1 Security of Key Agreement

We consider a key agreement scheme defined by

- one PPT algorithm $\text{setup}(1^s) \rightarrow \text{pp}$ which generates public parameters $\text{pp};$
- two probabilistic polynomially bounded interactive machines $A$ and $B$ with input $\text{pp}$ and producing a secret output $K$ (denoted by $K_A$ for $A$ and by $K_B$ for $B$).

Correctness implies that the following game outputs 1 with probability 1.

1: $\text{setup}(1^s) \rightarrow \text{pp}$
2: make $A(\text{pp})$ and $B(\text{pp})$ interact with each other and output $K_A$ and $K_B$
3: output $1_{K_A=K_B}$

Q.1 Give a formal definition for the security against key recovery under passive attacks.
Q.2 Formalize how to define the Diffie-Hellman protocol under this setting.
Q.3 Formally prove that the Diffie-Hellman protocol is secure in the sense of the previous question if and only if the computational Diffie-Hellman problem is hard.
Q.4 We now consider security against Alice’s key recovery under active attacks as defined by the following game:

1: $\text{setup}(1^s) \rightarrow \text{pp}$
2: $\text{st}_A \leftarrow \text{pp}$, $\text{finished}_A \leftarrow \text{false}$
3: $\text{st}_B \leftarrow \text{pp}$, $\text{finished}_B \leftarrow \text{false}$
4: run $A^{\text{OA}_{\text{OB}}}(\text{pp}) \rightarrow K$
5: output $1_{K=K_A}$ and $\text{finished}_A$

$\text{OA}(x)$:
6: if $\text{finished}_A$ then return
7: $\text{st}_A \leftarrow (\text{st}_A, x)$
8: run $A(\text{st}_A)$ to get private output $\text{st}_A$ and next message $y$
9: if $y$ non-final then return $y$
10: $\text{finished}_A \leftarrow \text{true}$
11: $K_A \leftarrow \text{st}_A$
12: return $y$
And the same for oracle OB. Prove that the Diffie-Hellman protocol is insecure in this sense.

Q.5 Based on some attacks seen in the course, formalize security against key recovery under active attacks making $K_A = K_B$. Prove that Diffie-Hellman is secure by assuming that the problem defined by the following game is hard:

1: setup$(1^s) \rightarrow pp = (q, g)$
2: pick $x, y \in \mathbb{Z}_q^*$
3: $B(pp, g^x, g^y) \rightarrow (u, v, w)$
4: return $1_{u^x = v^y = w}$ and $u, v, w \in \langle g \rangle$ and $w \neq 1$

where $g$ generates $\langle g \rangle$ of order $q$, with neutral element 1.

2 Advantage Amplification

Let $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ be $2n$ independent Boolean variables. We assume that $X_1, \ldots, X_n$ are identically distributed and that $Y_1, \ldots, Y_n$ are identically distributed. We assume that the statistical distance between the distributions of $X_i$ and $Y_j$ is $\varepsilon$. Given distinguisher, i.e. a Boolean algorithm $A$ (with unbounded complexity), we define $X = A(X_1, \ldots, X_n)$ and $Y = A(Y_1, \ldots, Y_n)$. We are interested in $A$ which maximizes the statistical distance between the distributions of $X$ and $Y$. We denote by $d$ the statistical distance and we identify random variables by their distributions when computing distances, by abuse of notation.

Q.1 Prove that $d(X, Y) = d((X_1, \ldots, X_n), (Y_1, \ldots, Y_n))$.
Q.2 Assume that $\Pr[X_i = 1] = 0$.
   Q.2a Give the distributions of $X_i$ and $Y_j$.
   Q.2b Compute $d(X, Y)$ in terms of $\varepsilon$ and $n$.
   Q.2c Give an asymptotic equivalent of the minimal $n$ such that $d(X, Y) \geq \frac{1}{2}$ in terms of $\varepsilon$, when $\varepsilon \to 0$.

Q.3 Assume now that $\Pr[X_i = 1] = \frac{1}{2}(1 - \varepsilon)$ and $\Pr[Y_i = 1] = \frac{1}{2}(1 + \varepsilon)$.
Q.3a Show that $A(z_1, \ldots, z_n) = 1_{z_1 + \cdots + z_n < \frac{\varepsilon}{2}}$ makes $d(X, Y)$ maximal.
Q.3b Given that $\Pr[X_1 + \cdots + X_n < \frac{n}{2}] = \Pr[Y_1 + \cdots + Y_n > \frac{n}{2}]$, prove that for $n$ odd, we have $d(X, Y) = |1 - 2 \Pr[X_1 + \cdots + X_n < \frac{n}{2}]|$. 
Q.3c Compute the expected value and the variance of $X_1 + \cdots + X_n$.
Q.3d By approximating $X_1 + \cdots + X_n$ to a normal distribution, give an asymptotic equivalent to $n$ so that $d(X, Y)$ is a constant.