The Computation of the Capacity Region of the Discrete MAC is a Rank-One Non-Convex Optimization Problem

Eduard Calvo^{*,0}, Daniel P. Palomar†, Javier R. Fonollosa^{*,0}, and Josep Vidal^{*,0}
e-mail: {eduard,fono,pepe}@gps.tsc.upc.edu, palomar@ust.hk
*SPCOM Group, Dpt. Signal Theory and Communications, Technical University of Catalonia (UPC), Spain.
†Dpt. of Electronic and Computer Engineering, Honk Kong University of Science and Technology (HKUST), China.

***Abstract*—The computation of the channel capacity of discrete memoryless channels is a convex problem that can be efficiently solved using the Arimoto-Blahut (AB) iterative algorithm. However, the extension of this algorithm to the computation of capacity regions of multiterminal networks is not straightforward since its computation gives rise to non-convex problems. In this context, the AB algorithm has been only successfully extended to the calculation of the sum-capacity of the discrete memoryless multiple-access channel. However, the computation of the capacity region still requires the use of computationally demanding random search algorithms or brute force (full search) methods.**

In this paper, we first give an alternative reformulation of the problem that identifies the non-convexity as a rank-one constraint. We then propose an efficient algorithm to compute outer and inner bounds on the capacity region by relaxing the original problem and then by projecting the relaxed solution onto the original space variable via a minimum divergence criterion.

There exists a class of channels for which the proposed algorithm can be shown to compute exactly the capacity region. As an illustration, we analyze two particular channels, the binary adder MAC and the binary switching MAC, in detail. In the general case, the algorithm is able to compute very tight bounds as shown by simulation.

I. INTRODUCTION

The characterization of the capacity of an arbitrary single-user memoryless channel is a problem that admits a single-letter representation in the form of a maximization of a concave function over a convex set. This is a convex problem that can be numerically evaluated efficiently in practice (i.e., with a polynomial time worst-case complexity) [4]. For the continuous Gaussian channel, for example, the solution admits a simple closed-form characterization. For the DMC there is no closed-form expression but we have the popular practical Arimoto-Blahut (AB) algorithm [2] [3].

In contrast, for the general multiuser case we do not even have a characterization of the capacity region. Fortunately, for the multiple-access channel (MAC) we also have a single-letter representation of the capacity region [1] [9]. However, the characterization is not in the form of a convex optimization problem as happened in the single-user case. Again, for the continuous Gaussian channel, the characterization simplifies to a convex problem which can always be numerically evaluated in an efficient way. For the discrete memoryless MAC (DMAC), however, the characterization is not in the form of a convex problem and, as a consequence, there is no efficient algorithm to compute the capacity region in practice. In this context, many authors have recently contributed toward the computation of the sum-capacity (total capacity) of an arbitrary DMAC [11] [12], and an algorithm for its exact computation has been found [13]. Other applications of generalizations of the AB algorithm can be found in the context of the computation of channel capacity with side information [7].

Regarding the computation of the whole capacity region of the DMAC, not much work has been done due to the intractability of the problem because of its non-convexity. As a consequence of the non-convexity, brute-force algorithms or random search methods seem to be the only alternative to compute inner bounds on the capacity region with no quantification on the suboptimality.

In this work, we show that the key difficulty in computing the capacity region of an arbitrary DMAC can be identified as a rank-one constraint (a non-convex constraint) in an optimization problem. Such type of problems cannot be solved optimally in polynomial time and suboptimal methods must be used. In particular, the use of relaxation methods provides near optimal solutions in polynomial time (e.g., [10]). Relaxation methods are based on (1) replacing the rank-one constrained problem by an approximate (not equivalent) tractable convex problem; (2) generating a potential solution belonging to the variable space of the original problem from the solution of the relaxed problem.

After having identified the source of the non-convexity as a rank-one constraint, we propose a relaxation method for the computation of the capacity region of an arbitrary DMAC that provides us with an inner and an outer bound of it. Therefore, the proposed method quantifies how suboptimal the computed solution is with the difference between the outer and the inner bound.

There exists a class $\mathcal{K}$ of channels for which the algorithm is able to compute exactly the capacity region. It comprises

---

[0] This work has been partially funded by the European Commission, the Spanish Ministry of Education and Science, the Catalan Government and FEDER funds under contracts TEC2006-06481, TEC2004-04526, 27187 SURFACE and 2005SGR-00639, and grant FPU-AP-2004-3549.

the subclass of channels with identical inner and outer bounds and the subclass of channels with strict outer bounds and tight inner bounds. For illustrative purposes, we analyze a representative example of each subclass: the Binary Switching MAC (BS-MAC) and the Binary Adder MAC (BA-MAC, also called Binary Erasure MAC), respectively. Additionally, we provide a non-deterministic example DMAC for which the proposed algorithm is also optimal to show that $\mathcal{K}$ is not a subset of the class of deterministic DMACs. Unfortunately, we have not been able to characterize $\mathcal{K}$ analytically. However, computer simulations of the proposed algorithm for various channels show that it performs indistinguishably to the optimal solution obtained with a computationally intensive brute-force full search approach.

The structure of this paper is as follows. Section II introduces the problem of the computation of the capacity region of the DMAC and reformulates it as a rank-one constrained optimization problem. Section III describes the proposed relaxation-based algorithm for the computation of inner and outer bounds on the capacity region. Then, Section IV provides the results on the optimality of the proposed algorithm for the BA-MAC and the BS-MAC. The performance of the proposed algorithm is showed in Section V. Finally, Section VI concludes the paper.

II. THE CAPACITY REGION AS A RANK-ONE CONSTRAINED OPTIMIZATION PROBLEM

The computation of the capacity region of an arbitrary DMAC (a convex set) is a non-convex problem. This problem can be formulated in a matrix form that identifies the non-convexity as a rank-one constraint.

A. The problem of the capacity region

The capacity region $\mathcal{C}$ of the two-user DMAC is the convex hull of the set of rate pairs (R_1, R_2) satisfying

$$0 \le R_1 \le I(X_1;Y|X_2) \tag{1}$$
$$0 \le R_2 \le I(X_2;Y|X_1) \tag{2}$$
$$R_1 + R_2 \le I(X_1X_2;Y) \tag{3}$$

for some choice of the distribution $P_{X_1X_2Y} = P_{X_1}P_{X_2}P_{Y|X_1X_2}$ on $\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}$, where the input alphabets can be characterized as

$$\mathcal{X}_k = \{x_k^{(1)}, \ldots, x_k^{(|\mathcal{X}_k|)}\},\ k = 1, 2, \tag{4}$$

P_{X_k} is the input probability distribution of the k-th user ($k = 1, 2$), and $P_{Y|X_1X_2}$ is the given conditional distribution that depends on the nature of the channel. It is well known that $\mathcal{C}$ is a convex set [6, Thm. 14.3.2] and hence, by applying the supporting hyperplane theorem [4, Sec. 2.5.2], the computation of the capacity region can be parameterized for $\theta \in [0, 1]$ as

$$\begin{aligned}
\underset{\{R_1,R_2,P_{X_1},P_{X_2}\}}{\text{maximize}} \quad & \theta R_1 + (1-\theta)R_2 & (5)\\
\text{subject to} \quad & 0 \le R_1 \le I_1(P_{X_1}, P_{X_2}, P_{Y|X_1X_2}) & (6)\\
& 0 \le R_2 \le I_2(P_{X_1}, P_{X_2}, P_{Y|X_1X_2}) & (7)\\
& R_1 + R_2 \le I_{12}(P_{X_1}, P_{X_2}, P_{Y|X_1X_2}) & (8)\\
& \sum_{x_k} P_{X_k}(x_k) = 1,\ P_{X_k}(x_k) \ge 0\ \forall x_k \in \mathcal{X}_k, & (9)
\end{aligned}$$

where expressions (6)-(8) correspond to the right hand sides of (1)-(3), respectively, instantiated for the DMC[1]:

$$I_1 \triangleq \sum_{x_1,x_2,y} P_{X_1X_2Y}(x_1,x_2,y) \log \frac{P_{Y|X_1X_2}(y|x_1x_2)}{\sum_{x_1'} P_{X_1}(x_1')P_{Y|X_1X_2}(y|x_1'x_2)} \tag{10}$$

$$I_2 \triangleq \sum_{x_1,x_2,y} P_{X_1X_2Y}(x_1,x_2,y) \log \frac{P_{Y|X_1X_2}(y|x_1x_2)}{\sum_{x_2'} P_{X_2}(x_2')P_{Y|X_1X_2}(y|x_1x_2')} \tag{11}$$

$$I_{12} \triangleq \sum_{x_1,x_2,y} P_{X_1X_2Y}(x_1,x_2,y) \log \frac{P_{Y|X_1X_2}(y|x_1x_2)}{\sum_{x_1',x_2'} P_{X_1X_2Y}(x_1',x_2',y)}. \tag{12}$$

Note that the solutions $P^\star_{X_1}(x_1;\theta)$ and $P^\star_{X_2}(x_2;\theta)$ generally depend on θ. For each θ, the problem (5)-(9) computes the intersection between the contour of the capacity region and a tangent hyperplane with normal vector $\mathbf{n} = [\theta,\ 1-\theta]^T$. Hence, the capacity region is obtained when (5)-(9) is solved for all $\theta \in [0, 1]$; in other words, the solutions $(R_1^\star(\theta), R_2^\star(\theta))$ are samples of the boundary of $\mathcal{C}$.

B. A rank-one constrained optimization problem

The problem (5)-(9) of the computation of the capacity region is non-convex because the constraints (6)-(8) are not jointly convex in P_{X_1} and P_{X_2}. For instance, the RHS of the constraint in (8) is not concave (note that it should be concave for the problem to be convex). To see this, observe that even though $x\log(1/x)$ is concave, the composition with a linear combination of terms xy is not[2]. Similar reasonings may be applied to the constraints (6) and (7) to obtain again that the lack of convexity follows from the product terms $P_{X_1}(x_1)P_{X_2}(x_2)$.

Although the problem (5)-(9) is not jointly convex in (P_{X_1}, P_{X_2}), it is separately convex in each of the input probability distributions. This would allow us to perform an alternate optimization procedure: $P_{X_1}^{(0)} \to P_{X_2}^{(0)} \to P_{X_1}^{(1)} \to \ldots$, where $P_{X_k}^{(n)}$ denotes the optimal solution P_{X_k} at the n-th iteration. However, alternate optimization procedures applied to non-convex problems only converge to local maxima of the cost function.

Interestingly, if we allow the variables X_1 and X_2 to be dependent with joint distribution $P_{X_1X_2}$, then problem (5)-(9) becomes convex (recall that $x\log(1/x)$ is a concave function).

---

[1] It is implicitly assumed $P_{X_1X_2Y} = P_{X_1}P_{X_2}P_{Y|X_1X_2}$.

[2] It is sufficient to verify that the Hessian of $f(x,y) = xy\log(xy)$ has one positive and one negative eigenvalue at $(x,y) = (1/\sqrt{2}, 1/\sqrt{2})$.

We will now reformulate the problem with a matrix-vector notation. Each of the input probability distributions P_{X_k} admits a vector representation of the form

$$\mathbf{p}_k \equiv P_{X_k} \Longleftrightarrow [\mathbf{p}_k]_i = P_{X_k}(x_k^{(i)}), 1 \leq i \leq |\mathcal{X}_k|, k = 1, 2, \quad (13)$$

while the joint distribution admits the matrix representation

$$\mathbf{P} \equiv P_{X_1X_2} \Longleftrightarrow [\mathbf{P}]_{i,j} = P_{X_1X_2}(x_1^{(i)}, x_2^{(j)}), \quad (14)$$

for $1 \leq i \leq |\mathcal{X}_1|$ and $1 \leq j \leq |\mathcal{X}_2|$.

Then, define $\mathcal{P}_{\mathsf{prod}}$ as the subset containing all the *product* distributions (P_{X_1}, P_{X_2}) of X_1 and X_2,

$$\mathcal{P}_{\mathsf{prod}} = \Big\{ \mathbf{P} \in \mathbb{R}^{|\mathcal{X}_1| \times |\mathcal{X}_2|} \mid [\mathbf{P}]_{i,j} = P_{X_1}(x_1^{(i)}) P_{X_2}(x_2^{(j)}) \text{ for some feasible } (P_{X_1}, P_{X_2}) \text{ on } \mathcal{X}_1 \times \mathcal{X}_2 \Big\}. \quad (15)$$

Lemma 1: For any probability matrix $\mathbf{P} \in \mathbb{R}^{|\mathcal{X}_1| \times |\mathcal{X}_2|}$, i.e., with non-negative entries and such that $\sum_{i,j} [\mathbf{P}]_{i,j} = 1$,

$$\mathbf{P} \in \mathcal{P}_{\mathsf{prod}} \Leftrightarrow \text{rank}(\mathbf{P}) = 1. \quad (16)$$

Proof: This follows from the fact that $\mathbf{P} = \mathbf{p}_1 \mathbf{p}_2^T$ if and only if $\mathbf{P} \in \mathcal{P}_{\mathsf{prod}}$. ■

Using Lemma 1, the following simpler equivalent description of $\mathcal{P}_{\mathsf{prod}}$ can be given

$$\mathcal{P}_{\mathsf{prod}} = \Big\{ \mathbf{P} \in \mathbb{R}^{|\mathcal{X}_1| \times |\mathcal{X}_2|} \Big| \text{rank}(\mathbf{P}) = 1, \mathbf{P} \succeq \mathbf{0}, \mathbf{1}^T \mathbf{P} \mathbf{1} = 1 \Big\}, \quad (17)$$

where $\succeq$ denotes component-wise inequality and $\mathbf{1}$ is an all-one column vector of appropriate length. The original problem (5)-(9) can now be expressed in an equivalent matrix form formulated in terms of $\mathbf{P} \in \mathcal{P}_{\mathsf{prod}}$, the joint distribution of X_1 and X_2, and its marginals $\mathbf{p}_1$ and $\mathbf{p}_2$, making use of expression (17) and the fact that

$$\mathbf{P1} = \mathbf{p}_1, \quad \mathbf{P}^T \mathbf{1} = \mathbf{p}_2, \quad (18)$$

i.e., that P_{X_1} and P_{X_2} are the marginal distributions of $P_{X_1X_2}$. The following reformulation of the problem is the key point of the identification of (5)-(9) as a rank-one non-convex optimization problem.

Proposition 1: The problem (5)-(9) of the computation of the capacity region of an arbitrary two-user DMAC is equivalent to the following *rank-one non-convex optimization problem*

$$\begin{array}{lll}
\underset{\{R_1, R_2, \mathbf{P}, \mathbf{p}_1, \mathbf{p}_2\}}{\text{maximize}} & \theta R_1 + (1-\theta) R_2 & (19) \\
\text{subject to} & 0 \leq R_1 \leq f_1(\mathbf{P}, \mathbf{p}_2) & (20) \\
 & 0 \leq R_2 \leq f_2(\mathbf{P}, \mathbf{p}_1) & (21) \\
 & R_1 + R_2 \leq f_{12}(\mathbf{P}) & (22) \\
 & \mathbf{P1} = \mathbf{p}_1 \ , \ \mathbf{P}^T \mathbf{1} = \mathbf{p}_2 & (23) \\
 & \mathbf{P} \succeq \mathbf{0} \ , \ \mathbf{1}^T \mathbf{P1} = 1 & (24) \\
 & \text{rank}(\mathbf{P}) = 1, & (25)
\end{array}$$

where the functions f_1 (26), f_2 (27), and f_{12} (28) are concave in $(\mathbf{P}, \mathbf{p}_2)$, $(\mathbf{P}, \mathbf{p}_1)$, and $\mathbf{P}$, respectively.

Proof: See the Appendix. ■

While (24) ensures that $\mathbf{P}$ is a feasible probability matrix, (25) constrains it to $\mathcal{P}_{\mathsf{prod}}$ (see Lemma 1), and (23) relates $\mathbf{P}$ with its marginals.

III. THE PROPOSED RELAXATION METHOD

The only non-convexity that makes the computation of the capacity region a hard problem is the rank-one constraint in (25) that forces the joint probability distribution $\mathbf{P}$ to be a product distribution. If (25) is removed, the resulting convex problem is equivalent to solving the capacity region of a DMAC with arbitrarily dependent codewords. This relaxation will be adopted in order to obtain good approximations of the solutions to (19)-(25).

By removing the rank-one constraint, the region obtained with the relaxed problem, that we will denote by $\mathcal{R}^{\circ}$, becomes an outer bound on the true capacity region. Two situations may occur depending on the optimal solution of the relaxed problem (19)-(24) for a given θ, $\mathbf{P}^{\star}(\theta)$.

1) $\text{rank}(\mathbf{P}^{\star}(\theta)) = 1$. The relaxation method solves optimally the original problem (19)-(25) at the given θ. If $\text{rank}(\mathbf{P}^{\star}(\theta)) = 1$ holds $\forall \theta \in [0, 1]$, $\mathcal{R}^{\circ} = \mathcal{C}$ and the whole capacity region is optimally computed.
2) $\text{rank}(\mathbf{P}^{\star}(\theta)) \neq 1$ (this will be the case in general). We will need to *project* $\mathbf{P}^{\star}(\theta)$ onto $\mathcal{P}_{\mathsf{prod}}$.

In this work, we will apply an projection based on the Kullback-Leibler divergence $D(\cdot \| \cdot)$ to approximate an arbitrary joint distribution $\mathbf{P}$ by a product distribution of the form $\mathbf{q}_1 \mathbf{q}_2^T$ defined on $\mathcal{X}_1 \times \mathcal{X}_2$. The divergence has been used in [5] and [8] as the criterion for approximating joint discrete probability distributions given a dependence tree, although

$$f_1(\mathbf{P}, \mathbf{p}_2) \triangleq \sum_{i,j,y} [\mathbf{P}]_{i,j} P_{Y|X_1X_2}(y|x_1^{(i)} x_2^{(j)}) \log \frac{P_{Y|X_1X_2}(y|x_1^{(i)} x_2^{(j)}) [\mathbf{p}_2]_j}{\sum_{i'} [\mathbf{P}]_{i',j} P_{Y|X_1X_2}(y|x_1^{(i')} x_2^{(j)})} \quad (26)$$

$$f_2(\mathbf{P}, \mathbf{p}_1) \triangleq \sum_{i,j,y} [\mathbf{P}]_{i,j} P_{Y|X_1X_2}(y|x_1^{(i)} x_2^{(j)}) \log \frac{[\mathbf{p}_1]_i P_{Y|X_1X_2}(y|x_1^{(i)} x_2^{(j)})}{\sum_{j'} [\mathbf{P}]_{i,j'} P_{Y|X_1X_2}(y|x_1^{(i)} x_2^{(j')})} \quad (27)$$

$$f_{12}(\mathbf{P}) \triangleq \sum_{i,j,y} [\mathbf{P}]_{i,j} P_{Y|X_1X_2}(y|x_1^{(i)} x_2^{(j)}) \log \frac{P_{Y|X_1X_2}(y|x_1^{(i)} x_2^{(j)})}{\sum_{i',j'} [\mathbf{P}]_{i',j'} P_{Y|X_1X_2}(y|x_1^{(i')} x_2^{(j')})} \quad (28)$$

here the purpose is different. The use of the divergence as the measure that quantifies the quality of the approximation offers several advantages, the most useful one being that, for fixed $\mathbf{P}$, the problem of finding the pair $(\mathbf{q}_1, \mathbf{q}_2)$ that minimizes $D(\mathbf{P}||\mathbf{q}_1\mathbf{q}_2^T)$ is easily solved by showing

$$\begin{aligned} D(\mathbf{P}||\mathbf{q}_1\mathbf{q}_2^T) &= D(\mathbf{P}||\mathbf{p}_1\mathbf{p}_2^T) + D(\mathbf{p}_1||\mathbf{q}_1) + D(\mathbf{p}_2||\mathbf{q}_2) \quad (29) \\ &\geq D(\mathbf{P}||\mathbf{p}_1\mathbf{p}_2^T), \quad (30) \end{aligned}$$

from where it follows $(\mathbf{q}_1^\star, \mathbf{q}_2^\star) = (\mathbf{p}_1, \mathbf{p}_2)$. Hence, in order to get an approximation of $\mathcal{C}$, it is sufficient to solve (19)-(24), take the marginal distributions of the solution, plug them into (5-9), and solve it in (R_1, R_2) (the problem (5)-(9) is convex for fixed input probability distributions). The solution of (5)-(9) in terms of (R_1, R_2) defines an inner bound $\mathcal{R}^{\mathsf{i}}$, that, together with the outer bound $\mathcal{R}^{\circ}$, is the output of the proposed relaxation method.

IV. ANALYTICAL RESULTS

Let us denote by $\mathcal{K}$ the class of discrete MACs for which the relaxation method of Section III computes optimally the capacity region. For some channels in $\mathcal{K}$, optimality of the proposed relaxation method can be proved by showing $\mathcal{R}^{\mathsf{i}} = \mathcal{R}^{\circ} = \mathcal{C}$. A representative example of this class of channels is the Binary Switching MAC (BS-MAC). For others, we may have $\mathcal{R}^{\mathsf{i}} = \mathcal{C} \subset \mathcal{R}^{\circ}$, as it is the case of the Binary Adder MAC (BA-MAC).

A. The Binary Switching MAC [14] [15]

The binary switching MAC is a binary-inputs, ternary-output ($\mathcal{Y} = \{0, 1, \infty\}$), deterministic multiple-access channel whose input-output relationship is

$$Y = X_2/X_1 = \begin{cases} 0 & \text{if } (X_1, X_2) = (1, 0) \\ 1 & \text{if } (X_1, X_2) = (1, 1) \\ \infty & \text{if } X_1 = 0 \end{cases} . \quad (31)$$

Note that, given the channel output Y, the value of X_1 is always decoded without ambiguity ($H(X_1|Y) = 0$). On the other hand, the information carried by X_2 can only be conveyed when $X_1 = 1$, since otherwise $Y = \infty$ independently of X_2. Therefore, a hierarchy is established among senders: sender 1 can always convey information to Y and it is also responsible for allowing sender 2 for effectively transmitting information to Y.

Proposition 2: The achievable rate region of the proposed algorithm equals the capacity region for the BS-MAC [14] [15], i.e., $\mathcal{R}^{\mathsf{i}} = \mathcal{C}$. Furthermore, the inner and outer bounds on the capacity region provided by the proposed algorithm coincide, $\mathcal{R}^{\mathsf{i}} = \mathcal{C} = \mathcal{R}^{\circ}$.

B. The Binary Adder MAC

The binary adder MAC [9] (or binary erasure MAC, as named in [6, Example 14.3.3]), BA-MAC, is the binary-inputs, ternary output ($\mathcal{Y} = \{0, 1, 2\}$), deterministic multiple-access channel whose input-output relationship is given by

$$Y = X_1 + X_2, \quad (32)$$

where the sum is taken over the natural numbers. In this case, none of the users has access to the channel in privileged conditions. There are two input combinations that can be correctly decoded without ambiguity ($(X_1, X_2) \in \{(0, 0), (1, 1)\}$), and two input combinations where the decoder has ambiguity about both users. Therefore, for this channel the capacity region must be symmetric with respect to R_1 and R_2.

Proposition 3: The achievable rate region of the proposed algorithm equals the capacity region for the BA-MAC [9]. Furthermore, the inner and outer bounds on the capacity region provided by the proposed algorithm do not coincide, i.e., $\mathcal{R}^{\mathsf{i}} = \mathcal{C} \subset \mathcal{R}^{\circ}$.

V. SIMULATION RESULTS

In this section we validate the performance of the relaxation method proposed in Section III over a non-deterministic channel in order to show that the class $\mathcal{K}$ is not a subset of the class of deterministic DMACs.

Let us analyze the bynary-inputs ternary-output DMAC characterized by the transition probability distribution

$$P_{Y|X_1X_2} \equiv \begin{bmatrix} 0.2 & 0.3 & 0.5 \\ 0.7 & 0.2 & 0.1 \\ 0.5 & 0.1 & 0.4 \\ 0.3 & 0.4 & 0.3 \end{bmatrix}, \quad (33)$$

where the columns represent the different elements of $\mathcal{Y} = \{0, 1, 2\}$, and the rows correspond to the natural ordering of the inputs. The channel (33) is used in [13] to illustrate the performance of the algorithm for the computation of the sum-capacity.

We apply the proposed relaxation method to compute the inner and outer bounds, $\mathcal{R}^{\mathsf{i}}$ and $\mathcal{R}^{\circ}$, respectively. Additionally, we consider the achievable region of a brute force random search algorithm, denoted by $\mathcal{R}$, as a benchmark. The comparison is shown in Figure 1, where it can be seen that $\mathcal{R}^{\mathsf{i}}$ and $\mathcal{R}^{\circ}$ coincide. This means that the capacity region in this case has been effectively computed.

Numerical simulations of the proposed algorithm for various other channels show that $\mathcal{R}^{\mathsf{i}}$ is undistinguishable from $\mathcal{R}$, the solution obtained with a computationally intensive brute-force full search approach.

VI. CONCLUSIONS

We have given a matrix formulation for the problem of the computation of the capacity region of the DMAC, $\mathcal{C}$. With this new formulation, all the non-convexity is identified as a rank-one constraint. Since problems with this constraint cannot be solved optimally, we have proposed a suboptimal yet practical method for the computation of $\mathcal{C}$. This method, which provides us with an inner and an outer bound of $\mathcal{C}$, is optimal for a class $\mathcal{K}$ of DMACs, for which the capacity region can be efficiently computed. Finally, three illustrative example channels belonging to $\mathcal{K}$ have been analyzed: two deterministic DMACs (the BS-MAC and the BA-MAC) and one non-deterministic DMAC (33). Future work will focus on the analytical characterization of the class $\mathcal{K}$.

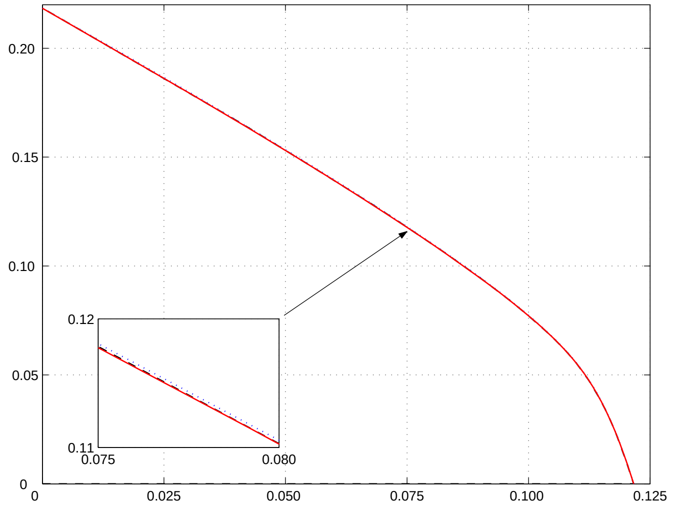

Fig. 1. The regions $\mathcal{R}$ (dashed black), $\mathcal{R}^{\mathrm{i}}$ (solid red), and $\mathcal{R}^{\circ}$ (dotted blue) for the channel (33).

APPENDIX: PROOF OF PROPOSITION 1

First, note that the functions $f_1(\mathbf{P},\mathbf{p}_2)$ in (26), $f_2(\mathbf{P},\mathbf{p}_1)$ in (27), and $f_{12}(\mathbf{P})$ in (28) simplify to I_1 (10), I_2 (11) and I_{12} (12), respectively, when the equivalences (13)-(14) are used for some $\mathbf{P}\in\mathcal{P}_{\mathsf{prod}}$ with marginal distributions $\mathbf{p}_1$ and $\mathbf{p}_2$. Equivalence between (5)-(9) and (19)-(25) is hence proved thanks to constraint (23) and the equivalence of (24)-(25) and $\mathbf{P}\in\mathcal{P}_{\mathsf{prod}}$, as stated in Lemma 1.

Regarding the concavity of the function $f_1(\mathbf{P},\mathbf{p}_2)$, we can rewrite (26) as

$$\begin{aligned} f_1(\mathbf{P},\mathbf{p}_2) &= \sum_{i,j}[\mathbf{P}]_{i,j}\times\\ &\underbrace{\times\Big(\sum_y P_{Y|X_1X_2}(y|x_1^{(i)}x_2^{(j)})\log P_{Y|X_1X_2}(y|x_1^{(i)}x_2^{(j)})\Big)}_{-H(Y|X_1=x_1^{(i)},X_2=x_2^{(j)})}\\ &+\sum_{j,y}\underbrace{\Big(\sum_i[\mathbf{P}]_{i,j}P_{Y|X_1X_2}(y|x_1^{(i)}x_2^{(j)})\Big)}_{\equiv P_{X_2Y}(x_2^{(j)},y)}\times\\ &\times\log\frac{[\mathbf{p}_2]_j}{\sum_{i'}[\mathbf{P}]_{i',j}P_{Y|X_1X_2}(y|x_1^{(i')}x_2^{(j)})}, \end{aligned}\tag{34}$$

which reduces to

$$\begin{aligned} f_1(\mathbf{P},\mathbf{p}_2) &= -H(Y|X_1X_2) &(35)\\ &-\sum_{j,y}P_{X_2Y}(x_2^{(j)},y)\log\frac{P_{X_2Y}(x_2^{(j)},y)}{[\mathbf{p}_2]_j}. &(36) \end{aligned}$$

The term $H(Y|X_1,X_2)$ is linear in $\mathbf{P}$ and thus concave. On the other hand, the RHS of (36) is jointly concave in $(P_{X_2Y},\mathbf{p}_2)$ (by the same arguments that ensure the convexity of the divergence [6, Thm. 2.7.2]), but since P_{X_2Y} is linear in $\mathbf{P}$, it is also concave in $(\mathbf{P},\mathbf{p}_2)$, so is the function $f_1(\mathbf{P},\mathbf{p}_2)$.

By symmetry, similar arguments can be used show concavity of $f_2(\mathbf{P},\mathbf{p}_1)$.

Finally, the function $f_{12}(\mathbf{P})$ (28) can be properly expanded similarly as

$$\begin{aligned} f_{12}(\mathbf{P}) &= \sum_{i,j}[\mathbf{P}]_{i,j}\times\\ &\underbrace{\times\Big(\sum_y P_{Y|X_1X_2}(y|x_1^{(i)}x_2^{(j)})\log P_{Y|X_1X_2}(y|x_1^{(i)}x_2^{(j)})\Big)}_{-H(Y|X_1=x_1^{(i)},X_2=x_2^{(j)})}\\ &+\sum_y\underbrace{\Big(\sum_{i,j}[\mathbf{P}]_{i,j}P_{Y|X_1X_2}(y|x_1^{(i)}x_2^{(j)})\Big)}_{\equiv P_Y(y)}\times\\ &\times\log\frac{1}{\sum_{i',j'}[\mathbf{P}]_{i',j'}P_{Y|X_1X_2}(y|x_1^{(i')}x_2^{(j')})}\\ &= -H(Y|X_1X_2)+\sum_y P_Y(y)\log\frac{1}{P_Y(y)}\\ &= -H(Y|X_1X_2)+H(Y). \end{aligned}\tag{37}$$

While $H(Y|X_1X_2)$ is linear (and hence concave) in $\mathbf{P}$, $H(Y)$ is concave in P_Y, which, in turn, is linear (concave) in $\mathbf{P}$. Hence, $f_{12}(\mathbf{P})$ is concave in $\mathbf{P}$. ■

REFERENCES

[1] R. Ahlswede, “Multi-way communication channels”, in *Proc. 2nd Int. Symp. Inform. Theory*, Tsakhadsor, Armenian SSR, 1971.

[2] S. Arimoto, “An algorithm for computing the capacity of arbitrary discrete memoryless channels”, *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 14-20, Jan 1972.

[3] R. E. Blahut, “Computation of channel capacity and rate-distortion functions”, *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 460-473, Jul 1972.

[4] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[5] 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.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[7] Frédéric Dupuis, W. Yu, and F. M. J. Willems, “Blahut-Arimoto algorithms for computing channel capacity and rate-distortion with side information”, in *Proc. IEEE Int. Symp. Information Theory*, Chicago, USA, June/July 2004, p. 181.

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

[9] H. J. Liao, “Multiple access channels”, Ph.D. dissertiation, Dep. Elec. Eng., Univ. of Hawaii, Honolulu 1972.

[10] W. Ma, T. N. Davidson, K. M. Wong, Z. Luo, and P. Ching, “Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA”, *IEEE Trans. Signal Processing*, vol. 50, pp. 912-922, Apr 2002.

[11] 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.

[12] Y. Watanabe and K. Kamoi, “The total capacity of multiple-access channel”, in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, June/July 2002, p. 308.

[13] 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.

[14] P. Vanroose and E. C. van der Meulen, “Coding for the binary switching multiple access channel”, in *Proc. 7th Symp. Inform. Theory in the Benelux*, Noordwijkerhout, 1986, pp. 183-189.

[15] A. J. Vinck, “On the multiple access channel”, in *Proc. 2nd Joint Swedish-Soviet Int. Workshop Inform. Theory*, Gränna, 1985, pp.24-29.