IERG5154 Final (72 hours)

Open notes, open book (Cover and Thomas), no internet (won't really help), no collaboration (for fairness). Hard-copy answer-sheet preferable, but if you're not on campus on Monday, soft-copy emailed to us is also ok.

1. **Secrecy for the erasure channel** *(8 points)*: Alice wishes to send Bob a message M over a binary erasure channel with erasure probability p. However, each bit X_i that she transmits to Bob is also overheard by evil eavesdropper Calvin, who hears a "degraded" version of the message Bob hears with erasure probability p'. Specifically, the bits Calvin overhears are a subset of the bits Bob hears, and the end-to-end channels from Alice to Bob and from Alice to Calvin are respectively BEC(p) and BEC(p'). Alice wants to ensure that her message to Calvin is "secret", *i.e.,* the mutual information between Alice's message M and Calvin's observations Z^n is at most ϵn.

   (a) *(2 points):* Use information theory inequalities to prove that Alice's optimal rate of secret transmission is no more than $p' - p$ if $p' > p$, and zero otherwise.

   (b) Show that random linear codes achieve such performance. Show that such codes have good computational complexity for Alice and Bob. Choose X^n to be a random *binary linear* code (known to both Bob and Calvin) of the message's $Rn = (p' - p - \epsilon)n$ bits, and $n(1-p')$ random bits denoted by K (K is known to neither Bob nor Calvin).[1] *Hint: A "fact" that is useful to know (and that you may use without proof) is that with high probability over the choice of random $m \times n$ binary matrices, the probability that it has full rank over $\mathbb{F}_2$ (the binary field) is at least $1 - 2^{-c|m-n|}$, for a universal constant $c > 0$.*[2] Proceed as in the following two parts.

      i. *(3 points):* Prove that with high probability over the choice of random linear codes Bob can indeed decode M. What is Alice's encoding complexity, and Bob's decoding complexity?

      ii. *(3 points):* Prove that Calvin has mutual information at most $\mathcal{O}(\epsilon n)$ with M, *i.e.,* prove that over the randomness in the channel, Calvin's observations are "almost independent" of M. *Hint: Can you show that, with high probability over erasure patterns and your random linear code, for any (M, K) pair giving a particular observation Z^n to Calvin, and any $M' \neq M$, there exists a K' such that the (M', K') pair produces the same Z^n?*

2. **Rate-distortion curve under a "different" distortion measure** *(4 points)*: A zero-mean σ^2-variance Gaussian source is required to be compressed. The per-symbol distortion measure, however, is given by $2(x - \hat{x} + 1)^2 + 2$. Compute the rate-distortion function for this source. *Hint: This is closely related to Problem 10.18 from Cover and Thomas (which you'll need to solve to solve this problem), but there's an important difference – be sure to point it out in your answer.*

[1] A binary linear code takes linear combinations of the source message bits over $\mathbb{F}_2$ to generate the codeword's bits.

[2] This fact is not hard to prove, but for the interested reader [1] has a more sophisticated result.

3. **Concatenated codes against "omniscient" adversaries** *(6 points)*: A certain binary-input binary-output channel has an "omniscient" (meaning "knowing everything") adversary. The description of the channel is as follows. Let the input to the channel be X^n. The adversary can flip up to any pn bits of the channel by adding a binary vector Z^n (of Hamming weight at most pn) to X^n. This Z^n may be a function of X^n. Based on $Y^n = X^n \oplus Z^n$, the receiver is required to decode X^n with zero error.[3] For such a channel, describe a concatenated coding scheme that enables the encoder and decoder to computationally efficiently encode and decode at as high a rate as possible. Formulate your answer as the solution to maximization problem. What's the highest value of p for which your codes achieve a strictly positive rate?[4] *Hint: Remember, these are "worst-case" channels, and hence you cannot expect that errors will behave randomly. However, recall that both Gilbert-Varshamov codes and Reed-Solomon codes can handle worst-case errors.*

4. **Random walk in two dimensions** *(7 points)*: A drunken man is walking on a square grid. With each step, he has probability $p_1 = 4/10$ of moving one step in the positive x direction, probability $p_2 = 1/10$ of moving one step in the negative x direction, probability $p_3 = 3/10$ of moving one step in the positive y direction, and probability $p_4 = 2/10$ of moving one step in the negative y direction.

   (a) *(1 point):* After n steps, what is his expected position?

   (b) *(2 points):* After n steps, what is the probability that he has taken exactly $k_1 n$ steps in the positive x-direction, $k_2 n$ in the negative x-direction, $k_3 n$ steps in the positive x-direction, and $k_4 n$ in the negative y-direction ($k_1 + k_2 + k_3 + k_4 = 1$, $k_i \geq 0\ \forall i$)? Use Stirling's approximation to write this overall probability in the form $\doteq 2^{-cn}$ for some c that depends on (p_1, p_2, p_3, p_4) and (k_1, k_2, k_3, k_4).

   (c) *(4 points):* To first order in exponent, what is the probability that after n steps the drunken man is outside the box given by $0.2n \leq x \leq 0.4n$, $0 \leq y \leq 0.2n$? That is, compute this probability $\doteq 2^{-c'n}$ for some c'. To get points for this question you need to find the *exact* value of c', depending only on the (p_1, p_2, p_3, p_4) values given in this problem. *Hint: Use the answer of the previous part to compute the probability for the "likeliest" tuple (k_1, k_2, k_3, k_4) outside this box, and then note that there's at most a polynomial number of possible (k_1, k_2, k_3, k_4) tuples.*

References

[1] Colin Cooper, "On the distribution of rank of a random matrix over a finite field," *Random Structures and Algorithms 17 (2000), 197–212.*

[3] These are exactly the "coding theory" channels we considered in class, for which we studied Gilbert-Varshamov codes, and the Plotkin and Hamming bounds.

[4] GV codes are currently the codes with the highest known rates against such a channel, but they are not computationally efficient. (Why?) However, there are no currently known codes that computationally efficiently achieve the same rates $(1 - H(2p))$ that GV codes do. This can act as a sanity check on your answer. (Alternatively, if you can design codes with rates equaling or exceeding $1 - H(2p)$, you're guaranteed an $A+++$ in the course...)