be solved using any of the methods including Interior point method, cutting plane method, randomised algorithms etc.
- The randomised algorithms are shown to be near optimal and hence they will be our focus for this project.
- During projection the criteria of minimum divergence is very important.
- This gives us bound on the capacity region. However exact capacity region can be calculated for
  1. the sub-class of channels with identical inner and outer bounds
  2. sub-class of channels with tight lower bound and strict upper bound.

Work Done till now

- Literature review.
- Analysis of earlier implementation
- Implementation of BA Algorithm for single user arbitary alphabet.
- Implementation of generalised BA Algorithm for n-user arbitrary alphabet

API implemented

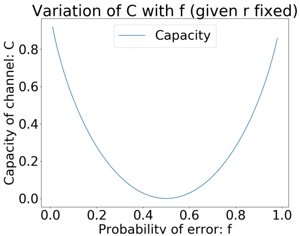
- **init:**
  arguments: inputDistribution, transmissionMatrix)
  This method initializes the input distribution and transmission matrix with the input provided.
- **decToBin**:
  arguments: x, repSize
  This method converts a decimal number to its binary representation of fixed size.
- **binToDec**:
  arguments: rep
  This method converts a binary representation to its decimal equivalent.

- **capacity**:
  arguments: outputDistribution, inputDistribution, transmissionMatrix
  This method computes the capacity of the channel from the input distribution, the output distribution and the transmission matrix. This is done by using the definition of mutual information.
- **marginalization**:
  arguments: mutuaInfo, outputDistribution, inputDistribution, transmissionMatrix
  This method marginalizes the mutual information to each symbol of each input distribution.
- **probAdjustment**:
  arguments: transmissionMatrix, inputDistribution, mutualInfo)
  This method adjusts the probability distribution of the input using the marginalized mutual information and transmission matrix.

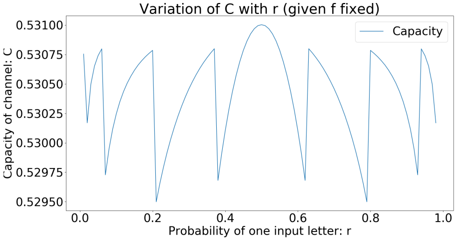
- **expectationMaximization**:
  arguments: inputDistribution, transmissionMatrix, error, iterations
  This implements the expectation maximization of the iterative Blahut-Arimoto algorithm. It is fed with an error to which we wish to regulate and the maximum number of iterations that we wish this procedure to run.

Results

- These are the results for binary symmetric channel.



Figure



Figure

- We have also verified our implementation for Point to Point channel (n-ary alphabet).

- These are the results for n-user case.

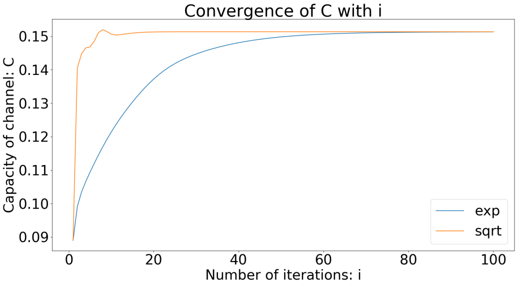


Figure: Caption

Input alphabet distributions:

$$P_X(X_1) = \begin{bmatrix} 0.3 & 0.7 \end{bmatrix}$$
$$P_X(X_2) = \begin{bmatrix} 0.4 & 0.6 \end{bmatrix}$$

Results

Transition Matrix:

$$P_{Y|X_M}(Y|X_M) = \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.7 & 0.2 & 0.1 \\ 0.5 & 0.1 & 0.4 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}$$

The graph shows that different functions used in (12) lead to different convergence rates. Square root function being faster than exponential function.

Future Plan

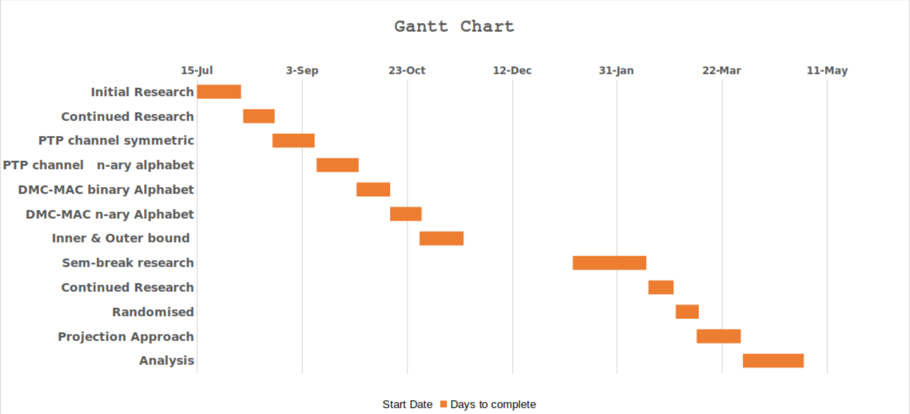


Figure: Gantt Chart

Conclusion

We have verified our code for Binary symmetric and Point to point channel. We have also implemented DMC-MAC for binary alphabet and DMC-MAC for n-ary Alphabet. Currently we are researching about calculating inner and outer bound of capacity region. Next, randomised algorithm to calculate the capacity will be implemented followed by analysis of minimum divergence criteria.

Thank You

References

R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Information Theory*, vol. 18, pp. 460–473, 1972.

Y. Watanabe and K. Kamoi, "The total capacity of multiple access channel," in *Int. Symp. Information Theory, Lausanne,Switzerland*, p. 308, IEEE, 2002.

M. Rezaeian and A. Grant, "Computation of total capacity for discrete memoryless multiple-access channels," *IEEE Transactions on information theory*, vol. 50, no. 11, pp. 2779–2784, 2004.

M. Rezaeian and A. Grant, "A generalization of arimoto-blahut algorithm," in *Information Theory, 2004. ISIT 2004. Proceedings. International Symposium on*, p. 181, IEEE, 2004.