the encoder does not require knowledge of the realization of Y^N to implement this scheme.)

Decoding: Having received $u_{E_{X|Y}}$ and observed the realization $Y^N = y^N$, the decoder sequentially builds an estimate $\hat{u}^N$ of u^N by the rule

$$\hat{u}_i = \begin{cases} u_i & \text{if } i \in E_{X|Y} \\ 0 & \text{if } i \in E^c_{X|Y} \text{ and } L_N^{(i)}(y^N, \hat{u}^{i-1}) \geq 1 \\ 1 & \text{else} \end{cases}$$

where

$$L_N^{(i)}(y^N, \hat{u}^{i-1}) = \frac{\Pr(U_i = 0 | Y^N = y^N, U^{i-1} = \hat{u}^{i-1})}{\Pr(U_i = 1 | Y^N = y^N, U^{i-1} = \hat{u}^{i-1})}$$

is a likelihood ratio, which can be computed recursively using the formulas:

$$L_N{}^{(2i-1)}(y^N, u^{2i-2}) = \frac{L^{(i)}_{N/2}(y^{N/2}, u_o^{2i-2} \oplus u_e^{2i-2}) L^{(i)}_{N/2}(y^N_{N/2+1}, u_e^{2i-2}) + 1}{L^{(i)}_{N/2}(y^{N/2}, u_o^{2i-2} \oplus u_e^{2i-2}) + L^{(i)}_{N/2}(y^N_{N/2+1}, u_e^{2i-2})}$$

and

$$L_N^{(2i)}(y^N, u^{2i-1}) = L^{(i)}_{N/2}(y^{N/2}, u_o^{2i-2} \oplus u_e^{2i-2})^{\delta_i} L^{(i)}_{N/2}(y^N_{N/2+1}, u_e^{2i-2})$$

where u_o^{2i-2} and u_e^{2i-2} denote, respectively, the parts of u^{2i-2} with odd and even indices, and δ_i equals 1 or -1 according to u_{2i-1} being 0 or 1, respectively. Having constructed $\hat{u}^N$, the decoder outputs $\hat{x}^N = \hat{u}^N G_N^{-1}$ as the estimate of x^N. (It is easy to verify that $G_N^{-1} = G_N$.)

Performance: The performance of the decoder is measured by the probability of error

$$P_e = \Pr(\hat{U}^N \neq U^N) = \Pr(\hat{U}_{E^c_{X|Y}} \neq U_{E^c_{X|Y}}),$$

which can be upper-bounded by standard (union-bound) techniques as

$$P_e \leq \sum_{i \in E^c_{X|Y}(N,R)} Z(U_i | Y^N, U^{i-1}). \tag{7}$$

The following is a simple corollary to Theorem 2 and (7).

Theorem 3. *For any fixed $R > H(X|Y)$ and $\beta < \frac{1}{2}$, the probability of error for the above polar source coding method is bounded as $P_e = O(2^{-N^\beta})$.*

Complexity: The complexity of encoding and that of decoding are both $O(N \log N)$.

IV. Application to channel coding: Duality

The above source coding scheme can be used to design a capacity-achieving code for any binary-input memoryless channel. Let such a channel be defined by the transition probabilities $W(y|x)$, $x \in \mathcal{X} = \{0, 1\}$ and $y \in \mathcal{Y}$. Consider the block coding scheme shown in Fig. 3, where signals flow from right to left. Here, $N = 2^n$, $n \geq 1$, is the code block length; U^N denotes the message vector, $X^N = U^N G_N$ the channel input vector, and Y^N the channel output vector. Due

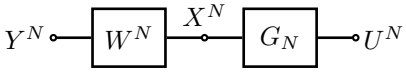

Fig. 3. Channel coding.

to memorylessness, $W^N(y^N|x^N) = \prod_{i=1}^N W(y_i|x_i)$ for any $x^N \in \mathcal{X}^N$, $y^N \in \mathcal{Y}^N$.

We turn the triple (U^N, X^N, Y^N) into a joint ensemble of random vectors by assigning the probabilities $\Pr(X^N = x^N) = 2^{-N}$ for all $x^N \in \{0, 1\}^N$. Under this assignment, (X^N, Y^N) may be regarded as independent samples from a source $(X, Y) \sim Q(x)W(y|x)$ where Q is the uniform distribution on $\{0, 1\}$. We let $I(W) = I(X; Y)$ denote the symmetric channel capacity and fix $R < I(W)$. This implies that $1 - R > H(X|Y)$. Let $E_{X|Y} = E_{X|Y}(N, 1 - R)$ denote a high-entropy set of rate $(1 - R)$ for the source (X, Y). The following coding scheme achieves reliable communication at rate R over the channel W.

Encoding: Prepare a binary source vector U^N as follows. Pick the pattern $U_{E_{X|Y}}$ at random from the uniform distribution and make it available to the decoder ahead of the session. In each round, fill $U_{E^c_{X|Y}}$ with uniformly chosen data bits. (Thus, $\lfloor NR \rfloor$ bits are sent in each round, for a data transmission rate of roughly R.) Encode U^N into a channel codeword by computing $X^N = U^N G_N$ and transmit X^N over the channel W.

Decoding: Having received Y^N, use the source decoder of the previous section to produce an estimate $\hat{U}_{E^c_{X|Y}}$ of the data bits $U_{E^c_{X|Y}}$.

Analysis: The error probability $\Pr(\hat{U}_{E^c_{X|Y}} \neq U_{E^c_{X|Y}})$ is bounded as $O(2^{-N^\beta})$ for any fixed $\beta < \frac{1}{2}$ since the source coding rate is $1 - R > H(X|Y)$. The complexity of the scheme is bounded as $O(N \log N)$.

Remark. The above argument reduces the channel coding problem for achieving the symmetric capacity $I(W)$ of a binary-input channel W to a source coding problem for a source $(X, Y) \sim QW$ where Q is uniform on $\{0, 1\}$. This reduction exploits the duality of the two problems. This dual approach provides an alternative proof of the channel coding results of [1]. It also complements the duality arguments in [2] and [3], where the source coding problem for a Ber(p) source was reduced to a channel coding problem for a binary symmetric channel with cross-over probability p.

V. Slepian-Wolf Coding

The above source coding method can be easily extended to the Slepian-Wolf setting [5]. Suppose $\{(X_i, Y_i)\}_{i=1}^\infty$ are independent samples from a source (X, Y) where both X and Y are binary RVs. In the Slepian-Wolf scenario, there are two encoders and one decoder. Fix a block-length $N = 2^n$, $n \geq 1$, and rates R_x and R_y for the two encoders. Encoder 1 observes X^N only and maps it to an integer $i_x \in [1, 2^{NR_x}]$, encoder 2 observes Y^N only and maps it to an integer $i_y \in [1, 2^{NR_y}]$. The decoder in the system observes (i_x, i_y)