Question 1

In this question you will analyze the following bit commitment protocol based on a pseudorandom generator $G\colon \{0,1\}^k \to \{0,1\}^{3k}$. First, receiver picks a random string $R \in \{0,1\}^{3k}$ and shares it with sender. To commit to a bit s, sender chooses a random X and sends $G(X) + s \cdot R$ (i.e., $G(X)$ when $s = 0$ and $G(X) + R$ when $s = 1$). To reveal, sender reveals s and X and receiver checks that his commitment C equals $G(X) + s \cdot R$.

(a) Prove that if G is a pseudorandom generator then the commitment is hiding. Work out the parameters.

Solution: If G is (s, ε)-pseudorandom then both commitments of 0 and 1 are $(s - O(k), \varepsilon)$-simulatable by a truly random string. Receiver's view is $(R, G(X) + s \cdot R)$. Suppose these can be distinguished by a pair of independent random strings (R, R'). If $s = 0$ then $G(X)$ is distinguished from R' with the same advantage. If $s = 1$ $G(X)$ can be distinguished by XORing R with the second half.

(b) Show that with probability $1 - 2^{-k}$ over the choice of R there does not exist a pair of inputs X and X' such that $G(X) + G(X') = R$. (**Hint:** Take a union bound over all pairs.)

Solution: For a fixed pair (X, X') the probability that $G(X) + G(X')$ equals R is 2^{-3k}. There are 2^k choices for X and as many for X'. By a union bound, the probability that some pair satisfies the equation is at most $2^{2k} \cdot 2^{-3k} = 2^{-k}$.

(c) Prove that the commitment is binding. Work out the parameters.

Solution: The commitment is $(\infty, 2^{-k})$-binding. Assume some C can be decommitted to both 0 and 1. Then C must equal both $G(X) = C$ and $G(X') + R$ for some X, X', so $G(X) + G(X') = R$. By part (b) this can happen with probability at most 2^{-k}.

Question 2

Let $f\colon \{0,1,2\} \times \{0,1,2\} \to \{0,1\}$ is the equality function $f(x,y) = 1$ if $x = y$ and 0 if $x \neq y$. Consider the following key exchange protocol based on a two-party protocol for f: Alice and Bob choose random inputs x and y from $\{0,1,2\}$ and run the protocol. After Bob obtains $f(x,y)$ he forwards this value to Alice. If $f(x,y) = 1$ each party outputs their input, and otherwise they repeat.

(a) Show that Alice's and Bob's output are equal and uniformly random with probability 1. What is the expected number of repetitions?

Solution: Conditioned on $f(x,y) = 1$, Alice's and Bob's input are $(0,0)$, $(1,1)$, and $(2,2)$ with probability a third each. Since the repetitions are independent the outputs will be uniform and equal at termination.

In any given round x and y are equal with probability $1/3$. The number of repetitions R is a geometric random variable with $1/3$ success probability, so the expected number of repetitions/ is 3.

(b) In question 4 of the midterm you showed that if a two-party protocol for f is simulatable against honest-but-curious then any two transcripts of the protocol are (s, ε)-indistinguishable. Assuming this, show that the key exchange protocol is secure, namely that the the key and the transcript are indistinguishable from a pair of independent random variables. Work out the parameters.

Solution: There are a couple of ways to do this simulation giving slightly different bounds. Here is one. Simulate the transcript and the key by the random variable

$$(T_1, \dots, T_{R-1}, S, K)$$

where R is the above geometric random variable, T_1 up to T_{R-1} are f-protocol transcripts on independent random *distinct* inputs, S is a simulated transcript, and K is a simulated random key, i.e. an independent and uniform $\{0, 1, 2\}$ random variable.

We will show that this random variable is (s, ε)-indistinguishable from the key agreement transcript and the real key. Suppose there is a distinguisher D with advantage ε. Then there is some value of R for which D distinguishes between the two. Conditioned on this choice of R, the first $R-1$ f-protocol transcripts are identically distributed in both random variables and independent of the rest. So there is some choice of transcripts for which D distinguishes (S, K) from the R-th f-protocol transcript and its output (i.e. the key). Since the real key and simulated key are identically distributed, there is a choice of the key for which D distinguishes S from the R-th f-protocol transcript in which both parties' inputs equals the key. This contradicts (s, ε)-simulatability of the protocol transcript.

Conditioned on R, the simulator needs to run $R - 1$ f-protocols on simulated inputs, one simulated protocol, and sample a random key. So its running time is $Rt_f + t_{key}$, where t_f is the time of a real/simulated protocol run and t_{key} is the time to sample a key.[1] Since the probability that R exceeds r is at most $(2/3)^r$, we can also say that the simulator size is at most $rt_f + t_{key}$ except with probability $(2/3)^r$. So we can conclude that the transcript is $(s, \varepsilon + (2/3)^r)$-simulatable in size $rt_f + t_{key}$ for every r.

Question 3

(20 points) Let *Com* be a bit commitment scheme. Consider the following variant *Com*′: To commit to a bit x, Sender chooses a random bit r and sends $Com'(x) = (Com(r), Com(x + r))$ as his commitment. Here + stands for XOR.

(a) Describe the revealment and the verification procedures for *Com*′.

Solution: To reveal x, Sender reveals both r and $x + r$. The verifier checks both commitments and accepts if they XOR to x.

(b) Prove that if *Com* is perfectly binding then so is *Com*′.

Solution: Given a commitment $C' = (C_1, C_2)$, because *Com* is perfectly binding there is a unique pair of values v_1 and v_2 that C_1 and C_2 can represent. Then $v_1 + v_2$ is the only possible decommitment of C'.

[1]In fact a ternary random variable cannot be perfectly sampled in finite size, so t_{key} should also be a constant times a geometric random variable.

(c) Prove that if Com is hiding then Com' is also hiding. Work out the parameters.

Solution: Suppose Com has size t_{Com} and is (s, ε)-simulatable in size t_{Sim}. Then Com' is (s, ε)-simulatable in size $t_{Sim} + t_{Com}$. The simulator Sim' outputs $(Com(r), Sim)$ where r is random. If $(Com(r), Sim)$ and $(Com(r), Com(x + r))$ can be distinguished by D with advantage ε, then they can be distinguished for some fixed r. Conditioned on r, $Com(r)$ and $Com(x+r)$ are independent so D has advantage ε even when $Com(r)$ is fixed to some value. So Com cannot be (s, ε)-simulatable.

Now Alice has committed to two bits x and x' using Com' and wants to prove to Bob that the two are equal. Their commitments are

$$Com'(x) = (Com(r), Com(x + r)) \quad \text{and} \quad Com'(x') = (Com(r'), Com(x' + r')).$$

Consider the following proof system:

1. Alice sends Bob the value $s = r + r'$.
2. Bob sends Alice a random bit b.
3. If $b = 0$, Alice reveals r and r'. If $b = 1$, Alice reveals $x + r$ and $x' + r'$.
4. Bob verifies the values revealed by Alice and accepts if their XOR equals s.

Assume that Com is perfectly binding and show the following.

(d) Completeness: If x equals x' then Bob accepts with probability 1.

Solution: If Alice chooses $b = 0$ then Bob reveals r and r' and their XOR equals s by assumption, so Bob accepts. If Alice chooses $b = 1$ then the XOR of $x + r$ and $x' + r'$ also equals $r + r' = s$, so Bob again accepts.

(e) Soundness: If x does not equal x' then upon interacting with a cheating Alice, Bob accepts with probability at most half.

Solution: Regardless of Alice's choice of s, either $r + r' \neq s$ or $(x + r) + (x' + r') \neq s$. In the first case, if Bob chooses $b = 0$ he will reject Alice's decommitments. In the second case, the same happens when $b = 1$.

(f) Zero-knowledge: If x equals x' and Com is hiding then the view of a cheating Bob (consisting of $Com'(x)$, $Com'(x')$, his randomness, and Alice's messages) is efficiently simulatable. Work out the parameters.

Solution: Bob's view consists of the four commitments $Com(r), Com(x+r), Com(r'), Com(x'+r')$, the value $s = r + r'$ and the revealed values in Step 3.

The simulator guesses the value of b at random. If $b = 0$ the simulator challenges Bob on $(Com(y), S_2, Com(y'),$ where $Com(y)$ and $Com(y')$ are true commitments to random bits and S_1 to S_4 are simulated commitments. If $b = 1$ the challenge is $(S_1, Com(y), S_3, Com(y'))$. If Bob's response b^* equals to b the simulator reveals y and y'. Otherwise the simulator tries again up to r times.

Assume Com is (s, ε)-hiding. Conditioned on $b^* = b$, the real and simulated views are $(s - t, 2\varepsilon)$-indistinguishable where t is the size of (cheating) Bob by the usual hybrid argument.

On the other hand, the simulator's challenge is (s, ε)-indistinguishable from (S_1, S_2, S_3, S_4). Since this is independent of b, the probability that Bob responds by b upon seeing this challenge is exactly $1/2$. Assuming $t \leq s$, it follows that $b^* = b$ is at least $1/2 + \varepsilon$. So the probability that all r simulation attempts fail is at most $(1/2 + \varepsilon)^r$, so the protocol is $(s - t, 2\varepsilon + (1/2 + \varepsilon)^r)$-zero knowledge with simulation overhead t.