Problem 1.

Part A. Suppose H_1, H_2 are sampled independently from universal hash function $\mathcal{H}$, and X_1, X_2 are drawn independently from the same distribution as X.

$$\begin{aligned}
\mathrm{Col}(H, H(X)) &= \Pr[H_1 = H_2, H_1(X_1) = H_2(X_2)] \\
&= \sum_{h \in \mathcal{H}} \Pr_{H_1, H_2}[H_1 = H_2 = h] \Pr_{X_1, X_2}[h(X_1) = h(X_2)] \\
&= \sum_{h} |\mathcal{H}|^{-2} (\Pr[X_1 = X_2] + \Pr[h(X_1) = h(X_2), X_1 \neq X_2]) \\
&\leq |\mathcal{H}|^{-1} (\max_x \Pr[X = x] + \Pr[h(X_1) = h(X_2) | X_1 \neq X_2]) \\
&\leq |\mathcal{H}|^{-1} (2^{-k} + 2^{-\ell})
\end{aligned}$$

Part B. We use h to denote a possible function in $\mathcal{H}$, and s to denote a string in $\{0, 1\}^\ell$.

$$\begin{aligned}
&\|(H, H(X)) - (H, U)\|_2^2 \\
&= \sum_{h,s} \Pr[(H, H(X)) = (h, s)]^2 + \sum_{h,s} \Pr[(H, U) = (h, s)]^2 \\
&\quad - 2 \sum_{h,s} \Pr[(H, H(X)) = (h, s)] \Pr[(H, U) = (h, s)] \\
&= \mathrm{Col}(H, H(X)) + |\mathcal{H}|^{-1} 2^{-\ell} - 2|\mathcal{H}|^{-1} 2^{-\ell} \sum_{h,s} \Pr[(H, H(X)) = (h, s)] \\
&= \mathrm{Col}(H, H(X)) - |\mathcal{H}|^{-1} 2^{-\ell}
\end{aligned}$$

Since $\ell = k - 2\log(1/\epsilon) - O(1)$, using the result in Part A we have

$$\mathrm{Col}(H, H(X)) - |\mathcal{H}|^{-1} 2^{-\ell} \leq |\mathcal{H}|^{-1} 2^{-k} \leq \frac{\epsilon^2}{|\mathcal{H}| 2^\ell}.$$

Part C. By Cauchy Schwartz inequality,

$$4\Delta^2 \leq \|(H, H(X)) - (H, U)\|_2^2 \cdot |\mathcal{H}| 2^\ell \leq \epsilon^2.$$

Problem 2.

If Hybrid 0 and Hybrid 1 are distinguishable, an adversary can sample r, x, and distinguish $\mathbf{s}^T\mathbf{A} + \mathbf{e}^T$ and $\mathbf{b}$. This violates the decisional LWE assumption.

Define $\mathbf{A}' = \begin{bmatrix} \mathbf{A} \\ \mathbf{b}^T \end{bmatrix}$. Define a family of hash function $h_{A'} : \{0,1\}^m \to \mathbb{Z}_p^{n+1}$ as $h(\mathbf{r}) = \mathbf{A}'\mathbf{r}$. h is a universal hash function because for any distinct vectors $\mathbf{x}, \mathbf{y} \in \{0,1\}^m$,

$$\Pr[h(\mathbf{x}) = h(\mathbf{y})] = \Pr[\mathbf{A}'(\mathbf{x} - \mathbf{y}) = 0] = p^{-(n+1)}.$$

The last step is because if $\mathbf{x}, \mathbf{y}$ are distinct at the i^{th} bit and $\mathbf{A}'(\mathbf{x} - \mathbf{y}) = 0$, the entry of $\mathbf{A}'$ on row i is determined after sampling the value on other rows. Recall that $\mathbf{A}'$ is a $(n+1) \times m$ matrix with all its entries uniformly sampled from $\mathbb{Z}_p$, the probability of $\mathbf{A}'(\mathbf{x} - \mathbf{y}) = 0$ should be $p^{-(n+1)}$.

We use the result in Problem 1: set $\ell = (n+1)\log p,\ k = m$, and ϵ is chosen so that $\ell = k - 2\log(1/\epsilon)$, then

$$\Delta\big((H, H(X)),\ (H, U)\big) \le \frac{\epsilon}{2} = 2^{-(m-(n+1)\log p)/2}.$$

If $m \ge 3(n+1)\log p$, the distinguish probability of any adversary is less than $p^{-(n+1)}$.

Problem 3.

Part A. Notice that $(1+N)^k = 1 + kN \mod N^2$. Since N is odd, we have $1 + N = (1+N)^{1+N}$ is the square of $(1+N)^{\frac{1+N}{2}}$. So $1 + N \in \mathbb{QR}_{N^2}$.

We can see that $\text{ord}(1+N) = N$. Here $\text{ord}(g)$ denotes the order of g, which is the smallest positive integer k satisfying $g^k = 1$. So $1+N$ generates a group of size N, which must be $\mathbb{G}_N$. ($\mathbb{G}_N$ is the only subgroup of $\mathbb{QR}_{N^2}$ that is of size N. This relies on the fact that $p' \neq q$ and $q' \neq p$.)

Part B. Suppose $g = (1+N)^x = 1 + xN$, then $g^a = (1+N)^y = 1 + yN$. We can calculate x, y then find $k = x^{-1}y \mod N$ using Euclidean algorithm.

Remark: We can not simply calculate the inverse of x because $\phi(N)$ is unknown.

Part C. Sample random $x \in \mathbb{Z}_{N^2}^*$, then x^2 is uniformly sampled in $\mathbb{QR}_{N^2}$. Then there exist unique $(g, h) \in \mathbb{G}_N \times \mathbb{H}_N$ such that $gh = x^2$ and (g, h) is uniformly distributed in $\mathbb{G}_N \times \mathbb{H}_N$.

Let $y = x^{2N}$. We prove y is uniformly sampled in $\mathbb{H}_N$. Let h_0 be a generator of $\mathbb{H}_N$ and $h = h_0^a$, then $y = g^N h^N = h^N = h_0^{aN}$. Since $\gcd(N, p'q') = 1$, and a is uniformly chosen from $\mathbb{Z}_{p'q'}$, we have aN uniformly chosen from $\mathbb{Z}_{p'q'}$.

Part D. $\mathsf{Dec}(sk, c) = sk^{-1} \log_{1+N}(c^{sk}) \mod N$. Part B shows $\log_{1+N}$ is efficiently computable.

Since $c^{sk} = h^{sk}(1+N)^{mp'q'} = (1+N)^{mp'q'}$, we have $\mathsf{Dec}(sk, c) = sk^{-1}mp'q' = m$.

Under DCR assumption, h is indistinguishable from a random element in $\mathbb{QR}_{N^2}$, thus multiply h to $(1+N)^m$ could act as a one time pad.

Problem 4.

Part A. $pk = (G, g, g^x)$, $c = (c_1, c_2) = (g^y, g^{xy} \cdot m)$
Rerandomization: sample z from $\{0, 1, ..., |G| - 1\}$

$$c' = (g^z \cdot c_1, (g^x)^z \cdot c_2) = (g^{y+z}, g^{xy} g^{xz} \cdot m)$$

Homomorphic evaluation:
Denote $c = (c_1, c_2) = (g^y, g^x y \cdot m_1)$, $c' = (c_1', c_2') = (g^{y'}, g^{xy'} \cdot m_2)$
Message m_1, m_2 is in Abelian group G

$$\text{Eval}(c, c') = (c_1 c_1', c_2 c_2') = (g^{y+y'}, g^{x(y+y')m_1 m_2})$$

Part B. $pk = N$, $c = h(1 + N)^m$
Rerandomization: sample h' from $\mathbb{H}_N$.

$$c' = c \cdot h' = hh'(1 + N)^m$$

Homomorphic evaluation: Denote $c = h(1 + N)^{m_1}$, $c' = h'(1 + N)^{m_2}$.
Message m_1, m_2 is in Abelian group $\mathbb{Z}_N$.

$$\text{Eval}(c, c') = c \cdot c' = hh'(1 + N)^{m_1 + m_2} \mod N^2$$

Part C. $pk = (\mathbf{A}, \mathbf{b})$, $c = (c_1, c_2) = (\mathbf{Ar}, \mathbf{b}^T \mathbf{r} + \lfloor p/q \rfloor m)$
Rerandomization: sample $r' \in \{0, 1\}^m$

$$c' = c + (\mathbf{Ar}', \mathbf{b}^T \mathbf{r}') = (\mathbf{A}(\mathbf{r} + \mathbf{r}'), \mathbf{b}^T (\mathbf{r} + \mathbf{r}') + \lfloor p/q \rfloor m)$$

Homomorphic evaluation:
Denote $c = (\mathbf{Ar}, \mathbf{b}^T \mathbf{r} + \lfloor p/q \rfloor m_1)$, $c' = (\mathbf{Ar}', \mathbf{b}^T \mathbf{r}' + \lfloor p/q \rfloor m_2)$.
Message m_1, m_2 is in Abelian group $\mathbb{Z}_q$

$$\text{Eval}(c, c') = c + c' = (\mathbf{A}(\mathbf{r} + \mathbf{r}'), \mathbf{b}^T (\mathbf{r} + \mathbf{r}') + \lfloor p/q \rfloor (m_1 + m_2))$$

Problem 5.

Part A. Let $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be a CPA-secure public-key encryption scheme, construct another encryption scheme that consists of

- $\widetilde{\mathsf{Gen}}(1^\lambda)$: return $(pk, sk) \leftarrow \mathsf{Gen}(1^\lambda)$.
- $\widetilde{\mathsf{Enc}}(pk, m) := \begin{cases} \mathsf{Enc}(pk, m) \| m & \text{if } \mathsf{Dec}(m, \mathsf{Enc}(pk, 0)) = 0 \\ \mathsf{Enc}(pk, m) \| 0 & \text{otherwise} \end{cases}$
- $\widetilde{\mathsf{Dec}}(sk, c_1 \| c_2) := \mathsf{Dec}(sk, c_1)$

It is not circularly secure since $\mathsf{Enc}(sk, sk)$ leaks sk. However, CPA security preserves since the original scheme is CPA secure and it's hard for an adversary to find some m s.t. $\mathsf{Dec}(m, \mathsf{Enc}(pk, 0)) = 0$.

Part B. The CPA security follows directly from the binary-secret LWE assumption. Note that

$$\mathsf{Enc}(\mathbf{s}, \mathbf{s}) = \left(\mathbf{R}, \mathbf{s}^T\mathbf{R} + \mathbf{e}^T + \left\lfloor \frac{q}{2} \right\rfloor \mathbf{s}^T\right) = \left(\mathbf{R}, \mathbf{s}^T\left(\mathbf{R} + \left\lfloor \frac{q}{2} \right\rfloor \mathbf{I}_n\right) + \mathbf{e}^T\right),$$

which is identically distributed to $\mathsf{Enc}(\mathbf{s}, 0^n) - \left(\left\lfloor \frac{q}{2} \right\rfloor \mathbf{I}_n, 0\right)$. Hence it's circularly secure.

Problem 6.

Part A.

$$\begin{aligned}
&\Delta\big(\big(\widetilde{pk}, \mathsf{Enc}(\widetilde{pk}, 0)\big),\ \big(\widetilde{pk}, \mathsf{Enc}(\widetilde{pk}, 1)\big)\big) \\
&= \sum_{\widetilde{pk}} \Pr\left[\mathsf{Gen}(1^\lambda, \mathsf{lossy}) = \widetilde{pk}\right] \Delta\big(\mathsf{Enc}(\widetilde{pk}, 0), \mathsf{Enc}(\widetilde{pk}, 1)\big) \\
&\leq \sum_{\widetilde{pk}} \Pr\left[\mathsf{Gen}(1^\lambda, \mathsf{lossy}) = \widetilde{pk}\right] \mathrm{negl}(\lambda) \\
&\leq \mathrm{negl}(\lambda)
\end{aligned}$$

Part B. Any lossy encryption scheme is CPA-secure under lossy mode since $\mathsf{Enc}(\widetilde{pk}, 0)$ and $\mathsf{Enc}(\widetilde{pk}, 1)$ are statistically indistinguishable.

By key indistinguishability, any adversary cannot distinguish which mode the scheme runs under, it is therefore CPA-secure under real mode after a simple hybrid.

Part C. Let $\mathsf{Gen}(1^\lambda, \mathsf{lossy})$ first run $\mathsf{Gen}(1^\lambda, \mathsf{real})$ to obtain (N, p, q), then sample $z \xleftarrow{\$} \mathcal{QR}_N$ uniformly, output $\widetilde{pk} = (N, z)$.

Key indistinguishability follows from Quadratic Residuosity assumption, note that $\mathsf{Enc}(\widetilde{pk}, 0)$ and $\mathsf{Enc}(\widetilde{pk}, 1)$ are both uniformly random in $\mathcal{QR}_N$ hence lossy encryption holds.