Polar Codes for Multi-terminal Communications

Aria G. Sahebi and S. Sandeep Pradhan

Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, MI 48109, USA.
Email: ariaghs@umich.edu, pradhanv@umich.edu

Abstract—It is shown that polar coding schemes achieve the known achievable rate regions for several multi-terminal communications problems including lossy distributed source coding, multiple access channels, broadcast channels and multiple description.

I. INTRODUCTION

Polar codes were recently proposed by Arikan [1] to achieve the symmetric capacity of binary input channels. This result was later generalized to arbitrary discrete memoryless channels [2]–[5]. Polar coding schemes were also developed to achieve the symmetric rate-distortion function for arbitrary discrete memoryless sources [6]–[8]. Among the existing works on the application of polar codes for multi-terminal cases we note [9], [10] for distributed source coding, [11], [12] for the multiple access channels and [13] for broadcast channels.

In a parallel work, we showed that nested polar codes can be used to achieve the Shannon capacity of arbitrary discrete memoryless channels and the Shannon rate-distortion function for discrete memoryless sources [14]. In this paper, we show that nested polar codes can achieve the best known achievable rate regions for several multi-terminal communication systems. We present several examples in this paper, including the distributed source coding problem, multiple access channels, computation over MAC, broadcast channels, multiple description coding, to illustrate how these codes can be employed to have an optimal performance for multi-terminal cases. The results of this paper are general regarding the size of alphabets using the approach of [4].

This paper is organized as follows: In Section II we state some preliminaries. In Section III we consider the distributed source coding problem and show that polar codes achieve the Berger-Tung rate region. In Section IV we consider a distributed source coding problem in which the decoder is interested in decoding the sum of auxiliary random variables and show that polar codes have the optimal performance (this scheme has the Korner-Marton scheme as a special case). In Section V we show that polar codes achieve the capacity region for multiple access channels. In Section VI we show that polar codes have an optimal performance for the problem of computation over MAC where the decoder is interested in the sum of variables. In Section VII we study the performance of polar codes for broadcast channels. In Section VIII we show that polar codes are optimal for the multiple description problem. Finally, in Section IX we discuss briefly other possible problems and extensions to multiple user (more that two) cases.

II. PRELIMINARIES

1) Channel Parameters: For a channel \((X, Y, W)\), assume \(X\) is equipped with the structure of a group \((G, +)\). The symmetric capacity is defined as \(I(W) = I(X; Y)\) where the channel input \(X\) is uniformly distributed over \(X\) and \(Y\) is the output of the channel. For \(d \in G\), we define

\[ Z_d(W) = \frac{1}{d} \sum_{x \in G} \sum_{y \in Y} W(y|x)W(y|x+d) \]

and for \(H \leq G\) define \(Z^H(W) = \sum_{d \in H} Z_d(W)\).

2) Binary Polar Codes: For any \(N = 2^n\), a polar code of length \(N\) designed for the channel \((Z_2, Y, W)\) is a linear (coset) code characterized by a generator matrix \(G_N\) and a set of indices \(A \subseteq \{1, \cdots, N\}\) of almost perfect channels. The generator matrix for polar codes is defined as \(G_N = B_N F^{\otimes n}\) where \(B_N\) is a permutation of rows, \(F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\) and \(\otimes\) denotes the Kronecker product. The set \(A\) is a function of the channel. The decoding algorithm for polar codes is a specific form of successive cancellation [1].

3) Polar Codes Over Abelian Groups: For any discrete memoryless channel, there always exists an Abelian group of the same size as that of the channel input alphabet. Polar codes for arbitrary discrete memoryless channels (over arbitrary Abelian groups) are introduced in [4].

III. DISTRIBUTED SOURCE CODING: THE BERGER-TUNG PROBLEM

In the distributed source coding problem, two separate sources \(X\) and \(Y\) communicate with a centralized decoder. Let \(X, Y\) and \(U, V\) be the source and the reconstruction alphabets of the two terminals and assume \(X\) and \(Y\) have the joint distribution \(p_{XY}\). Let \(d_1 : X \times U \rightarrow R^+\) and \(d_2 : Y \times V \rightarrow R^+\) be the distortion measures for terminals \(X\) and \(Y\) respectively. We denote this source by \((X, Y, U, V, p_{XY}, d_1, d_2)\). Let \(U\) and \(V\) be auxiliary random variables taking values from \(U\) and \(V\) respectively such that \(U \leftrightarrow X \leftrightarrow Y \leftrightarrow V\), \(E(d_1(X, U)) \leq D_1\) and \(E(d_2(Y, V)) \leq D_2\) for some distortion levels \(D_1, D_2 \in R^+\). It is known by the Berger-Tung
Theorem III.1. For a source \((X,Y,U,V,p_{XY},d_1,d_2)\), assume \(U\) and \(V\) are finite. Then the Berger-Tung rate region is achievable using nested polar codes.

It suffices to show that the following rates are achievable:

\[ R_1 = I(X;U) - I(U;V), \quad R_2 = I(Y;V) \]

Let \(G\) be an Abelian group of the size larger than or equal to the size of both \(U\) and \(V\). Note that for the source \(Y\), we can use a nested polar codes as introduced in \([14]\) to achieve the rate \(I(Y;V)\). Furthermore, we have access to the outcome \(v^N_1\) of \(V^N_1\) at the decoder with high probability. It remains to show that the rate \(R_1 = I(X;U) - I(U;V)\) is achievable when the sequence \(v^N_1 \in G^N\) with \(d_2(y^N_1, v^N_1) \leq D_2\) is available at the decoder.

Given the test channel \(p_{X|U}\), define the artificial channels \((G,G^2,W_c)\) and \((G,X \times G,W_s)\) such that for \(s,z \in G\) and \(x \in X\),

\[ W_c(v,z|s) = p_{V|U}(v,z-s), \quad W_s(x,z|s) = p_{XV}(x,z-s) \]

These channels have been depicted in Figures [1 and 2](#fig1).

Let \(G\) be a random variable uniformly distributed over \(G\) which is independent from \(X\) and \(U\). It is straightforward to show that in this case, \(Z\) is also uniformly distributed over \(G\). Similarly to the point-to-point result \([14]\), we can show that the symmetric capacities of the channels \(W_c\) and \(W_s\) are given by \(\bar I(W_c) = \log q - H(U|V)\) and \(\bar I(W_s) = \log q - H(U|X)\). We employ a nested polar code in which the inner code is a good channel code for the channel \(W_c\) and the outer code is a good source code for \(W_s\). The rate of this code is equal to \(R = \bar I(W_s) - \bar I(W_c) = I(X;U) - I(U;V)\). The rest of this section is devoted to some general definitions and lemmas which are used in the proofs.

Let \(N = 2^n\) for some positive integer \(n\) and let \(G\) be the corresponding \(N \times N\) generator matrix for polar codes. For \(i = 1,\ldots,N\), and for \(z^N_1, a^N_1 \in G^N\), \(v^N_1 \in V^N\) and \(x^N_1 \in X^N\), let

\[ W^{(i)}_{c,N}(v^N_1, a^N_1|z^N_1) = \sum_{a^N_{i+1} \in G^N \setminus \{a^N_1\}} \frac{1}{q^{N-1}} W_{c,N}(z^N_1, v^N_1|a^N_1) \]

\[ W^{(i)}_{s,N}(x^N_1, z^N_1|a^N_1) = \sum_{a^N_{i+1} \in G^N \setminus \{a^N_1\}} \frac{1}{q^{N-1}} W_{s,N}(x^N_1, z^N_1|a^N_1) \]

Let the random vectors \(X^N_1, Y^N_1, U^N_1, V^N_1\) be distributed according to \(p_{XUV}^N\) and let \(Z^N_1\) be a random variable uniformly distributed over \(G^N\) which is independent of \(X^N_1, Y^N_1, U^N_1, V^N_1\). Let \(S^N_1 = Z^N_1 - U^N_1\) and \(A^N_1 = S^N_1 G^{-1}\) (Here, \(G^{-1}\) is the inverse of the mapping \(G : G^N \to G^N\)). In other words, the joint distribution of the random vectors is given by

\[ p_{A^N_1 S^N_1 U^N_1 V^N_1 X^N_1 Z^N_1}(a^N_1, s^N_1, u^N_1, v^N_1, x^N_1, z^N_1) \]

\[ = \frac{1}{q^N} p_{XUV}(x^N_1, u^N_1, v^N_1) \mathbb{1}\{s^N_1 = a^N_1 G, u^N_1 = z^N_1 - a^N_1 G\} \]

A. Sketch of the proof

The following theorems state the standard channel coding and source coding polarization phenomena for the general case.

Theorem III.2. For any \(\epsilon > 0\) and \(0 < \beta < \frac{1}{2}\), there exist a large \(N = 2^n\) and a partition \(\{A_H | H \leq G\}\) of \(\{0,1\}\) such that for \(H \leq G\) and \(i \in A_H\),

\[ |\bar I(W^{(i)}_{c,N}) - \log G| < \epsilon \]

\[ Z^H(W^{(i)}_{c,N}) < 2^{-N^\beta} \]

Moreover, as \(\epsilon \to 0\) (and \(N \to \infty\)), \(\frac{|A_H|}{N} \to p_H\) for some probabilities \(p_H, H \leq G\) adding up to one with \(\sum_{H \leq G} p_H \log G = \bar I(W_c)\).

Theorem III.3. For any \(\epsilon > 0\) and \(0 < \beta < \frac{1}{2}\), there exist a large \(N = 2^n\) and a partition \(\{B_H | H \leq G\}\) of \(\{0,1\}\) such that for \(H \leq G\) and \(i \in A_H\),

\[ |\bar I(W^{(i)}_{s,N}) - \log G| < \epsilon \]

\[ Z^H(W^{(i)}_{s,N}) < 2^{-N^\beta} \]

Moreover, as \(\epsilon \to 0\) (and \(N \to \infty\)), \(\frac{|B_H|}{N} \to q_H\) for some probabilities \(q_H, H \leq G\) adding up to one with \(\sum_{H \leq G} q_H \log G = \bar I(W_s)\).

For \(H \leq G\), define

\[ A_H = \{i \in [1,N] | Z^H(W^{(i)}_{c,N}) < 2^{-N^\beta} \} \]

\[ K \leq H : Z^K(W^{(i)}_{c,N}) < 2^{-N^\beta} \}

\[ B_H = \{i \in [1,N] | Z^H(W^{(i)}_{s,N}) < 1 - 2^{-N^\beta} \} \]

\[ K \leq H : Z^K(W^{(i)}_{s,N}) < 1 - 2^{-N^\beta} \}

For \(H \leq G\) and \(K \leq G\), define \(A_{H,K} = A_H \cap B_K\). Note that for large \(N\), \(2^{-N^\beta} < 1 - 2^{-N^\beta}\). This implies for \(i \in A_H\), we have \(Z^H(W^{(i)}_{s,N}) < 1 - 2^{-N^\beta}\) and hence \(i \in \cup_{K \leq H} B_K\).

Therefore, for \(K \not\leq H\), we have \(A_{H,K} = \emptyset\). This means \(\{A_{H,K} | K \leq H \leq G\}\) forms a partition of \([1,N]\). Note that

Lemma III.1. The channel \(W_c\) is degraded with respect to the channel \(W_s\) in the sense of \([15]\) Definition III.1).

Proof: In the Definition \([15]\) Definition III.1), let the channel \((X \times G, G^2, W)\) be such that for \(v, z, z' \in G\) and \(x \in X\), \(W(v,z|x,z') = p_{V|X}(v|x) \mathbb{1}\{z = z'\}\).
as \( N \) increases, \( \frac{|A_H|}{N} \to p_H \) and \( \frac{|B_H|}{N} \to q_H \).

The encoding and decoding rules are as follows: Let \( z_i^N \in G^N \) be an outcome of the random variable \( Z_i^N \) known to both the encoder and the decoder. Given \( K \leq H \leq G \), let \( T_K \) be a transversal of \( H \) in \( G \) and let \( T_{K \leq H} \) be a transversal of \( K \) in \( H \). Any element \( g \) of \( G \) can be represented by \( g = [g]_K + [g]_{T_K \leq H} + [g]_{T_H} \) for unique \([g]_K, [g]_{T_K \leq H}, [g]_{T_H} \in T_K \leq H \) and \([g]_{T_H} \in T_H \). Also note that \( T_{K \leq H} + T_H \) is a transversal \( T_K \) of \( K \) in \( G \) so that \( g \) can be uniquely represented by \([g]_{T_K} = [g]_{T_K \leq H} + [g]_{T_H} \).

Given a source sequence \( x_i^N \in X^N \), the encoding rule is as follows: For \( i \in [1, N] \), if \( i \in A_{H,K} \) for some \( K \leq H \leq G \), \([a_i]_K \) is uniformly distributed over \( K \) and is known to both the encoder and the decoder. Each \( a_i \) is chosen randomly so that for \( g \in G \),

\[
P(a_i = g) = \frac{P_{A_i|X_i^N} Z_i^{N \cdot A_i} \cdot 1 (g | x_i^N, z_i^N, a_i^{-1})}{P_{A_i|X_i^N} Z_i^{N \cdot A_i} \cdot 1 (a_i^{-1} + T_K [x_i^N, z_i^N, a_i^{-1}])}
\]

Note that \( a_i^N \) can be decomposed as \( a_i^N = [v_i^N]_K + [a_i^N]_{T_K \leq H} + [a_i^N]_{T_H} \) in which \([a_i^N]_K \) is known to the decoder. The decoder sends \([a_i^N]_{T_K \leq H} \) to the decoder and the decoder uses the channel code to recover \([a_i^N]_{T_K \leq H} \). The decoding rule is as follows: Given \( z_i^N, v_i^N, [a_i]_K \) and \([a_i]_{T_K \leq H} \), and for \( i \in A_{H,K} \), let

\[
\hat{a}_i = \arg \max_{g \in [a_i]_K + [a_i]_{T_K \leq H} + T_H} W^{(i)}_{c,N} (z_i^N, v_i^N, a_i^{-1} | g)
\]

Finally, the decoder outputs \( z_i^N - \hat{a}_i^N G \). Note that the rate of this code is equal to

\[
R = \sum_{K \leq H \leq G} \frac{|A_{H,K}|}{N} \log \frac{|H|}{|K|}
= \sum_{K \leq H \leq G} \frac{|A_{H,K}|}{N} \log \frac{|G|}{|H|} - \sum_{K \leq H \leq G} \frac{|A_{H,K}|}{N} \log \frac{|G|}{|H|}
\]

\[
\to I(W_s) - I(W_c) = I(X; U) - I(U; V)
\]

IV. DISTRIBUTED SOURCE CODING: DECODING THE SUM OF VARIABLES

For a distributed source \( (X \times Y, p_{X,Y}) \) let the random variable \( U \) and \( V \) take values from a group \( G \). Assume that \( U \) and \( V \) satisfy the Markov chain \( U \leftrightarrow X \leftrightarrow Y \leftrightarrow V \) and assume \( E \{d(X, Y, g(U + V)) \leq D \} \) for some function \( g \). For \( W = U + V \), we show that the following rates are achievable:

\[
R_1 = H(W) - H(U|X), \quad R_2 = H(W) - H(V|Y)
\]

The source \( X \) employs a nested polar codes whose inner code is a good channel code for the channel \((G, G, W_{c,X})\) and whose outer code is a good code for the test channel \((G, X \times G, W_{s,X})\) where for \( s, t, q, z \in G \) and \( x \in X \),

\[
W_{c,X}(q|s+t) = p_W(q-s-t), W_{s,X}(x,z|s) = p_{X,U}(x,z-s)
\]

Similarly, the source \( Y \) employs a nested polar code whose inner code is a good channel code for the channel \((G, G, W_{c,Y})\) and whose outer code is a good source code for the test channel \((G, Y \times G, W_{s,Y})\) where for \( s, t, q, r \in G \) and \( y \in Y \),

\[
W_{c,Y}(q|s+t) = p_W(q-s-t), W_{s,Y}(y,r|t) = p_{Y,V}(y,r-t)
\]

These channels are depicted in Figures 3 and 4.

V. MULTIPLE ACCESS CHANNELS

Let the finite sets \( X \) and \( Y \) be the input alphabets of a two-user MAC and let \( Z \) be the output alphabet. In order to show that nested polar codes achieve the capacity of a MAC, it suffices to show that the rates \( R_1 = I(X; Z|Y) = H(X) - H(X|Z) \) and \( R_2 = I(Y; Z) \) are achievable. It is known from the point-to-point result [14] that the \( Y \) terminal can communicate with the decoder with rate \( I(Y; Z) \) so that \( y_i^N \) is available at the decoder with high probability. It remains to show that the rate \( R_1 \) is achievable for the \( X \) terminal when \( y_i^N \) is available at the decoder. Let \( G \) be an Abelian group with \( |G| = |X| \). Define the artificial channels \((G, G, W_s)\) and \((G, Y \times G, W_{c})\) such that for \( u, z \in G \) and \( y \in Y \),

\[
W_s(u|s) = p_X(s-u), W_c(y, z, s|u) = p_{X,Y}(s-u, y, z)
\]

These channels have been depicted in Figures 7 and 8.

Similarly to previous cases, one can show that the symmetric capacities of the channels are equal to \( I(W_s) = \)
log \( q - H(X) \) and \( \bar{I}(W_c) = \log q - H(X|YZ) \). We employ a nested polar code in which the inner code is a good source code for the test channel \( W_s \) and the outer code is a good channel code for \( W_c \). The rate of this code is equal to \( R = \bar{I}(W_c) - \bar{I}(W_s) = I(X;Z|Y) \). Here, we only give a sketch of the proof. First note that the channel \( W_c \) is degraded with respect to \( W_c \) so that the source code is contained in the channel code. For \( s_1^N \in G^N, y_1^N \in Y^N \) and \( z_1^N \in Z^N \), let

\[
W_{s,N}(a_1^{-1}|a_i) = \frac{1}{q^{N-1}} W_{s}(s_1^N | a_1^N G)
\]

\[
W_{c,N}(y_1^N, z_1^N, a_1^{-1}|a_i) = \frac{1}{q^{N-1}} W_{c}(y_1^N, z_1^N | a_1^N G)
\]

Let the random vectors \( X_1^N, Y_1^N, U_1^N, V_1^N \) be distributed according to \( p_{X,Y,U,V}^N \) and let \( S_1^N \) be a random variable uniformly distributed over \( G^N \) which is independent of \( X_1^N, Y_1^N, U_1^N, V_1^N \). The encoding and decoding rules are similar to those of the point-to-point channel coding result; i.e., at the encoder, the distribution \( p_{A_1|S_1^N} p_{X_1^N|A_1^{-1}} \) is used for soft encoding and at the decoder, \( W_{c,N}^{(1)}(y_1^N, z_1^N, a_1^{-1}|a_i) \) is used in the successive cancelation decoder to decode \( a_1^N \). The final decoder output is equal to \( z_1^N - a_1^N G \). Note that since \( y_1^N \) is known to the decoder with high probability, it can be used as the channel output for \( W_c \).

VI. COMPUTATION OVER MAC

In this section, we consider a simple computation problem over a MAC with input alphabets \( \mathcal{X}, \mathcal{Y} \) and output alphabet \( \mathcal{Z} \). The two input terminals of a MAC, \( X \) and \( Y \) are trying to communicate with a centralized decoder which is interested in the sum of the inputs \( S = X + Y \) where + is summation over a group \( G \). We show that the rate \( R = \min(H(X), H(Y)) - H(S) \) is achievable using polar codes. The terminal \( X \) employs a nested polar code whose inner code is a good source code for the test channel \( (G, G, W_{s,X}) \) and whose outer code is a good channel code for the channel \( (G, Z \times G, W_{c,X}) \) where for \( u, v, r, z \in G \) and \( z \in \mathcal{Z} \),

\[
W_{s,X}(r|u) = p_X(r-u), W_{c,X}(z, q|u+v) = p_{SZ}(q-u-v, z)
\]

Similarly, the terminal \( Y \) employs a nested polar code whose inner code is a good source code for the test channel \( (G, G, W_{s,Y}) \) and whose outer code is a good channel code for the channel \( (G, Z \times G, W_{c,Y}) \) where for \( u, v, t, z \in G \) and \( z \in \mathcal{Z} \),

\[
W_{s,Y}(t|v) = p_Y(t-v), W_{c,Y}(z, q|u+v) = p_{SZ}(q-u-v, z)
\]

Note that the two terminals use the same channel code. These channels are depicted in Figures 9 and 10 for inner code. For the outer code, see Figure 11.

Similarly to previous cases, one can show that the symmetric capacities of the channels are equal to \( \bar{I}(W_s) = \log q - H(X) \) and \( \bar{I}(W_c) = \log q - H(X|YZ) \). We employ a nested polar code in which the inner code is a good source code for both test channels \( W_{s,X} \) and \( W_{s,Y} \) and the outer code is a good channel code for \( W_{c,X} = W_{c,Y} \). The rate of this code is equal to \( R = \bar{I}(W_{c,X}) = \max(I(W_{s,X}), I(W_{s,Y})) = \min(H(X), H(Y)) - H(S|Z) \). It is worth noting that it can be shown that the intersection of the two source codes is contained in the common channel code.

VII. THE BROADCAST CHANNEL

In this section, we show that polar codes achieve the capacity of a broadcast channel \( (\mathcal{X}, \mathcal{Y} \times \mathcal{Z}, W) \) when \( \mathcal{X} = G \) for some arbitrary Abelian group \( G \).

Let \( X \) be a random variable over \( \mathcal{X} \) such that \( E\{w(X)\} \leq D \) and let \( Y, Z \) be the corresponding channel outputs. Let \( U, V \) be random variable over \( G \) satisfying the Markov chain \( UV \leftrightarrow X \leftrightarrow YZ \) such that there exists a function \( g : G^2 \to \mathcal{X} \) with \( g(U,V) = X \). It suffices to show that the following rates are achievable

\[
R_1 = I(U;Y) - I(U;V) = H(U|Y) - H(U|V), R_2 = I(V;Z)
\]

Note that the \( Z \) terminal can use a point-to-point channel code to achieve the desired rate. It remains to show that the rate \( R_2 \) is achievable when \( v_1^N \) is available at the encoder.

Define the artificial channels \( (G, G^2, W_s) \) and \( (G, \mathcal{Y} \times G, W_c) \) such that for \( s, v, z \in G \) and \( y \in \mathcal{Y} \),

\[
W_s(v, z|s) = p_{UV}(z-s, v), W_c(y, z|s) = p_{UY}(z-s, y)
\]

These channels have been depicted in Figures 13 and 14.

Similarly to previous cases, one can show that the symmetric capacities of the channels are equal to \( \bar{I}(W_s) = \log q - H(U|V) \) and \( \bar{I}(W_c) = \log q - H(U|Y) \). Note that to guarantee that \( W_s \) is degraded with respect to \( W_c \), we need
an additional condition on the auxiliary random variables. It is sufficient to assume that the Markov chain $U \leftrightarrow X \leftrightarrow V$ holds.

We employ a nested polar code in which the inner code is a good source code for the test channel $W_a$ and the outer code is a good channel code for $W_c$. The rate of this code is equal to $R = I(W_c) - I(W_x) = I(U; Y) - I(U; V)$.

**VIII. MULTIPLE DESCRIPTION CODING**

Consider a multiple description problem in which a source $X$ is to be reconstructed at three terminals $U$, $V$ and $W$. There are two encoders and three decoders. Terminals $U$ and $V$ have access to the output of their corresponding encoders and terminal $W$ has access to the output of both encoders. The goal is to find all achievable tuples $(R_1, R_2, D_1, D_2, D_3)$ where $R_1$ and $R_2$ are the rates of encoders $U$ and $V$ respectively and $D_1$, $D_2$ and $D_3$ are the distortion levels corresponding to decoders $U$, $V$ and $W$ respectively. $D_1$, $D_2$ and $D_3$ are measured as the average of distortion measures $d_1(\cdot, \cdot)$, $d_2(\cdot, \cdot)$ and $d_3(\cdot, \cdot)$ respectively. Let $U$, $V$ and $W$ be random variables such that $E[d_1(X, U)] \leq D_1$, $E[d_2(X, V)] \leq D_2$ and $E[d_3(X, W)] \leq D_3$. We show that the tuple $(R_1, R_2, D_1, D_2, D_3)$ is achievable if

$$R_1 \geq I(X; U)$$

$$R_2 \geq I(X; V)$$

$$R_1 + R_2 \geq I(X; UVW) + I(U; V)$$

It suffices to show that the rates $R_1 = I(X; UVW) - I(X; V) + I(U; V)$, $R_2 = I(X; V)$ are achievable. The point-to-point source coding result implies that with $R_2 = I(X; V)$ we can have $v_1^N$ at the output of the second decoder with high probability. To achieve the rate $R_1$ when $v_1^N$ is available, first we note that $R_1 = H(U) - H(U|V X) + H(W|UV) - H(W|UX)$. We use a code with rate $R_{11} = H(U) - H(U|V X)$ for sending $U$ and another code $R_{12} = H(W|UV) - H(W|UX)$ for sending $W$. The corresponding channels are depicted in Figures 15 and 16.

**IX. OTHER PROBLEMS AND DISCUSSION**

In this paper, we studied the main multi-terminal communication problems in their simplest forms (e.g., no time sharing etc.). The approach of this paper can be extended to the more general formulations and to other similar problems. The approach presented in this paper can also be extended to multiple user (more than two) cases in a straightforward fashion. We briefly discuss the 3-user MAC as an example. Consider a 3-user MAC with inputs $W$, $X$ and $Y$ and output $Z$. We have seen in Section IV that with rates $R_X = I(X; Y|Z)$ and $R_Y = I(Y; Z)$, we can access to $x_1^N$ and $y_1^N$ at the decoder with high probability. The channels $W_s, W$ and $W_c, W$ depicted in Figures 19 and 20 can be used to design a nested polar code of rate $R_M = I(W; Z|X, Y)$ for terminal $W$.

![Fig. 19: Channel for inner code.](image1)

![Fig. 20: Channel for outer code.](image2)

**REFERENCES**

[1] E. Arikan, “Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels,” IEEE Transactions on Information Theory, vol. 55, no. 7, pp. 3051–3073, 2009.

[2] E. Sasoglu, E. Telatar, and E. Arikan, “Polarization for arbitrary discrete memoryless channels,” IEEE Information Theory Workshop, Dec. 2009, Lausanne, Switzerland.

[3] R. Mori and T. Tanaka, “Channel Polarization on $q$-ary Discrete Memoryless Channels by Arbitrary Kernels,” Proceedings of IEEE International Symposium on Information Theory, July 2010, Austin, TX.

[4] A. Sahebi and S. Pradhan, “Multilevel Channel Polarization for Arbitrary Discrete Memoryless Channels,” Information Theory, IEEE Transactions on, vol. 59, no. 12, pp. 7839–7857, 2013.

[5] W. Park and A. Barg, “Polar codes for $q$-ary channels, $q = 2^t$,” 2012, Online: http://arxiv.org/abs/1107.4965.

[6] S. Korada and R. Urbanke, “Polar codes are optimal for lossy source coding,” in Information Theory Workshop, 2009. ITW 2009. IEEE, 2009, pp. 149–153.

[7] M. Karzand and I. Telatar, “Polar codes for $q$-ary source coding,” in Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on, 2010, pp. 909–912.

[8] A. Sahebi and S. Pradhan, “Polar codes for sources with finite reconstruction alphabets,” in Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on, 2012, pp. 580–586.

[9] E. Arikan, “Source polarization,” in Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on, 2010, pp. 899–903.

[10] ——, “Polar coding for the slippian-wolf problem based on monotone chain rules,” in Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on, 2012, pp. 566–570.

[11] E. Sasoglu, E. Telatar, and E. Yeh, “Polar codes for the two-user binary-input multiple-access channel,” in Information Theory Workshop (ITW), 2010 IEEE, 2010, pp. 1–5.

[12] E. Abbe and E. Telatar, “Polar Codes for the $m$-User MAC,” 2010, Online: http://arxiv.org/abs/1002.0777.

[13] N. Goela, E. Abbe, and M. Gastpar, “Polar Codes For Broadcast Channels,” 2013, Online: http://arxiv.org/abs/1301.6150.

[14] A. Sahebi and S. Pradhan, “Nested Polar Codes Achieve the Shannon Rate-Distortion Function and the Shannon Capacity,” 2014, submitted to ISIT 2014.

[15] ——, “Nested Polar Codes Achieve the Shannon Rate-Distortion Function and the Shannon Capacity,” 2014, Online: http://arxiv.org/abs.