Name:

Points: /75

Time: 2h30

Exam ACCQ202

Exercise 1 (Question, 4pt)**.** *The minimum expected number of yes/no questions needed to identify random variable X is* 4.004*. Show that the number of possible realizations of X is larger than 16.*

Exercise 2 (Universal Compression, 15pt)**.** *Let $X^n = (X_1, \ldots, X_n)$ be a sequence of i.i.d.* Bernoulli(p) *random variables, where $p \in [0, 1]$ is unknown. Consider the following compression scheme.*

First, encode the number $k = k(x^n)$ of ones in x^n. Next, list all binary sequences of length n containing exactly k ones in lexicographic order, and encode the position of x^n in this list. The decoder can recover x^n exactly from these two descriptions.

1. *(4pt) Show that the total description length satisfies*

$$\ell(x^n) \leq \lceil \log_2(n+1) \rceil + \left\lceil \log_2 \binom{n}{k} \right\rceil .$$

2. *(2pt) Let $q = k/n$. Show that*

$$\binom{n}{k} q^k (1-q)^{n-k} \leq 1.$$

3. *(3pt) Deduce that*

$$\log_2 \binom{n}{k} \leq n H_B(k/n).$$

Hence, $\ell(x^n) \leq n H_B(k/n) + \lceil \log_2(n+1) \rceil + 1$.

4. *(3pt) Using a well-known inequality, deduce that*

$$\mathbb{E}[\ell(X^n)] \leq n H_B(p) + \lceil \log_2(n+1) \rceil + 1.$$

5. *(3pt) Conclude that, for every* $p \in [0,1]$,

$$\lim_{n\to\infty} \frac{\mathbb{E}[\ell(X^n)]}{n} = H_B(p).$$

Exercise 3 (Zero-error communication, 6pt)**.** *Find an example of a channel* $Q(y|x)$ *with* k *inputs,* $m > k$ *outputs such that it is possible to communicate at capacity with zero-error. This channel should be such that for any output* y *we have* $Q(y|x) > 0$ *for at least one input* x*.*

Exercise 4 (Channels with Memory Have Higher Capacity. 12pt)**.** *Consider a binary symmetric channel with $Y_i = X_i \oplus Z_i$, where $\oplus$ is mod 2 addition, and $X_i, Y_i \in \{0, 1\}$. Here, the X_i's, Y_i's, and Z_i's denote the channel inputs, channel outputs, and channel noises.*

Throughout this exercise, we assume that all $\{Z_i\}$ have constant marginal probabilities

$$\Pr\{Z_i = 1\} = p = 1 - \Pr\{Z_i = 0\}, \tag{1}$$

but that $Z_1, Z_2, \ldots, Z_n$ are not necessarily independent. Assume that $Z^n = Z_1, \ldots, Z_n$ is independent of the input X^n.

- *(3pt). Show that*

$$I(X^n; Y^n) = H(Y_1, Y_2, \ldots, Y_n) - H(Z_1, Z_2, \ldots, Z_n).$$

- *(5pt.) Show that for any Z^n that satisfies (1) we have*

$$\max_{p(x^n)} \frac{I(X^n; Y^n)}{n} \geq C(p),$$

with equality if and only if the Z_i's are independent—here $C(p)$ denotes the capacity of the BSC with cross-over parameter p.

- *(4pt). Show that for any Z^n that satisfies* (1) *we have*

$$\max_{p(x^n)} I(X^n; Y^n) \leq n - H_B(p),$$

and specify a Z^n that satisfies (1) *for which we have equality.*

Exercise 5 (Shift channel, 32pt)**.** *Consider the following channel. Let X^n be an i.i.d. Bernoulli$(1/2)$ sequence. Let T be a random variable independent of X^n that is uniformly distributed over $\{0, 1, 2, \ldots, n-1\}$. Let $\tilde{X}^n$ be a T-right-circular shift of X^n (for example if $X^5 = (11001)$ and $T = 3$, then $\tilde{X}^5 = 00111$).*

Alice observes only X^n and Bob observes only $\tilde{X}^n$. Alice sends Bob a message M of k bits. Bob uses $(M, \widetilde{X}^n)$ to produce an estimate $\hat{T}$ of T. Let $P_e^n = \Pr(\hat{T} \neq T)$.

1. *(1pt) Show that $H(M) \leq k$.*

2. *(4pt) Deduce that*
$$k \geq H(T \mid \widetilde{X}^n) - H(T \mid M, \widetilde{X}^n).$$

3. *(3pt) Use a well-known inequality to show that*
$$H(T \mid M, \widetilde{X}^n) \leq 1 + P_e^n \log_2 n.$$

4. *(2pts) Compute $H(T \mid \widetilde{X}^n)$.*

5. *(1pt) Deduce that*

$$k \geq (1 - P_e^n) \log_2 n - 1.$$

We now construct a communication scheme that achieves this bound asymptotically. For any binary sequence x^n, define

$$f(x^n) = \sum_{i=1}^{n} x_i 2^{i-1}.$$

For $t \in \{0, \ldots, n-1\}$, let $S_t(x^n)$ denote the t-right-circular shift of x^n.

Alice computes $f(S_t(X^n))$ for every $t \in \{0, \ldots, n-1\}$ and sends Bob an index A attaining the maximum. Bob similarly computes $f(S_t(\widetilde{X}^n))$ for all t and determines an index B attaining the maximum. Bob declares

$$\hat{T} \equiv A - B \pmod{n},$$

6. *(2pt) Show that f is one-to-one on $\{0,1\}^n$.*

Probability of error. *We now show that Bob's estimate is correct with probability tending to one as $n \to \infty$. To do this, it suffices to show that all circular shifts of X^n are distinct with high probability (hence A and B are unique and $\hat{T} = T$).*

7. *(4pt) Fix $t \in \{0, \ldots, n-1\}$. Show that for any nonnegative integer i*

$$S_t(x^n) = x^n \quad \Longrightarrow \quad S_{i \cdot t \bmod n}(x^n) = x^n.$$

8. *(6pt) Show that any sequence satisfying $S_t(x^n) = x^n$ is completely determined by its first $r = gcd(n, t)$ bits (gcd = greatest common divisor). Hint: first consider the case where t divides n.*

9. *(2pt) For any $t \in \{1, \ldots, n-1\}$, deduce that*

$$\Pr(S_t(X^n) = X^n) \leq 2^{-\frac{n}{2}}$$

10. *(3pt) Deduce that the probability that two distinct circular shifts of X^n coincide tends to zero as $n \to \infty$.*

11. *(2pt) Deduce that Alice can communicate $k = \lceil \log_2 n \rceil$ bits and Bob can recover T with probability tending to one.*

12. *(2pt) Let $Min(n, \varepsilon)$ be the minimum number of bits Alice should communicate to Bob, among all schemes such that $P_e^n \leq \varepsilon$. Conclude that*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{Min(n, \varepsilon)}{\log_2 n} = 1.$$

Exercise 6 (Capacity, 6pt)**.** *Consider a discrete memoryless channel with transition probabilities $Q(y \mid x)$, where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.*

Suppose we construct a new channel by replacing an output symbol y_0 with two symbols y_a and y_b, such that

$$Q'(y_a \mid x) = \alpha Q(y_0 \mid x), \qquad Q'(y_b \mid x) = (1-\alpha)Q(y_0 \mid x),$$

for every $x \in \mathcal{X}$, where $\alpha \in [0,1]$. All other transition probabilities remain unchanged.

Show that the capacity of the new channel is equal to that of the original channel.

Hint: *Let Y and Y' denote the outputs of the original and modified channels, respectively. Show that both $X - Y - Y'$ and $X - Y' - Y$ form Markov chains, then use Data Processing Inequality.*