Forty-Eighth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
September 29 - October 1, 2010

Numerical Computation of Sum Capacity for Discrete Multiple Access Channels with Causal Side Information at the Transmitter

Assadallah Sahebalam
Department of Electrical Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
a_sahebalam@yahoo.com

Mohammad Osmani Bojd
Department of Electrical Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
mohammad.osmani@gmail.com

Ghosheh Abed Hodtani
Department of Electrical Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
ghodtani@gmail.com

Abstract— **The computation of capacity for discrete memoryless channels can be efficiently solved using the Arimoto-Blahut (AB) iterative algorithm. However, the extension of this algorithm to compute the capacity for channels with causal side information (SI) at the transmitter is not straightforward, because generally it is hard to evaluate the rates and capacities having auxiliary random variables. In this paper, we use an alternative reformulation of differential evolution optimization method to compute the capacity for channels with causal side information and introduce efficient algorithm to compute the capacity of these channels. Also we extend this algorithm to compute the sum capacity of discrete multiple access channels with causal side information at the transmitter.**

Keywords- Arimoto-Blahut algorithm, differential evolution algorithm, causal and non causal side information, discrete multiple access channel

I. INTRODUCTION

A numerical algorithm for the computation of the capacity for a discrete memoryless channel has been introduced in [1], [2], [3] and [4]. This celebrated algorithm is known as the Arimoto-Blahut algorithm. This algorithm has been successfully extended to the calculation of the sum capacity of discrete multiple access channels [5], [6], [7]. In [8] an algorithm for computing channel capacity and rate-distortion with non causal side information is introduced.

A. Our work

In this paper, first, we modify the Arimoto-Blahut algorithm [1], [2] using the improved version of differential evolution optimization method [9] for channel with causal side information at the transmitter. And then, we apply the modified Arimoto-Blahut algorithm and the partial improvement of the method used in [5], [6] and [7], to compute the sum capacity of discrete multiple access channels with causal side information at the transmitter and illustrate the algorithm with three examples.

B. Notation

We use **x**, **y** and **s** to denote the input vector $(x_1 \dots x_n)$, output vector $(y_1 \dots y_n)$ and state vector $(s_1 \dots s_n)$, respectively, and allow $\mathbf{x}^i$, $\mathbf{y}^i$ and $\mathbf{s}^i$ to denote $(x_1 \dots x_i)$, $(y_1 \dots y_i)$ and $(s_1 \dots s_i)$, respectively. $P(s)$, $P(x)$, $P(x|y)$ represent probability distribution functions with the specific values at the i^{th} channel use, $P(s_i)$, $P(x_i)$ and $P(x_i|y_i)$. Also, $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{S}$ represent the input alphabet, output alphabet and state alphabet, respectively.

C. The Arimoto-Blahut Algorithm

Using the lemma in [10], we can write the following expressions,

$$P^*(x|y) = \frac{P(x)P(y|x)}{\sum_x P(x)P(y|x)} \quad (1)$$

$$P^*(x) = \frac{e^{\sum_y P(y|x)\log p(x|y)}}{\sum_x e^{\sum_y P(y|x)\log p(x|y)}} \quad (2),$$

and we start with a guess of maximizing distribution $P(x)$ and find the best conditional distribution $P^*(x|y)$. The closed form expression for $P^*(x)$ is as follows.

$$P^{t+1}(x) = \frac{P^t(x)e^{D^t(P(y|x)||P(y))}}{\sum_x P^t(x)e^{D^t(P(y|x)||P(y))}} \quad (3),$$

where $D^t(\cdot)$ and $P^t(x)$ are Kullback-Leibler distance and input distribution after 't' iteration, respectively. For simplicity, we rewrite (3) as:

$$P^{t+1}(x) = \frac{P^t(x)e^{D^t}}{\sum_x P^t(x)e^{D^t}} \quad (4)$$

For accelerating the convergence, we use a coefficient, η, in (4) [11], [12]:

$$P^{t+1}(x) = \frac{P^t(x)e^{\eta D^t}}{\sum_x P^t(x)e^{\eta D^t}} \quad (5)$$

where,

$$1 \le \eta \le \frac{1}{1-\sum_y \min_X P(y|x)} \quad (6).$$

Then the channel capacity is computed with following expression:

$$C = \sum_{x,y} P(x)P(y|x)\log\frac{P(x|y)}{P(x)} \qquad (7)$$

D. Causal Side Information

We begin with a brief review of discrete memoryless channel with causal side information, along with the appropriate definitions. Consider the channel depicted in Fig. 1. The channel is discrete with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, and state alphabet $\mathcal{S}$, all of which are finite sets. The channel states are i.i.d with distribution $P(\mathbf{s})$ and independent of the input sequence. Furthermore, given the states, the channel is memoryless with transition distribution $P(\mathbf{y}|\mathbf{x},\mathbf{s})$. Hence, the conditional distribution of **y** and **s** given **x** can be written as:

$$P(\mathbf{y},\mathbf{s}|\mathbf{x})=P(\mathbf{s})P(\mathbf{y}|\mathbf{x},\mathbf{s})=\prod_{i=1}^{n} P(s_i)P(y_i|x_i,s_i) \qquad (8)$$

The state sequence in this model plays the role of 'side information'. The encoder maps the message $\omega \in \{1,2,\dots,2^{nR}\}$ into $\mathcal{X}^n$ using functions

$$\mathbf{x}^i = f_i(\omega,\mathbf{s}^i),\ \ 1 \le i \le n \qquad (9),$$

where $\mathbf{s}^\mathbf{i} = (s_1,s_2,\dots,s_i)$ is the state information at the encoder before the i^{th} transmission; namely, the encoder operates causally with respect to the state sequence. The receiver decodes the message ω from received vector **y** as $\hat{\omega}$=g(**y**), where **y** denotes the whole received sequence. This constitutes a code. The notions of achievable rates of transmission and capacity are defined analogously with the ordinary discrete memoryless channel.

Shannon [13] showed that this capacity is equal to the regular capacity of the derived discrete memoryless channel shown in Fig. 2. The input alphabet of the derived channel, denoted by $\mathcal{T}$, is the set of all possible mappings $t: \mathcal{S} \to \mathcal{X}$, which we refer to as strategies or strategy functions. We may describe each strategy $t_j(s) \in \mathcal{T}$ by the $|\mathcal{S}|$ -tuple $(x_j^1, x_j^2, \dots, x_j^{|\mathcal{S}|})$, i.e, $t_j(s) \in x_j^s$ for $s = 1,\dots,|\mathcal{S}|$ and j=1,…,n. Therefore, $|\mathcal{T}| = |\mathcal{X}|^{|\mathcal{S}|}$. The device shown in Fig. 2, called "transducer", just takes t and s as inputs and produces x as its output by computing $x = t(s)$. Therefore, the boxed section constitutes a discrete memoryless channel defined by,

$$P(y|t) = \sum_s P(s)P(y|x(t,s),s) \qquad (10),$$

and also,

$$P(\mathbf{y}|\mathbf{t})=\prod_{i=1}^{n} P(y_i|t_i) \qquad (11)$$

thus, the capacity with causal side information at the transmitter is given by,

$$C = \max_{P(t)} I(t;y) \qquad (12),$$

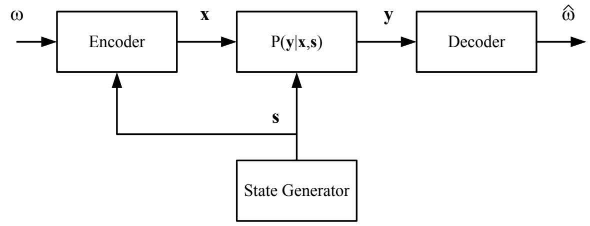

Figure 1. Channel configuration

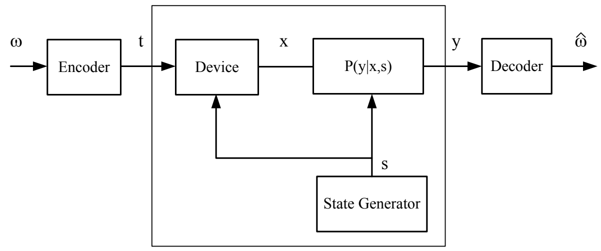

Figure 2. Shannon's Equivalent channel

where the maximization is taken over the distribution $P(t)$ of the random strategy variable $t_j \in \mathcal{T}$. Note that:

1- The strategies employed by the derived channel are functions of the current state alone. The operational meaning of this structure is that it is possible to achieve capacity using a code of the form $x_i = f_i(\omega, S_i)$.

2- At most $|\mathcal{Y}|$ of the strategies need be given positive probability in order to achieve capacity. Therefore if the cardinalities of $\mathcal{X}$ and $\mathcal{Y}$ are equal (in particular if $\mathcal{X} = \mathcal{Y}$), then at most $|\mathcal{X}|$ of the strategies need be given positive probability.

In general, one does not know in advance which function of the strategies is to be used to achieve capacity. Fading coefficients, channel interference levels, states of a markov channel and channel gains are some examples and adaptive rate/power control over Rayleigh fading channels, MIMO beam-forming, precoding and multi-tone water filling are some scenarios and applications for this case [14], [15].

E. Differential Evolution Optimization Algorithm

Differential evolution algorithms are commonly used for global optimization. It has emerged as one of the techniques most favored by engineers for solving continuous optimization problems [9], [16]. This method has several attractive features. Besides being an exceptionally simple evolutionary strategy, it is significantly faster and robust for solving numerical optimization problems and is more likely to find the true global optimum function. Also, it is worth mentioning that differential evolution has a compact structure with a small computer code and has fewer control parameters in comparison to other evolutionary algorithms. Differential evolution has been successfully applied to a wide range of problems [17], [18], [19], [20], [21], [22], [22], [23], [24], [25].

The working with differential evolution depends on the manipulation and efficiency of three main steps; mutation, reproduction and selection. We give an alternative reformulation of differential evolution optimization method to compute the capacity for channels with causal side information at the transmitter and extend it to sum capacity of discrete multiple access channels with causal side information at the transmitter.

II. The Proposed Algorithm

A. Computing the capacity for channel with causal side information at the transmitter

As we said before in introduction, for channel in Fig. 3 we need $|\mathcal{X}|^{|\mathcal{T}|.|\mathcal{S}|}$ functions to take t and s as inputs and produce x as its output by $x_i = f_i(\omega, s_i)$. For example, the channel in which $\mathbf{s} = (0,1)$ is the state vector, $\mathbf{x} = (0,1)$ is the input vector and $\mathbf{y} = (0,1)$ is the output vector, strategies with at most binary alphabet need be given to achieve capacity; therefore, we have 16 ($2^{2\times 2}$) functions. Fig. 3 shows the two types of them as a graph. The first figure (a) provides xor function and the other provides xnor function.

We can make a matrix for every function. For xor function in the above example, this matrix is A and for xnor function this matrix is B.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{13}$$

$$B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{14}$$

The rows of matrix are possible values for auxiliary random variable 't' and state 's' and columns are possible values for input 'x'. For channel with cardinality equal to 3 in the input and output of channel, the number of functions and matrixes can be 81 or 729. The former is related to channel the strategy of which has two alphabets and the latter for the maximum number of alphabet, i.e. three.

We name every row of matrix with a random vector $X_i, i = 1, \dots, |\mathcal{S}| \times |\mathcal{T}|$, these random vector are members of a set which consists of all possible permutations of one in $V = (1,0,\dots,0)$. The dimension of V and the number of members of this set is $|\mathcal{X}|$. In previous example for xor function this set is $\{(0,1),(1,0)\}$ and $X_1 = (1,0)$, $X_2 = (0,1)$, $X_3 = (0,1)$ and $X_4 = (1,0)$.

Only one or some functions achieve capacity, but which of them? We should choose the best function. The set $\{X_1, \dots, X_{|\mathcal{S}|\times|\mathcal{T}|}\}$ optimization leads to optimize functions. $\{X_1, \dots, X_{|\mathcal{S}|\times|\mathcal{T}|}\}$ plays the role of a population of $|\mathcal{S}| \times |\mathcal{T}|$ candidate solutions in differential evolution algorithm. Optimizing is done in parallel form in three main steps; mutation, reproduction and selection, but we need some changes in these steps.

This algorithm for computing the capacity for channel with causal side information at the transmitter goes through these steps:

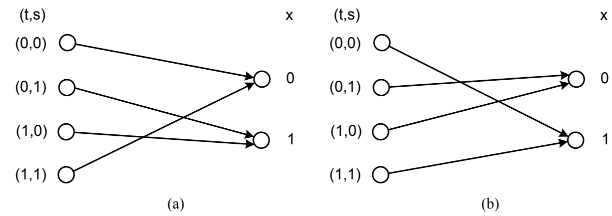

Figure 3. Two types of functions from t, s to x: a) x=xor(t,s), b) x=xnor(t,s)

Step 1: The first step is the random initialization of the parent population ($X_i, i = 1, \dots, |\mathcal{S}| \times |\mathcal{T}|$). Randomly generate a population of $|\mathcal{S}| \times |\mathcal{T}|$ arrays. The size of every X_i is $|\mathcal{X}|$ and then form matrix X as the following form,

$$X_G = \begin{bmatrix} X_1 \\ X_2 \\ . \\ . \\ . \\ X_{|\mathcal{S}|\times|\mathcal{T}|} \end{bmatrix} \tag{15}$$

Matrix X makes a function from t and s to x.

Step 2: Calculating the objective function value C(X) using (10), (11), (12) and the Arimoto-Blahut algorithm.

Step 3: Generating perturbed individual $V_i (i = 1, \dots, |\mathcal{S}| \times |\mathcal{T}|)$. This array is randomly generated from the population and then form perturbed matrix V as the following form,

$$V_{G+1} = \begin{bmatrix} V_{1,G+1} \\ V_{2,G+1} \\ . \\ . \\ . \\ V_{|\mathcal{S}|\times|\mathcal{T}|,G+1} \end{bmatrix} \tag{16}$$

Step 4: Recombining each target array X_i with perturbed individual generated in step 3 to generate a trial vector U_i using the following form,

$$U_{i,G+1} = \begin{cases} V_{i,G+1} & \text{if } rand_i < C_r \text{or } i = k \\ X_{i,G} & \text{otherwise} \end{cases} \tag{17},$$

where $k \in \{1, \dots, |\mathcal{S}| \times |\mathcal{T}|\}$, $i = 1, \dots, |\mathcal{S}| \times |\mathcal{T}|$ and $rand_i \in \{1, \dots, |\mathcal{X}|\}$ is a random number, chosen once for each i. And also, $C_r \in \{1, \dots, |\mathcal{X}|\}$ which regulates the convergence rate of proposed algorithm and ensures at least one of the randomly selected U_i to be V_i; therefore, the matrix U_{G+1} is as the following form,

$$U_{G+1} = \begin{bmatrix} U_{1,G+1} \\ U_{2,G+1} \\ . \\ . \\ . \\ U_{|\mathcal{S}|\times|\mathcal{T}|,G+1} \end{bmatrix} \tag{18}$$

Step 5: Calculating the objective function value $C(U_{G+1})$ using (10), (11) and (12) and the Arimoto-Blahut algorithm.

Step 6: Choosing the best between these two (function value at target and trial point), using the following form for next generation,

$$X_{G+1} = \begin{cases} U_{G+1} & \text{if } C(U_{G+1}) > C(X_G) \\ X_G & \text{otherwise} \end{cases} \qquad (19)$$

Step 7: Checking whether convergence criterion is met, if yes then stop; otherwise go to step 3.

B. Computing the sum capacity of discrete multiple access channel with causal side information at the transmitter

The total capacity, C_{total} , of discrete multiple access channel is the solution for the following optimization problem:

$$\max_{P(x_1)p(x_2)\ldots P(x_m)} I(x_1, x_2, \ldots; y) \qquad (20)$$

As a new method, we can make a vector the elements of which are the initial input joint distribution, $P(x_1, \ldots, x_m)$, as follows,

$$P(x_1, \ldots, x_m) \Leftrightarrow \begin{bmatrix} P(x_1 = i_{1,1}, \ldots, x_m = i_{m,1}) \\ \cdot \\ \cdot \\ \cdot \\ P(x_1 = i_{1,|\mathcal{X}_1|}, \ldots, x_m = i_{m,|\mathcal{X}_m|}) \end{bmatrix} \qquad (21),$$

and then apply it to (5) as follows,

$$P^{t+1}(x_1, \ldots, x_m) = \frac{P^t(x_1,\ldots,x_m) e^{\eta D^t}}{\sum_{x_1,\ldots,x_m} P^t(x_1,\ldots,x_m) e^{\eta D^t}} \qquad (22)$$

After the convergence of the Arimoto-Blahut algorithm, we can change the optimal $P(x_1, \ldots, x_m)$ to the joint distribution which is named $P^*(x_1, \ldots, x_m)$. According to the rank of $P^*(x_1, \ldots, x_m)$, we have two situations:

- $\mathrm{Rank}(P^*(x_1, \ldots, x_m)) = 1$. At this case we can compute the sum capacity by following expression,

$$C_{total} = \sum P^*(x_1, \ldots, x_m) P(y|x_1, \ldots, x_m) \log \frac{P(x_1,\ldots,x_m|y)}{P^*(x_1,\ldots,x_m)} \qquad (23),$$

- $\mathrm{Rank}(P^*(x_1, \ldots, x_m)) \neq 1$. This will be the case in general. We need to project $P(x_1, \ldots, x_m)$ as a product distribution ($P_{product}(x_1, \ldots, x_m)$) [26], [27]. Then we can compute the sum capacity with following expression,

$$C_{total} = \sum P_{product}(x_1, \ldots, x_m) P(y|x_1, \ldots, x_m) \log \frac{P(x_1,\ldots,x_m|y)}{P_{product}(x_1,\ldots,x_m)} \qquad (24),$$

We can extend the proposed algorithm to compute the sum capacity of discrete multiple access channels with causal side information at the transmitter [28]. Fig. 4 shows a discrete multiple access channels with causal side information at the transmitter. s_i represents the state of the channel that is revealed to all transmitters just before each transmission instant. For this channel the sum capacity is the solution of the following optimization problem:

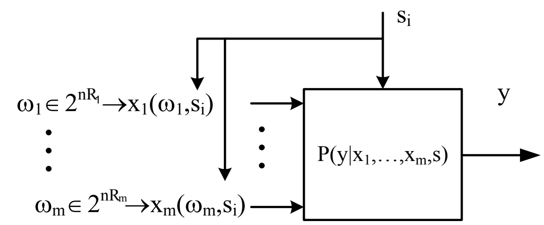

Figure 4. Discrete multiple access channel with causal side information at the transmitter

$$\max_{P(x_1)p(x_2)\ldots P(x_m)} I(x_1(\omega_1, s_i), \ldots, x_m(\omega_m, s_i); y) \qquad (25)$$

III. EVALUATION RESULTS

In this section we validate the performance of the algorithm proposed in section II over two channels and one discrete multiple access channel. All of them are with causal side information at the transmitter.

Example 1. Consider the channel with following transition probability distribution; this channel has two states, **s** = (0, 1) with $P(s = 0) = \frac{1}{4}$ and $P(s = 1) = \frac{3}{4}$, the inputs of this channel are 0 or 1.

$$P(y|x, s) = \begin{bmatrix} 0.55 & 0.45 \\ 0.65 & 0.35 \\ 0.45 & 0.55 \\ 0.90 & 0.10 \end{bmatrix}$$

where the columns represent the different elements of $\mathcal{Y} = \{0,1\}$ and the rows correspond to the natural ordering of the x, s. The $|\mathcal{T}|$ is equal to $|\mathcal{Y}|$. We have applied the algorithm to this channel. Fig. 5 shows the convergence of this algorithm to C = 0.04 bits per channel use after 3 iterations.

The optimal function (matrix) obtained by this evaluation is:

$$\hat{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Example 2. In this example, $\mathcal{Y} = \mathcal{X} = \{0,1,2\}$ and states are the same as previous example, but we want to choose $|\mathcal{T}|$ equal to two, not the maximum value, i.e. three. For this channel transition probability distribution is:

$$P(y|x, s) = \begin{pmatrix} 0.60 & 0.30 & 0.10 \\ 0.70 & 0.20 & 0.10 \\ 0.70 & 0.20 & 0.10 \\ 0.10 & 0.80 & 0.10 \\ 0.25 & 0.25 & 0.50 \\ 0.50 & 0.40 & 0.10 \end{pmatrix}$$

Fig. 6 Shows convergence of this algorithm to C = 0.2559 bits per channel use after 10 iterations and the optimal function (matrix) obtained by this evaluation is:

$$\widehat{A} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Example 3. Consider two-user binary discrete multiple access channel with two states, **s** = (0,1), with $P(s=0) = \frac{1}{4}$ and $P(s=1) = \frac{3}{4}$, the transition probability distribution is:

$$P(y|x_1, x_2, s) = \begin{pmatrix} 0.25 & 0.75 \\ 0.20 & 0.80 \\ 0.30 & 0.70 \\ 0.10 & 0.90 \\ 0.40 & 0.60 \\ 0.50 & 0.50 \\ 0.65 & 0.35 \\ 0.85 & 0.15 \end{pmatrix}$$

where the columns represent the different elements of $\mathcal{Y} = \{0,1\}$ and the rows correspond to the natural ordering of the x_1, x_2 and s. The $|\mathcal{T}_1|$ and $|\mathcal{T}_2|$ are equal to $|\mathcal{Y}|$. This algorithm converges to C_{total} = 0.3324 bits per channel use after 23 iterations.

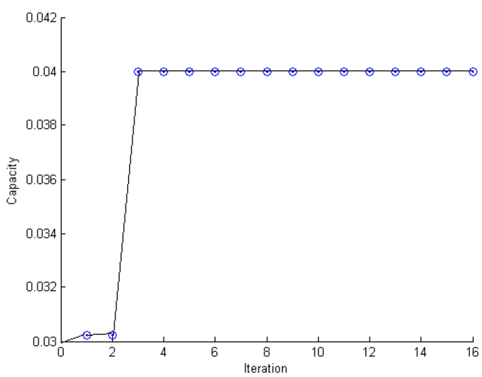

Figure 5. Optimized result for example 1

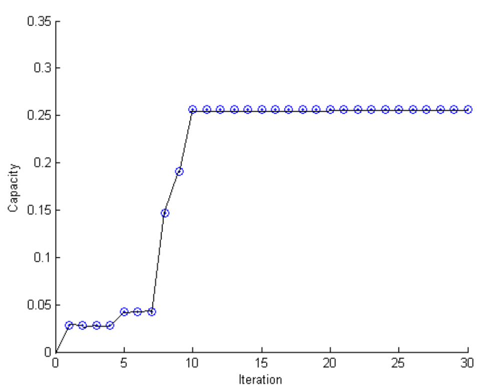

Figure 6. Optimized result for example 2

IV. CONCLUSION

We have given a new algorithm based on differential evolution optimization algorithm and the Arimoto-Blahut algorithm to compute the capacity for discrete memoryless channel and discrete multiple access channel with causal side information at the transmitter. This algorithm is fast, robust and needs small computer codes for optimization and simulation. With using the searching algorithm, the example 1 needs 16 iterations and the example 2 needs 81 iterations (these results are obtained by brute force algorithm), but we have optimized these two examples with 3 and 10 iterations, respectively. Also, we have obtained the sum capacity of discrete multiple access channels with causal side information at the transmitter with fewer iterations.

REFERENCES

[1] S. Arimoto, "An Algorithm for Computing the Capacity of Arbitrary Discrete Memoryless Channel," IEEE Trans. Inform. Theory, vol. IT-18, pp. 14-20, Jan. 1972.

[2] R. Blahut, "Computation of Channel Capacity and Rate-Distortion Functions," IEEE Trans. Inform. Theory, vol. IT-18, pp. 460-472, Jul. 1972.

[3] I. Csiszar and G. Tusnady, "Information Geometry and Alternating Minimization Procedures," Statistics and Decisions, Supplement Issue, pp. 205-237, 1984.

[4] M. Rezaeian and A. Grant, "A Generalization of the Arimoto-Blahut Algorithm," ISIT 2004.

[5] Y. Watanabe, "The Total Capacity of Two-User Multiple-Access Channel with Binary Output", IEEE Trans. Inform. Theory, vol. 42, pp. 1453-1465, Sept. 1996.

[6] Y. Watanabe and K. Kamoi, "The Total Capacity of Multiple-Access Channel," ISIT 2002.

[7] M. Rezaeian and A. Grant, "Computation of Total Capacity for Discrete Memoryless Multiple-Access Channels", IEEE Trans. Inform. Theory, vol.50, pp. 2779-2784, Nov. 2004.

[8] Fr´ed´eric Dupuis, W. Yu, and F. M. J. Willems, "Blahut-Arimoto Algorithms for Computing Channel Capacity and Rate-Distortion with Side Information," ISIT 2004.

[9] R. Storn, K. Price, "Differential Evolution - A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Space," Technical Report TR-95-012, ICSI, March 1995. Available via the Internet: ftp.icsi.berkeley.edu/pub/techreports/ 1995/tr-95-012.ps.Z

[10] T. M. Cover and J. A. Thomas, Elements of Information Theory, John Wiley & Sons, 2006.

[11] Y. Yu, "Squeezing the Arimoto-Blahut Algorithm for Faster Convergence," ITW 2010 (Cairo, Egypt).

[12] G. Matz and P. Duhamel, "Information Geometric Formulation and Interpretation of Accelerated Blahut-Arimoto-Type Algorithms," ITW 2004.

[13] C. E. Shannon, "Channels with Side Information at the Transmitter," IBM Journal Research and Development, vol. 2, pp. 289-293, Oct. 1958.

[14] G. Keshet, Y. Steinberg, and N. Merhav, Channel Coding in the Presence of Side Information: Subject Review, Now Publisher, pp. 9-15.

[15] U. Erez and R. Zamir, "Capacity and Coding for Symmetric Channels with Side Information at the Transmitter," Tech. Report EE-S-97-10, pp. 1-5, Oct. 1997.

[16] R. Storn and K. Price, "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Space," Journal of Global Optimization, vol. 11, pp. 41-359, 1997.

[17] F. S. Wang and W. M. Cheng, "Simultaneous Optimization of Feeding Rate and Operation Parameters for Fed-Batch Fermentation Processes," Biotechnology Progress 15, pp. 949-952, 1999.

[18] B. V. Babu and S. A. Munawar, "Differential Evolution for the Optimal Design of Heat Exchangers," In Proceedings of All-India

Seminar on Chemical Engineering Progress on Resource Development: A Vision 2010 and Beyond, Bhuvaneshwar, 2000.

[19] B. V. Babu and R. P. Singh, "Synthesis & Optimization of Heat Integrated Distillation Systems Using Differential Evolution," In Proceedings of All-India Seminar on Chemical Engineering Progress on Resource Development: A Vision 2010 and Beyond, Bhuvaneshwar, 2000.

[20] R Angira and B. V. Babu, "Optimization of Non-Linear Chemical Processes Using Modified Differential Evolution (MDE)," In Proceedings of the 2nd Indian International Conference on Artificial Intelligence, Pune, India, pp. 911-923, 2005.

[21] R. Angira and B. V. Babu, "Optimization of Process Synthesis and Design Problems: A Modified Differential Evolution Approach," Chemical Engineering Science 61, pp. 4707-4721, 2006.

[22] B. V. Babu and R. Angira, "Optimization of Thermal Cracker Operation using Differential Evolution," In Proceedings of International Symposium & 54th Annual Session of IIChE, Chennai, December 19-22, 2001.

[23] B. V. Babu and R. Angira, "Optimization of Water Pumping System Using Differential Evolution Strategies," In Proceedings of The Second International Conference on Computational Intelligence, Robotics, and Autonomous Systems, Singapore, 2003.

[24] M. H. Lee, C. Han, and K. S. Chang, "Dynamic Optimization of a Continuous Polymer Reactor Using a Modified Differential Evolution Algorithm," Industrial and Engineering Chemistry Research 38, pp. 4825-4831, 1999.

[25] J. C. Lu and F. S. Wang, "Optimization of Low Pressure Chemical Vapor Deposition Reactors Using Hybrid Differential Evolution," Canadian Journal of Chemical Engineering 79, pp. 246-254, 2001.

[26] C. K. Chow and C. N. Liu, "Approximating discrete probability distributions with dependence trees," IEEE Trans. Inform. Theory, vol. IT-14, pp. 462-467, May 1968.

[27] H. Ku and S. Kullback, "Approximating discrete probability distributions," IEEE Trans. Inform. Theory, vol. IT-15, pp. 444-447, July 1969.

[28] S. Sigurjonsson and Y. H. Kim, "On Multiple User Channels with State Information at the Transmitters," ISIT 2005.