

Constant Weight Polar Codes through Periodic Markov Processes

Boaz Shuval, Ido Tal

The Andrew and Erna Viterbi Faculty of Electrical and Computer Engineering,
Technion, Haifa 32000, Israel.

Email: {bshuval@, idotal@ee.}technion.ac.il

Abstract—Constant weight codes can arise from an input process sampled from a periodic Markov chain. A previous result showed that, in general, polarization does *not* occur for input-output processes with an underlying *periodic* Markov chain. In this work, we show that if we fix the initial state of an underlying periodic Markov chain, polarization *does* occur. Fixing the initial state is aligned with ensuring a constant weight code.

I. INTRODUCTION

Polar codes have been shown to achieve the capacity or information rate of many channel models [1]–[12]. Specifically, in [4] and [5], polar codes were shown to achieve the information rate of input-output processes with an underlying finite-state, *aperiodic*, and irreducible Markov (FAIM) process. Moreover, in [4, Theorem 4] it was shown by a counterexample that aperiodicity is a necessary condition. In this paper, we prove that polarization does occur for periodic processes, provided that the initial state is fixed. This setting is natural in the context of codewords constrained to be of a constant weight.

An example of an irreducible Markov chain of period $p = 4$ for producing sequences of constant weight $N/2$ (i.e., normalized weight $1/2$) is given in Figure 1. This holds if the initial state is ε and the sequence length N is a multiple of 4. A few things to note:

- We produce only a subset of the sequences of weight $N/2$.
- All the sequences of length 4 have equal probability, $1/6$, where $6 = \binom{4}{2}$ is the number of sequences of length 4 and weight 2.
- The number of states can be drastically reduced, while producing the same set of sequences at the same probabilities, see Figure 2.
- Enlarging the period p from 4 to a larger number will allow us to produce a larger subset of sequences of weight $N/2$, at the expense of more states.
- If the sequence length N is not a multiple of the period p , we must also constrain the final state to be in a certain set. For example, in Figure 1 and $N = 6$, the final state must be either ‘10’ or ‘01’. That is, of weight 1.
- For $\alpha = a/b$ rational, a similar graph can be constructed for sequences of weight $\lfloor \alpha N \rfloor$, when N is not a multiple of b . The same holds for weight $\lceil \alpha N \rceil$. This again requires constraining the final state to be of a certain weight.
- For $\alpha \leq \beta$ rational, a similar graph can be constructed for sequences of weight between αN and βN . Now, generally, the final state is constrained to have a range of weights.
- For positive integers a and b , the aperiodic graph in Figure 3 can ensure sequences whose weight modulo b

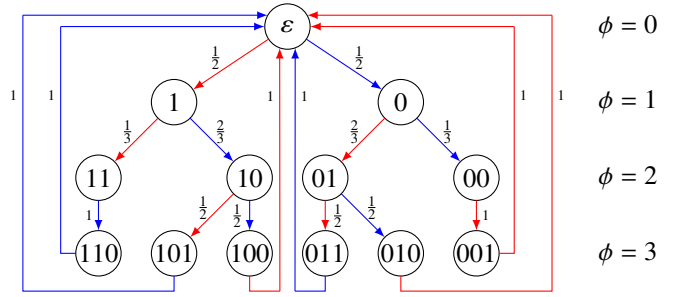


Fig. 1. Markov chain for producing sequences of weight $N/2$. Labels for blue and red arrows are ‘0’ and ‘1’, respectively. Transition probabilities between states are shown on the arrows. The state ε is the start and end state, at which the output is ensured to have weight $N/2$. The chain has period $p = 4$; each horizontal layer of the graph is a different phase ϕ .

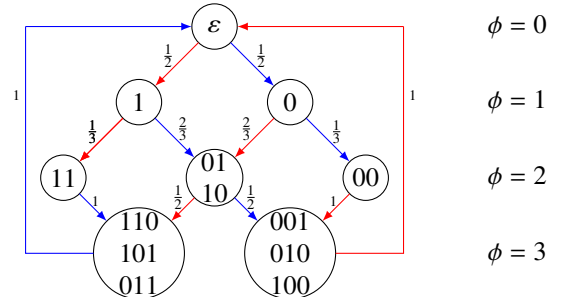


Fig. 2. Condensed Markov chain for producing sequences of weight $N/2$. Here, states from Figure 1 of the same phase and weight are merged. The probability of a blue (red) edge from a node with phase ϕ and weight w to a node with phase $\phi + 1$ and weight $w + 1$ is the number of valid sequences, i.e. length 4 and weight 2, whose weight in the prefix of length $\phi + 1$ is w ($w + 1$) divided by the number of valid sequences whose weight in the prefix of length ϕ is w .

is a . This is obtained by fixing the initial state to be ‘0’ and the final state to be ‘ a ’.

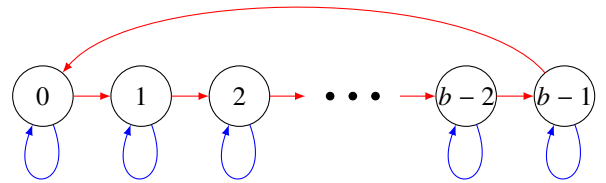


Fig. 3. Markov chain for producing sequences whose weight modulo b is constrained to some integer $0 \leq a < b$. Blue and red arrows are labelled ‘0’ and ‘1’, respectively. All edge probabilities are $1/2$. The chain is aperiodic. To construct the sequences, we constrain the initial state to ‘0’ and the final state to ‘ a ’.

II. SETTING AND MAIN THEOREM

Although we will ultimately fix the initial state, we start by considering a strictly stationary process, $(X_j, Y_j, S_j)_{j \in \mathbb{Z}}$. Here, X_j is binary, $Y_j \in \mathcal{Y}$, and $S_j \in \mathcal{S}$, with $\mathcal{Y}$ and $\mathcal{S}$ finite alphabets. We call S_j the state at time j . The conditional distribution of the process is $P_{X_j, Y_j, S_j | S_{j-1}}$, independent of j by stationarity. The distribution has a Markov property: conditioned on S_{j-1} , the random variables X_k, Y_k, S_k are independent of X_l, Y_l, S_{l-1} for any $l < j \leq k$.

The state sequence S_j is a finite-state irreducible Markov (FIM) process. By the Perron-Frobenius theorem [13, Theorem 8.4.4 and Corollary 8.1.30], it indeed has a unique stationary distribution π , which is positive. That is, for any state $s \in \mathcal{S}$, $\pi(s) > 0$. We denote the period of the state sequence by $\mathbf{p}$. That is, we can partition $\mathcal{S}$ into $\mathbf{p}$ classes, which we call *phases*. The phase of $s \in \mathcal{S}$ is denoted $\phi(s)$, an integer between 0 and $\mathbf{p}-1$. This partition is such that $\mathbb{P}(S_{j+1} = s' | S_j = s)$ is positive only if $\phi(s') \equiv \phi(s) + 1 \pmod{\mathbf{p}}$. For example, in Figure 1, each horizontal layer corresponds to a different phase.

The conditional entropy, Bhattacharyya, and total variation parameters are key to proving polarization theorems. For a binary random variable U , general random variables W, Q , and an event A ,

$$\begin{aligned} H(W|Q, A) &= - \sum_{w, q} P_{W, Q|A}(w, q) \log_2 P_{W|Q, A}(w|q), \\ Z(U|Q, A) &= 2 \sum_q \sqrt{P_{U, Q|A}(0, q) P_{U, Q|A}(1, q)}, \\ K(U|Q, A) &= \sum_q |P_{U, Q|A}(0, q) - P_{U, Q|A}(1, q)|. \end{aligned}$$

The entropy rate of our process, excluding the states, is

$$\mathcal{H}_{X|Y} = \lim_{N \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N). \quad (1)$$

We also denote

$$\mathcal{H}_X = \lim_{N \rightarrow \infty} \frac{1}{N} H(X_1^N). \quad (2)$$

These limits exist, see [4, Section II].

Following is our main theorem. By the Honda-Yamamoto scheme [14], it implies that a polar coding scheme with rate approaching $\mathcal{H}_X - \mathcal{H}_{X|Y}$ and vanishing error probability exists, if we condition on either the initial state or the initial state and the set of final states.

Theorem 1. *Let $(X_j, Y_j, S_j)_{j \in \mathbb{Z}}$ be a FIM process, where $(S_j)_{j \in \mathbb{Z}}$ has period $\mathbf{p}$. For $N = 2^n$, denote by U_1^N the polar transform of X_1^N , and fix $0 < \beta < 1/2$. Fix $0 \leq \varphi < \mathbf{p}$, and a non-empty set $\Psi_0 \subseteq \mathcal{S}$ such that all states $s_0 \in \Psi_0$ have phase $\phi(s_0) = \varphi$. Furthermore, let Ψ_N be a non-empty subset of $\mathcal{S}$ such that all states $s_N \in \Psi_N$ have phase $\phi(s_N) = \varphi + N \bmod \mathbf{p}$. Denote the event A_N as*

$$A_N \triangleq \{S_0 \in \Psi_0 \text{ and } S_N \in \Psi_N\}.$$

Then,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : Z(U_i | U_1^{i-1}, Y_1^N, A_N) < 2^{-N^\beta} \right\} \right| &= 1 - \mathcal{H}_{X|Y}, \\ \lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : K(U_i | U_1^{i-1}, A_N) < 2^{-N^\beta} \right\} \right| &= \mathcal{H}_X, \end{aligned}$$

We note that the encoding and decoding complexity of our coding scheme is $O(|\mathcal{S}|^3 N \log N)$, utilizing successive-cancellation trellis decoding [15], [16] and the Honda-Yamamoto scheme [14]. The only difference is that rather than marginalizing S_0, S_N over all states in $|\mathcal{S}|$, we marginalize S_0 only over Ψ_0 and S_N only over Ψ_N .

Throughout this paper, we assume the setting of Theorem 1. We denote by U_1^N, V_1^N the polar transforms of X_1^N and X_{N+1}^{2N} , respectively. We further denote $Q_i = (U_1^{i-1}, Y_1^N)$ and $R_i = (V_1^{i-1}, Y_{N+1}^{2N})$, for $1 \leq i \leq N$.

Our proof of Theorem 1 will follow by first showing slow polarization and then showing fast polarization. We factor the period $\mathbf{p}$ into

$$\mathbf{p} = \mathbf{d} \cdot \mathbf{q}, \quad (3)$$

where $\mathbf{q}$ is odd and $\mathbf{d}$ is a power of 2. In essence, we show that both slow and fast polarization occur if we fix the phase modulo $\mathbf{d}$ of the initial state. Our proofs are extensions of the ideas presented in [1], [17], [4],[5], and [6], while also taking into account the above.

For $0 \leq \Delta < \mathbf{d}$, define

$$\mathcal{S}(\Delta) = \{s \in \mathcal{S} : \phi(s) \equiv \Delta \pmod{\mathbf{d}}\}.$$

We denote the event

$$D(\Delta) \triangleq \{S_0 \in \mathcal{S}(\Delta)\} = \{\phi(S_0) \equiv \Delta \pmod{\mathbf{d}}\}.$$

Since $N = 2^n$ and $\mathbf{d}$ are both powers of 2, we have

$$D(\Delta) = \{S_N \in \mathcal{S}(\Delta)\}, \quad \text{for } N \geq \mathbf{d}. \quad (4)$$

Observe that by stationarity, we have for $0 \leq \Delta \leq \mathbf{d}$

$$\mathbb{P}(D(\Delta)) = 1/\mathbf{d}. \quad (5)$$

For $N = 2^n \geq \mathbf{d}$ denote

$$\begin{aligned} \mathcal{S}_N^2(\Delta) &= \{(s_0, s_N) \in \mathcal{S}^2 : s_0 \in \mathcal{S}(\Delta) \text{ and} \\ &\quad \phi(s_N) - \phi(s_0) \equiv N \pmod{\mathbf{p}}\}, \end{aligned} \quad (6)$$

and

$$\begin{aligned} \mathcal{S}_N^3(\Delta) &= \{(s_0, s_N, s_{2N}) \in \mathcal{S}^3 : (s_0, s_N) \in \mathcal{S}_N^2(\Delta) \text{ and} \\ &\quad (s_N, s_{2N}) \in \mathcal{S}_N^2(\Delta)\}. \end{aligned} \quad (7)$$

Observe that for $N = 2^n \geq \mathbf{d}$,

$$\begin{aligned} D(\Delta) \text{ holds} &\Leftrightarrow \{S_0 \in \mathcal{S}(\Delta)\} \Leftrightarrow \{S_N \in \mathcal{S}(\Delta)\} \Leftrightarrow \{S_{2N} \in \mathcal{S}(\Delta)\} \\ &\Leftrightarrow \{(S_0, S_N) \in \mathcal{S}_N^2(\Delta)\} \Leftrightarrow \{(S_0, S_N, S_{2N}) \in \mathcal{S}_N^3(\Delta)\}. \end{aligned}$$

The following definition and the observation and lemma that follow will be of use in the sequel.

Definition 1. Denote $f(N) \simeq g(N)$ if $f(N) - g(N)$ can be upper and lower bounded by expressions independent of N . It is easily shown that $\simeq$ is a transitive relation.

Observe that if $f(N) \simeq g(N)$ and the limit $\lim_{N \rightarrow \infty} f(N)/N$ exists, then $\lim_{N \rightarrow \infty} g(N)/N$ exists, and the limits coincide.

Lemma 2. *If V, Q , and W are random variables, A is an event, and the support of W is finite, then*

$$H(V, W|Q, A) \simeq H(V|Q, A) \simeq H(V|Q, W, A), \quad (8)$$

Proof: This is by the chain rule. For the left $\approx$, by the chain rule, $H(V, W|Q, A) = H(V|Q, A) + H(W|V, Q, A)$. Since the support of W is finite, the claim follows. For the right $\approx$, we use the chain rule on $H(W, V|Q, A)$ both ways and obtain

$$H(V|W, Q, A) - H(V|Q, A) = H(W|V, Q, A) - H(W|Q, A).$$

Since both terms on the right-hand side are bounded between 0 and the base-2 logarithm of the support of W , we conclude that $H(V|Q, W, A) \approx H(V|Q, A)$. ■

Finally, as a shorthand, for random variables X, Y we denote $P_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y)$. The extension for multiple random variables is immediate.

III. PROOF OF SLOW POLARIZATION

The proof of slow polarization comprises two steps: showing a convergence to a limiting random variable and then showing that this random variable is either 0 or 1, with probability 1.

For a given $0 \leq \Delta < \mathbf{d}$, $N = 2^n$, and $1 \leq i \leq N$, denote

$$\hat{H}_n(\Delta) = H(U_i|Q_i, S_0, S_N, D(\Delta)), \quad (9)$$

$$\hat{H}_n^-(\Delta) = H(U_i + V_i|Q_i, R_i, S_0, S_{2N}, D(\Delta)). \quad (10)$$

$$H_n(\Delta) = H(U_i|Q_i, D(\Delta)). \quad (11)$$

Lemma 3. Fix $0 \leq \Delta < \mathbf{d}$. Consider the random process

$$J_n = 1 + \langle B_0, B_1, \dots, B_n \rangle_2 = 1 + \sum_{j=0}^{n-1} B_j 2^{n-j},$$

where the B_j are i.i.d. Bernoulli 1/2 random variables. Then, the random processes $(\hat{H}_n(\Delta))_{n=1}^\infty$ and $(H_n(\Delta))_{n=1}^\infty$ where $i = J_n$ converge a.s., in L_1 , and in probability to limiting random variables $\hat{H}_\infty(\Delta) \in [0, 1]$ and $H_\infty(\Delta) \in [0, 1]$, respectively.

Proof: Take n_0 large enough such that $2^{n_0} \geq \mathbf{d}$. The proof follows by showing that the processes $(\hat{H}_n(\Delta))_{n=n_0}^\infty$ and $(H_n(\Delta))_{n=n_0}^\infty$ are a bounded sub-martingale and a bounded a super-martingale, respectively, with respect to $(J_n)_{n=1}^\infty$, see [18, Theorems 7.3.1 and 14.2.1 and Proposition 5.2.3]. Since we are considering entropies of binary random variables, both $\hat{H}_n(\Delta)$ and $H_n(\Delta)$ are clearly bounded between 0 and 1. It remains to show that $\mathbb{E}[\hat{H}_{n+1}(\Delta)|J_1, J_2, \dots, J_n] \geq \hat{H}_n(\Delta)$ and $\mathbb{E}[H_{n+1}(\Delta)|J_1, J_2, \dots, J_n] \leq H_n(\Delta)$, for $n \geq n_0$. Indeed, let $i = J_n$ be the random index corresponding to $\hat{H}_n(\Delta)$. Then,

$$\begin{aligned} \mathbb{E}[\hat{H}_{n+1}(\Delta)|J_1, J_2, \dots, J_n] &\stackrel{(a)}{=} \frac{1}{2}H(U_i + V_i|Q_i, R_i, S_0, S_{2N}, D(\Delta)) \\ &\quad + \frac{1}{2}H(V_i|U_i + V_i, Q_i, R_i, S_0, S_{2N}, D(\Delta)) \\ &\stackrel{(b)}{=} \frac{1}{2}H(U_i, V_i|Q_i, R_i, S_0, S_{2N}, D(\Delta)) \\ &\stackrel{(c)}{\geq} \frac{1}{2}H(U_i, V_i|Q_i, R_i, S_0, S_N, S_{2N}, D(\Delta)) \\ &\stackrel{(d)}{=} \frac{1}{2}H(U_i|Q_i, R_i, S_0, S_N, S_{2N}, D(\Delta)) \\ &\quad + \frac{1}{2}H(V_i|U_i, Q_i, R_i, S_0, S_N, S_{2N}, D(\Delta)) \\ &\stackrel{(e)}{=} \frac{1}{2}H(U_i|Q_i, S_0, S_N, D(\Delta)) + \frac{1}{2}H(V_i|R_i, S_N, S_{2N}, D(\Delta)) \\ &\stackrel{(f)}{=} \hat{H}_n(\Delta). \end{aligned}$$

Above, (a) is since B_{n+1} is Bernoulli 1/2, (b) is by the chain rule, (c) is since conditioning reduces entropy, (d) is again by the chain rule, (e) is by the Markov property and (4), and (f) is by stationarity.

For $H_n(\Delta)$ we have

$$\begin{aligned} \mathbb{E}[H_{n+1}(\Delta)|J_1, J_2, \dots, J_n] &\stackrel{(a)}{=} \frac{1}{2} \left(H(U_i + V_i|Q_i, R_i, D(\Delta)) + H(V_i|U_i + V_i, Q_i, R_i, D(\Delta)) \right) \\ &\stackrel{(b)}{=} \frac{1}{2}H(U_i, V_i|Q_i, R_i, D(\Delta)) \\ &\stackrel{(c)}{=} \frac{1}{2}H(U_i|Q_i, R_i, D(\Delta)) + \frac{1}{2}H(V_i|U_i, Q_i, R_i, D(\Delta)) \\ &\stackrel{(d)}{\leq} \frac{1}{2}H(U_i|Q_i, D(\Delta)) + \frac{1}{2}H(V_i|R_i, D(\Delta)) \\ &\stackrel{(e)}{=} H_n(\Delta). \end{aligned}$$

Above, (a) is since B_{n+1} is Bernoulli 1/2, (b) is by the chain rule, (c) is again by the chain rule, (d) is since conditioning reduces entropy and (4), and (e) is by stationarity. ■

For $N > 0$, denote

$$\begin{aligned} \mathcal{S}_N^3 &= \{(s_0, s_N, s_{2N}) \in \mathcal{S}^3 : \\ &\quad \phi(s_N) - \phi(s_0) \equiv \phi(s_{2N}) - \phi(s_N) \equiv N \pmod{\mathbf{p}}\}. \end{aligned} \quad (12)$$

Further denote, for fixed N clear from the context,

$$\pi(s_0, s_N, s_{2N}) \triangleq \mathbb{P}(S_0 = s_0, S_N = s_N, S_{2N} = s_{2N}). \quad (13)$$

Lemma 4. There exists N_0 and $\mu > 0$ such that for all $N \geq N_0$ and all $(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3$ we have

$$\pi(s_0, s_N, s_{2N}) > \mu.$$

Proof: See Appendix A. ■

For a given $0 \leq \Delta < \mathbf{d}$, $N = 2^n \geq \mathbf{d}$, and $1 \leq i \leq N$, we follow [6, sec. VI.B] and define for $(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)$,

$$\begin{aligned} \alpha(s_0, s_N, s_{2N}) &\triangleq \\ &H(U_i + V_i|Q_i, R_i, S_0 = s_0, S_N = s_N, S_{2N} = s_{2N}) \end{aligned} \quad (14)$$

and

$$\beta(s_0, s_N, s_{2N}) \triangleq \frac{\gamma(s_0, s_N) + \gamma(s_N, s_{2N})}{2}, \quad (15)$$

where

$$\gamma(s_0, s_N) \triangleq H(U_i|Q_i, S_0 = s_0, S_N = s_N). \quad (16)$$

By stationarity,

$$\gamma(s_N, s_{2N}) = H(V_i|R_i, S_N = s_N, S_{2N} = s_{2N}).$$

The following is the analog of [6, Lemma 16] for our setting.

Lemma 5. Let $N = 2^n \geq \max\{N_0, \mathbf{d}\}$, for N_0 as in Lemma 4. For $0 \leq \Delta < \mathbf{d}$ and $1 \leq i \leq N$ fixed

$$\hat{H}_n^-(\Delta) \geq \mathbf{d} \cdot \sum_{(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)} \pi(s_0, s_N, s_{2N}) \cdot \alpha(s_0, s_N, s_{2N}), \quad (17)$$

and

$$\hat{H}_n(\Delta) = \mathbf{d} \cdot \sum_{(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)} \pi(s_0, s_N, s_{2N}) \cdot \beta(s_0, s_N, s_{2N}). \quad (18)$$

Furthermore, for all $(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)$,

$$\alpha(s_0, s_N, s_{2N}) \geq \beta(s_0, s_N, s_{2N}). \quad (19)$$

Proof: Note that under the stationary distribution, we have

$$\mathbb{P}(S_0 = s_0, S_N = s_N, S_{2N} = s_{2N} | D(\Delta)) = \mathbf{d} \cdot \pi(s_0, s_N, s_{2N}).$$

Hence, by (10), (14), and since conditioning reduces entropy,

$$\begin{aligned} \hat{H}_n^-(\Delta) &= H(U_i + V_i | Q_i, R_i, S_0, S_{2N}, D(\Delta)) \\ &\geq H(U_i + V_i | Q_i, R_i, S_0, s_N, S_{2N}, D(\Delta)) \\ &= \sum_{(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)} \mathbf{d} \cdot \pi(s_0, s_N, s_{2N}) \cdot \alpha(s_0, s_N, s_{2N}), \end{aligned}$$

which establishes (17).

To see (18), observe that by stationarity we have

$$\begin{aligned} \hat{H}_n(\Delta) &= H(U_i | Q_i, S_0, s_N, D(\Delta)) \\ &= \frac{H(U_i | Q_i, S_0, s_N, D(\Delta)) + H(V_i | R_i, S_N, S_{2N}, D(\Delta))}{2} \\ &= \sum_{(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)} \mathbf{d} \cdot \pi(s_0, s_N, s_{2N}) \cdot \frac{\gamma(s_0, s_N) + \gamma(s_N, s_{2N})}{2} \\ &= \sum_{(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)} \mathbf{d} \cdot \pi(s_0, s_N, s_{2N}) \cdot \frac{\beta(s_0, s_N, s_{2N})}{2}, \end{aligned}$$

where in the last two equalities we used (16) and (15).

Finally,

$$\begin{aligned} &\alpha(s_0, s_N, s_{2N}) \\ &\stackrel{(a)}{\geq} H(U_i + V_i | V_i, Q_i, R_i, S_0 = s_0, S_N = s_N, S_{2N} = s_{2N}) \\ &= H(U_i | V_i, Q_i, R_i, S_0 = s_0, S_N = s_N, S_{2N} = s_{2N}) \\ &\stackrel{(b)}{=} H(U_i | Q_i, S_0 = s_0, S_N = s_N) \\ &= \gamma(s_0, s_N), \end{aligned}$$

where (a) is since conditioning by V_i reduces entropy, and (b) is by the Markov property. Similarly, by conditioning on U_i , we have

$$\alpha(s_0, s_N, s_{2N}) \geq H(V_i | R_i, S_N = s_N, S_{2N} = s_{2N}) = \gamma(s_N, s_{2N}).$$

Combining these two inequalities and recalling (15) yields (19), which completes the proof. ■

Recall from (3) the factorization $\mathbf{p} = \mathbf{d} \cdot \mathbf{q}$, where $\mathbf{d}$ is a power of 2 and $\mathbf{q}$ is odd. We have already used the special structure of $\mathbf{d}$ previously. The entire factorization of $\mathbf{p}$ is the key for proving the periodic analog of [6, Lemma 17], below.

Lemma 6. *For every $0 < \epsilon < 1$ there exists $\Theta = \Theta(\epsilon) > 0$, for which the following holds. Let $N = 2^n \geq \max\{N_0, \mathbf{d}\}$, where the Euler totient of $\mathbf{q}$ divides n , and N_0 is as promised in Lemma 4. Fix $0 \leq \Delta < \mathbf{d}$ and $1 \leq i \leq N$. Then, if $\epsilon \leq \hat{H}_n(\Delta) \leq 1 - \epsilon$, there exist $(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)$ such that*

$$\alpha(s_0, s_N, s_{2N}) > \beta(s_0, s_N, s_{2N}) + \Theta. \quad (20)$$

Proof: By definition of γ in (16) and stationarity,

$$\begin{aligned} \hat{H}_n(\Delta) &= \mathbf{d} \sum_{(s_0, s_N) \in \mathcal{S}_N^2(\Delta)} \mathbb{P}(S_0 = s_0, S_N = s_N) \cdot \gamma(s_0, s_N) \\ &= \mathbf{d} \sum_{(s_N, s_{2N}) \in \mathcal{S}_N^2(\Delta)} \mathbb{P}(S_N = s_N, S_{2N} = s_{2N}) \cdot \gamma(s_N, s_{2N}). \end{aligned} \quad (21)$$

Our aim is to show that there exists a triplet $(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)$ for which

$$\min\{\gamma(s_0, s_N), \gamma(s_N, s_{2N})\} \leq 1 - \epsilon \quad (22)$$

and

$$\max\{\gamma(s_0, s_N), \gamma(s_N, s_{2N})\} \geq \epsilon. \quad (23)$$

The above would imply (20), due to [4, Lemma 11] and part (i) of [17, Lemma 2.2]. Note that the latter is required since the former does not involve conditioning. We further note that [4, Lemma 11] is stated for strict inequalities, but actually holds also for non-strict inequalities, as can be seen in its proof.

We will show the existence of a triplet $(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)$ for which

$$\gamma(s_0, s_N) \leq \hat{H}_n(\Delta) \quad \text{and} \quad \gamma(s_N, s_{2N}) \geq \hat{H}_n(\Delta) \quad (24)$$

or

$$\gamma(s_0, s_N) \geq \hat{H}_n(\Delta) \quad \text{and} \quad \gamma(s_N, s_{2N}) \leq \hat{H}_n(\Delta) \quad (25)$$

Note that by our assumption that $\epsilon \leq \hat{H}_n(\Delta) \leq 1 - \epsilon$, the above would indeed imply (22) and (23). The above can be written succinctly as

$$(\hat{H}_n(\Delta) - \gamma(s_0, s_N)) \cdot (\hat{H}_n(\Delta) - \gamma(s_N, s_{2N})) \leq 0.$$

The proof is by contradiction. Namely, assume that for all $(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)$,

$$(\hat{H}_n(\Delta) - \gamma(s_0, s_N)) \cdot (\hat{H}_n(\Delta) - \gamma(s_N, s_{2N})) > 0. \quad (26)$$

Let phase $0 \leq \varphi < \mathbf{p} = \mathbf{d} \cdot \mathbf{q}$ be such that $\varphi \equiv \Delta \pmod{\mathbf{d}}$ and $\varphi \equiv -2 \pmod{\mathbf{q}}$. Such a φ indeed exists, by the Chinese remainder theorem [19, Theorem 2.12.6]. Fix s_0, s_N such that $\phi(s_0) = \varphi$ and $\phi(s_N) \equiv \varphi + N \pmod{\mathbf{p}}$. Note that $(s_0, s_N) \in \mathcal{S}_N^2(\Delta)$. For $s \in \mathcal{S}$, denote $\phi'(s) = \phi(s) \pmod{\mathbf{q}}$. Thus, $\phi'(s_0) \equiv -2 \pmod{\mathbf{q}}$ and $\phi'(s_N) \equiv -1 \pmod{\mathbf{q}}$, where the second equality is by Euler's theorem [19, Theorem 2.14.1] and our assumption that the totient of $\mathbf{q}$ divides n . Since (26) is assumed to hold, w.l.o.g. further assume that $\gamma(s_0, s_N) < \hat{H}_n(\Delta)$.

Our aim is to show that for any $(s, s') \in \mathcal{S}_N^2(\Delta)$, we have $\gamma(s, s') < \hat{H}_n(\Delta)$. This would establish the contradiction, by (21). Indeed, by stationarity, for every $0 \leq \Delta < \mathbf{d}$, $\sum_{(s, s') \in \mathcal{S}_N^2(\Delta)} \mathbb{P}(S_0 = s, S_N = s') = 1/\mathbf{d}$. We do this by showing that for any $0 \leq j < \mathbf{q}$ and any $(s, s') \in \mathcal{S}_N^2(\Delta)$ such that $\phi'(s) = j$, we have $\gamma(s, s') < \hat{H}_n(\Delta)$.

- 1) We have already established that $\gamma(s_0, s_N) < \hat{H}_n(\Delta)$. Recall that $\phi'(s_0) = -2$ and $\phi'(s_N) \equiv -1 \pmod{\mathbf{q}}$.
- 2) We claim that for any $a_0 \in \mathcal{S}$ such that $(s_N, a_0) \in \mathcal{S}_N^2(\Delta)$, we have $\gamma(s_N, a_0) < \hat{H}_n(\Delta)$. Indeed, this follows by (26), and noting that $(s_0, s_N, a_0) \in \mathcal{S}_N^3(\Delta)$. Note that $\phi'(a_0) \equiv 0 \pmod{\mathbf{q}}$.
- 3) We now claim that for any a_0, a_1 such that $(s_N, a_0, a_1) \in \mathcal{S}_N^3(\Delta)$, we have $\gamma(a_0, a_1) < \hat{H}_n(\Delta)$. Indeed, this follows by (26) and item 2. Note that $\phi'(a_0) \equiv 0 \pmod{\mathbf{q}}$ and $\phi'(a_1) \equiv 1 \pmod{\mathbf{q}}$.
- 4) We conclude from item 3 that for any $(a_0, a_1) \in \mathcal{S}_N^2(\Delta)$ such that $\phi'(a_0) \equiv 0 \pmod{\mathbf{q}}$ we have $\gamma(a_0, a_1) < \hat{H}_n(\Delta)$. This will be the basis of our induction on j . Note that $\phi'(a_1) \equiv 1 \pmod{\mathbf{q}}$.

5) For the induction step, assume that for some $j \geq 0$, for any $(a_j, a_{j+1}) \in \mathcal{S}_N^2(\Delta)$ such that $\phi'(a_j) \equiv j \pmod{q}$, it holds that $\gamma(a_j, a_{j+1}) < \hat{H}_n(\Delta)$. We now show that for any $(a_{j+1}, a_{j+2}) \in \mathcal{S}_N^2(\Delta)$ such that $\phi'(a_{j+1}) \equiv j+1 \pmod{q}$, it holds that $\gamma(a_{j+1}, a_{j+2}) < \hat{H}_n(\Delta)$. Indeed this holds since $(a_j, a_{j+1}, a_{j+2}) \in \mathcal{S}_N^3(\Delta)$, the induction hypothesis, and (26).

This completes the proof. $\blacksquare$

The following is a corollary of Lemmas 4 to 6.

Corollary 7. Fix $0 < \epsilon < 1$. There exists $\tau(\epsilon) > 0$ such that the following holds. If $N = 2^n$ is as in Lemma 6, $0 \leq \Delta < d$, $1 \leq i \leq N$, and $\epsilon \leq \hat{H}_n(\Delta) \leq 1 - \epsilon$, then

$$\hat{H}_n^-(\Delta) - \hat{H}_n(\Delta) > \tau(\epsilon).$$

Proof: Take $\tau(\epsilon) = \mu \cdot d \cdot \Theta(\epsilon)$, where μ is as promised in Lemma 4 and $\Theta(\epsilon)$ is as promised in Lemma 6. $\blacksquare$

The following lemma is essentially [20, Theorem 1], adapted to our case.

Lemma 8. The random variable $\hat{H}_\infty(\Delta)$ is a $\{0, 1\}$ random variable with probability 1.

Proof: Consider the random process $(\hat{H}_n(\Delta))_{n=1}^\infty$, as in Lemma 3, where it was shown to converge a.s., in L_1 , and in probability to $\hat{H}_\infty(\Delta) \in [0, 1]$. To show that $\hat{H}_\infty(\Delta) \in \{0, 1\}$ with probability 1, assume to the contrary that there exists $\epsilon > 0$ such $\mathbb{P}(\hat{H}_\infty(\Delta) \in (2\epsilon, 1 - 2\epsilon)) > \rho$, for some $\rho > 0$. Next, note that

$$\begin{aligned} & \mathbb{P}(\hat{H}_n(\Delta) \in (\epsilon, 1 - \epsilon)) \\ & \geq \mathbb{P}(\hat{H}_n(\Delta) \in (\epsilon, 1 - \epsilon) \text{ and } |\hat{H}_n(\Delta) - \hat{H}_\infty(\Delta)| < \epsilon) \\ & \geq \mathbb{P}(\hat{H}_\infty(\Delta) \in (2\epsilon, 1 - 2\epsilon) \text{ and } |\hat{H}_n(\Delta) - \hat{H}_\infty(\Delta)| < \epsilon) \\ & = \mathbb{P}(\hat{H}_\infty(\Delta) \in (2\epsilon, 1 - 2\epsilon)) \\ & \quad - \mathbb{P}(\hat{H}_\infty(\Delta) \in (2\epsilon, 1 - 2\epsilon) \text{ and } |\hat{H}_n(\Delta) - \hat{H}_\infty(\Delta)| \geq \epsilon) \\ & \geq \mathbb{P}(\hat{H}_\infty(\Delta) \in (2\epsilon, 1 - 2\epsilon)) - \mathbb{P}(|\hat{H}_n(\Delta) - \hat{H}_\infty(\Delta)| \geq \epsilon) \\ & > \rho - \mathbb{P}(|\hat{H}_n(\Delta) - \hat{H}_\infty(\Delta)| \geq \epsilon), \end{aligned}$$

where the last inequality follows from our contradictory assumption. By convergence in probability of $\hat{H}_n(\Delta)$, we obtain

$$\liminf_{n \rightarrow \infty} \mathbb{P}(\hat{H}_n(\Delta) \in (\epsilon, 1 - \epsilon)) \geq \rho.$$

Thus, from Corollary 7, recalling our constraint on n from Lemma 6, there exist an infinite sequence $n = n_1, n_2, \dots$ and $\tau > 0$ such that $\mathbb{P}(|\hat{H}_{n+1}^-(\Delta) - \hat{H}_n(\Delta)| > \tau) > \rho/2$. Since $\hat{H}_{n+1}(\Delta)$ equals $\hat{H}_{n+1}^-(\Delta)$ with probability 1/2, we have $\mathbb{P}(|\hat{H}_{n+1}(\Delta) - \hat{H}_n(\Delta)| > \tau) > \rho/4$. This implies that $\hat{H}_n(\Delta)$ cannot converge in probability. This is a contradiction, and we conclude that $\hat{H}_\infty(\Delta) \in \{0, 1\}$ with probability 1. $\blacksquare$

The following lemma states that fixing the initial phase does not change the entropy rate. Note that in the following we have already established the existence of the limit on the left-hand side, and claim that the right-hand side limit exists.

Lemma 9. For $0 \leq \varphi < p$, we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N) = \lim_{N \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N, \phi(S_0) = \varphi).$$

Proof: By (8), we have $H(X_1^N | Y_1^N) \simeq H(X_1^N | Y_1^N, \phi(S_0))$, by taking A as the always true event. Thus, our aim is to show that for any $0 \leq \varphi < p$,

$$H(X_1^N | Y_1^N, \phi(S_0)) \simeq H(X_1^N | Y_1^N, \phi(S_0) = \varphi).$$

By stationarity, we have $\mathbb{P}(\phi(S_0) = -\varphi' \pmod{p}) = 1/p$ for any $0 \leq \varphi' < p$. Thus,

$$H(X_1^N | Y_1^N, \phi(S_0)) = \sum_{\varphi'=0}^{p-1} \frac{1}{p} H(X_1^N | Y_1^N, \phi(S_0) = -\varphi' \pmod{p}). \quad (27)$$

Recall that φ is fixed. For a given $0 \leq \varphi' \leq p$, denote $\psi = \varphi' + \varphi$. Note that $\psi \geq 0$. We have

$$\begin{aligned} & H(X_1^N | Y_1^N, \phi(S_0) \equiv -\varphi' \pmod{p}) \\ & \stackrel{(a)}{\simeq} H(X_1^N | Y_1^N, S_\psi, S_N, \phi(S_0) \equiv -\varphi' \pmod{p}) \\ & = H(X_1^N | Y_1^N, S_\psi, S_N, \phi(S_\psi) = \varphi) \\ & \stackrel{(b)}{=} H(X_1^N | Y_1^{N+\psi}, S_\psi, S_N, \phi(S_\psi) = \varphi) \\ & \stackrel{(c)}{\simeq} H(X_1^{N+\psi} | Y_1^{N+\psi}, S_\psi, S_N, \phi(S_\psi) = \varphi) \\ & \stackrel{(d)}{\simeq} H(X_{\psi+1}^{N+\psi} | Y_1^{N+\psi}, S_\psi, S_N, \phi(S_\psi) = \varphi) \\ & \stackrel{(e)}{=} H(X_{\psi+1}^{N+\psi} | Y_{\psi+1}^{N+\psi}, S_\psi, S_N, \phi(S_\psi) = \varphi) \\ & \stackrel{(f)}{\simeq} H(X_{\psi+1}^{N+\psi} | Y_{\psi+1}^{N+\psi}, \phi(S_\psi) = \varphi) \\ & \stackrel{(g)}{=} H(X_1^N | Y_1^N, \phi(S_0) = \varphi), \end{aligned}$$

where (a) is by the RHS of (8) since the number of states is finite; (b) is by the Markov property; (c) is by the LHS of (8) since we have added a finite number ψ of binary random variables; (d) is again by the LHS of (8), now removing a finite number ψ of binary random variables; (e) is by the Markov property; (f) is by the RHS of (8); and (g) is by stationarity.

Thus, we have shown that

$$H(X_1^N | Y_1^N, \phi(S_0) \equiv -\varphi' \pmod{p}) \simeq H(X_1^N | Y_1^N, \phi(S_0) = \varphi).$$

Using this in (27) we obtain

$$H(X_1^N | Y_1^N, \phi(S_0)) \simeq H(X_1^N | Y_1^N, \phi(S_0) = \varphi),$$

completing the proof. $\blacksquare$

The following proposition establishes slow polarization. It is in a similar vein to [6, Theorem 20].

Proposition 10. For $0 \leq \Delta < d$, $0 < \epsilon < 1$, and $N = 2^n$, the following holds:

$$\lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : H(U_i | U_1^{i-1}, Y_1^N, D(\Delta)) < \epsilon \right\} \right| \quad (28a)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : H(U_i | U_1^{i-1}, Y_1^N, S_0, S_N, D(\Delta)) < \epsilon \right\} \right| \quad (28b)$$

$$= 1 - \lim_{n \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N, S_0, S_N, D(\Delta)) \quad (28c)$$

$$= 1 - \lim_{n \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N, D(\Delta)) \quad (28d)$$

$$= 1 - \lim_{n \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N) \quad (28e)$$

and

$$\lim_{n \rightarrow \infty} \frac{1}{N} |\{i : H(U_i|U_1^{i-1}, Y_1^N, D(\Delta)) > 1 - \epsilon\}| \quad (29a)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{N} |\{i : H(U_i|U_1^{i-1}, Y_1^N, S_0, S_N, D(\Delta)) > 1 - \epsilon\}| \quad (29b)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N, S_0, S_N, D(\Delta)) \quad (29c)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N, D(\Delta)) \quad (29d)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N). \quad (29e)$$

Proof: We first prove (29e)=(29d), which implies (28e)=(28d). Note that $D(\Delta)$ is a union of the events $\phi(S_0) = \varphi$, where $\varphi \equiv \Delta \pmod{\mathbf{d}}$. Given $D(\Delta)$, each such event has probability $1/\mathbf{q}$, by stationarity. Hence,

$$H(X_1^N | Y_1^N, D(\Delta)) = \sum_{\varphi \equiv \Delta \pmod{\mathbf{d}}} \frac{1}{\mathbf{q}} H(X_1^N | Y_1^N, \phi(S_0) = \varphi).$$

Note that the number of summands is $\mathbf{q}$, by the Chinese remainder theorem [19, Theorem 2.12.6]. Multiplying both sides by $1/N$ and taking limits with $n \rightarrow \infty$, we use Lemma 9 to obtain the desired equality.

We next prove (29d)=(29c), which implies (28d)=(28c). We do so by using the notation $\simeq$ from Definition 1. By the RHS of (8), and noting that the number of states is finite, we have $H(X_1^N | Y_1^N, D(\Delta)) \simeq H(X_1^N | Y_1^N, S_0, S_N, D(\Delta))$, from which the equality follows.

We next prove that (29c)=(29b) and (28c)=(28b). First, note that for the random process $(\hat{H}_n(\Delta))_{n=1}^\infty$, as in Lemma 3,

$$\begin{aligned} \mathbb{E} [\hat{H}_n(\Delta)] &= \frac{1}{N} \sum_{i=1}^N H(U_i | U_1^{i-1}, Y_1^N, S_0, S_N, D(\Delta)) \\ &= \frac{1}{N} H(U_1^N | Y_1^N, S_0, S_N, D(\Delta)) \\ &= \frac{1}{N} H(X_1^N | Y_1^N, S_0, S_N, D(\Delta)). \end{aligned}$$

Thus, $\lim_{n \rightarrow \infty} \mathbb{E} [\hat{H}_n(\Delta)]$ equals (29c). By the bounded convergence theorem [18, Theorem 7.3.1] and Lemma 3,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E} [\hat{H}_n(\Delta)] &= \mathbb{E} \left[\lim_{n \rightarrow \infty} \hat{H}_n(\Delta) \right] = \mathbb{E} [\hat{H}_\infty(\Delta)] \\ &= \mathbb{P} (\hat{H}_\infty(\Delta) = 1) = \mathbb{P} (\hat{H}_\infty(\Delta) > 1 - \epsilon), \quad (30) \end{aligned}$$

where the latter two equalities are since $\hat{H}_\infty(\Delta) \in \{0, 1\}$ by Lemma 5. Hence,

$$\begin{aligned} (29c) &= \mathbb{P} (\hat{H}_\infty(\Delta) > 1 - \epsilon) \\ &\stackrel{(a)}{=} \lim_{n \rightarrow \infty} \mathbb{P} (\hat{H}_n(\Delta) > 1 - \epsilon) = (29b), \quad (31) \end{aligned}$$

where (a) is by [18, Theorem 10.1.1(3)] and since convergence in probability of $\hat{H}_n(\Delta)$ implies convergence in distribution [18, Theorem 10.2.1]. This establishes (29b)=(29c). The equality (28b) = (28c) follows similarly, for the process $(1 - \hat{H}_n(\Delta))_n$.

Finally, we prove that (29c)=(29a) and (28c)=(28a). To this end, we define the random sequence $(H_n(\Delta))_{n=1}^\infty$, where $H_n(\Delta)$

was defined in (11) and the index i is drawn in the same manner as in Lemma 3. Now, utilizing the chain rule we obtain that

$$\begin{aligned} \mathbb{E} [H_n(\Delta)] &= \frac{1}{N} \sum_{i=1}^N H(U_i | U_1^{i-1}, Y_1^N, D(\Delta)) \\ &= \frac{1}{N} H(U_1^N | Y_1^N, D(\Delta)) = \frac{1}{N} H(X_1^N | Y_1^N, D(\Delta)). \end{aligned}$$

Taking limits on both sides yields that

$$\lim_{n \rightarrow \infty} \mathbb{E} [H_n(\Delta)] = \lim_{n \rightarrow \infty} \frac{1}{N} H(X_1^N | Y_1^N, D(\Delta)) = \mathbb{E} [\hat{H}_\infty(\Delta)],$$

where the latter equality is since we have shown above that (29d) = (29c) = (29b) = $\mathbb{E} [\hat{H}_\infty(\Delta)]$, see (30) and (31). Next, since conditioning reduces entropy we have $H_n(\Delta) \geq \hat{H}_n(\Delta)$. By [6, Lemma 19], this implies that $H_n(\Delta)$ converges in L_1 to $\hat{H}_\infty(\Delta)$. Since $H_n(\Delta)$ is also bounded, then L_1 convergence implies convergence in probability and thus in distribution. Now, using the same argument as for (31),

$$(29c) = \mathbb{P} (\hat{H}_\infty(\Delta) > 1 - \epsilon) = \lim_{n \rightarrow \infty} \mathbb{P} (H_n(\Delta) > 1 - \epsilon) = (29a).$$

To see that (28c) = (28a), we repeat the above for the process $(1 - H_n(\Delta))_n$. ■

IV. PROOF OF FAST POLARIZATION

To prove fast polarization, we first define $M(\Delta)$ for all $0 \leq \Delta < \mathbf{d}$,

$$M(\Delta) = \frac{1}{\mathbf{d} \cdot \min\{\pi_s : s \in \mathcal{S}(\Delta)\}}. \quad (32)$$

Recall that for any s , we have by irreducibility and stationarity that $\pi_s > 0$, and hence $M(\Delta)$ is well defined (the denominator is positive). The utility of $M(\Delta)$ is seen in the next lemma, which allows us to bound the probability of two dependent events by their marginal probabilities.

Lemma 11. Fix $0 \leq \Delta < \mathbf{d}$ and $N = 2^n \geq \mathbf{d}$. Let $A, B \subseteq \mathcal{X}^N \times \mathcal{Y}^N$. Then,

$$\begin{aligned} \mathbb{P} \left((X_1^N, Y_1^N) \in A \quad \text{and} \quad (X_{N+1}^{2N}, Y_{N+1}^{2N}) \in B \middle| D(\Delta) \right) \\ \leq M(\Delta) \cdot \mathbb{P} \left((X_1^N, Y_1^N) \in A \middle| D(\Delta) \right) \\ \cdot \mathbb{P} \left((X_{N+1}^{2N}, Y_{N+1}^{2N}) \in B \middle| D(\Delta) \right). \quad (33) \end{aligned}$$

Proof: See Appendix A. ■

Recall the definitions of U_i, V_i, Q_i, R_i from Section II. For a given $0 \leq \Delta < \mathbf{d}$, $N = 2^n$, and $1 \leq i \leq N$, denote

$$\begin{aligned} Z_n(\Delta) &= Z(U_i | Q_i, D(\Delta)) \\ Z_n^-(\Delta) &= Z(U_i + V_i | Q_i, R_i, D(\Delta)) \\ Z_n^+(\Delta) &= Z(V_i | U_i + V_i, Q_i, R_i, D(\Delta)) \end{aligned}$$

and

$$\begin{aligned} \hat{K}_n(\Delta) &= K(U_i | Q_i, S_0, S_N, D(\Delta)) \\ \hat{K}_n^-(\Delta) &= K(U_i + V_i | Q_i, R_i, S_0, S_{2N}, D(\Delta)) \\ \hat{K}_n^+(\Delta) &= K(V_i | U_i + V_i, Q_i, R_i, S_0, S_{2N}, D(\Delta)). \end{aligned}$$

Lemma 12. Fix $0 \leq \Delta < \mathbf{d}$, $N = 2^n \geq \mathbf{d}$, and $1 \leq i \leq N$. Then,

$$Z_n^-(\Delta) \leq 2 \cdot M(\Delta) \cdot Z_n(\Delta) \quad \text{and} \quad Z_n^+(\Delta) \leq M(\Delta) \cdot (Z_n(\Delta))^2.$$

This is an extension of [4, Section VI] for our scenario.

Proof: We first prove that $Z_n^-(\Delta) \leq 2 \cdot M(\Delta) \cdot Z_n(\Delta)$. As a shorthand, denote for an i clear from the context,

$$\begin{aligned} p(u, q) &= \mathbb{P}(U_i = u, Q_i = q | D(\Delta)) \\ p(v, r) &= \mathbb{P}(V_i = v, R_i = r | D(\Delta)) \\ p(u, v, q, r) &= \mathbb{P}(U_i = u, V_i = v, Q_i = q, R_i = r | D(\Delta)). \end{aligned}$$

We further denote

$$\begin{aligned} a_q &= p(0, q), \quad b_q = p(1, q), \\ c_r &= p(0, r), \quad d_r = p(1, r). \end{aligned} \quad (34)$$

Observe that by Lemma 11, we have

$$p(u, v, q, r) \leq M(\Delta) p(u, q) p(v, r). \quad (35)$$

Indeed,

$$\begin{aligned} Z_n^-(\Delta) &= 2 \sum_{q,r} \sqrt{\mathbb{P}(U_i + V_i = 0, Q_i = q, R_i = r | D(\Delta))} \\ &\quad \cdot \sqrt{\mathbb{P}(U_i + V_i = 1, Q_i = q, R_i = r | D(\Delta))} \\ &= 2 \sum_{q,r} \sqrt{(p(0, 0, q, r) + p(1, 1, q, r))} \\ &\quad \cdot \sqrt{(p(0, 1, q, r) + p(1, 0, q, r))} \\ &\stackrel{(a)}{=} 2M(\Delta) \sum_{q,r} \sqrt{(a_q c_r + b_q d_r)(a_q d_r + b_q c_r)} \\ &= 2M(\Delta) \sum_{q,r} \sqrt{(a_q^2 + b_q^2) c_r d_r + (c_r^2 + d_r^2) a_q b_q} \\ &\stackrel{(b)}{\leq} 2M(\Delta) \sum_{q,r} \left(\sqrt{(a_q^2 + b_q^2) c_r d_r} + \sqrt{(c_r^2 + d_r^2) a_q b_q} \right) \\ &\stackrel{(c)}{\leq} 2M(\Delta) \sum_{q,r} \left((a_q + b_q) \sqrt{c_r d_r} + (c_r + d_r) \sqrt{a_q b_q} \right) \\ &\stackrel{(d)}{=} M(\Delta) \left(2 \sum_r \sqrt{c_r d_r} + 2 \sum_q \sqrt{a_q b_q} \right) \\ &\stackrel{(e)}{=} 2M(\Delta) Z_n(\Delta), \end{aligned}$$

where (a) is by (34) and (35); (b) and (c) follow from the inequality $\sqrt{\alpha + \beta} \leq \sqrt{\alpha} + \sqrt{\beta}$, which holds for $\alpha, \beta \geq 0$; (d) holds since $a_q + b_q = \mathbb{P}(Q_i = q)$, $c_r + d_r = \mathbb{P}(R_i = r)$ and we sum over all q and r , respectively; finally, (e) is by stationarity.

Next,

$$\begin{aligned} Z_n^+(\Delta) &= 2 \sum_{q,r,t} \sqrt{\mathbb{P}(V_i = 0, U_i + V_i = t, Q_i = q, R_i = r | D(\Delta))} \\ &\quad \cdot \sqrt{\mathbb{P}(V_i = 1, U_i + V_i = t, Q_i = q, R_i = r | D(\Delta))} \\ &= 2 \sum_{q,r,t} \sqrt{\mathbb{P}(U_i = t, V_i = 0, Q_i = q, R_i = r | D(\Delta))} \\ &\quad \cdot \sqrt{\mathbb{P}(U_i = 1 + t, V_i = 1, Q_i = q, R_i = r | D(\Delta))} \\ &= 2 \sum_{q,r,t} \sqrt{p(t, 0, q, r) p(1 + t, 1, q, r)} \\ &\stackrel{(a)}{\leq} 2M(\Delta) \sum_{q,r,t} \sqrt{p(t, q) p(1 + t, q) p(0, r) p(1, r)} \\ &\stackrel{(b)}{=} 2M(\Delta) \sum_{q,r,t} \sqrt{p(0, q) p(1, q) p(0, r) p(1, r)} \\ &\stackrel{(c)}{=} 2M(\Delta) \cdot 2 \sum_{q,r} \sqrt{a_q b_q c_r d_r} \end{aligned}$$

$$\begin{aligned} &= M(\Delta) \cdot 2 \sum_q \sqrt{a_q b_q} \cdot 2 \sum_r \sqrt{c_r d_r} \\ &\stackrel{(d)}{=} M(\Delta) (Z_n(\Delta))^2, \end{aligned}$$

where (a) is by Lemma 11, (b) and (c) are since t can be either 0 or 1, and (d) is by stationarity. ■

Lemma 13. Fix $0 \leq \Delta < \mathbf{d}$, $N = 2^n \geq \mathbf{d}$, and $1 \leq i \leq N$. Then,

$$\hat{K}_n^-(\Delta) \leq \mathbf{d} \cdot M(\Delta) \cdot (\hat{K}_n(\Delta))^2 \quad \text{and} \quad \hat{K}_n^+(\Delta) \leq 2 \cdot \hat{K}_n(\Delta).$$

This is an extension of [5, Proposition 12] to our scenario.

Proof: Recalling (13), we denote

$$\lambda(s_N) = \pi(s_0, s_N, s_{2N}),$$

where we have deliberately omitted the dependence on s_0, s_{2N} . By summing over s_0 or s_{2N} in $\mathcal{S}(\Delta)$ we obtain

$$\sum_{s_0 \in \mathcal{S}(\Delta)} \lambda(s_N) = P_{S_N, s_{2N}}(s_N, s_{2N}), \quad (35a)$$

$$\sum_{s_{2N} \in \mathcal{S}(\Delta)} \lambda(s_N) = P_{S_0, S_N}(s_0, s_N). \quad (35b)$$

For $(s_0, s_N, s_{2N}) \in \mathcal{S}_N^3(\Delta)$, we have by (32) and the Markov property

$$\lambda(s_N) \leq \mathbf{d} \cdot M(\Delta) P_{S_0, S_N}(s_0, s_N) P_{S_N, s_{2N}}(s_N, s_{2N}). \quad (36)$$

We further define for $s_0, s_N, s_{2N} \in \mathcal{S}_N^3(\Delta)$ and i clear from the context the following shorthand

$$\begin{aligned} p_{s_0}^{s_N}(u, q) &= P_{U_i, Q_i | S_0, s_N}(u, q | s_0, s_N) \\ p_{s_0}^{s_N}(q) &= \sum_u p_{s_0}^{s_N}(u, q) \\ p_{s_N}^{s_{2N}}(v, r) &= P_{V_i, R_i | S_N, s_{2N}}(v, r | s_N, s_{2N}) \\ p_{s_N}^{s_{2N}}(r) &= \sum_v p_{s_N}^{s_{2N}}(v, r) \\ p_{s_0}^{s_{2N}}(u, v, q, r) &= P_{U_i, V_i, Q_i, R_i | S_0, s_{2N}}(u, v, q, r | s_0, s_{2N}) \\ \bar{p}_{s_0}^{s_{2N}}(t, v, q, r) &= P_{U_i + V_i, V_i, Q_i, R_i | S_0, s_{2N}}(t, v, q, r | s_0, s_{2N}) \\ \bar{p}_{s_0}^{s_{2N}}(t, q, r) &= \sum_v \bar{p}_{s_0}^{s_{2N}}(t, v, q, r) \\ K_{s_0}^{s_N}(U_i | Q_i) &= \sum_q |p_{s_0}^{s_N}(0, q) - p_{s_0}^{s_N}(1, q)| \\ K_{s_N}^{s_{2N}}(V_i | R_i) &= \sum_r |p_{s_N}^{s_{2N}}(0, r) - p_{s_N}^{s_{2N}}(1, r)|. \end{aligned} \quad (37)$$

Denote

$$\pi(s_0, *, s_{2N}) = P_{S_0, s_{2N}}(s_0, s_{2N}).$$

We then have

$$\begin{aligned} \pi(s_0, *, s_{2N}) p_{s_0}^{s_{2N}}(u, v, q, r) &= \pi(s_0) P_{S_{2N} | S_0}(s_{2N} | s_0) p_{s_0}^{s_{2N}}(u, v, q, r) \\ &= \pi(s_0) P_{U_i, V_i, Q_i, R_i, S_{2N} | S_0}(u, v, q, r, s_{2N} | s_0) \\ &= \pi(s_0) \sum_{s_N \in \mathcal{S}(\Delta)} P_{U_i, V_i, Q_i, R_i, S_N, s_{2N} | S_0}(u, v, q, r, s_N, s_{2N} | s_0) \\ &= \pi(s_0) \sum_{s_N \in \mathcal{S}(\Delta)} P_{U_i, Q_i, s_N | S_0}(u, q, s_N | s_0) \\ &\quad \cdot P_{V_i, R_i, s_{2N} | S_N}(v, r, s_{2N} | s_N) \\ &= \pi(s_0) \sum_{s_N \in \mathcal{S}(\Delta)} P_{S_N | S_0}(s_N | s_0) p_{s_0}^{s_N}(u, q) \\ &\quad \cdot P_{S_{2N} | S_N}(s_{2N} | s_N) p_{s_N}^{s_{2N}}(v, r) \\ &= \sum_{s_N \in \mathcal{S}(\Delta)} \lambda(s_N) p_{s_0}^{s_N}(u, q) p_{s_N}^{s_{2N}}(v, r), \end{aligned} \quad (38)$$

where the latter equality is since $\pi(s_0, s_N, s_{2N}) = \pi(s_0)P_{S_N|S_0}(s_N|s_0)P_{S_{2N}|S_N}(s_{2N}|s_N)$.

Denote the set $\mathcal{S}_\star = \{s_0, s_{2N} : (s_0, s_{2N}) \in \mathcal{S}(\Delta)\}$. Then,

$$\hat{K}_n^-(\Delta) = \sum_{\mathcal{S}_\star} \sum_{q,r} \pi(s_0, *, s_{2N}) |\bar{p}_{s_0}^{s_{2N}}(0, q, r) - \bar{p}_{s_0}^{s_{2N}}(1, q, r)|.$$

We first bound the inner sum for some fixed $(s_0, s_{2N}) \in \mathcal{S}_\star$.

$$\begin{aligned} & \sum_{q,r} \pi(s_0, *, s_{2N}) |\bar{p}_{s_0}^{s_{2N}}(0, q, r) - \bar{p}_{s_0}^{s_{2N}}(1, q, r)| \\ &= \sum_{q,r} |\pi(s_0, *, s_{2N}) \bar{p}_{s_0}^{s_{2N}}(0, q, r) - \pi(s_0, *, s_{2N}) \bar{p}_{s_0}^{s_{2N}}(1, q, r)| \\ &\stackrel{(a)}{=} \sum_{q,r} \left| \sum_{s_N \in \mathcal{S}(\Delta)} \lambda(s_N) \sum_v p_{s_N}^{s_{2N}}(v, r) (p_{s_0}^{s_N}(v, q) - p_{s_0}^{s_N}(v+1, q)) \right| \\ &\stackrel{(b)}{\leq} \sum_{q,r, s_N \in \mathcal{S}(\Delta)} \lambda(s_N) \left| \sum_v p_{s_N}^{s_{2N}}(v, r) (p_{s_0}^{s_N}(v, q) - p_{s_0}^{s_N}(v+1, q)) \right| \\ &= \sum_{q,r, s_N \in \mathcal{S}(\Delta)} \lambda(s_N) |p_{s_0}^{s_N}(0, q) - p_{s_0}^{s_N}(1, q)| \cdot |p_{s_N}^{s_{2N}}(0, r) - p_{s_N}^{s_{2N}}(1, r)| \\ &= \sum_{s_N \in \mathcal{S}(\Delta)} \lambda(b) K_{s_0}^{s_N}(U_i|Q_i) K_{s_N}^{s_{2N}}(V_i|R_i) \\ &\stackrel{(c)}{\leq} \mathbf{d} \cdot M(\Delta) \sum_{s_N \in \mathcal{S}(\Delta)} P_{S_0, S_N}(s_0, s_N) K_{s_0}^{s_N}(U_i|Q_i) \\ &\quad \cdot P_{S_N, S_{2N}}(s_N, s_{2N}) K_{s_N}^{s_{2N}}(V_i|R_i) \\ &\stackrel{(d)}{\leq} \mathbf{d} \cdot M(\Delta) \sum_{s_N \in \mathcal{S}(\Delta)} P_{S_0, S_N}(s_0, s_N) K_{s_0}^{s_N}(U_i|Q_i) \\ &\quad \cdot \sum_{s'_N \in \mathcal{S}(\Delta)} P_{S_N, S_{2N}}(s'_N, s_{2N}) K_{s'_N}^{s_{2N}}(V_i|R_i), \end{aligned}$$

where (a) is by (37) and (38), (b) is by the triangle inequality, (c) is by (36), and (d) is by adding nonnegative terms to the sum. Finally, by summing over $\mathcal{S}_\star$ we obtain

$$\hat{K}_n^-(\Delta) = \mathbf{d} \cdot M(\Delta) \cdot (\hat{K}_n(\Delta))^2.$$

Next,

$$\hat{K}_n^+(\Delta) = \sum_{\mathcal{S}_\star} \sum_{t,q,r} \pi(s_0, *, s_{2N}) |\bar{p}_{s_0}^{s_{2N}}(t, 0, q, r) - \bar{p}_{s_0}^{s_{2N}}(t, 1, q, r)|.$$

Similar to the above, we first bound the inner sum for some fixed $(s_0, s_{2N}) \in \mathcal{S}_\star$,

$$\begin{aligned} & \sum_{t,q,r} \pi(s_0, *, s_{2N}) |\bar{p}_{s_0}^{s_{2N}}(t, 0, q, r) - \bar{p}_{s_0}^{s_{2N}}(t, 1, q, r)| \\ &= \sum_{t,q,r} |P_{S_0, S_{2N}}(s_0, s_{2N}) \bar{p}_{s_0}^{s_{2N}}(t, 0, q, r) \\ &\quad - P_{S_0, S_{2N}}(s_0, s_{2N}) \bar{p}_{s_0}^{s_{2N}}(t, 1, q, r)| \\ &\stackrel{(a)}{=} \sum_{t,q,r} \left| \sum_{s_N \in \mathcal{S}(\Delta)} \lambda(s_N) \left(p_{s_0}^{s_N}(t, q) p_{s_N}^{s_{2N}}(0, r) \right. \right. \\ &\quad \left. \left. - p_{s_0}^{s_N}(t+1, q) p_{s_N}^{s_{2N}}(1, r) \right) \right| \\ &\stackrel{(b)}{=} \frac{1}{2} \sum_{t,q,r} \left| \sum_{s_N \in \mathcal{S}(\Delta)} \lambda(s_N) \left(p_{s_0}^{s_N}(q) (p_{s_N}^{s_{2N}}(0, r) - p_{s_N}^{s_{2N}}(1, r)) \right) \right. \\ &\quad \left. + \sum_{s_N \in \mathcal{S}(\Delta)} \lambda(s_N) \left(p_{s_N}^{s_{2N}}(r) (p_{s_0}^{s_N}(0, q) - p_{s_0}^{s_N}(1, q)) \right) \right| \end{aligned}$$

$$\begin{aligned} & \stackrel{(c)}{\leq} \sum_{q, s_N \in \mathcal{S}(\Delta)} \lambda(s_N) p_{s_0}^{s_N}(q) \left(\sum_r |p_{s_N}^{s_{2N}}(0, r) - p_{s_N}^{s_{2N}}(1, r)| \right) \\ &\quad + \sum_{r, s_N \in \mathcal{S}(\Delta)} \lambda(s_N) p_{s_0}^{s_N}(r) \left(\sum_q |p_{s_0}^{s_N}(0, q) - p_{s_0}^{s_N}(1, q)| \right) \\ &= \sum_{s_N \in \mathcal{S}(\Delta)} \lambda(s_N) K_{s_N}^{s_{2N}}(V_i|R_i) + \sum_{s_N \in \mathcal{S}(\Delta)} \lambda(s_N) K_{s_0}^{s_N}(U_i|Q_i), \end{aligned}$$

where (a) is by (38), (b) is by the following algebraic identity that holds for any four numbers a, b, c, d ,

$$ab - cd = \frac{(a+c)(b-d) + (b+d)(a-c)}{2},$$

and (c) is by the triangle inequality and since t is binary.

It remains to sum over $\mathcal{S}_\star$. Recalling (36), stationarity implies that

$$\begin{aligned} \sum_{\mathcal{S}_N^3(\Delta)} \lambda(s_N) K_{s_N}^{s_{2N}}(V_i|R_i) &= \sum_{s_N, s_{2N} \in \mathcal{S}_N^2(\Delta)} P_{S_N, S_{2N}}(s_N, s_{2N}) K_{s_N}^{s_{2N}}(V_i|R_i) \\ &= \hat{K}_n(\Delta) \\ \sum_{\mathcal{S}_N^3(\Delta)} \lambda(s_N) K_{s_0}^{s_N}(U_i|Q_i) &= \sum_{s_0, s_N \in \mathcal{S}_N^2(\Delta)} P_{S_0, S_N}(s_0, s_N) K_{s_0}^{s_N}(U_i|Q_i) \\ &= \hat{K}_n(\Delta). \end{aligned}$$

Thus, we obtain

$$\hat{K}_n^+(\Delta) \leq 2\hat{K}_n(\Delta).$$

This completes the proof. $\blacksquare$

The following proposition establishes fast polarization. Note that this is not yet Theorem 1, since that requires conditioning on a finer event than $D(\Delta)$.

Proposition 14. Fix $0 < \beta < 1/2$ and $0 \leq \Delta < \mathbf{d}$. Then, for $N = 2^n$,

$$\lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : Z(U_i|Q_i, D(\Delta)) < 2^{-N^\beta} \right\} \right| = 1 - \mathcal{H}_{X|Y}, \quad (39)$$

$$\lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : K(U_i|Q_i, S_0, S_N, D(\Delta)) < 2^{-N^\beta} \right\} \right| = \mathcal{H}_{X|Y}. \quad (40)$$

Proof: We prove fast polarization using [21, Proposition 49]. We first concentrate on (39). By Lemma 12, the process $Z_n(\Delta)$ satisfies [21, Equation 167], with $\kappa = 2M(\Delta)$. This also holds if we offset our process to start at some n' instead of 0. As a shorthand, we write Z_n in place of $Z_n(\Delta)$ in the sequel.

Fix $\delta > 0$. Thus, [21, Proposition 49] implies that there exist $\eta > 0$ and n_0 such that for $\beta' = (\beta + 1/2)/2$,

$$\mathbb{P} \left(Z_{n+n'} \leq 2^{-2n\beta'} \text{ for all } n \geq n_0 \mid Z_{n'} \leq \eta \right) \geq 1 - \delta/2.$$

We now show that we can choose n' such that

$$\mathbb{P}(Z_{n'} \leq \eta) > 1 - \mathcal{H}_{X|Y} - \delta/2. \quad (41)$$

This will imply that

$$\mathbb{P} \left(Z_{n+n'} \leq 2^{-2n\beta'} \text{ for all } n \geq n_0, Z_{n'} \leq \eta \right) > 1 - \mathcal{H}_{X|Y} - \delta.$$

Hence,

$$\mathbb{P} \left(Z_{n+n'} \leq 2^{-2n\beta'} \text{ for all } n \geq n_0 \right) > 1 - \mathcal{H}_{X|Y} - \delta.$$

This implies that

$$\lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : Z(U_i|Q_i, D(\Delta)) < 2^{-N^\beta} \right\} \right| \geq 1 - \mathcal{H}_{X|Y} \quad (42)$$

since $\beta' > \beta$ and $\delta > 0$ is arbitrary. All that remains to prove (39), is to show that the above “ $\geq$ ” is in fact an equality. We will come back to this point later.

To prove (41), we call upon Proposition 10 with $\epsilon = \sqrt{\eta}$. By, (28a), (28e), and (1),

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_n(\Delta) < \sqrt{\eta}) = 1 - \mathcal{H}_{X|Y}.$$

Thus, there exists n' such that

$$\mathbb{P}(H_{n'}(\Delta) < \sqrt{\eta}) = 1 - \mathcal{H}_{X|Y} - \delta/2.$$

By [5, Lemma 1], $H_{n'}(\Delta) < \sqrt{\eta}$ implies $Z_{n'} < \eta$, which establishes (41).

We now concentrate on (40). By Lemma 13, the process $\hat{K}_n(\Delta)$ satisfies [21, Equation 167], with $\kappa = \max\{\mathbf{d}M(\Delta), 2\}$. As before, this also holds if we offset our process to start at some n' instead of 0. As a shorthand, we write $\hat{K}_n$ in place of $\hat{K}_n(\Delta)$ in the sequel.

Fix $\delta > 0$. Thus, [21, Proposition 49] implies that there exist $\eta > 0$ and n_0 such that for $\beta' = (\beta + 1/2)/2$,

$$\mathbb{P}\left(\hat{K}_{n+n'} \leq 2^{-2^{n\beta'}} \text{ for all } n \geq n_0 \mid \hat{K}_{n'} \leq \eta\right) \geq 1 - \delta/2.$$

We now show that we can choose n' such that

$$\mathbb{P}(\hat{K}_{n'} \leq \eta) > \mathcal{H}_{X|Y} - \delta/2. \quad (43)$$

This will imply that

$$\mathbb{P}\left(\hat{K}_{n+n'} \leq 2^{-2^{n\beta'}} \text{ for all } n \geq n_0, \hat{K}_{n'} \leq \eta\right) > \mathcal{H}_{X|Y} - \delta.$$

Hence,

$$\mathbb{P}\left(\hat{K}_{n+n'} \leq 2^{-2^{n\beta'}} \text{ for all } n \geq n_0\right) > \mathcal{H}_{X|Y} - \delta.$$

This implies that

$$\lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : \hat{K}(U_i|Q_i, S_0, S_N, D(\Delta)) < 2^{-N\beta'} \right\} \right| \geq \mathcal{H}_{X|Y} \quad (44)$$

since $\beta' > \beta$ and $\delta > 0$ is arbitrary. All that remains to prove (40), is to show that the above “ $\geq$ ” is in fact an equality. We will come back to this point later.

To prove (43), we call upon Proposition 10 with $\epsilon = \sqrt{1 - \eta^2}$. By, (29b), (29e), and (1),

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_n(\Delta) > \sqrt{1 - \eta^2}) = \mathcal{H}_{X|Y}.$$

Thus, there exists n' such that

$$\mathbb{P}(H_{n'}(\Delta) > \sqrt{1 - \eta^2}) = \mathcal{H}_{X|Y} - \delta/2.$$

By [5, Lemma 1], $H_{n'}(\Delta) > \sqrt{1 - \eta^2}$ implies $\hat{K}_{n'} < \eta$, which establishes (43).

To complete the proof, we now show that the weak inequalities in (42) and (44) are in fact equalities. To see this, first note by [5, Lemma 2] that $Z(U_i|Q_i, S_0, S_N, D(\Delta)) \leq Z(U_i|Q_i, D(\Delta))$. Hence, (42) implies that

$$\lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : Z(U_i|Q_i, S_0, S_N, D(\Delta)) < 2^{-N\beta} \right\} \right| \geq 1 - \mathcal{H}_{X|Y}, \quad (45)$$

and a strict inequality in (42) implies a strict inequality in (45). Now, assume to the contrary a strict inequality in either (42) or (44), or both. This implies a strict inequality in either (45)

or (44), or both. Thus, for all N large enough, there must exist an index $0 \leq i < N$, such that

$$\begin{aligned} Z(U_i|Q_i, S_0, S_N, D(\Delta)) &< 2^{-N\beta}, \\ K(U_i|Q_i, S_0, S_N, D(\Delta)) &< 2^{-N\beta}. \end{aligned}$$

This cannot hold, since $K + Z \geq 1$, by [5, Lemma 1]. ■

V. PROOF OF THEOREM 1

The following technical lemma will be used in our proof of Theorem 1. Its proof is given in Appendix A.

Lemma 15. *Let U be a binary random variable; W, Q general random variables; and $A_1, A_2, \dots, A_\ell$ mutually exclusive events. Denote $C = A_1 \cup A_2 \cup \dots \cup A_\ell$. Then,*

$$Z(U|Q, C) \geq \sum_{j=1}^{\ell} \mathbb{P}(A_j|C) Z(U|Q, A_j), \quad (46)$$

$$K(U|Q, C) \leq \sum_{j=1}^{\ell} \mathbb{P}(A_j|C) K(U|Q, A_j), \quad (47)$$

$$H(W|Q, C) \geq \sum_{j=1}^{\ell} \mathbb{P}(A_j|C) H(W|Q, A_j). \quad (48)$$

Proof of Theorem 1: Recall the factorization (3), $\mathbf{p} = \mathbf{d} \cdot \mathbf{q}$, where $\mathbf{q}$ is odd and $\mathbf{d}$ is a power of 2. Recall that φ is the common phase of all the vertices in Ψ_0 and denote $\Delta = \varphi \bmod \mathbf{d}$. Next, note that by (6) that for $N = 2^n \geq \mathbf{d}$ we have for all $(s_0, s_N) \in \Psi_0 \times \Psi_N$ that $(s_0, s_N) \in \mathcal{S}_N^2(\Delta)$.

Proposition 14 with $\beta' = (\beta + 1/2)/2$ in place of β implies that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : Z(U_i|U_1^{i-1}, Y_1^N, D(\Delta)) < 2^{-N\beta} \right\} \right| &= 1 - \mathcal{H}_{X|Y}, \\ \lim_{n \rightarrow \infty} \frac{1}{N} \left| \left\{ i : K(U_i|U_1^{i-1}, S_0, S_N, D(\Delta)) < 2^{-N\beta} \right\} \right| &= \mathcal{H}_X. \end{aligned}$$

Note that the second equality is obtained by considering the case in which the Y are constant, and hence nothing is gained by conditioning on them. The proof of Theorem 1 will follow by showing that there is some $\xi > 0$ independent of N and $0 \leq i < N$ such that

$$Z(U_i|U_1^{i-1}, Y_1^N, A) \leq \xi \cdot Z(U_i|U_1^{i-1}, Y_1^N, D(\Delta)) \quad (49)$$

and

$$K(U_i|U_1^{i-1}, A) \leq \xi \cdot K(U_i|U_1^{i-1}, S_0, S_N, D(\Delta)), \quad (50)$$

for $N = 2^n$ large enough. Indeed, we claim that $\xi = \frac{1}{\mu \cdot \mathbf{d}}$ suffices for $N \geq N_0$, where μ and N_0 are as promised by Lemma 4.

For compactness, denote $\Psi_{0,N} \triangleq \Psi_0 \times \Psi_N$. To show (49) we apply Lemma 15 with

$$\begin{aligned} C &= \{(S_0, S_N) \in \mathcal{S}_N^2(\Delta)\}, \\ A_1 &= \{(S_0, S_N) \in \Psi_{0,N}\}, \\ A_2 &= \{(S_0, S_N) \in \mathcal{S}_N^2(\Delta) \setminus \Psi_{0,N}\}. \end{aligned}$$

Note that by definition C and A_1 are simply $D(\Delta)$ and A , respectively. Thus,

$$\begin{aligned}
& Z(U_i|U_1^{i-1}, Y_1^N, D(\Delta)) \\
& \stackrel{(a)}{\geq} \mathbb{P}(A_1|D(\Delta)) Z(U_i|U_1^{i-1}, Y_1^N, A_1) \\
& \quad + \mathbb{P}(A_2|D(\Delta)) Z(U_i|U_1^{i-1}, Y_1^N, A_2) \\
& \stackrel{(b)}{\geq} \mathbb{P}(A|D(\Delta)) Z(U_i|U_1^{i-1}, Y_1^N, A) \\
& = \mathbb{P}((S_0, S_N) \in \Psi_{0,N}|D(\Delta)) \cdot Z(U_i|U_1^{i-1}, Y_1^N, A) \\
& = \frac{\mathbb{P}((S_0, S_N) \in \Psi_{0,N})}{\mathbb{P}(D(\Delta))} Z(U_i|U_1^{i-1}, Y_1^N, A) \\
& \stackrel{(c)}{\geq} \mu \cdot \mathbf{d} \cdot Z(U_i|U_1^{i-1}, Y_1^N, A) \\
& = \xi^{-1} \cdot Z(U_i|U_1^{i-1}, Y_1^N, A)
\end{aligned}$$

where (a) is by (46), (b) is since $A_1 = A$ and probabilities are nonnegative, and (c) is by (5) and Lemma 4. This yields (49).

To show (50), we apply Lemma 15 with

$$\begin{aligned}
C &= \{(S_0, S_N) \in \Psi_{0,N}\}, \\
A_{s_0, s_N} &= \{(S_0, S_N) = (s_0, s_N)\},
\end{aligned}$$

where in the above (s_0, s_N) range over $\Psi_{0,N}$. Note that in this case, $C = A$, and clearly $C = \cup_{(s_0, s_N) \in \Psi_{0,N}} A_{s_0, s_N}$. Thus,

$$\begin{aligned}
& K(U_i|U_1^{i-1}, A) \\
& = K(U_i|U_1^{i-1}, C) \\
& \stackrel{(a)}{\leq} \sum_{s_0, s_N} \mathbb{P}(A_{s_0, s_N}|C) K(U_i|U_1^{i-1}, S_0=s_0, S_N=s_N) \\
& \stackrel{(b)}{=} \sum_{s_0, s_N} \frac{\mathbb{P}(A_{s_0, s_N})}{\mathbb{P}(C)} K(U_i|U_1^{i-1}, S_0=s_0, S_N=s_N) \\
& \stackrel{(c)}{=} \sum_{s_0, s_N} \frac{\mathbb{P}(A_{s_0, s_N}, D(\Delta))}{\mathbb{P}(C)} K(U_i|U_1^{i-1}, S_0=s_0, S_N=s_N) \\
& = \sum_{s_0, s_N} \frac{\mathbb{P}(A_{s_0, s_N}|D(\Delta))}{\mathbb{P}(D(\Delta))^{-1} \mathbb{P}(C)} K(U_i|U_1^{i-1}, S_0=s_0, S_N=s_N) \\
& \stackrel{(d)}{\leq} \frac{1}{\mu \cdot \mathbf{d}} \sum_{s_0, s_N} \mathbb{P}(A_{s_0, s_N}|D(\Delta)) K(U_i|U_1^{i-1}, S_0=s_0, S_N=s_N), \\
& = \xi \cdot K(U_i|U_1^{i-1}, S_0, S_N, D(\Delta)),
\end{aligned}$$

where (a) is by (47), (b) is since $A_{s_0, s_N} \subseteq C$, (c) is since $A_{s_0, s_N} \subseteq D(\Delta)$, and (d) is by (5) and Lemma 4. This yields (50), completing the proof. $\blacksquare$

APPENDIX A AUXILIARY PROOFS

Proof of Lemma 4: By the Markov property, we must show that there exist an N_0 and a μ such that for all $N \geq N_0$ and all s_0, s_N, s_{2N} satisfying (12),

$$\begin{aligned}
& \pi(s_0, s_N, s_{2N}) = \\
& \pi(s_0) \mathbb{P}(S_N = s_N | S_0 = s_0) \mathbb{P}(S_{2N} = s_{2N} | S_N = s_N) > \mu.
\end{aligned}$$

By irreducibility and stationarity, we have for all s_0 that

$$\pi(s_0) \geq \pi_{\min} = \min_{s \in \mathcal{S}} \pi(s) > 0.$$

Hence, our claim will follow by showing that there exists N_0 and $\mu' > 0$ such that for all $N \geq N_0$ and all s_0, s_N satisfying $\phi(s_N) - \phi(s_0) \equiv N \pmod{\mathfrak{p}}$,

$$\mathbb{P}(S_N = s_N | S_0 = s_0) \geq \mu'.$$

Namely, we can take $\mu = \pi_{\min} \cdot (\mu')^2/2$.

By [22, Theorem 1.3], there exists N_0 such that for all $\sigma, \omega \in \mathcal{S}$ such that $\phi(\omega) - \phi(\sigma) \equiv N_0 \pmod{\mathfrak{p}}$, we have $\mathbb{P}(S_{N_0} = \omega | S_0 = \sigma) > 0$. We take μ' as the minimum of this probability over all σ and ω as above. Now, for $N \geq N_0$ define $\mathcal{S}' = \{\sigma \in \mathcal{S} : \phi(\sigma) - \phi(s_0) \equiv N - N_0 \pmod{\mathfrak{p}}\}$. We have

$$\begin{aligned}
& \mathbb{P}(S_N = s_N | S_0 = s_0) \\
& = \sum_{\sigma \in \mathcal{S}'} \mathbb{P}(S_{N-N_0} = \sigma | S_0 = s_0) \mathbb{P}(S_N = s_N | S_{N-N_0} = \sigma) \\
& \geq \sum_{\sigma \in \mathcal{S}'} \mathbb{P}(S_{N-N_0} = \sigma | S_0 = s_0) \mu' \\
& = \mu'.
\end{aligned}$$

The equalities follow by periodicity of the Markov chain and the definitions of $\mathcal{S}'$ and ϕ ; the inequality follows since $\phi(s_N) - \phi(\sigma) \equiv N_0 \pmod{\mathfrak{p}}$ for any $\sigma \in \mathcal{S}'$, by our definition of $\mathcal{S}'$, and our assumption that $\phi(s_N) - \phi(s_0) \equiv N \pmod{\mathfrak{p}}$. $\blacksquare$

Proof of Lemma 11: For $(x_1^N, y_1^N), (x_{N+1}^{2N}, y_{N+1}^{2N}) \in \mathcal{X}^N \times \mathcal{Y}^N$, denote by $a(x_1^N, y_1^N)$ and $b(x_{N+1}^{2N}, y_{N+1}^{2N})$ the following events:

$$\begin{aligned}
a(x_1^N, y_1^N) &= \{(X_1^N, Y_1^N) = (x_1^N, y_1^N)\}, \\
b(x_{N+1}^{2N}, y_{N+1}^{2N}) &= \{(X_{N+1}^{2N}, Y_{N+1}^{2N}) = (x_{N+1}^{2N}, y_{N+1}^{2N})\}.
\end{aligned}$$

We start by noting the following: it suffices to prove that for all $(x_1^N, y_1^N), (x_{N+1}^{2N}, y_{N+1}^{2N}) \in \mathcal{X}^N \times \mathcal{Y}^N$ we have

$$\begin{aligned}
& \mathbb{P}\left(a(x_1^N, y_1^N), b(x_{N+1}^{2N}, y_{N+1}^{2N}) \middle| D(\Delta)\right) \leq M(\Delta) \\
& \cdot \mathbb{P}\left(a(x_1^N, y_1^N) \middle| D(\Delta)\right) \cdot \mathbb{P}\left(b(x_{N+1}^{2N}, y_{N+1}^{2N}) \middle| D(\Delta)\right). \quad (51)
\end{aligned}$$

Indeed, this would imply (33), since

$$\begin{aligned}
& \mathbb{P}\left((X_1^N, Y_1^N) \in A \text{ and } (X_{N+1}^{2N}, Y_{N+1}^{2N}) \in B \middle| D(\Delta)\right) \\
& = \sum_{(x_1^N, y_1^N) \in A} \sum_{(x_{N+1}^{2N}, y_{N+1}^{2N}) \in B} \mathbb{P}\left(a(x_1^N, y_1^N), b(x_{N+1}^{2N}, y_{N+1}^{2N}) \middle| D(\Delta)\right) \\
& \leq M(\Delta) \sum_{(x_1^N, y_1^N) \in A} \mathbb{P}\left(a(x_1^N, y_1^N) \middle| D(\Delta)\right) \\
& \quad \cdot \sum_{(x_{N+1}^{2N}, y_{N+1}^{2N}) \in B} \mathbb{P}\left(b(x_{N+1}^{2N}, y_{N+1}^{2N}) \middle| D(\Delta)\right) \\
& = M(\Delta) \cdot \mathbb{P}\left((X_1^N, Y_1^N) \in A \middle| D(\Delta)\right) \\
& \quad \cdot \mathbb{P}\left((X_{N+1}^{2N}, Y_{N+1}^{2N}) \in B \middle| D(\Delta)\right).
\end{aligned}$$

To prove (51), we start with its LHS and note that

$$\begin{aligned}
& \mathbb{P}\left(a(x_1^N, y_1^N), b(x_{N+1}^{2N}, y_{N+1}^{2N}) \middle| D(\Delta)\right) \\
& = \sum_{s_N \in \mathcal{S}(\Delta)} \mathbb{P}\left(a(x_1^N, y_1^N), S_N = s_N, b(x_{N+1}^{2N}, y_{N+1}^{2N}) \middle| D(\Delta)\right) \\
& \stackrel{(a)}{=} \sum_{s_N \in \mathcal{S}(\Delta)} \mathbb{P}(S_N = s_N | D(\Delta)) \cdot \mathbb{P}\left(a(x_1^N, y_1^N) \middle| S_N = s_N, D(\Delta)\right) \\
& \quad \cdot \mathbb{P}\left(b(x_{N+1}^{2N}, y_{N+1}^{2N}) \middle| S_N = s_N, D(\Delta)\right)
\end{aligned}$$

$$\begin{aligned}
& \stackrel{(b)}{=} \sum_{s_N \in \mathcal{S}(\Delta)} \mathbb{P}(S_N = s_N | D(\Delta)) \cdot \frac{\mathbb{P}(a(x_1^N, y_1^N), S_N = s_N | D(\Delta))}{\mathbb{P}(S_N = s_N | D(\Delta))} \\
& \quad \cdot \frac{\mathbb{P}(b(x_{N+1}^{2N}, y_{N+1}^{2N}), S_N = s_N | D(\Delta))}{\mathbb{P}(S_N = s_N | D(\Delta))} \\
& = \sum_{s_N \in \mathcal{S}(\Delta)} \frac{1}{\mathbb{P}(S_N = s_N | D(\Delta))} \cdot \mathbb{P}(a(x_1^N, y_1^N), S_N = s_N | D(\Delta)) \\
& \quad \cdot \mathbb{P}(b(x_{N+1}^{2N}, y_{N+1}^{2N}), S_N = s_N | D(\Delta)) \\
& \stackrel{(c)}{=} \sum_{s_N \in \mathcal{S}(\Delta)} \frac{1/d}{\pi_{s_N}} \cdot \mathbb{P}(a(x_1^N, y_1^N), S_N = s_N | D(\Delta)) \\
& \quad \cdot \mathbb{P}(b(x_{N+1}^{2N}, y_{N+1}^{2N}), S_N = s_N | D(\Delta)) \\
& \stackrel{(d)}{\leq} M(\Delta) \sum_{s_N \in \mathcal{S}(\Delta)} \cdot \mathbb{P}(a(x_1^N, y_1^N), S_N = s_N | D(\Delta)) \\
& \quad \cdot \mathbb{P}(b(x_{N+1}^{2N}, y_{N+1}^{2N}), S_N = s_N | D(\Delta)) \\
& \stackrel{(e)}{\leq} M(\Delta) \sum_{s_N \in \mathcal{S}(\Delta)} \sum_{s'_N \in \mathcal{S}(\Delta)} \mathbb{P}(a(x_1^N, y_1^N), S_N = s_N | D(\Delta)) \\
& \quad \cdot \mathbb{P}(b(x_{N+1}^{2N}, y_{N+1}^{2N}), S_N = s'_N | D(\Delta)) \\
& = M(\Delta) \sum_{s_N \in \mathcal{S}(\Delta)} \mathbb{P}(a(x_1^N, y_1^N), S_N = s_N | D(\Delta)) \\
& \quad \cdot \sum_{s'_N \in \mathcal{S}(\Delta)} \mathbb{P}(b(x_{N+1}^{2N}, y_{N+1}^{2N}), S_N = s'_N | D(\Delta)) \\
& = M(\Delta) \cdot \mathbb{P}(a(x_1^N, y_1^N) | D(\Delta)) \cdot \mathbb{P}(b(x_{N+1}^{2N}, y_{N+1}^{2N}) | D(\Delta)),
\end{aligned}$$

where (a) is by the Markov property; (b) is valid since the denominator is positive by irreducibility and stationarity; (c) is since for $s_N \in \mathcal{S}_N$, we have

$$\frac{1}{\mathbb{P}(S_N = s_N | D(\Delta))} = \frac{\mathbb{P}(D(\Delta))}{\mathbb{P}(S_N = s_N, D(\Delta))} = \frac{1/d}{\pi_{s_N}},$$

where the latter equality holds since $D(\Delta) \subseteq \{S_N = s_N\}$ and by (5); (d) is by (32); and (e) is since the sum over s'_N adds nonnegative terms to the RHS. This completes the proof. ■

Proof of Lemma 15: Since the events $(A_j)_j$ are mutually exclusive and their union is C , we have $\mathbb{P}(A_j | C) = \mathbb{P}(A_j) / \mathbb{P}(C)$. To keep the notation light, for a general event A , we denote $\mathbb{P}(U = u, Q = q | A) = p(u, q | A)$ and $\mathbb{P}(U = u, Q = q, A) = p(u, q, A)$. Observe that $p(u, q | A) = p(u, q, A) / \mathbb{P}(A)$.

Recall that by the Cauchy-Schwarz inequality, for sequences $(a_j)_j$ and $(b_j)_j$ of the same length we have

$$\sqrt{\left(\sum_j a_j^2\right)\left(\sum_{j'} b_{j'}^2\right)} \geq \sum_j a_j b_j.$$

Using the above inequality with $a_j = \sqrt{p(0, q, A_j)}$ and $b_j = \sqrt{p(1, q, A_j)}$, we obtain

$$\begin{aligned}
Z(U|Q, C) &= 2 \sum_q \sqrt{p(0, q|C)p(1, q|C)} \\
&= \frac{2}{\mathbb{P}(C)} \sum_q \sqrt{p(0, q, C)p(1, q, C)}
\end{aligned}$$

$$\begin{aligned}
&= \frac{2}{\mathbb{P}(C)} \sum_q \sqrt{\left(\sum_{j=1}^{\ell} p(0, q, A_j)\right)\left(\sum_{j'=1}^{\ell} p(1, q, A_{j'})\right)} \\
&\geq \frac{2}{\mathbb{P}(C)} \sum_q \sum_{j=1}^{\ell} \sqrt{p(0, q, A_j) \cdot p(1, q, A_j)} \\
&= 2 \sum_{j=1}^{\ell} \frac{\mathbb{P}(A_j)}{\mathbb{P}(C)} \sum_q \sqrt{p(0, q|A_j) \cdot p(1, q|A_j)} \\
&= \sum_{j=1}^{\ell} \mathbb{P}(A_j | C) Z(U|Q, A_j).
\end{aligned}$$

This yields (46).

To prove (47), we have with the above notation and using the triangle inequality,

$$\begin{aligned}
K(U|Q, C) &= \sum_q |p(0, q|C) - p(1, q|C)| \\
&= \frac{1}{\mathbb{P}(C)} \sum_q |p(0, q, C) - p(1, q, C)| \\
&= \frac{1}{\mathbb{P}(C)} \sum_q \left| \sum_{j=1}^{\ell} (p(0, q, A_j) - p(1, q, A_j)) \right| \\
&\leq \frac{1}{\mathbb{P}(C)} \sum_q \sum_{j=1}^{\ell} |p(0, q, A_j) - p(1, q, A_j)| \\
&= \sum_{j=1}^{\ell} \frac{1}{\mathbb{P}(C)} \sum_q |p(0, q, A_j) - p(1, q, A_j)| \\
&= \sum_{j=1}^{\ell} \frac{\mathbb{P}(A_j)}{\mathbb{P}(C)} \sum_q |p(0, q|A_j) - p(1, q|A_j)| \\
&= \sum_{j=1}^{\ell} \mathbb{P}(A_j | C) K(U|Q, A_j).
\end{aligned}$$

To prove (48), we define an indicator random variable J that under C equals j if we are under event A_j . Since conditioning reduces entropy,

$$\begin{aligned}
H(W|Q, C) &\geq H(W|Q, C, J) = \sum_{j=1}^{\ell} \mathbb{P}(A_j | C) H(W|Q, C, J = j) \\
&= \sum_{j=1}^{\ell} \mathbb{P}(A_j | C) H(W|Q, A_j).
\end{aligned}$$

This completes the proof. ■

REFERENCES

- [1] E. Arkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.
- [2] —, "Source polarization," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2010)*, Austin, Texas, 2010, pp. 899–903.
- [3] E. Şaşıoğlu, E. Telatar, and E. Arkan, "Polarization for arbitrary discrete memoryless channels," *arXiv:0908.0302v1*, 2009.
- [4] E. Şaşıoğlu and I. Tal, "Polar coding for processes with memory," *IEEE Trans. Inform. Theory*, vol. 65, no. 4, pp. 1994–2003, April 2019.
- [5] B. Shuval and I. Tal, "Fast polarization for processes with memory," *IEEE Trans. Inform. Theory*, vol. 65, no. 4, pp. 2004–2020, April 2019.
- [6] I. Tal, H. D. Pfister, A. Fazeli, and A. Vardy, "Polar codes for the deletion channel: weak and strong polarization," *IEEE Trans. Inform. Theory*, vol. 68, no. 4, pp. 2239–2265, April 2022.
- [7] H. D. Pfister and I. Tal, "Polar codes for channels with insertions, deletions, and substitutions," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2021)*, Melbourne, Victoria, Australia, 2021, pp. 2554–2559.

- [8] D. Arava and I. Tal, "Stronger polarization in the deletion channel," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2023)*, Taipei, Taiwan, 2023, pp. 1711–1716.
- [9] H. Mahdaviyar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inform. Theory*, vol. 57, pp. 6428–6443, 2011.
- [10] H. Mahdaviyar, "Polar coding for non-stationary channels," *IEEE Trans. Inform. Theory*, vol. 66, pp. 6920–6938, 2020.
- [11] E. Hof and S. Shamai, "Secrecy-achieving polar-coding for binary-input memoryless symmetric wire-tap channels," *arXiv:1005.2759v2*, 2010.
- [12] E. Hof, I. Sason, S. Shamai (Shitz), and C. Tian, "Capacity-achieving polar codes for arbitrarily permuted parallel channels," *IEEE Trans. Inform. Theory*, vol. 59, March 2013.
- [13] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.
- [14] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric channels," *IEEE Trans. Inform. Theory*, vol. 59, no. 12, pp. 7829–7838, December 2012.
- [15] R. Wang, J. Honda, H. Yamamoto, R. Liu, and Y. Hou, "Construction of polar codes for channels with memory," in *Proc. IEEE Inform. Theory Workshop (ITW'2015)*, Jeju Island, Korea, 2015, pp. 187–191.
- [16] J. Honda and H. Yamamoto, "Variable-to-fixed length homophonic coding with a modified shannon-fano-elias code," *arXiv:1607.06914v1*, 2016.
- [17] E. Şaşıoğlu, "Polarization and polar codes," in *Found. and Trends in Commun. and Inform. Theory*, vol. 8, no. 4, 2012, pp. 259–381.
- [18] J. S. Rosenthal, *A first look at rigorous probability theory*. New Jersey, London, Singapore, Beijing, Shanghai, Hong Kong, Taipei, Chennai: World Scientific, 2006.
- [19] H. Paley and P. Weichsel, *A first course in abstract algebra*. Holt, Rinehart and Winston, Inc., 1966.
- [20] E. Şaşıoğlu, "Polar coding theorems for discrete systems," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, 2011.
- [21] B. Shuval and I. Tal, "Universal polarization for processes with memory," *Accepted for publication in IEEE Trans. Inform. Theory*, 2025, available online at <https://ieeexplore.ieee.org/document/10836796>.
- [22] E. Seneta, *Non-negative matrices and Markov chains*, 2nd ed. New York, NY: Springer, 1981.