

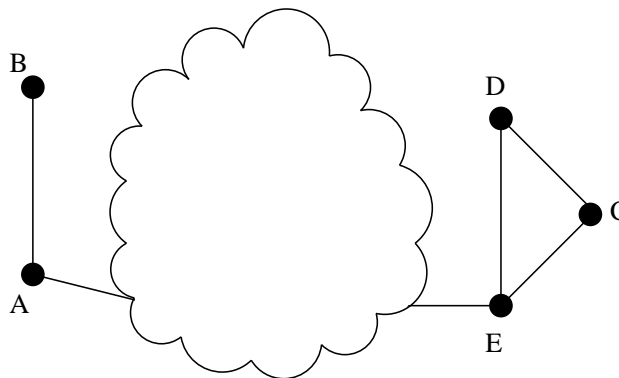
### Final Examination

1. (10 points) **Huffman code**

Find the Huffman binary codeword lengths for the probability assignment

$$\mathbf{p} = (1/45, 1/45, 6/45, 6/45, 7/45, 8/45, 8/45, 8/45).$$

2. (20 points) **Random walk obscured by clouds**



A random walk takes place on this undirected connected graph. Part of the graph is obscured. We observe that  $\frac{1}{6}$ th of the time, in the limit as  $n \rightarrow \infty$ , is spent in  $A$ .

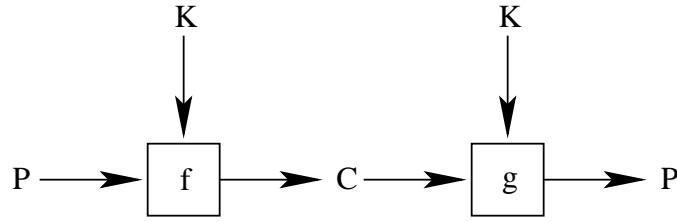
- (a) What fraction of the time is spent in the cloud?
- (b) What is the entropy rate of the random walk?

3. (20 points) **Entropy inequality**

Let  $(X, Y) \sim p(x, y)$  be discrete real valued random variables.

- (a) Compare  $H(X|X \cdot Y)$  and  $H(X, Y) - H(X \cdot Y)$ . Is the left-hand side  $\leq, \geq$ , or  $=$  to the right-hand side?
- (b) Find conditions for equality.

4. (20 points) **Shannon cryptography**



In cryptography, we want to know how large the random key must be in order to obscure the text. Consider the following relations between random variables  $P$ ,  $K$ , and  $C$ :

- i.* The plaintext  $P$  and the key  $K$  are independent of each other.
- ii.* The ciphertext is a deterministic function of the plaintext and key:  $C = f(P, K)$ .
- iii.* The plaintext is a deterministic function of the ciphertext and key:  $P = g(C, K)$ .

Note we do not assume that  $f$  and  $g$  are invertible functions.

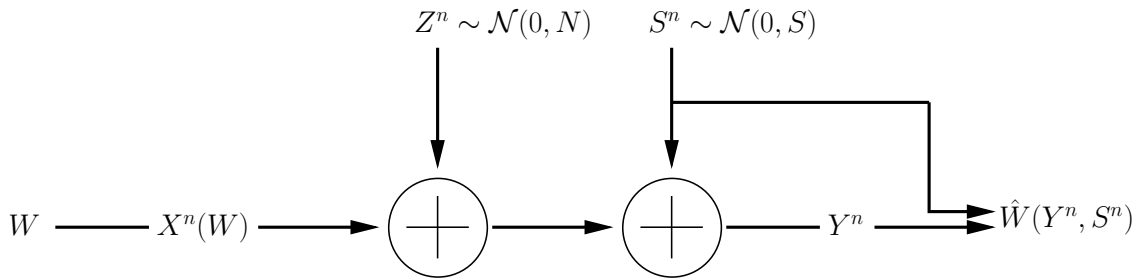
Show the following:

$$\text{If } P \text{ and } C \text{ are independent, then } H(K) \geq H(P).$$

That is, if we wish the ciphertext and plaintext to be independent, and we wish to be able to decode the plaintext from the ciphertext, then the entropy of the secret key  $K$  must be greater than the entropy of the plaintext  $P$  it is encrypting.

5. (20 points) **Noise partially known**

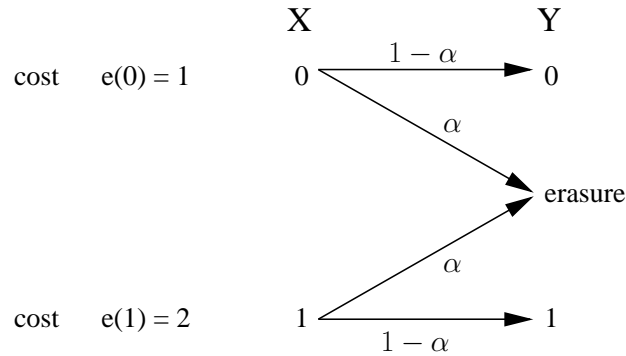
What is the capacity of the Gaussian channel with part of the noise known at the receiver? *You are not required to give proofs of achievability and converse.*



Here  $Z, S$  are i.i.d. Gaussian, independent of  $X$  and each other.  $X, S$ , and  $Z$  are also jointly independent.  $Y^n = X^n + Z^n + S^n$ . There is a power  $P$  constraint on the input  $X$ . That is,  $\frac{1}{n} \sum_{i=1}^n x_i^2(w) \leq P$  for all  $w \in \{1, 2, \dots, 2^{nR}\}$ .

6. (20 points) **Ink is expensive**

Consider the memoryless binary erasure channel



with cost  $e(0) = 1$  for  $X = 0$ , and cost  $e(1) = 2$  for  $X = 1$ , and cost constraint

$$\frac{1}{n} \sum_{i=1}^n e(x_i(w)) \leq E, \quad \text{for all } w \in \{1, 2, \dots, 2^{nR}\}$$

Thus each codeword  $x^n(w)$  of a  $(2^{nR}, n)$  code must satisfy the cost constraint  $E$ .

Find the capacity  $C(E)$  as a function of  $E$ .

*You are not required to give proofs of achievability and a converse.*

7. (20 points) **Side information**

Consider a horse race  $X \sim p(x)$  with  $m$ -for-1 fair odds. Let  $(X, Y, Z) \sim p(x, y, z)$ . Gambler  $A$  has side information  $Y$ , and gambler  $B$  has side information  $Z$ .

- (a) Suppose  $X \rightarrow Z \rightarrow Y$  forms a Markov chain. Which gambler does best?
- (b) Now, gambler  $A$  is given side information  $Z$  in addition to his information  $Y$ , where  $p(x, y, z)$  is arbitrary. How much does this combined information  $(Z, Y)$  increase his growth rate of wealth beyond his rate from  $Y$  alone?

8. (20 points) **Poisoned wine**

You have eight wine bottles, and exactly one is poisoned, according to the following probabilities:

Bottle:	1	2	3	4	5	6	7	8
Probability of poison:	1/45	1/45	6/45	6/45	7/45	8/45	8/45	8/45

You intend to determine which bottle is poisoned by testing the wines on lab rats. If a rat imbibes poison, it instantly dies. Since we don't want to deal with drunken rats, there is one constraint:

**No rat can be used more than once.**

*The following two problems should be considered separately.*

- (a) Determine which bottle is poisoned in a way that ensures *no more than one rat is killed*.
- (b) Determine which bottle is poisoned in a way that minimizes the average number of rats used.

9. (30 points) **Memory chip**

A memory chip has one million memory cells, each storing one bit. However, due to the manufacturing process, each cell is defective with probability  $p = .01$ . Cell defects are independent.

We have to live with the defective chips because otherwise, we would have to wait a long time to get a chip that doesn't have any defects.

So one option is to measure each cell before the chip is used, identify the defects, and not use those cells.

- (a) What is the average usable memory size of a chip?

Now suppose we don't make these measurements ahead of time, and instead decide to simply use error correction codes to combat the defects, under several different assumptions.

- (b) First, assume that defective cells give illegitimate voltage readings (erasures) and therefore can be identified as soon as the memory is read. What is the effective memory size of the chip?
- (c) Now assume that the defective cells give the opposite reading from that which was written in memory. What is the effective memory size of the chip?