and tries to recover (X^N, Y^N) with vanishing probability of error. The well-known Slepian-Wolf theorem states that this is possible provided $R_x \geq H(X|Y)$, $R_y \geq H(Y|X)$, and $R_x + R_y \geq H(X,Y)$.

It is straightforward to design a polar coding scheme that achieves the corner point $(H(X|Y), H(Y))$ of the Slepian-Wolf rate region. Fix $R_y > H(Y)$ and $R_x > H(X|Y)$. For $N = 2^n$, $n \geq 1$, consider a pair of high-entropy sets $E_Y = E_Y(N, R_y)$ and $E_{X|Y} = E_{X|Y}(N, R_x)$.

Encoding: Given a realization $X^N = x^N$, encoder 1 calculates $u^N = x^N G_N$ and sends $u_{E_{X|Y}}$ to the common decoder. Given a realization $Y^N = y^N$, encoder 2 calculates $v^N = y^N G_N$ and sends v_{E_Y}.

Decoding: The decoder first applies the decoding algorithm of Section III to obtain an estimate $\hat{y}^N$ of y^N from v_{E_Y}. Next, the decoder applies the same algorithm to obtain an estimate of x^N using $\hat{y}^N$ (as a substitute for the actual realization y^N) and $u_{E_{X|Y}}$.

We omit the analysis of this scheme since it essentially consists of two single-user source coding schemes of the type treated in Section III.

It is clear that polar coding can achieve all points of the Slepian-Wolf region by time-sharing between the corner points $(H(X), H(X|Y))$ and $(H(X|Y), H(Y))$.

We should remark that polar coding for Slepian-Wolf problem was first studied in [6], [2], and [3] under the assumptions that $X, Y \sim \text{Ber}(\frac{1}{2})$, and $X \oplus Y \sim \text{Ber}(p)$.

The above approach to Slepian-Wolf coding reduces the problem to single-user source coding problems. A direct appoach would be to have each encoder apply polar transforms locally, with encoder 1 computing $U^N = X^N G_N$ and encoder 2 computing $V^N = Y^N G_N$. Preliminary analyses show that such local operations polarize X_1^N and Y_1^N not only individually but also in a joint sense. A detailed study of such schemes is left for future work.

VI. Polarization of non-binary memoryless sources

Theorem 4. *Let $X \sim P_X$ be a memoryless source over $\mathcal{X} = \{0, 1, \ldots, q-1\}$ for some prime $q \geq 2$. For $n \geq 1$ and $N = 2^n$, let $X^N = (X_1, \ldots, X_N)$ be N independent drawings from the source X. Let $U^N = X^N G_N$ where G_N is as defined in* (3) *but the matrix operation is now carried out in GF(q). Then, the polarization limits in Theorem 1 remain valid provided the entropy terms are calculated with respect to base-q logarithms.*

If q is not prime, the theorem may fail. Consider X over $\{0, 1, 2, 3\}$ with $P_X(0) = P_X(2) = \frac{1}{2}$. Then, it is straightforward to check that U^N has the same distribution as X^N for all N. On closer inspection, we realize that X is actually a binary source under disguise. More precisely, X is already polarized over $\{0, 2\}$, which is a subfield of $GF(4)$, and vectors over this subfield are closed under multiplication by G_N.

The preceding example illustrates the difficulties in making a general statement regarding source polarization over arbitrary alphabets. If we introduce some randomness into the construction as in [7], it is possible to polarize sources over arbitrary alphabets, still maintaining the $O(N \log N)$ complexity of the construction.

Acknowledgment

Helpful discussions with E. Şaşoğlu and S. B. Korada are gratefully acknowledged. This work was supported in part by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under contract no. 107E216, and in part by the European Commission FP7 Network of Excellence NEWCOM++ (contract no. 216715).

VII. Appendix

A. *Proof of Inequality* (4)

First we prove that $Z(X)^2 \leq H(X)$ for any $X \sim \text{Ber}(p)$ with equality if and only if $p \in \{0, \frac{1}{2}, 1\}$. Let $F(p) = H(Z) - Z(X)^2 = -p \log_2(p) - (1-p)\log_2(1-p) - 4p(1-p)$, and compute

$$\frac{dF}{dp} = \frac{1}{\ln 2}\left[-\ln p + \ln(1-p)\right] - 4 + 8p,$$

$$\frac{d^2F}{dp^2} = \frac{1}{\ln 2}\left[-\frac{1}{p} - \frac{1}{1-p}\right] + 8,$$

$$\frac{d^3F}{dp^3} = \frac{1}{\ln 2}\left[\frac{1}{p^2} - \frac{1}{(1-p)^2}\right].$$

Inspection of the third order derivative shows that dF/dp is strictly convex for $p \in [0, \frac{1}{2})$ and strictly concave for $p \in (\frac{1}{2}, 1]$. Thus, $dF/dp = 0$ can have at most one solution in each interval $[0, \frac{1}{2})$ and $(\frac{1}{2}, 1]$. Since $dF/dp = 0$ at $p = \frac{1}{2}$, the number of zeros of dF/dp over $[0,1]$ is at most three. Thus, $F(p)$ can have at most three zeros over $[0,1]$. Since $F(p) = 0$ for $p \in \{0, \frac{1}{2}, 1\}$, there can be no other zeros.

Thus, for any pair of random variables (X, Y) with X binary, if we condition on $Y = y$, we have

$$Z(X|Y=y)^2 \leq H(X|Y=y).$$

Averaging over Y, and by Jensen's inequality, we obtain (4).

B. *Proof of Inequality* (5)

Recall that the Rényi entropy of order α ($\alpha > 0$, $\alpha \neq 1$) for a RV X is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \sum_x P_X(x)^\alpha$$

and has the following properties [8].

- $H_\alpha(X)$ is strictly decreasing in α unless P_X is uniform on its support $\text{Supp}(X) = \{x : P_X(x) > 0\}$.
- $H(X) = \lim_{\alpha \to 1} H_\alpha(X)$.

Now suppose $X \sim \text{Ber}(p)$ and note that

$$H_{\frac{1}{2}}(X) = \log\left[\sum_x \sqrt{P_X(x)}\right]^2 = \log(1 + Z(X)).$$