MAC Polar Codes and Matroids

Emmanuel Abbe, Emre Telatar
Information Processing group, EPFL
Lausanne 1015, Switzerland
Email: {emmanuel.abbe,emre.telatar}@epfl.ch

***Abstract*—In this paper, a polar code for the m-user multiple access channel (MAC) with binary inputs is constructed. In particular, Arıkan's polarization technique applied individually to each user polarizes any m-user binary input MAC into a finite collection of extremal MACs. The extremal MACs have a number of desirable properties: (i) the 'uniform sum rate'[1] of the original channel is not lost, (ii) the extremal MACs have rate regions that are not only polymatroids but matroids and thus (iii) their uniform sum rate can be reached by each user transmitting either uncoded or fixed bits; in this sense they are easy to communicate over. A polar code can then be constructed with an encoding and decoding complexity of $O(n \log n)$ (where n is the block length), a block error probability of $o(\exp(-n^{1/2-\varepsilon}))$, and capable of achieving the uniform sum rate of any binary input MAC with arbitrary many users. An application of this polar code construction to a coding scheme for the AWGN channel is also discussed.**

I. Introduction

In [2], Arıkan shows that a single-user binary input channel can be "polarized" by a simple process that coverts n independent uses of this channel into n successive uses of channels which are almost "extremal". The extremal single-user channels are either perfect or pure noise channels, i.e., with a uniform mutual information equal either to 1 or to 0. Furthermore, the fraction of almost perfect channels resulting from the polarization process is close to the uniform mutual information (called 'symmetric capacity' in [2]) of the original channel. This process allows then the deterministic construction of a code with low encoding and decoding complexity, and capable of achieving the uniform mutual information of any binary input channel. By definition of the uniform mutual information, this code achieves the capacity of any channel whose capacity achieving input distribution is the uniform distribution. Moreover, [11] generalizes the original polar code construction to channels with input alphabet of arbitrary prime cardinality, allowing to approach the capacity of any discrete memoryless channels with polar codes.

For a 2-user binary input MAC, by applying Arikan's construction to each user's input separately, [12] shows that a similar phenomenon appears: the n independent uses of the MAC are converted into n successive uses of MACs which are close to"extremal MACs". These extremal MACs are of four kinds: (i) each user sees a pure noise channel, (ii) one of the user sees a pure noisy channel and the other sees a perfect channel, (iii) both users see a perfect channel, (iv) a pure contention channel: a channel whose uniform rate region is the triangle with vertices (0,0), (0,1), (1,0). Note that for this channel, either of the two user communicates at zero rate, the other user sees a perfect channel. Moreover [12] shows that the uniform sum rate of the original MAC is preserved during the polarization process, and that the polarization to the extremal MACs occurs fast enough, so as to ensure a vanishing block error probability. This allows the construction of a polar code to achieve reliable communication at uniform sum rate on binary input 2-user MACs.

In this paper, we investigate the case where the number of users m is arbitrary. In the two user case, the extremal MACs are not just MACs for which each users sees either a perfect or pure noise channel, as there is also the pure contention channel. However, the uniform rate region of the 2-user extremal MACs are all polyhedron with integer valued constraints. We will see in this paper that the approach used for the 2-user case faces a new phenomenon when the number of users reaches 4, and the extremal MACs are no longer in a one to one correspondence with the polyhedron having integer valued constraints. To characterize the extremal MACs, we first show how an unusual relationship between random variables defined in terms of mutual information falls precisely within the independence notion of the matroid theory. This relationship is used to show that the extremal MACs are in a one to one correspondence with binary matroids, and are "equivalent" (in a sense which will be defined later) to linear deterministic channels. This is then used to conclude the construction of a polar code ensuring reliable communication on binary input MACs for arbitrary values of m. Finally, we discuss the implications of taking m arbitrary large, for the purpose of coding over an additive white Gaussian noise channel.

II. The Polarization Construction

We consider a m-user multiple access channel with binary input alphabets (BMAC) and arbitrary output alphabet $\mathcal{Y}$. The channel is specified by the conditional probabilities

$$P(y|\bar{x}), \quad \text{for all } y \in \mathcal{Y} \text{ and } \bar{x} = (x[1], \dots, x[m]) \in \mathbb{F}_2^m.$$

Let $E_m := \{1, \dots, m\}$ and let $X[1], \dots, X[m]$ be mutually independent and uniformly distributed binary random variables. Let $\bar{X} := (X[1], \dots, X[m])$. We denote by Y the

[1] In this paper all mutual informations are computed when the inputs of a MAC are distributed independently and uniformly. The resulting rate regions, sum rates, etc., are prefixed by 'uniform' to distinguish them from the capacity region, sum capacity, etc.

output of $\bar{X}$ through the MAC P. For $J \subseteq E_m$, we define

$$X[J] := \{X[i] : i \in J\},$$
$$I[J](P) := I(X[J]; Y X[J^c]),$$

where J^c denotes the complement set of J in E_m, and

$$I(P) : 2^m \to \mathbb{R}$$
$$J \mapsto I[J](P)$$

where 2^m denotes the power set of E_m and where $I[\emptyset](P) = 0$. Note that

$$\mathcal{I}(P) := \{(R_1, \ldots, R_m) : 0 \leq \sum_{i \in J} R_i \leq I[J](P), \forall J \subseteq E_m\}$$

is an inner region to the capacity region of the MAC P. We refer to $\mathcal{I}(P)$ as the uniform rate region and to $I[E_m](P)$ as the uniform sum rate. We now consider two independent uses of such a MAC. We define

$$\bar{X}_1 := (X_1[1], \ldots, X_1[m]), \quad \bar{X}_2 := (X_2[1], \ldots, X_2[m]),$$

where $X_1[i], X_2[i]$, with $i \in E_m$, are mutually independent and uniformly distributed binary random variables. We denote by Y_1 and Y_2 the respective outputs of $\bar{X}_1$ and $\bar{X}_2$ through two independent uses of the MAC P:

$$\bar{X}_1 \xrightarrow{P} Y_1, \quad \bar{X}_2 \xrightarrow{P} Y_2. \tag{1}$$

We define two additional binary random vectors

$$\bar{U}_1 := (U_1[1], \ldots, U_1[m]), \quad \bar{U}_2 := (U_2[1], \ldots, U_2[m])$$

with mutually independent and uniformly distributed components, and we put $\bar{X}_1$ and $\bar{X}_2$ in the following one to one correspondence with $\bar{U}_1$ and $\bar{U}_2$:

$$\bar{X}_1 = \bar{U}_1 + \bar{U}_2, \qquad \bar{X}_2 = \bar{U}_2,$$

where the addition in the above is the modulo 2 component wise addition.

Definition 1. Let $P : \mathbb{F}^m \to \mathcal{Y}$ be a m-user BMAC. We define two new m-user BMACs, $P^- : \mathbb{F}_2^m \to \mathcal{Y}^2$ and $P^+ : \mathbb{F}_2^m \to \mathcal{Y}^2 \times \mathbb{F}_2^m$, by

$$P^-(y_1, y_2|\bar{u}_1) := \sum_{\bar{u}_2 \in \mathbb{F}_2^m} \frac{1}{2^m} P(y_1|\bar{u}_1 + \bar{u}_2) P(y_2|\bar{u}_2),$$
$$P^+(y_1, y_2, \bar{u}_1|\bar{u}_2) := \frac{1}{2^m} P(y_1|\bar{u}_1 + \bar{u}_2) P(y_2|\bar{u}_2),$$

for all $\bar{u}_i \in \mathbb{F}_2^m$, $y_i \in \mathcal{Y}$, $i = 1, 2$.

That is, we have now two new m-user BMACs with extended output alphabets:

$$\bar{U}_1 \xrightarrow{P^-} (Y_1, Y_2), \quad \bar{U}_2 \xrightarrow{P^+} (Y_1, Y_2, \bar{U}_1) \tag{2}$$

which also defines $I[J](P^-)$ and $I[J](P^+)$, $\forall J \subseteq E_m$.

This construction is the natural extension of the construction for $m = 1, 2$ in [2], [12]. Here again, we are comparing two independent uses of the same channel P (cf. (1)) with two successive uses of the channels P^- and P^+ (cf. (2)). Note that

$$I[J](P^-) \leq I[J](P) \leq I[J](P^+), \quad \forall J \subseteq E_m.$$

Definition 2. Let $\{B_n\}_{n \geq 1}$ be i.i.d. uniform random variables valued in $\{-, +\}$. Let the random processes $\{P_n,\ n \geq 0\}$ be defined by

$$P_0 := P,$$
$$P_n := P_{n-1}^{B_n}, \quad \forall n \geq 1. \tag{3}$$

A. Discussion

When $m = 1$, we have $2I(P) = I(P^-) + I(P^+)$, which implies that $I(P_n)$ (which in this case denoting a sequence of scalar random variable and not of functions) is a martingale. However when $m \geq 2$, it is not clear that the same will hold for $I[J](P_n)$ for different J's. Moreover, it has been shown that, when $m = 1$, I_n tends to either 0 or 1. Which means that the extremal channels of the single-user polarization scheme are either pure noise or perfect channels. This then facilitates the design of a code: Provided that successive cancellation decoding is used, frozen bits are set on the very noisy channels and uncoded information bits are sent and recovered on the almost perfect channels. When $m \geq 2$, several new points need to be investigated. In particular, one needs to check what happens to the martingale property for different J's. Then, if the convergence of each $I[J](P_n)$ can be proved, one has to examine whether the obtained limiting MACs are also extremal MACs, along the spirit of creating pure noise and perfect channels as in the single-user polarization. Moreover, in order to obtain a coding theorem, one has to ensure that the convergence of these mutual informations is taking place fast enough, so as to ensure a block error probability that tends to zero for the coding scheme.

III. Results

Summary: In Section III-A, we show that $I(P_n)$ tends a.s. to a matroid rank function (cf. Definition 5). We then see in Section III-B that the extreme points of a uniform rate region with matroidal constraints can be achieved by each user sending uncoded or frozen bits; in particular the uniform sum rate can be achieved by such strategies. We then show in Section III-D, that for arbitrary m, $I(P_n)$ does not tend to an arbitrary matroid rank function but to the rank function of a binary matroid (cf. Definition 6). This is used to show that the convergence to the extremal MACs happens fast enough, and that the construction of previous section leads to a polar code having a low encoding and decoding complexity and achieving the uniform sum rate on any binary MAC.

A. The extremal MACs

Lemma 1. *$\{I[J](P_n), n \geq 0\}$ is a bounded super-martingale when $J \subsetneq E_m$ and a bounded martingale when $J = E_m$.*

Proof: For any $J \subseteq E_m$, $I[J](P_n) \le m$ and

$$\begin{aligned} 2I[J](P) &= I(X_1[J]X_2[J]; Y_1Y_2X_1[J^c]X_2[J^c]) \\ &= I(U_1[J]U_2[J]; Y_1Y_2U_1[J^c]U_2[J^c]) \\ &= I(U_1[J]; Y_1Y_2U_1[J^c]U_2[J^c]) \\ &\quad + I(U_2[J]; Y_1Y_2U_1[J^c]U_2[J^c]U_1[J]) \\ &\ge I(U_1[J]; Y_1Y_2U_1[J^c]) \\ &\quad + I(U_2[J]; Y_1Y_2\bar{U}_1U_2[J^c]) \\ &= I[J](P^-) + I[J](P^+), \end{aligned} \tag{4}$$

where equality holds above, if $J^c = \emptyset$, i.e., if $J = E_m$. ■

Note that the inequality in the above are only due to the bounds on the mutual informations of the P^- channel. Because of the equality when $J = E_m$, our construction preserves the uniform sum rate. As a corollary of previous Lemma, we have the following result.

Theorem 1. *The process $\{I[J](P_n), J \subseteq E_m\}$ converges a.s..*

Note that for a fixed n, $\{I[J](P_n), J \subseteq E_m\}$ denotes the collection of the 2^m random variables $I[J](P_n)$, for $J \subseteq E_m$. When the convergence takes place (a.s.), let us define

$$I_\infty[J] := \lim_{n\to\infty} I[J](P_n)$$

and I_∞ to be the function $J \mapsto I_\infty[J]$.

From previous theorem, $I_\infty[J]$ is a random variable valued in $[0, |J|]$. We will now further characterize these random variables.

The following Lemma is proved in [2].

Lemma 2. *For any $\varepsilon > 0$, there exists $\delta > 0$ such that $I(A_2; B_1B_2A_1) - I(A_2; B_2) < \delta$ implies*

$$I(A_2; B_2) \in [0, \varepsilon) \cup (1-\varepsilon, 1],$$

whenever (A_1, A_2, B_1, B_2) are random variables valued in $\mathbb{F}_2 \times \mathbb{F}_2 \times \mathcal{B} \times \mathcal{B}$, with $\mathcal{B}$ any set, and

$$\mathbb{P}_{A_1A_2B_1B_2}(a_1, a_2, b_1, b_2) = \frac{1}{4} Q(b_1|a_1 + a_2) Q(b_2|a_2),$$

for any $a_i \in \mathbb{F}_2$, $b_i \in \mathcal{B}$, $i = 1, 2$, where Q is a binary input $\mathcal{B}$-output channel.

This Lemma is used to prove the following.

Lemma 3. *For any $\varepsilon > 0$ and any m-user BMAC P, there exists $\delta > 0$, such that for any $J \subseteq E_m$, if $I[J](P^+) - I[J](P) < \delta$, we have*

$$I[J](P) - I[J \setminus i] \in [0, \varepsilon) \cup (1-\varepsilon, 1], \quad \forall i \in J,$$

where $I[\emptyset] = 0$.

Proof: Let $i \in J$. Note that

$$\begin{aligned} &I[J](P^+) - I[J](P) \\ &= I(U_2[J]; Y_1Y_2\bar{U}_1U_2[J^c]) - I(U_2[J]; Y_2U_2[J^c]) \\ &= I(U_2[J]; Y_1\bar{U}_1|Y_2U_2[J^c]) \\ &\ge I(U_2[i]; Y_1U_1[i]U_1[J^c]|Y_2U_2[J^c]) \\ &= I(U_2[i]; Y_1U_1[J^c]Y_2U_2[J^c]U_1[i]) - I(U_2[i]; Y_2U_2[J^c]) \\ &= I(U_2[i]; Y_1X_1[J^c]Y_2X_2[J^c]U_1[i]) - I(U_2[i]; Y_2X_2[J^c]). \end{aligned}$$

Using Lemma 2 with $A_k = U_k[i]$, $B_k = Y_kX_k[J^c]$, $k = 1, 2$, and

$$Q(y, x[J^c]|x[i]) = \frac{1}{2^{m-1}} \sum_{x[j]\in\mathbb{F}_2, j\notin J^c\cup\{i\}} P(y|\bar{x}),$$

we conclude that we can take δ small enough, so that $I[J](P^+) - I[J](P) < \delta$ implies $I(U_2[i]; Y_2X_2[J^c]) \in [0, \varepsilon/2) \cup (1-\varepsilon/2, 1]$. Moreover, we have

$$I[J](P) = I[J \setminus i](P) + I(U_2[i]; Y_2X_2[J^c]).$$

■

Lemma 4. *With probability one, $I_\infty[J] - I_\infty[J \setminus i] \in \{0, 1\}$, $\forall J \subseteq E_m, i \in J$, where $I_\infty[\emptyset] := 0$.*

Proof: Let $i \in E_m$. From Lemma 1, we have that $I[i](P_n)$ converges a.s., hence $\lim_{n\to\infty} |I[i](P_{n+1}) - I[i](P_n)| = 0$ a.s. Moreover, $|I[i](P_{n+1}) - I[i](P_n)|$ is equal to $I[i](P_n^+) - I[i](P_n)$ w.p. half and $I[i](P_n) - I[i](P_n^-)$ w.p. half. Hence, from (4), $\mathbb{E}|I[i](P_{n+1}) - I[i](P_n)| \ge \frac{1}{2}(I[i](P_n^+) - I[i](P_n))$. But $|I[i](P_{n+1}) - I[i](P_n)|$ is bounded by 1, hence $\lim_{n\to\infty} \mathbb{E}|I[i](P_{n+1}) - I[i](P_n)| = 0$ and $\lim_{n\to\infty} I[i](P_n^+) - I[i](P_n) = 0$. Finally, we conclude using Lemma 3 that $I_\infty[i] \in \{0, 1\}$. The other claim in the Lemma's statement is proved analogously using Lemma 3. ■

Note that Lemma 4 implies in particular that $\{I_\infty[J], J \subseteq E_m\}$ is a.s. a discrete random vector.

Definition 3. We denote by $\mathcal{A}_m$ the support of $\{I_\infty[J], J \subseteq E_m\}$ (when the convergence takes place, i.e., a.s.). This is a subset of $\{0, \ldots, m\}^{2^m}$.

We have already seen that not every element in $\{0, \ldots, m\}^{2^m}$ can belong to $\mathcal{A}_m$. We will now further characterize the set $\mathcal{A}_m$.

Definition 4. A polymatroid is a set E_m, called a ground set, equipped with a function $f : 2^m \to \mathbb{R}$ (where 2^m denotes the power set of E_m), called a rank function, which satisfies

$$\begin{aligned} &f(\emptyset) = 0, \\ &f[J] \le f[K], \quad \forall J \subseteq K \subseteq E_m, \\ &f[J \cup K] + f[J \cap K] \le f[J] + f[K], \quad \forall J, K \subseteq E_m. \end{aligned}$$

A proof of the following result can be found in [13].

Theorem 2. *For any MAC and any distribution of the inputs $X[E]$, we have that $\rho(S) = I(X[S]; YX[S^c])$ is a rank function on E, where we denote by Y the output of the MAC with input $X[E]$. Hence, (E, ρ) is a polymatroid.*

Therefore, any realization of $I(P_n)$ is a rank function and the elements of $\mathcal{A}_m$ are the image of a polymatroid rank function.

Definition 5. A matroid is a polymatroid whose rank function is integer valued and satisfies $f(J) \le |J|$, $\forall J \subseteq E_m$. We denote by MAT_m the set of all matroids with ground state E_m. We use the notation $r_{\mathbb{B}}$ to refer to the rank function of a matroid $\mathbb{B}$. We will sometimes identify a matroid with its rank function image, in which case, we consider an element of MAT_m as a 2^m integer valued vector. We also define a basis of a matroid by the collection of maximal subsets of E_m for which $f(J) = |J|$. One can show that a matroid is equivalently defined from its bases.

Using Lemma 4 and the definition of a matroid, we have the following result.

Theorem 3. *For every $m \geq 1$, $\mathcal{A}_m \subseteq \mathrm{MAT}_m$, i.e., I_∞ is a matroid rank function.*

We will see that the inclusion is strict for $m \geq 4$.

B. Communicating on Matroids

We have shown that, when n tends to infinity, the MACs that we create with the polarization construction of Section II are particular MACs: the mutual informations $I_\infty[J]$ are integer valued (and satisfy the other matroid properties). A well-known result of matroid theory (cf. Theorem 22 of [4]) says that the vertices of a polymatroid given by a rank function f are the vectors of the following form:

$$\begin{aligned} &x_{j(1)} = f(A_1), \\ &x_{j(i)} = f(A_i) - f(A_{i-1}), \quad \forall 2 \leq i \leq k \\ &x_{j(i)} = 0, \quad \forall k < i \leq m, \end{aligned}$$

for some $k \leq m$, $j(1), j(2), \ldots, j(m)$ distinct in E_m and $A_i = \{j(1), j(2), \ldots, j(i)\}$, where the vertices strictly in the positive orthant are given for $k = m$.

Therefore, we have the following corollary.

Corollary 1. *The uniform rate region defined by an element of $\mathcal{A}_m$ has vertices on the hypercube $\{0,1\}^m$. In particular, when operating at a vertex of a MAC having such uniform rate region, each user sees either a perfect or pure noise channel.*

C. Convergence Speed and Representation of Matroids

Convention: for a given m, we write the collection $\{I_\infty[J], J \subseteq E_m\}$ by skipping the empty set (since $I_\infty[\emptyset] = 0$) and as follows: when $m = 2$, we order the sequence as $(I_\infty[1], I_\infty[2], I_\infty[1,2])$, and when $m = 3$, as $(I_\infty[1], I_\infty[2], I_\infty[3], I_\infty[1,2], I_\infty[1,3], I_\infty[2,3], I_\infty[1,2,3])$, etc.

When $m = 2$, [12] shows that $\{I_\infty[J], J \subseteq E_m\}$ belongs a.s. to $\{(0,0,0), (0,1,1), (1,0,1), (1,1,1), (1,1,2)\}$. These are precisely all the matroids with two elements. The speed of convergence to these matroids is shown to be fast in [12] through the following steps. The main idea is to deduce the convergence speed of $I[J](P_n)$ from the convergence speed result obtained in the single-user setting, which is fast, namely $o(2^{-n^\beta})$, for any $\beta < 1/2$, cf. [3]. We do not need to check the speed convergence to $(0,0,0)$. For $(1,0,1)$, the speed convergence can be deduced from the $m = 1$ speed convergence result as follows. First note that $I(X[1];Y) \leq I[1](P) = I(X[1]; YX[2])$. Then, it is shown that, if $I[1](P_n)$ tends to 1, it must be that along those paths of the B_n process, $\hat{I}[1](P_n)$ tends to 1 as well, where $\hat{I}[i](P) = I(X[i]; Y)$. Now, since $\hat{I}[1](P_n)$ tends to 1, one can show that it must tend fast from the single-user result. A similar treatment can be done for $(0,1,1)$ and $(1,1,2)$. However, for $(1,1,1)$, another step is required. Indeed, for this case, $\hat{I}[1](P_n)$ and $\hat{I}[2](P_n)$ tend to zero. Hence, $\hat{I}[1,2](P) = I(X[1] + X[2]; Y)$ is introduced and it is shown that $\hat{I}[1,2](P_n)$ tends to 1. Moreover, if we denote by Q the single-user channel between $X[1]+X[2]$ and Y, we have that $\hat{I}[1,2](P) = I(Q)$, $\hat{I}[1,2](P^-) = I(Q^-)$ and $\hat{I}[1,2](P^+) = I(U_2[1] + U_2[2]; Y_1 Y_2 U_1[1] U_1[2]) \geq I(U_2[1] + U_2[2]; Y_1 Y_2 U_1[1] + U_1[2]) = I(Q^+)$. Hence, using the single-user channel result, $\hat{I}[1,2](P_n)$ tends to 1 fast.

Note that a property of the matroids $\{(0,0,0), (0,1,1), (1,0,1), (1,1,1), (1,1,2)\}$ is that we can express any of them as the uniform rate region of a linear deterministic channel: $(1,0,1)$ is in particular the uniform rate region of the MAC whose output is $Y = X[1]$, $(0,1,1)$ corresponds to $Y = X[2]$, $(1,1,1)$ to $Y = X[1] + X[2]$ and $(1,1,2)$ to $(Y_1, Y_2) = (X[1], X[2])$. Indeed, this is related to the fact that any matroid with a two element ground state can be represented in the binary field. Let us introduce the definition of binary matroids.

Definition 6. *Linear matroids:* let A be a $k \times m$ matrix over a field. Let E_m be the index set of the columns in A. The rank of $J \subseteq E_m$ is defined by the rank of the sub-matrix with columns indexed by J.
Binary matroids: A matroid is binary if it is a linear matroid over the binary field. We denote by BMAT_m the set of binary matroids with m elements.

1) The $m = 3$ Case: MAT_3 is given by 8 unlabeled matroids (16 labeled ones). Moreover, they are all binary representable (there are 16 labeled binary matroids). For example, it is clear that the deterministic MAC whose output is $X[1] + X[2] + X[3]$ has a uniform rate region given by $(1,1,1,1,1,1,1)$. Similarly, all matroids for $m = 3$ correspond to the rate region of a linear deterministic MAC. However, one can also show that any 3-user binary MAC with uniform rate region given by a matroid is equivalent to a linear deterministic channel in the following sense. A MAC with output Y and uniform rate region given by $(1,1,1,1,1,1,1)$ must satisfy $I(X[1] + X[2] + X[3]; Y) = 1$, and similarly for other matroids (with $m = 3$), where the linear forms of inputs which can be recovered from the output are dictated by the binary representation of the matroid. However, the above claim is not quite sufficient to show that, if $\{I[J](P_n), J \subset E_m\}$ tends to $(1,1,1,1,1,1,1)$, we have along this path that $\hat{I}[1,2,3](P^n)$ tends to 1, where $\hat{I}[1,2,3](P) = I(X[1] + X[2] + X[3]; Y)$. For this, one can show a stronger version of the claim which says that if a MAC has a uniform rate region "close to" $(1,1,1,1,1,1,1)$, it must be that $I(X[1]+X[2]+X[3];Y)$ is close to 1. In any case, a similar technique as for the $m = 2$ case lets one show that the convergence to the matroids in $\mathcal{A}_3$ must take place fast enough, using the single-user result.

2) The $m = 4$ Case: We have that MAT_4 contains 17 unlabeled matroids (68 labeled ones). However, there are only 16 unlabeled binary matroids with ground state 4. Hence, there must be a matroid which does not have a binary representation.

This matroid is given by $(1,1,1,1,2,2,2,2,2,2,2,2,2,2,2)$ (one can easily check that this is not a binary matroid). It is denoted $U_{2,4}$ and is the uniform matroid of rank 2 with 4 elements (for which any 2 elements set is a basis). Of course, that this matroid is not binary, does not imply that an hypothetic convergence to it must be slow. It means that we will not be able to use our previous technique with linear deterministic channels and single-user convergence result.

Luckily, one can show that there is no MAC leading to $U_{2,4}$ and the following holds.

Lemma 5. $\mathcal{A}_4 \subset \mathrm{BMAT}_4 \subsetneq \mathrm{MAT}_4$.

Hence, the $m = 4$ case can be treated in a similar manner as the previous cases.
We conclude this section by proving the following result, which implies Lemma 5.

Lemma 6. *$U_{2,4}$ cannot be the uniform rate region of any MAC with four users and binary inputs.*

Proof: Assume that $U_{2,4}$ is the uniform rate region of a MAC. We then have

$$I(X[i,j];Y) = 0, \quad (5)$$

$$I(X[i,j];YX[k,l]) = 2, \quad (6)$$

for all i,j,k,l distinct in $\{1,2,3,4\}$. Let y_0 be in the support of Y. For $x \in \mathbb{F}_2^4$, define $\mathbb{P}(x|y_0) = W(y_0|x)/\sum_{z\in\mathbb{F}_2^4} W(y_0|z)$. Then from (6), $\mathbb{P}(0,0,*,*|y_0) = 0$ for any choice of $*,*$ which is not $0,0$ and $\mathbb{P}(0,1,*,*|y_0) = 0$ for any choice of $*,*$ which is not $1,1$. On the other hand, from (5), $\mathbb{P}(0,1,1,1|y_0)$ must be equal to p_0. However, we have form (6) that $\mathbb{P}(1,0,*,*|y_0) = 0$ for any choice of $*,*$ (even for $1,1$ since we now have $\mathbb{P}(0,1,1,1|y_0) > 0$). At the same time, this implies that the average of $\mathbb{P}(1,0,*,*|y_0)$ over $*,*$ is zero. This brings a contradiction, since from (5), this average must equal to p_0. ∎

Moreover, a similar argument can be used to prove a stronger version of Lemma 6 to show that no sequence of MACs can have a uniform rate region that converges to $U_{2,4}$.

3) Arbitrary values of m: We have seen in previous section that for $m = 2,3,4$, the extremal MACs have uniform rate region that are not any matroids but binary matroids. This allows us to conclude that for $m = 2,3,4$, $\{I[J](P_n), J \subseteq E_m\}$ must tend fast enough to $\{I_\infty[J], J \subseteq E_m\}$. Indeed, by working with the linear deterministic representation of the MACs, the problem of showing that the convergence speed is fast in the MAC setting is a consequence of the result shown in [2] for the single-user setting.

We now show that this approach can be used for any value of m.

Definition 7. A matroid is BUMAC if its rank function can be expressed as $r(J) = I(X[J];YX[J^c])$, $J \subseteq E_m$, where $X[E]$ has independent and binary uniformly distributed components, and Y is the output of a binary input MAC with input $x[E]$.

Theorem 4. *A matroid is BUMAC if and only if it is binary.*

The converse of this theorem is easily proved and the direct part uses the following steps, which are detailed in [1]. First the following theorem on the representation of binary matroids due to Tutte, whose proof can be found in [10].

Theorem 5 (Tutte)**.** *A matroid is binary if and only if it has no minor that is $U_{2,4}$.*

A minor of matroid is a matroid which is either a restriction or a contraction of the original matroid to a subset of the ground set. A contraction can be defined as a restriction on the dual matroid, which is another matroid whose bases are the complement set of the bases of the original matroid. Using Lemma 5, we already know that $U_{2,4}$ is not a restriction of any BUMAC matroid. To show that a BUMAC matroid cannot have $U_{2,4}$ as a contraction, Lemma 5 can be used in a dual manner, since one can show that the rank function of the dual of a BUMAC matroid is given by $r^*(J) = |J| - I(X[J];Y)$.

In the following theorem, we connect extremal MACs to linear deterministic channels.

Theorem 6. *Let $X[E]$ have independent and binary uniformly distributed components. Let Y be the output of a MAC with input $X[E]$ and for which $f(J) = I(X[J];YX[J^c])$ is integer valued, for any $J \subseteq E_m$. From Theorem 4, $f(\cdot)$ is also the rank function of a binary matroid, so let A be a matrix representation of this binary matroid. We then have*

$$I(AX[E];Y) = \mathrm{rank}A = f(E_m).$$

The proof of this theorem, together with further investigations on this subject can be found in [1]. Moreover, one can show a stronger version of these theorems for MACs having a uniform rate region which tends to a matroid.

Theorem 7. *Let $X[E]$ have independent and binary uniformly distributed components. Let Y be the output of a MAC with input $X[E]$ and for which $f : 2^m \ni J \mapsto I(X[J];YX[J^c])$ satisfies $\max_{J\in 2^m} d(f(J),\mathbb{Z}) < \varepsilon$ for $\varepsilon > 0$. Then, $\|f - r\|_{l_1} \le \delta(\varepsilon)$, where r is the rank function of a binary matroid and where $\delta(\varepsilon)$ tends to zero with ε. Moreover,*

$$|I(AX[E];Y) - f(E_m)| < \gamma(\varepsilon),$$

where $\gamma(\varepsilon)$ tends to zero with ε.

Theorem 4 says that an extremal MAC must have (with probability one) the same uniform rate region as the one of a linear deterministic channel, i.e., a channel whose output is a collection of linear forms of the inputs. However, Theorem 6, says something stronger, namely, that from the output of an extremal MAC, one can recover a collection of linear forms of the inputs and essentially nothing else. In that sense, extremal MACs are equivalent to linear deterministic channels. This also suggests that we could have started from the beginning by working with $S[J](P) := I(\sum_{i\in J} X_i;Y)$ instead of $I[J](P) = I(X[J];YX[J^c])$ to analyze the polarization of a MAC. The second measure is the natural one to study a MAC, since it characterizes the rate region. However, we have just shown that it is sufficient to work with the first measure to characterize the uniform rate regions of the polarized MACs. Indeed, one can show that $S[J](P_n)$ tends either to 0 or 1 and Eren Şaşoğlu has provided a direct argument showing that these measures fully characterize the uniform rate region of the extremal MACs. Moreover, the process of identifying which matroids can have a rank function derived from an information theoretic measure, such as the entropy, has been investigated in different works, cf. [14] and references therein. In particular, the problem of characterizing the entropic matroids has consequent applications in network information theory and network coding problems as described in [7].

Entropic matroids are defined as follows. Let E be a finite set and $X[E] = \{X_i\}_{i\in E}$ be a random vector with each component valued in a finite alphabet. Let $h(I) := h(X[I])$.

Theorem 8. *$h(\cdot)$ is a rank function. Hence, (E, h) is a polymatroid.*

A (poly)matroid is then called entropic, if its rank function can expressed as the entropy of a certain random vector, as above. A proof of previous theorem is available in [8]. The work of Han, Fujishige, Zhang and Yeung, [6], [5], [14] has resulted in the complete characterization of entropic matroids for $|E| = 2, 3$. However, the problem is open when $|E| \geq 4$. Note that in our case, where we have been interested in characterizing BUMAC matroids instead of entropic matroids, we have also faced a different phenomenon when $|E| \geq 4$. Other similar problems have been studied in [9].

In the next section, we collect the results of previous sections and describe the polar code construction for the MAC.

D. MAC Polar code construction

Let $n = 2^l$ for some $l \in \mathbb{Z}_+$ and let $G_n = \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}^{\otimes l}$ denote the lth Kronecker power of the given matrix. Let $U[k]^n := (U_1[k], \ldots, U_n[k])$ and

$$X[k]^n = U[k]^n G_n, \quad k \in E_m.$$

When $X[E_m]^n$ is transmitted over n independent uses of P to receive Y^n, define for any $i \in \{1, \ldots, n\}$ the channel $P_{(i)} : \mathbb{F}_2^m \to \mathcal{Y}^n \times \mathbb{F}_2^{m(i-1)}$ to be the channel whose inputs and outputs are $U_i[E_m] \to Y^n U^{i-1}[E_m]$. Let $\varepsilon > 0$ and let $A[k] \subset \{1, \ldots, n\}$ denote the sets of indices where information bits are transmitted by user k, which are chosen as follows: for a fixed $i \in \{1, \ldots, n\}$, if

$$\|I(P_{(i)}) - r_{\mathbb{B}}\|_{l_1} < \varepsilon$$

for the rank function $r_{\mathbb{B}}$ of a binary matroid $\mathbb{B}$, then pick a basis of $\mathbb{B}$ and include i in $A[k]$ if k belongs to that basis. If no such binary matroid exists, do not include i in $A[k]$ for all $k \in E_m$. Choose the bits indexed by $A[k]^c$, for all k, independently and uniformly at at random, and reveal their values to the transmitter and receiver.

For an output sequence Y^n, the receiver can then decode successively $U_1[E_m]$, then $U_2[E_m]$, etc., until $U_n[E_m]$. Moreover, since $I[E_m](P)$ is preserved through the polarization process (cf. the equality in (4)), we guarantee that for every $\delta > 0$, there exists a n_0 such that

$$\sum_{k=1}^{m} |A[k]| > n(I[E_m](P) - \delta), \quad \forall n \geq n_0.$$

Using the results of previous section, we can then show the following theorem, which ensures that the code described above allows reliable communication at sum rate.

Theorem 9. *For any $\beta < 1/2$, the block error probability of the code described above, under successive cancellation decoding, is $o(2^{-n^\beta})$.*

Moreover, this code has an encoding and decoding complexity of $O(n \log n)$ and a block error probability of $o(\exp(-n^{1/2-\varepsilon}))$, for any $\varepsilon > 0$.

E. Proof of Theorem 9

The following theorem summarizes the results achieved until Theorem 4.

Theorem 10. *Let P be a m-user binary input MAC. For any $\varepsilon > 0$, we have*

$$\lim_{l\to\infty} \frac{1}{2^l} |\{i \in \{1, \ldots, 2^l\} : d(I(P_{(i)}), \mathrm{BMAT}_m) \geq \varepsilon\}| = 0,$$

where

$$d(I(P_{(i)}), \mathrm{BMAT}_m) = \min_{r_{\mathbb{B}}:\, \mathbb{B}\in \mathrm{BMAT}_m} \|I(P_{(i)}) - r_{\mathbb{B}}\|_{l_1}.$$

Proof: When P_l is defined as in (3), we have

$$\Pr\{d(I(P_l), \mathrm{BMAT}_m) \geq \varepsilon\} = \frac{1}{2^l} |\{i \in \{1, \ldots, 2^l\} : d(I(P_{(i)}), \mathrm{BMAT}_m) \geq \varepsilon\}|$$

and we conclude using Theorem 3 and Theorem 4. ∎

Let E_i denote the event of making an error when decoding $U_i[E_m]$ given the output Y^n. Note that a transmission error when using previous code over a block length n is given by

$$\bigcup_{i=1}^{n} E_i = \bigsqcup_{i=1}^{n} \left(\cup_{j=1}^{i} E_{j-1}^c \cup E_i\right),$$

where $E_0^c = \emptyset$, and the block error probability is given by

$$\begin{aligned}\Pr\{\cup_{i=1}^n E_i\} &= \sum_{i=1}^{n} \Pr\{\cup_{j=1}^{i} E_{j-1}^c \cup E_i\}\\ &\leq \sum_{i=1}^{n} \Pr\{E_i | \cup_{j=1}^{i} E_{j-1}^c\}. \qquad (7)\end{aligned}$$

The ith term in (7) is the probability of making an error when decoding the inputs of channel $P_{(i)}$ given its output. We want to use the fact that, when l grows, this channel tends to an extremal channel, and hence the error probability must vanish when transmitting the information bits appropriately as done in (III-D). However, we must ensure that this happens fast enough, in order to guarantee that the summation in (7) tends to zero as well. For this purpose, we need to introduce new mathematical objects to bound the error probabilities appearing in (7).

Definition 8. For a m-user BMAC P with output alphabet $\mathcal{Y}$ and for $S \subseteq E_m$, we define $P^{[S]}$ to be the single-user binary input channel with output alphabet $\mathcal{Y}$, obtained from P by

$$P^{[S]}(y|s) = \frac{1}{2^{m-1}} \sum_{x[E_m] \in \mathbb{F}_2^m : \sum_{i \in S} x_i = s} P(y|x[E_m])$$

for all $y \in \mathcal{Y}$, $s \in \mathbb{F}_2$.

Schematically, if $P : X[E_m] \to Y$, we have $P^{[S]} : \sum_{i \in S} X_i \to Y$.

Definition 9. The *Bhattacharyya parameter* of a single-user channel Q with binary input and output alphabet $\mathcal{Y}$ is defined by

$$Z(Q) = \sum_{y \in \mathcal{Y}} \sqrt{Q(y|0)Q(y|1)}.$$

We know from [3], that for a single-user binary input channel Q, the convergence to extremal channels is fast:

$$\lim_{l \to \infty} \frac{1}{2^l} |\{i \in \{1, \ldots, 2^l\} : \|I(Q_{(i)}) - 1\| < \varepsilon, \quad Z(Q_{(i)}) \geq 2^{-2^{l\beta}}\}| = 0, \qquad (8)$$

where $\beta < 1/2$ and where $Q_{(1)}, \ldots, Q_{(n)}$ are the channels obtained from the polarization construction in the single-user setting. We would like to conclude an equivalent statement for $P_{(i)}^{[S]}$. However, $P_{(i)}^{[S]}$ is a channel with input $\sum_{s \in S} U_i[s]$ and output $U^{i-1}[E_m]Y^n$ and does not correspond to a channel obtained in a single-user polarization construction. Nevertheless, the output set contains "more information" than needed to fall back on an expression appearing for the single-user polarization construction, since we can process the output $U^{i-1}[E_m]Y^n$ to get $\sum_{s \in S} U_1[s], \ldots, \sum_{s \in S} U_{i-1}[s], Y^n$, and the following holds.

Lemma 7. *For $i \in \{1, \ldots, n\}$ and $S \subseteq E_m$, let $Q_{(i)}^{[S]}$ be the single-user channel with input $\sum_{s \in S} U_i[s]$ and output $\sum_{s \in S} U_1[s], \ldots, \sum_{s \in S} U_{i-1}[s], Y^n$. We then have,*

$$Z(P_{(i)}^{[S]}) \leq Z(Q_{(i)}^{[S]}).$$

The following lemma allows to conclude the proof of Theorem 9.

Lemma 8. *If $\|I(P_{(i)}) - r_{\mathbb{B}}\|_{l_1} < \varepsilon$ for some $\mathbb{B} \in \mathrm{BMAT_m}$, then*

$$\Pr\{E_i | \cup_{i=1}^{i-1} E_i^c\} \leq \sum_{S \in \mathcal{L}(\mathbb{B})} Z(P_{(i)}^{[S]}) \leq \delta(\varepsilon), \qquad (9)$$

where $\delta(\varepsilon)$ tends to zero with ε, and where $\mathcal{L}(\mathbb{B})$ denotes the collection of subsets of E_m that are indexed by the rows of an irreducible matrix representation of $\mathbb{B}$. Moreover, for any $\varepsilon > 0$ and $\beta < 1/2$,

$$\lim_{l \to \infty} \frac{1}{2^l} |\{i \in \{1, \ldots, 2^l\} : \|I(P_{(i)}) - r_{\mathbb{B}}\|_{l_1} < \varepsilon, \quad \mathbb{B} \in \mathrm{BMAT_m}, \sum_{S \in \mathcal{L}(\mathbb{B})} Z(P_{(i)}^{[S]}) \geq m2^{-2^{l\beta}}\}| = 0. \qquad (10)$$

Note: if $m = 3$ and $\mathbb{B}$ is the binary matroid corresponding to the matrix $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$, we have $\mathcal{L}(\mathbb{B}) = \{\{1, 3\}, \{2, 3\}\}$.

Proof: The proof of (9) requires two results. First Theorem 7, which guarantees that $\|I(P_{(i)}) - \mathbb{B}\| < \varepsilon$ implies that $P_{(i)}$ is "equivalent" to the collection of $P_{(i)}^{[S]}$, with $S \in \mathcal{L}(\mathbb{B})$. Then, the fact that the Bhattacharyya parameter of a single-user channel provides an upper bound to the error probability over the considered channel, cf. [2]. Finally, the proof of (10) follows from (8) and Lemma 7. ■

IV. Discussion

We have constructed a polar code for the MAC with arbitrary many users, which preserves the properties (complexity, error probability decay) of the original polar code construction in [2], [12]. The polar code constructed in this paper is shown to achieve only a portion of the dominant face of the MAC region, which is however sufficient to achieve the uniform sum rate on any binary input MAC. It is still an open problem to find a polar code construction that can achieve all rates in the uniform rate region. The connections presented in section III-C between matroid theory and the MAC mutual informations bring interesting problems regarding information theory inequalities, and possibly, regarding matroid theory itself. Moreover, the current work leads to a code construction for the AWGN channel, in the following way. Over an AWGN channel, by transmitting the standardized average of i.i.d. binary random variables, scaled to satisfy the power constraint, the receiver observes

$$Y = \frac{2\sqrt{p}}{\sqrt{m}} \sum_{i=1}^{m} (X_i - 1/2) + Z,$$

where Z is Gaussian distributed. We can look at this channel as being a m-user BMAC, $(X_1, \ldots, X_m) \to Y$, and the polar code constructed in this paper can be used to communicate over this channel. Moreover, from the central limit theorem, by taking m arbitrarily large, the output distribution of previous scheme is arbitrarily close to a Gaussian distribution, and hence, this coding scheme can achieve rates arbitrarily close to the AWGN capacity. Yet, to ensure that this scheme provides a 'low encoding and decoding complexity code' for the AWGN channel, further complexity considerations must be investigated when assuming m arbitrarily large. In particular, the decoder must recover a m-dimensional binary vector over each extremal MACs. The total (maximal) number of hypothesis is 2^m, however the decoder can proceed with each of the m user individually, by using the marginalized single-user channel between one user and the output, which is an extremal channel in the single-user sense. Finally, the search of a basis for a given matroid is tractable and achieved in at most m steps, by using a greedy algorithm that checks the independence of a given set and increases the set by one element at each step (starting with the empty set).

REFERENCES

[1] E. Abbe, *Mutual Information, Matroids and Extremal Channels.* Preprint, 2010.

[2] E. Arıkan, *Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,* IEEE Trans. Inform. Theory, vol. IT-55, pp. 3051–3073, July 2009.

[3] E. Arıkan and E. Telatar, *On the rate of channel polarization,* in Proc. 2009 IEEE Int. Symp. Inform. Theory, Seoul, pp. 1493–1495, 2009.

[4] J. Edmonds, *Submodular functions, matroids and certain polyhedra,* Lecture Notes in Computer Science, Springer, 2003.

[5] S. Fujishije, *Polymatroidal dependence structure of a set of random variables*, Information and Control, vol. 39, pp. 55-72, 1978.

[6] T. S. Han, *A uniqueness of shannon's information distance and related nonnegativity problems,* J. Comb., Inform. Syst. Sci., vol. 6, no. 4, pp. 320-331, 1981.

[7] B. Hassibi and S. Shadbakht, *A construction of entropic vectors*, ITA workshop at UCSD, San Diego, February 2007.

[8] L. Lovász, *Submodular functions and convexity,* in Mathematical Programming - The State of the Art, A. Bachem, M. Grtchel, and B. Korte, Eds. Berlin: Springer-Verlag, 1982, pp. 234257.

[9] F. Matúš, *Probabilistic conditional independence structures and matroid theory: background,* Int. J. of General Systems 22, pp. 185-196.

[10] J. Oxley, *Matroid Theory,* Oxford Science Publications, New York, 1992.

[11] E. Şaşoğlu, E. Telatar, E. Arıkan, *Polarization for arbitrary discrete memoryless channels*, August 2009, arXiv:0908.0302v1 [cs.IT].

[12] E. Şaşoğlu, E. Telatar, E. Yeh, *Quasi-polarization for the two user binary input multiple access channel*, IEEE Information Theory Workshop, Cairo, January 2010.

[13] D. Tse and S. Hanly, *Multi-access Fading Channels: Part I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities,* IEEE Trans. Inform. Theory, vol. IT-44, no. 7, pp. 2796-2815, November 1998.

[14] Z. Zhang and R. Yeung, *On characterization of entropy function via information inequalities,* IEEE Trans. on Information Theory, vol. 44, no. 4, pp. 1140-1452, 1998.