ISIT'2013 1569719107

Polarization theorems for Arbitrary DMCs

Rajai Nasser, Emre Telatar
Ecole Polytechnique Fédérale de Lausanne,
Lausanne, Switzerland
Email: {rajai.nasser, emre.telatar}@epfl.ch

***Abstract*—A polarization phenomenon in a special sense is shown for an arbitrary discrete memoryless channel (DMC) by imposing a quasigroup structure on the input alphabet. The same technique is used to derive a polarization theorem for an arbitrary multiple access channel (MAC) by using an appropriate Abelian group structure. These results can be used to construct capacity-achieving polar codes for arbitrary DMCs with a block error probability of $o(2^{-N^{1/2-\epsilon}})$, and an encoding/decoding complexity of $O(N \log N)$, where N is the block length.**

I. Introduction

Polar coding, invented by Arıkan [1], is the first low complexity coding technique that achieves the symmetric capacity of binary-input memoryless channels. Polar codes rely on a phenomenon called *polarization*, which is the process of converting a set of identical copies of a given single user binary-input channel to a set of "almost extremal channels", i.e., either "almost perfect channels" or "almost useless channels". The probability of error of successive cancellation decoding of polar codes was proven to be equal to $o(2^{-N^{1/2-\epsilon}})$ by Arıkan and Telatar [2].

Arıkan's technique was generalized by Şaşoğlu et al. for channels with an input alphabet of prime size [3]. Generalization to channels with arbitrary input alphabet size is not simple since it was shown in [3] that if we use any group operation for the polarization method, it is not guaranteed that polarization will happen as usual to "almost perfect (or useless) channels". Şaşoğlu [4] used a special type of quasigroup operations to ensure polarization.

Park and Barg [5] showed that polar codes can be constructed using the group structure $\mathbb{Z}_{2^r}$. Sahebi and Pradhan [6] showed that polar codes can be constructed using any Abelian group structure. In [5] and [6], polarization does not happen in the usual sense, indeed, it was already proven by Şaşoğlu et al. that it is not the case. It is shown in [5] and [6] that while it is true that we don't always have polarization to "almost perfect channels" or "almost useless channels" if a general Abelian operation is used, we always have polarization to "almost useful channels" (i.e., channels that are easy to be used for communication). [5] and [6] rely mainly on the properties of Battacharyya parameters to derive polarization results. In this paper, we adopt a different approach: we give a direct proof of polarization for the more general case of quasigroups using only information theoretic concepts (namely, mutual information).

In the case of multiple access channels (MAC), we find two main results in the literature: (i) Şaşoğlu et al. constructed polar codes for the two-user MAC with an input alphabet of prime size [7], (ii) Abbe and Telatar used matroid theory to construct polar codes for the m-user MAC with binary input [8]. The generalization of the results in [8] to MAC with arbitrary input alphabet size is not trivial even in the case of prime size since there is no known characterization for non-binary matroids. In this paper, we will see how we can construct polar codes for an arbitrary MAC where the input alphabet size is allowed to be arbitrary, and possibly different from one user to another.

II. Preliminaries

Definition 1. *A quasigroup is a pair $(Q, *)$, where $*$ is a binary operation on the set Q satisfying the following:*

- *For any two elements $a, b \in Q$, there exists a unique element $c \in Q$ (denoted by $c = b\backslash_* a$) such that $a = b * c$.*
- *For any two elements $a, b \in Q$, there exists a unique element $d \in Q$ (denoted by $d = a/^* b$) such that $a = d * b$.*

Remark 1. *$(Q, /^*)$ and $(Q, \backslash_*)$ are also quasigroups.*

Notation 1. *For any two subsets A and B of Q, we define:*

$$A * B := \{a * b : \ a \in A, b \in B\}.$$

If A and B are non-empty, then $|A * B| \geq \max\{|A|, |B|\}$.

Definition 2. *Let Q be any set. A partition $\mathcal{H}$ of Q is said to be a* balanced partition *if and only if all the elements of $\mathcal{H}$ have the same size. We denote the common size of its elements by $||\mathcal{H}||$. The number of elements in $\mathcal{H}$ is denoted by $|\mathcal{H}|$ as usual. Clearly, $|Q| = |\mathcal{H}| \times ||\mathcal{H}||$ for such a partition.*

Definition 3. *Let $\mathcal{H}$ be a balanced partition of $(Q, *)$. The* projection onto $\mathcal{H}$ *is the mapping $\text{Proj}_{\mathcal{H}} : Q \longrightarrow \mathcal{H}$, where $\text{Proj}_{\mathcal{H}}(x)$ is the unique element $H \in \mathcal{H}$ such that $x \in H$.*

Definition 4. *Let $(Q, *)$ be a quasigroup. A balanced partition $\mathcal{H}$ of Q is said to be a* stable partition *of $(Q, *)$ if and only if there exist n different balanced partitions $\mathcal{H}_1$, ..., $\mathcal{H}_n$ of Q such that:*

- *$\mathcal{H}_1 = \mathcal{H}$.*
- *$\mathcal{H}_{i+1} = \mathcal{H}_i^* := \{A * B : \ A, B \in \mathcal{H}_i\}$ for all $i \leq n-1$.*
- *$\mathcal{H}_n^* = \mathcal{H}$.*

n is called the degree *of $\mathcal{H}$. It is easy to see that if $\mathcal{H}$ is a stable partition of degree n, then $||\mathcal{H}_i|| = ||\mathcal{H}||$ for all $1 \leq i \leq n$.*

Example 1. *Let $Q = \mathbb{Z}_n \times \mathbb{Z}_n$, define $(x_1, y_1) * (x_2, y_2) = (x_1 + y_1 + x_2 + y_2, y_1 + y_2)$. For each $j \in \mathbb{Z}_n$, define $H_j =$*

$\{(j,k):\ k\in\mathbb{Z}_n\}$, *then the partition* $\mathcal{H}=\{H_j:\ j\in\mathbb{Z}_n\}$ *is a stable partition of* $(Q,*)$ *of degree* n.

Definition 5. *For any two partitions* $\mathcal{H}_1$ *and* $\mathcal{H}_2$, *we define:*

$$\mathcal{H}_1\wedge\mathcal{H}_2=\{A\cap B:\ A\in\mathcal{H}_1,\ B\in\mathcal{H}_2,\ A\cap B\neq\emptyset\}.$$

Lemma 1. *If* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *are stable then* $\mathcal{H}_1\wedge\mathcal{H}_2$ *is also a stable partition of* $(Q,*)$, *and* $(\mathcal{H}_1\wedge\mathcal{H}_2)^*=\mathcal{H}_1^*\wedge\mathcal{H}_2^*$.

Proof: See [9]. ∎

III. Polarization process

In this section, we deal with ordinary channels having a quasigroup structure on the input alphabet.

Definition 6. *Let* $(Q,*)$ *be an arbitrary quasigroup, and let* $P:Q\longrightarrow\mathcal{Y}$ *be a single user channel. We define the two channels* $P^-:Q\longrightarrow\mathcal{Y}\times\mathcal{Y}$ *and* $P^+:Q\longrightarrow\mathcal{Y}\times\mathcal{Y}\times Q$ *as follows:*

$$P^-(y_1,y_2|u_1)=\frac{1}{|Q|}\sum_{u_2\in Q}P(y_1|u_1*u_2)P(y_2|u_2),$$

$$P^+(y_1,y_2,u_1|u_2)=\frac{1}{|Q|}P(y_1|u_1*u_2)P(y_2|u_2).$$

For any $s=(s_1,...,s_n)\in\{-,+\}^n$, *we define*

$$P^s:=((P^{s_1})^{s_2}...)^{s_n}.$$

Remark 2. *Let* U_1 *and* U_2 *be two independent random variables uniformly distributed in* Q. *Set* $X_1=U_1*U_2$ *and* $X_2=U_2$, *then* X_1 *and* X_2 *are independent and uniform in* Q *since* $*$ *is a quasigroup operation. Let* Y_1 *and* Y_2 *be the outputs of the channel* P *when* X_1 *and* X_2 *are the inputs respectively. It is easy to see that* $I(P^-)=I(U_1;Y_1,Y_2)$ *and* $I(P^+)=I(U_2;Y_1,Y_2,U_1)$. *We have:*

$$\begin{aligned}I(P^-)+I(P^+)&=I(U_1;Y_1,Y_2)+I(U_2;Y_1,Y_2,U_1)\\&=I(U_1,U_2;Y_1,Y_2)=I(X_1,X_2;Y_1,Y_2)\\&=I(X_1;Y_1)+I(X_2;Y_2)=2I(P).\end{aligned}$$

It is clear that

$$I(P^+)=I(U_2;Y_1,Y_2,U_1)\geq I(U_2;Y_2)=I(X_2;Y_2)=I(P).$$

We conclude that $I(P^-)\leq I(P)\leq I(P^+)$.

Definition 7. *Let* $\mathcal{H}$ *be a stable partition of* $(Q,/^*)$, *we define the channel* $P[\mathcal{H}]:\mathcal{H}\longrightarrow\mathcal{Y}$ *by:*

$$P[\mathcal{H}](y|H)=\frac{1}{||\mathcal{H}||}\sum_{\substack{x\in Q\\ \mathrm{Proj}_{\mathcal{H}}(x)=H}}P(y|x).$$

Remark 3. *If* X *is a random variable uniformly distributed in* Q *and* Y *is the output of the channel* P *when* X *is the input, then it is easy to see that* $I(P[\mathcal{H}])=I(\mathrm{Proj}_{\mathcal{H}}(X);Y)$.

Definition 8. *Let* $\{B_n\}_{n\geq1}$ *be i.i.d. uniform random variables in* $\{-,+\}$. *We define the channel-valued process* $\{P_n\}_{n\geq0}$ *by:*

$$\begin{aligned}P_0&:=P,\\P_n&:=P_{n-1}^{B_n}\ \forall n\geq1.\end{aligned}$$

The main result of this paper is that almost surely P_n becomes a channel where the output is "almost equivalent" to the projection of the input onto a stable partition of $(Q,/^*)$:

Theorem 1. *Let* $(Q,*)$ *be a quasigroup with* $|Q|\geq2$, *and let* $P:Q\longrightarrow\mathcal{Y}$ *be an arbitrary channel. Then for any* $\delta>0$, *we have:*

$$\lim_{n\to\infty}\frac{1}{2^n}\Big|\Big\{s\in\{-,+\}^n:\ \exists\mathcal{H}_s\text{ a stable partition of }(Q,/^*),\ \big|I(P^s)-\log|\mathcal{H}_s|\big|<\delta,\big|I(P^s[\mathcal{H}_s])-\log|\mathcal{H}_s|\big|<\delta\Big\}\Big|=1.$$

In order to prove this theorem, we need several lemmas:

Lemma 2. *Let* $(Q,*)$ *be a quasigroup. If* A, B *and* C *are three non-empty subsets of* Q *such that* $|A|=|B|=|C|=|A*C|=|B*C|$, *then either* $A\cap B=\emptyset$ *or* $A=B$.

Proof: See [9]. ∎

Definition 9. *Let* Q *be a set, and let* A *be a subset of* Q. *We define the distribution* $\mathbb{I}_A$ *on* Q *as* $\mathbb{I}_A(x)=\frac{1}{|A|}$ *if* $x\in A$ *and* $\mathbb{I}_A(x)=0$ *otherwise.*

Definition 10. *Let* Q *and* $\mathcal{Y}$ *be two arbitrary sets. Let* $\mathcal{H}$ *be a set of subsets of* Q. *Let* (X,Y) *be a pair of random variables in* $Q\times\mathcal{Y}$. *We define:*

$$\mathcal{A}_{\mathcal{H},\delta}(X,Y)=\Big\{y\in\mathcal{Y}:\exists H_y\in\mathcal{H},||\mathrm{P}_{X|Y=y}-\mathbb{I}_{H_y}||_\infty<\delta\Big\},$$

$$\mathcal{P}_{\mathcal{H},\delta}(X;Y)=\mathrm{P}_Y\big(\mathcal{A}_{\mathcal{H},\delta}(X,Y)\big).$$

If $\mathcal{P}_{\mathcal{H},\delta}(X;Y)>1-\delta$ for a small enough δ, then Y is "almost equivalent" to $\mathrm{Proj}_{\mathcal{H}}(X)$. In the next lemma we will show that if $I(P^-)$ is close to $I(P)$, then the output of P is "almost equivalent" to the projection of the input onto a certain balanced partition $\mathcal{H}$.

Lemma 3. *Let* $(Q,*)$ *be a quasigroup with* $|Q|\geq2$, *and let* $\mathcal{Y}$ *be an arbitrary set. For any* $\delta>0$, *there exists* $\epsilon_1(\delta)>0$ *depending only on* Q *and* δ *such that for any two pairs of random variables* (X_1,Y_1) *and* (X_2,Y_2) *that are independent and identically distributed in* $Q\times\mathcal{Y}$ *in such a way that* X_1 *and* X_2 *are uniform in* Q, *if* $H(X_1*X_2|Y_1,Y_2)<H(X_1|Y_1)+\epsilon_1(\delta)$ *then there exists a balanced partition* $\mathcal{H}$ *of* Q *such that* $\mathcal{P}_{\mathcal{H},\delta}(X_1;Y_1)>1-\delta$, *and* $|H*H'|=|H|=|H'|\ \forall H,H'\in\mathcal{H}$.

Proof: Choose $\delta>0$, and let $\delta'=\min\Big\{\frac{\delta}{|Q|^2},\frac{1}{|Q|^4}\Big\}$. Define:

- $p_{y_1}(x_1):=\mathrm{P}_{X_1|Y_1}(x_1|y_1)$ and $p_{y_1,x_2}(x):=p_{y_1}(x/^*x_2)$.
- $q_{y_2}(x_2):=\mathrm{P}_{X_2|Y_2}(x_2|y_2)$ and $q_{y_2,x_1}(x):=q_{y_2}(x_1\backslash_*x)$.

We have:

$$\mathrm{P}_{X_1*X_2|Y_1,Y_2}(x|y_1,y_2)=\sum_{x_1\in Q}p_{y_1}(x_1)q_{y_2,x_1}(x)\qquad(1)$$

$$=\sum_{x_2\in Q}q_{y_2}(x_2)p_{y_1,x_2}(x).\qquad(2)$$

Due to the strict concavity of the entropy function, there exists $\epsilon'(\delta') > 0$ such that:

- If $\exists x_1, x_1' \in Q$ such that $p_{y_1}(x_1) \geq \delta'$, $p_{y_1}(x_1') \geq \delta'$ and $||q_{y_2,x_1} - q_{y_2,x_1'}||_\infty \geq \delta'$ then

$$H(X_1 * X_2|Y_1 = y_1, Y_2 = y_2) \geq H(X_2|Y_2 = y_2) + \epsilon'(\delta'), \quad (3)$$

(see (1)).

- If $\exists x_2, x_2' \in Q$ such that $q_{y_2}(x_2) \geq \delta'$, $q_{y_2}(x_2') \geq \delta'$ and $||p_{y_1,x_2} - p_{y_1,x_2'}||_\infty \geq \delta'$ then

$$H(X_1 * X_2|Y_1 = y_1, Y_2 = y_2) \geq H(X_1|Y_1 = y_1) + \epsilon'(\delta'), \quad (4)$$

(see (2)).

Define:

$$\mathcal{C}_1 = \Big\{(y_1, y_2) \in \mathcal{Y} \times \mathcal{Y} : \ \forall x_1, x_1' \in Q, (p_{y_1}(x_1) \geq \delta' \text{ and } p_{y_1}(x_1') \geq \delta') \Rightarrow ||q_{y_2,x_1} - q_{y_2,x_1'}||_\infty < \delta'\Big\}.$$

$$\mathcal{C}_2 = \Big\{(y_1, y_2) \in \mathcal{Y} \times \mathcal{Y} : \ \forall x_2, x_2' \in Q, (q_{y_2}(x_2) \geq \delta' \text{ and } q_{y_2}(x_2') \geq \delta') \Rightarrow ||p_{y_1,x_2} - p_{y_1,x_2'}||_\infty < \delta'\Big\}.$$

From (3) we have:

$$\begin{aligned} H(X_1 * X_2|Y_1, Y_2) &\geq H(X_2|Y_2) + \epsilon'(\delta')\mathrm{P}_{Y_1,Y_2}(\mathcal{C}_1^c) \\ &= H(X_1|Y_1) + \epsilon'(\delta')\mathrm{P}_{Y_1,Y_2}(\mathcal{C}_1^c) \end{aligned}$$

Similarly, from (4) we have

$$H(X_1 * X_2|Y_1, Y_2) \geq H(X_1|Y_1) + \epsilon'(\delta')\mathrm{P}_{Y_1,Y_2}(\mathcal{C}_2^c).$$

Let $\epsilon_1(\delta) = \epsilon'(\delta')\frac{\delta'^2}{2}$, and suppose that

$$H(X_1 * X_2|Y_1, Y_2) < H(X_1|Y_1) + \epsilon_1(\delta),$$

then we must have $\mathrm{P}_{Y_1,Y_2}(\mathcal{C}_1^c) < \frac{\delta'^2}{2}$ and $\mathrm{P}_{Y_1,Y_2}(\mathcal{C}_2^c) < \frac{\delta'^2}{2}$, which imply that $\mathrm{P}_{Y_1,Y_2}(\mathcal{C}) > 1 - \delta'^2$, where $\mathcal{C} = \mathcal{C}_1 \cap \mathcal{C}_2$.

Now for each $a, a', x \in Q$, define:

- $\pi_{a,a'}(x) := (x * a)/^* a'$, and $\gamma_{a,a'}(x) := a' \backslash_* (a * x)$.

And for each $(y_1, y_2) \in \mathcal{Y} \times \mathcal{Y}$, define:

- $A_{y_1} := \{x_1 \in Q, p_{y_1}(x_1) \geq \delta'\}$.
- $B_{y_2} := \{x_2 \in Q, q_{y_2}(x_2) \geq \delta'\}$.
- $a_{y_1} := \arg\max_{x_1} \ p_{y_1}(x_1)$. $b_{y_2} := \arg\max_{x_2} \ q_{y_2}(x_2)$.
- $H_{y_1,y_2} = \Big\{x_1 \in Q : \exists b_1, b_1', b_2, b_2', ..., b_n, b_n' \in B_{y_2}, x_1 = (\pi_{b_n,b_n'} \circ ... \circ \pi_{b_1,b_1'})(a_{y_1})\Big\}$.
- $K_{y_1,y_2} = \Big\{x_2 \in Q : \exists a_1, a_1', a_2, a_2', ..., a_n, a_n' \in A_{y_1}, x_2 = (\gamma_{a_n,a_n'} \circ ... \circ \gamma_{a_1,a_1'})(b_{y_2})\Big\}$.

Suppose that $(y_1, y_2) \in \mathcal{C}$. Let $x_1 \in H_{y_1,y_2}$, and let n be minimal such that there exists $b_1, b_1', b_2, b_2', ..., b_n, b_n' \in B_{y_2}$ satisfying $x_1 = (\pi_{b_n,b_n'} \circ ... \circ \pi_{b_1,b_1'})(a_{y_1})$. Define $a_1 := a_{y_1}$, and for $1 \leq i \leq n$ define $a_{i+1} = \pi_{b_i,b_i'}(a_i)$, so that $a_{n+1} = x_1$. We must have $a_i \neq a_j$ for $i \neq j$ since n was chosen to be minimal. Therefore, $n + 1 \leq |Q|$.

For any $1 \leq i \leq n$, we have $a_{i+1} = (a_i * b_i)/^* b_i'$. Let $x = a_i * b_i$, then $a_{i+1} = x/^* b_i'$ and $a_i = x/^* b_i$. We have $(y_1, y_2) \in \mathcal{C}$, $q_{y_2}(b_i) \geq \delta'$ and $q_{y_2}(b_i') \geq \delta'$, so we must have $||p_{y_1,b_i'} - p_{y_1,b_i}||_\infty < \delta'$, and $|p_{y_1,b_i'}(x) - p_{y_1,b_i}(x)| < \delta'$, which implies that $|p_{y_1}(a_{i+1}) - p_{y_1}(a_i)| < \delta'$. Therefore:

$$\begin{aligned} &|p_{y_1}(x_1) - p_{y_1}(a_{y_1})| \\ &= |p_{y_1}(a_{n+1}) - p_{y_1}(a_1)| \leq \sum_{i=1}^{n} |p_{y_1}(a_{i+1}) - p_{y_1}(a_i)| \\ &< n\delta' \leq (|Q| - 1)\delta' \leq \frac{|Q| - 1}{|Q|^4} < \frac{|Q| - 1}{|Q|^2}. \end{aligned} \quad (5)$$

Since $p_{y_1}(a_{y_1}) \geq \frac{1}{|Q|}$, we have $p_{y_1}(x_1) > \frac{1}{|Q|^2} > \delta'$ for every $x_1 \in H_{y_1,y_2}$. Therefore, $H_{y_1,y_2} \subset A_{y_1}$ $\forall (y_1, y_2) \in \mathcal{C}$. A similar argument yields $K_{y_1,y_2} \subset B_{y_2}$ $\forall (y_1, y_2) \in \mathcal{C}$.

Fix two elements $b, b' \in B_{y_2}$. We have $(x_1 * b)/^* b' \in H_{y_1,y_2}$ and so $x_1 * b \in H_{y_1,y_2} * b'$ for any $x_1 \in H_{y_1,y_2}$. Therefore, $H_{y_1,y_2} * b \subset H_{y_1,y_2} * b'$. But this is true for any two elements $b, b' \in B_{y_2}$, so $H_{y_1,y_2} * b = H_{y_1,y_2} * b'$ $\forall b, b' \in B_{y_2}$, and $|H_{y_1,y_2} * B_{y_2}| = |H_{y_1,y_2}|$. Similarly, we have $|A_{y_1} * K_{y_1,y_2}| = |K_{y_1,y_2}|$. If we also take into consideration the fact that $H_{y_1,y_2} \subset A_{y_1}$ and $K_{y_1,y_2} \subset B_{y_2}$ we conclude:

$$|B_{y_2}| \leq |H_{y_1,y_2} * B_{y_2}| = |H_{y_1,y_2}| \leq |A_{y_1}|,$$

$$|A_{y_1}| \leq |A_{y_1} * K_{y_1,y_2}| = |K_{y_1,y_2}| \leq |B_{y_2}|.$$

Therefore, $|A_{y_1}| = |H_{y_1,y_2}| = |B_{y_2}| = |K_{y_1,y_2}|$. We conclude that $H_{y_1,y_2} = A_{y_1}$ and $K_{y_1,y_2} = B_{y_2}$. Moreover, we have $|A_{y_1} * B_{y_2}| = |A_{y_1}| = |B_{y_2}|$.

Recall that $|p_{y_1}(x_1) - p_{y_1}(a_{y_1})| < (|Q| - 1)\delta'$ for all $x_1 \in A_{y_1}$ (see (5)) and $p_{y_1}(x_1) < \delta' \leq (|Q| - 1)\delta'$ for $x_1 \notin A_{y_1}$. It is easy to deduce that

$$||p_{y_1} - \mathbb{I}_{A_{y_1}}||_\infty < |Q|(|Q| - 1)\delta' < |Q|^2\delta'.$$

Therefore, $||p_{y_1} - \mathbb{I}_{A_{y_1}}||_\infty < \delta$ and $||p_{y_1} - \mathbb{I}_{A_{y_1}}||_\infty < \frac{1}{|Q|^2}$. Similarly, $||q_{y_2} - \mathbb{I}_{B_{y_2}}||_\infty < \delta$ and $||q_{y_2} - \mathbb{I}_{B_{y_2}}||_\infty < \frac{1}{|Q|^2}$.

Now define $\mathcal{C}_{Y_1} = \Big\{y_1 \in \mathcal{Y} : \ \mathrm{P}_{Y_2}\big((y_1, Y_2) \in \mathcal{C}\big) > 1 - \delta'\Big\}$, and for each $y_1 \in \mathcal{C}_{Y_1}$, define

$$\mathcal{K}_{y_1} = \Big\{y_2 \in \mathcal{Y} : \ (y_1, y_2) \in \mathcal{C}\Big\}.$$

Then we have:

$$1 - \delta'^2 < \mathrm{P}_{Y_1,Y_2}(\mathcal{C}) \leq \Big(1 - \mathrm{P}_{Y_1}(\mathcal{C}_{Y_1})\Big)(1 - \delta') + \mathrm{P}_{Y_1}(\mathcal{C}_{Y_1}),$$

from which we conclude that $\mathrm{P}_{Y_1}(\mathcal{C}_{Y_1}) > 1 - \delta'$. And by definition, we also have $\mathrm{P}_{Y_2}(\mathcal{K}_{y_1}) > 1 - \delta'$ for all $y_1 \in \mathcal{C}_{Y_1}$. Define $\mathcal{H}_{y_1} = \{B_{y_2} : \ y_2 \in \mathcal{K}_{y_1}\}$.

Fix $y_1 \in \mathcal{C}_{Y_1}$. Since $|A_{y_1} * B| = |A_{y_1} * B'| = |A_{y_1}| = |B| = |B'|$ for every $B, B' \in \mathcal{H}_{y_1}$, we conclude that the elements of $\mathcal{H}_{y_1}$ are disjoint and have the same size (lemma 2). Now since $\mathrm{P}_{Y_2}(\mathcal{K}_{y_1}) > 1 - \frac{1}{|Q|^4}$ and since X_2 is uniform in Q, it is easy to see that $\mathcal{H}_{y_1}$ covers Q and so it is a balanced partition of Q for all $y_1 \in \mathcal{C}_{Y_1}$. Moreover, since $\mathrm{P}_{Y_2}(\mathcal{K}_{y_1}) > 1 - \frac{1}{|Q|^4}$, we

can also conclude that all the balanced partitions $\mathcal{H}_{y_1}$ are the same. Let us denote this common balanced partition by $\mathcal{H}'$.

We have $|A * B| = |A| = |B|$ for all $A \in \mathcal{H}$ and all $B \in \mathcal{H}'$, where $\mathcal{H} = \{A_{y_1} : \ y_1 \in \mathcal{C}_{Y_1}\}$. By using a similar argument as in the previous paragraph, we can deduce that $\mathcal{H}$ is a balanced partition of Q. Moreover, since (X_1, Y_1) and (X_2, Y_2) are identically distributed, we can see that $\mathcal{H} = \mathcal{H}'$. We conclude the existence of a balanced partition $\mathcal{H}$ of Q satisfying $|A * B| = |A| = |B|$ for all $A, B \in \mathcal{H}$ and

$$P_{Y_1}\left(\left\{y \in \mathcal{Y} : \ \exists H_y \in \mathcal{H}, \ ||P_{X_1|Y_1=y} - \mathbb{I}_{H_y}||_\infty < \delta\right\}\right) \geq \mathrm{P}_{Y_1}(\mathcal{C}_{Y_1}) > 1 - \delta' > 1 - \delta.$$

■

Lemma 4. *Let X_1 and X_2 be two independent random variables in Q. If there exist two sets $A, B \subset Q$ satisfying $||\mathrm{P}_{X_1} - \mathbb{I}_A||_\infty < \delta$, $||\mathrm{P}_{X_2} - \mathbb{I}_B||_\infty < \delta$ and $|A*B| = |A| = |B|$, then we must also have $||\mathrm{P}_{X_1 * X_2} - \mathbb{I}_{A*B}||_\infty < 2\delta + |Q|\delta^2$.*

Proof: See [9]. ■

Lemma 5. *Let $(Q, *)$ be a quasigroup having at least two elements, and let $\mathcal{Y}$ be an arbitrary set. For any $\delta > 0$, there exists $\epsilon(\delta) > 0$ depending only on Q and δ such that for any channel $P : Q \to \mathcal{Y}$ satisfying $|I(P^-) - I(P)| < \epsilon(\delta)$ and $|I(P^{--}) - I(P^-)| < \epsilon(\delta)$, there exists a balanced partition $\mathcal{H}$ of Q such that $\mathcal{H}^{/^*} = \{H/^*H' : \ H, H' \in \mathcal{H}\}$ is also a balanced partition of Q, $\mathcal{P}_{\mathcal{H},\delta}(X_1; Y_1) > 1 - \delta$, $\mathcal{P}_{\mathcal{H},\delta}(U_2; Y_1, Y_2, U_1) > 1 - \delta$ and $\mathcal{P}_{\mathcal{H}^{/^*},\delta}(U_1; Y_1, Y_2) > 1 - \delta$. Where U_1 and U_2 are two independent random variables uniformly distributed in Q, $X_1 = U_1 * U_2$, $X_2 = U_2$, and Y_1 (resp. Y_2) is the output of the channel P when X_1 (resp. X_2) is the input.*

Proof: The proof follows from lemmas 3 and 4. See [9] for the details. ■

Now we are ready to prove theorem 1. In fact, we will prove a stronger theorem:

Theorem 2. *Let $(Q, *)$ be a quasigroup with $|Q| \geq 2$ and let $P : Q \longrightarrow \mathcal{Y}$ be an arbitrary channel. Then for any $\delta > 0$, we have:*

$$\lim_{n\to\infty} \frac{1}{2^n}\left|\left\{s \in \{-,+\}^n : \ \exists \mathcal{H}_s \text{ a stable partition of } (Q, /^*), \ \left|I(P^s[\mathcal{H}']) - \log \frac{|\mathcal{H}_s|.||\mathcal{H}_s \wedge \mathcal{H}'||}{||\mathcal{H}'||}\right| < \delta \text{ for all stable partitions } \mathcal{H}' \text{ of } (Q, /^*)\right\}\right| = 1.$$

Proof: We only give a sketch of the proof (see [9] for the details). Due to the continuity of the entropy function, and because of lemma 1, there exists $\gamma(\delta) > 0$ depending only on Q such that if (X, Y) is a pair of random variables in $Q \times \mathcal{Y}$ where X is uniform, and if there exists a stable partition of $\mathcal{H}$ such that $\mathcal{P}_{\mathcal{H},\gamma(\delta)}(X; Y) > 1 - \gamma(\delta)$, then $\left|I(X;Y) - \log|\mathcal{H}|\right| < \delta$ and $\left|I\big(\mathrm{Proj}_{\mathcal{H}'}(X); Y\big) - \log \frac{|\mathcal{H}|.||\mathcal{H}\wedge\mathcal{H}'||}{||\mathcal{H}'||}\right| < \delta$ for all stable partitions $\mathcal{H}'$ of $(Q, /^*)$. For any balanced partition $\mathcal{H}$ of $(Q, /^*)$, we define $\mathcal{H}^-$ and $\mathcal{H}^+$, by $\mathcal{H}^{/^*}$ and $\mathcal{H}$ respectively.

Let m be the number of different balanced partitions of Q and let $l > m$. Let P_n be as in definition 8. From remark 2 we can easily see that the process $\{I(P_n)\}_n$ is a martingale, and so it converges almost surely. From this, we can deduce that $|I(P^{(s_1,s_2,-)}) - I(P^{s_1,s_2})| < \epsilon\Big(\min\Big\{\gamma(\delta), \frac{1}{2|Q|^2}\Big\}\Big)$ for almost all $s_1 \in \{-,+\}^{n-l}$, for any $\delta > 0$, any $s_2 \in \{-,+\}^i$ $(0 \leq i \leq l+1)$ and any $l > m$, where $\epsilon(.)$ is as in lemma 5.

We can see from lemma 5 that for almost all $s_1 \in \{-,+\}^{n-l}$, there exists a balanced partition $\mathcal{H}_{s_1}$ such that the channel $P^{(s_1,s_2)}$ is "almost determined" by the projection onto $\mathcal{H}_{s_1}^{s_2}$ for any $s_2 \in \{-,+\}^i$, $0 \leq i \leq l$. For such an s_1, if $s_2 \in \{-,+\}^l$ contains $l' \geq m$ minus signs, then there exist $l'+1$ balanced partitions $\mathcal{H}_i$ $(0 \leq i \leq l')$ such that $\mathcal{H}_0 = \mathcal{H}_{s_1}$, $\mathcal{H}_{l'} = \mathcal{H}_{s_1}^{s_2}$, and $\mathcal{H}_{i+1} = \mathcal{H}_i^{/^*}$ for each $0 \leq i \leq l'-1$. Since m is the number of different balanced partitions of Q, there exist $i < j \leq l'$ such that $\mathcal{H}_i = \mathcal{H}_j$, so $\mathcal{H}_{s_1}^{s_2}$ is a stable partition. Therefore, the output of $P^{(s_1,s_2)}$ is "almost determined" by the projection onto a stable partition. By letting l tend to infinity, we guaranty that almost all $s_2 \in \{-,+\}^l$ contain at least m minus signs, and we get the result. ■

IV. The case of groups

Lemma 6. *Let $(G, *)$ be a group, and let $\mathcal{H}$ be a stable partition of $(G, /^*)$. There exists a normal subgroup H of G such that $\mathcal{H}$ is the quotient group of G by H (denoted by G/H), and $\mathrm{Proj}_{\mathcal{H}}(x) = x \bmod H$ for all $x \in G$.*

Proof: Let H be the element of $\mathcal{H}$ containing the neutral element e of G. For any $H' \in \mathcal{H}$, we have $H' = H'/^*e \subset H'/^*H$. Now because of the stability of $\mathcal{H}$, we have $|H'/^*H| = |H'|$ and so $H'/^*H = H'$ for all $H' \in \mathcal{H}$. This implies that $\mathcal{H}^{/^*} = \mathcal{H}$, from which we can easily deduce that $\mathcal{H}^* = \mathcal{H}$.

Now for any $H' \in \mathcal{H}$, we have $H' = e * H' \subset H * H' \in \mathcal{H}$, $H' = H' * e \subset H' * H \in \mathcal{H}$, and $|H'| = |H * H'| = |H' * H|$, from which we conclude that $H * H' = H' * H = H'$. This implies that $H * H = H$, and $k * H = H * k$ for any $k \in G$. Therefore, H is a normal subgroup of G, and $\mathcal{H}$ is the quotient subgroup of G by H. ■

By combining the last lemma and theorem 2 we get:

Theorem 3. *Let $P : G \longrightarrow \mathcal{Y}$ be a channel where the input alphabet G has a group structure. P_n converges almost surely to deterministic homomorphism channels:*

$$\lim_{n\to\infty} \frac{1}{2^n}\Big|\big\{s \in \{-,+\}^n : \exists H \text{ a normal subgroup of } G, \ \big|I(P^s) - \log|G/H|\big| < \epsilon, \big|I(P^s[H]) - \log|G/H|\big| < \epsilon\big\}\Big| = 1$$

for any $\epsilon > 0$. Where $P[H] : G/H \longrightarrow \mathcal{Y}$ is defined by:

$$P[H](y|a) = \frac{1}{|H|} \sum_{\substack{x \in G \\ x \bmod H = a}} P(y|x).$$

V. Polarization for arbitrary multiple access channels

In this section, we prove a polarization theorem for an arbitrary multiple access channel, where there is no constraint on the input alphabets' sizes: they can be arbitrary, and possibly different from one user to another.

If we have $|\mathcal{X}_k| = p_1^{r_1} p_2^{r_2} ... p_{n_k}^{r_{n_k}}$, where p_1, ..., p_{n_k} are prime numbers, we can assume that $\mathcal{X}_k = \mathbb{F}_{p_1}^{r_1} \mathbb{F}_{p_2}^{r_2} ... \mathbb{F}_{p_{n_k}}^{r_{n_k}}$, and so we can replace the k^{th} user by $r_1 + r_2 + ... + r_{n_k}$ virtual users having $\mathbb{F}_{p_1}$, $\mathbb{F}_{p_2}$, ..., or $\mathbb{F}_{p_{n_k}}$ as input alphabet respectively. Therefore, we can assume without loss of generality that $\mathcal{X}_k = \mathbb{F}_{q_k}$ for all k, where q_k is a prime number. Let p_1, p_2, ..., p_l be the distinct primes which appear in q_1, ..., q_m, and for each $1 \leq i \leq l$ let m_i be the number of times p_i appears in q_1, ..., q_m. The m_i users having their inputs in $\mathbb{F}_{p_i}$ will be indexed by $(i,1)$, ... , (i,j) , ... , (i,m_i), where $1 \leq i \leq l$ and $1 \leq j \leq m_i$. The input of the $(i,j)^{th}$ user is denoted by $X_{i,j} \in \mathbb{F}_{p_i}$. The vector $(X_{i,1}, ..., X_{i,m_i}) \in \mathbb{F}_{p_i}^{m_i}$ is denoted by $\vec{X}_i$.

Definition 11. *In order to simplify our notation, we will introduce the notion of generalized matrices:*

- *A generalized matrix $A = (A_1, ..., A_l) \in \prod_{i=1}^{l} \mathbb{F}_{p_i}^{m_i \times r_i}$ is a collection of l matrices. $\mathbb{F}_{p_i}^{m_i \times r_i}$ denotes the set of $m_i \times r_i$ matrices with coefficients in $\mathbb{F}_{p_i}$.*
- *A generalized vector $\vec{x} = (\vec{x}_1, ..., \vec{x}_l) \in \prod_{i=1}^{l} \mathbb{F}_{p_i}^{m_i}$ is a collection of l vectors.*
- *Additions, transpositions, and multiplications of generalized matrices and generalized vectors are defined naturally as component-wise operations (e.g., $A^T = (A_1^T, ..., A_l^T)$).*
- *A generalized matrix A is said to be full rank if and only if each matrix component of it is full rank.*
- *The logarithmic rank of a generalized matrix is defined by:* $\text{lrank}(A) = \sum_{i=1}^{l} \text{rank}(A_i). \log p_i$.

Definition 12. *Let $P : \prod_{i=1}^{l} \mathbb{F}_{p_i}^{m_i} \to \mathcal{Y}$ be an m-user MAC, let $A \in \prod_{i=1}^{l} \mathbb{F}_{p_i}^{m_i \times r_i}$ be a full rank generalized matrix. We define the $\left(\sum_{i=1}^{l} r_i\right)$-user MAC $P[A] : \prod_{i=1}^{l} \mathbb{F}_{p_i}^{r_i} \to \mathcal{Y}$ as follows:*

$$P[A](y|\vec{u}) = \frac{1}{\prod_{i=1}^{l} p_i^{m_i - r_i}} \sum_{\substack{\vec{x} \in \prod_{i=1}^{l} \mathbb{F}_{p_i}^{m_i} \\ A^T \vec{x} = \vec{u}}} P(y|\vec{x}).$$

The main result of this section is that, almost surely, P_n becomes an "almost deterministic generalized linear channel" (i.e., the output is "almost determined" by the application of a generalized matrix):

Theorem 4. *Let $P : \prod_{i=1}^{l} \mathbb{F}_{p_i}^{m_i} \to \mathcal{Y}$ be an m-user MAC. Then for every $\epsilon > 0$, we have:*

$$\lim_{n \to \infty} \frac{1}{2^n} \Big| \big\{ s \in \{-,+\}^n : \exists A_s \in \prod_{i=1}^{l} \mathbb{F}_{p_i}^{m_i \times r_{i,s}}, A_s \text{ is full rank}, |I(P^s) - \text{lrank}(A_s)| < \epsilon, |I(P^s[A_s]) - \text{lrank}(A_s)| < \epsilon \big\} \Big| = 1.$$

Proof: We can show that for any subgroup H of the group $\prod_{i=1}^{l} \mathbb{F}_{p_i}^{m_i}$, the modulo H operation is equivalent to the multiplication by a certain full rank generalized matrix. The result then follows from theorem 3. See [9] for the details. ■

VI. Discussion and Conclusion

It is possible to show that the polarized channels described in theorems 1 and 4 have a *Battacharyya parameter* that is less than $2^{-2^{\beta \cdot n}}$ where $\beta < \frac{1}{2}$ and n is the number of polarization steps. These results can be used to construct polar codes for arbitrary DMCs and arbitrary MACs with a probability of error that is less than $o(2^{-N^{1/2-\epsilon}})$, where N is the block length. See [9] for the details.

It was shown in this paper that being a quasi-group is a sufficient property for an operation to ensure polarization when it is used in the construction of polar codes. The determination of a more general property that is both necessary and sufficient remains an open problem.

References

[1] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *Information Theory, IEEE Transactions on*, vol. 55, no. 7, pp. 3051 –3073, 2009.

[2] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, 28 2009.

[3] E. Şaşoğlu, E. Telatar, and E. Arıkan, "Polarization for arbitrary discrete memoryless channels," in *Information Theory Workshop, 2009. ITW 2009. IEEE*, 2009, pp. 144 –148.

[4] E. Sasoglu, "Polar codes for discrete alphabets," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, 2012, pp. 2137–2141.

[5] W. Park and A. Barg, "Polar codes for q-ary channels,," *Information Theory, IEEE Transactions on*, vol. 59, no. 2, pp. 955–969, 2013.

[6] A. Sahebi and S. Pradhan, "Multilevel polarization of polar codes over arbitrary discrete memoryless channels," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, 2011, pp. 1718–1725.

[7] E. Şaşoğlu, E. Telatar, and E. Yeh, "Polar codes for the two-user multiple-access channel," *CoRR*, vol. abs/1006.4255, 2010. [Online]. Available: http://arxiv.org/abs/1006.4255

[8] E. Abbe and E. Telatar, "Polar codes for the -user multiple access channel," *Information Theory, IEEE Transactions on*, vol. 58, no. 8, pp. 5437 –5448, aug. 2012.

[9] R. Nasser and E. Telatar, "Polar coding for arbitrary DMCs," Tech. Rep., 2013. [Online]. Available: http://infoscience.epfl.ch/record/183335