Problem 1 (Random Coding).

(a) Let $D_{u,t} := \{d_h\big(Y^n, C_n(u)\big) = t,\ U = u\}$, i.e., the event that message u is selected and the Hamming distance between the received sequence Y^n and $C_n(u)$ is t. Now observe the following:

$$\begin{aligned}
\Pr\Big(u \notin \hat{U}(Y^n) \,\Big|\, D_{u,t}\Big) &= \Pr\Big(\exists \tilde{u} \in \mathcal{U} \setminus \{u\},\ d_h\big(Y^n, C_n(\tilde{u})\big) < t \,\Big|\, D_{u,t}\Big) \\
&= \Pr\left(\bigcup_{\tilde{u}\in\mathcal{U}\setminus\{u\}} d_h\big(Y^n, C_n(\tilde{u})\big) < t \,\middle|\, D_{u,t}\right) \\
&\overset{(i)}{\leq} \sum_{\tilde{u}\in\mathcal{U}\setminus\{u\}} \Pr\Big(d_h\big(Y^n, C_n(\tilde{u})\big) < t \,\Big|\, D_{u,t}\Big) \\
&\overset{(ii)}{=} \sum_{\tilde{u}\in\mathcal{U}\setminus\{u\}} \Pr\Big(d_h\big(Y^n, C_n(\tilde{u})\big) < t\Big).
\end{aligned}$$

The inequality (i) is justified by the union bound and the equality (ii) is follows from the fact that the codewords are selected independently, i.e., $C_n(\tilde{u})$ is independent of Y and $C_n(u)$.
Note that for any $\tilde{u}$ and Y^n, $\Pr\Big(d_h\big(Y^n, C_n(\tilde{u})\big) < t\Big) = \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n}$. Hence, as a final result we have

$$\Pr\Big(u \notin \hat{U}(Y^n) \,\Big|\, D_{u,t}\Big) \leq |\mathcal{U}| \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n} = \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n(1-R)}.$$

(b)

$$\begin{aligned}
&\Pr\Big(\mathrm{dec}_n(Y^n) \neq u,\ u \in \hat{U}(Y^n) \,\Big|\, D_{u,t}\Big) \\
&= \Pr\Big(\exists \tilde{u} \in \mathcal{U} \setminus \{u\},\ d_h\big(Y^n, C_n(\tilde{u})\big) = t,\ \mathrm{dec}_n(Y^n) = \tilde{u} \,\Big|\, D_{u,t}\Big) \\
&\overset{(i)}{\leq} \sum_{\tilde{u}\in\mathcal{U}\setminus\{u\},} \Pr\Big(d_h\big(Y^n, C_n(\tilde{u})\big) = t,\ \mathrm{dec}_n(Y^n) = \tilde{u} \,\Big|\, D_{u,t}\Big)
\end{aligned}$$

where (i) follows from union bound. Now, for any $\tilde{u} \in \mathcal{U} \setminus \{u\}$, observe the factorization below.

$$\begin{aligned}
&\Pr\Big(d_h\big(Y^n, C_n(\tilde{u})\big) = t,\ \mathrm{dec}_n(Y^n) = \tilde{u} \,\Big|\, D_{u,t}\Big) \\
&= \Pr\Big(d_h\big(Y^n, C_n(\tilde{u})\big) = t \,\Big|\, D_{u,t}\Big) \Pr\Big(\mathrm{dec}_n(Y^n) = \tilde{u} \,\Big|\, D_{u,t},\ d_h\big(Y^n, C_n(\tilde{u})\big) = t\Big) \\
&\overset{(ii)}{=} \binom{n}{t} 2^{-n} \Pr\Big(\mathrm{dec}_n(Y^n) = \tilde{u} \,\Big|\, D_{u,t},\ d_h\big(Y^n, C_n(\tilde{u})\big) = t\Big)
\end{aligned}$$

Again, (ii) is justified by the independent codeword selection procedure. To upper bound the probability of the remaining event, note that the conditioned events $D_{u,t}$ and $d_h\big(Y^n, C_n(\tilde{u})\big) = t$ imply there are at least 2 codewords in the decoding set $\hat{U}(Y^n)$. Therefore the probability of $\tilde{u}$ being selected randomly at uniform is less than $\frac{1}{2}$. i.e.,

$$\Pr\Big(\mathrm{dec}_n(Y^n) = \tilde{u} \,\Big|\, D_{u,t},\ d_h\big(Y^n, C_n(\tilde{u})\big) = t\Big) \le \frac{1}{2}.$$

Combining the results, finally we obtain

$$\Pr\Big(\mathrm{dec}_n(Y^n) \neq u,\ u \in \hat{U}(Y^n) \,\Big|\, D_{u,t}\Big) \le \frac{1}{2}\binom{n}{t}2^{-n(1-R)}.$$

(c) Let $D_u := \{U = u\}$. i.e., the true codeword is u. Then we have

$$\Pr\Big(\mathrm{dec}_n(Y^n) \neq u \,\Big|\, D_u\Big) = \sum_{t=0}^{n} \Pr\Big(\mathrm{dec}_n(Y^n) \neq u \,\Big|\, D_{u,t}\Big)\Pr\Big(d_h\big(Y^n, C_n(u)\big) = t \,\Big|\, D_u\Big).$$

and

$$\Pr\Big(d_h\big(Y^n, C_n(u)\big) = t \,\Big|\, D_u\Big) = \Pr\Big(\mathrm{BSC}(p) \text{ flips t bits}\Big) = \binom{n}{t}p^t(1-p)^{(n-t)}.$$

Furthermore,

$$\begin{aligned}
&\Pr\Big(\mathrm{dec}_n(Y^n) \neq u \,\Big|\, D_{u,t}\Big) \\
&= \Pr\Big(\mathrm{dec}_n(Y^n) \neq u,\ u \notin \hat{U}(Y^n) \,\Big|\, D_{u,t}\Big) + \Pr\Big(\mathrm{dec}_n(Y^n) \neq u,\ u \in \hat{U}(Y^n) \,\Big|\, D_{u,t}\Big) \\
&\overset{(i)}{\le} \sum_{i=0}^{t-1}\binom{n}{i}2^{-n(1-R)} + \frac{1}{2}\binom{n}{t}2^{-n(1-R)}
\end{aligned}$$

where (i) is obtained from the results of part (a) and (b). The above expression is a probability so it also has to be less than 1.

$$\Pr\Big(\mathrm{dec}_n(Y^n) \neq u \,\Big|\, D_{u,t}\Big) \le \min\left\{1, \sum_{i=0}^{t-1}\binom{n}{i}2^{-n(1-R)} + \frac{1}{2}\binom{n}{t}2^{-n(1-R)}\right\}.$$

Combining all the above, we obtain

$$\Pr\Big(\mathrm{dec}_n(Y^n) \neq u \,\Big|\, D_u\Big) \le \sum_{t=0}^{n}\binom{n}{t}p^t(1-p)^{n-t}\min\left\{1, \sum_{i=0}^{t-1}\binom{n}{i}2^{-n(1-R)} + \frac{1}{2}\binom{n}{t}2^{-n(1-R)}\right\}.$$

This expression does not depend on the choice of u. The conditioning can be removed.

$$\Pr(\mathrm{dec}_n(Y^n) \neq U) \le \sum_{t=0}^{n}\binom{n}{t}p^t(1-p)^{n-t}\min\left\{1, \sum_{i=0}^{t-1}\binom{n}{i}2^{-n(1-R)} + \frac{1}{2}\binom{n}{t}2^{-n(1-R)}\right\}.$$

Observe that $\Pr(\mathrm{dec}_n(Y^n) \neq U) = E_{C_n}[\Pr(\mathrm{dec}_n(Y^n) \neq U|C_n)]$, that is, the average error probability averaged over the selection of codebooks. Therefore, there must exist a codebook C_n that satisfies

$$\Pr(\mathrm{dec}_n(Y^n) \neq U|C_n) \le \sum_{t=0}^{n}\binom{n}{t}p^t(1-p)^{n-t}\min\left\{1, \sum_{i=0}^{t-1}\binom{n}{i}2^{-n(1-R)} + \frac{1}{2}\binom{n}{t}2^{-n(1-R)}\right\}.$$

(d) As in the hint, define $\rho := t/n$ and $\bar{\rho} := 1 - \rho$.

$$1 = \sum_{i=0}^{n} \binom{n}{i} \rho^i \bar{\rho}^{n-i} \geq \sum_{i=0}^{t} \binom{n}{i} \rho^i \bar{\rho}^{n-i}$$

For $t \leq n/2$, note that $\rho \leq \frac{1}{2}$ and $\bar{\rho} \geq \frac{1}{2}$. Thus $\rho^i \bar{\rho}^{n-i}$ is decreasing in i. Then we have

$$1 \geq \sum_{i=0}^{t} \binom{n}{i} \rho^i \bar{\rho}^{n-i} \geq \sum_{i=0}^{t} \binom{n}{i} \rho^t \bar{\rho}^{n-t} = \sum_{i=0}^{t} \binom{n}{i} 2^{-nh_2(\rho)},$$

from which we conclude $\sum_{i=0}^{t} \binom{n}{i} \leq 2^{nh_2(\rho)}$.

(e) It suffices to prove part (i). A similar proof follows for part (ii).
For part (i), we have

$$\sum_{i=0}^{\lfloor nq \rfloor} \binom{n}{i} p^i \bar{p}^{n-i} = \sum_{i=0}^{\lfloor nq \rfloor} \binom{n}{i} \frac{p}{q}^i \frac{\bar{p}}{\bar{q}}^{n-i} q^i \bar{q}^{n-i}.$$

We know $q < p$, therefore $\frac{p}{q} > 1$ and $\frac{\bar{p}}{\bar{q}} < 1$. $\frac{p}{q}^i \frac{\bar{p}}{\bar{q}}^{n-i}$ is then increasing in i and

$$\frac{p}{q}^i \frac{\bar{p}}{\bar{q}}^{n-i} \leq \frac{p}{q}^{nq} \frac{\bar{p}}{\bar{q}}^{n-nq} = 2^{-nD_2(q||p)}$$

for any $i \leq \lfloor nq \rfloor$. This yields

$$\sum_{i=0}^{\lfloor nq \rfloor} \binom{n}{i} p^i \bar{p}^{n-i} \leq \sum_{i=0}^{\lfloor nq \rfloor} \binom{n}{i} q^i \bar{q}^{n-i} 2^{-nD(q||p)} \leq 2^{-nD(q||p)}.$$

For part (ii), we have $\frac{p}{q} < 1$ and $\frac{\bar{p}}{\bar{q}} > 1$. Thus,

$$\frac{p}{q}^i \frac{\bar{p}}{\bar{q}}^{n-i} \leq 2^{-nD_2(q||p)}$$

for any $i \geq nq$. A similar proof follows from here.

(f) It is known that the capacity $C = 1 - h_2(p)$ for BSC(p). Recall that $1 - h_2(p)$ is a continuous and decreasing function on $[0, 1/2]$. Hence, for any $R < 1 - h_2(p)$, $p < 1/2$ we can find a $p < q < 1/2$ such that $R < 1 - h_2(q)$. In part (c), we proved

$$\Pr\Big(\mathrm{dec}_n(Y^n) \neq U\Big) \leq \sum_{t=0}^{n} \binom{n}{t} p^t (1-p)^{n-t} \min\left\{1, \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n(1-R)} + \frac{1}{2}\binom{n}{t} 2^{-n(1-R)}\right\}.$$

Taking such q described above, we can split the outer sum into two parts.

$$\sum_{t=0}^{n} \binom{n}{t} p^t (1-p)^{n-t} \min\left\{1, \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n(1-R)} + \frac{1}{2}\binom{n}{t} 2^{-n(1-R)}\right\} = S_1 + S_2,$$

$$S_1 := \sum_{t=0}^{\lfloor nq \rfloor} \binom{n}{t} p^t (1-p)^{n-t} \min\left\{1, \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n(1-R)} + \frac{1}{2}\binom{n}{t} 2^{-n(1-R)}\right\},$$

$$S_2 := \sum_{t=\lceil nq \rceil}^{n} \binom{n}{t} p^t (1-p)^{n-t} \min\left\{1, \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n(1-R)} + \frac{1}{2}\binom{n}{t} 2^{-n(1-R)}\right\}.$$

For S_1, we have

$$S_1 \leq \sum_{t=0}^{\lfloor nq \rfloor} \binom{n}{t} p^t (1-p)^{n-t} \sum_{i=0}^{t} \binom{n}{i} 2^{-n(1-R)}$$

Note that $t < nq < n/2$. Using the result of part (d), we have

$$S_1 \leq \sum_{t=0}^{\lfloor nq \rfloor} \binom{n}{t} p^t (1-p)^{n-t} \sum_{i=0}^{t} \binom{n}{i} 2^{-n(1-R)} \leq 2^{-n(1-R)} \sum_{t=0}^{\lfloor nq \rfloor} \binom{n}{t} p^t (1-p)^{n-t} 2^{nh_2(t/n)}$$
$$\overset{(i)}{\leq} 2^{-n(1-R)} 2^{nh_2(t/n)} \overset{(ii)}{\leq} 2^{-n(1-h_2(q)-R)}$$

where (i) follows from the fact that $\sum_{t=0}^{\lfloor nq \rfloor} \binom{n}{t} p^t (1-p)^{n-t} \leq 1$ and (ii) follows from $h_2(.)$ is increasing in $[0, \frac{1}{2}]$ and $\frac{t}{n} < q < \frac{1}{2}$.

For S_2, we have

$$S_2 \leq \sum_{t=\lceil nq \rceil}^{n} \binom{n}{t} p^t (1-p)^{n-t} \overset{(iii)}{\leq} 2^{-nD_2(q||p)}$$

where (iii) is obtained from part (e). Combining all the upper bounds, we obtain the final upper bound as

$$\Pr\Big(\mathrm{dec}_n(Y^n) \neq U\Big) \leq 2^{-n(1-h_2(q)-R)} + 2^{-nD_2(q||p)}$$

for any q satisfying $p < q < \frac{1}{2}$ and $1 - h_2(q) > R$. This shows that

$$\Pr\Big(\mathrm{dec}_n(Y^n) \neq U\Big) \to 0 \text{ as } n \to \infty.$$

In part (c), we have shown that such $\{C_n\}$ exists.

Problem 2 (Feinstein's Theorem).

(a)
$$Z(W) = \sqrt{W(0|0)W(0|1)} + \sqrt{W(1|0)W(1|1)} = 2\sqrt{p(1-p)}.$$

To easily calculate $V(W)$, let $\bar{p} := 1 - p$ and note that

$$i(X;Y) = \begin{cases} \log_2\big(2\bar{p}\big), & \text{w.p. } \bar{p} \\ \log_2\big(2p\big), & \text{w.p. } p \end{cases}$$

with $E[i(X;Y)] = I(X;Y) = 1 - h_2(p)$. Therefore, its variance is straightforwardly calculated as

$$\begin{aligned} V(W) &= \mathrm{Var}\big(i(X;Y)\big) = E[(i(X;Y) - I(X;Y))^2] \\ &= p\left[1 + \log_2 p - (1 + p\log_2 p + \bar{p}\log_2 \bar{p})\right]^2 + p\left[1 + \log_2 p - (1 + p\log_2 p + \bar{p}\log_2 \bar{p})\right]^2 \\ &= p\bar{p}^2\left(\log_2 \frac{p}{\bar{p}}\right)^2 + p^2\bar{p}\left(\log_2 \frac{\bar{p}}{p}\right)^2 \\ &= p\bar{p}\left(\log_2 \frac{\bar{p}}{p}\right)^2. \end{aligned}$$

(b) Suppose B_i, $i \geq 1$ are i.i.d. and bounded random variables which can take values on the interval $[a, b]$. For this case, given $t > 0$, Hoeffding's inequality can be expressed as

$$\Pr\left(\sum_{i=1}^{n} B_i \leq n(E[B] - t)\right) \leq \exp\left(\frac{-2nt^2}{(b-a)^2}\right).$$

For our case, we know that $i(X_1^n; Y_1^n) = \log_2\left(\frac{W(Y_1^n|X_1^n)}{W(Y_1^n)}\right) = \log_2\left(\prod_{i=1}^n \frac{W(Y_i|X_i)}{W(Y_i)}\right) = \sum_{i=1}^n \log_2\left(\frac{W(Y_i|X_i)}{W(Y_i)}\right) = \sum_{j=1}^n i(X_j;Y_j)$ since the channel is memoryless and X_is are i.i.d. Moreover, $i(X;Y)$ only takes the values $\log_2(2\bar{p})$ and $\log_2(2p)$. Applying Hoeffding's inequality, we obtain

$$\begin{aligned} \Pr\left(i(X_1^n; Y_1^n) < n(R+\delta)\right) &\leq \Pr\left(i(X_1^n; Y_1^n) \leq n\Big(I(X;Y) - \big(I(X;Y) - R - \delta\big)\Big)\right) \\ &\leq \exp\left(\frac{-2n(I(X;Y) - R - \delta)^2}{\log_2(\bar{p}/p)^2}\right) \\ &= \exp\left(\frac{-2n(I(X;Y) - R - \delta)^2 p\bar{p}}{p\bar{p}\log_2(\bar{p}/p)^2}\right) \\ &= \exp\left(\frac{-n(I(X;Y) - R - \delta)^2 Z_p^2}{2V_p}\right). \end{aligned}$$

(c) We do the variable change $\gamma = 1 - \frac{\delta}{I(X;Y)-R}$. Note that $0 < \gamma < 1$ provided that $\delta < I(X;Y) - R$. We also have $\delta = \bar{\gamma}\big(I(X;Y) - R\big)$. Substituting this new variable, we obtain

$$\begin{aligned} \epsilon &\leq \exp\left(\frac{-n(I(X;Y) - R - \bar{\gamma}(I(X;Y) - R))^2 Z_p^2}{2V_p}\right) + 2^{-n\bar{\gamma}(I(X;Y)-R)} \\ &= \exp\left(\frac{-n(\gamma(I(X;Y) - R))^2 Z_p^2}{2V_p}\right) + \exp\left(\frac{-n\bar{\gamma}(I(X;Y) - R)}{\log_2 e}\right). \end{aligned}$$

Since we are able to choose δ freely as long as $0 < \delta < I(X;Y) - R$, this corresponds to choosing γ freely as long as $0 < \gamma < 1$. Observe that the inequality is satisfied for all allowed choices of δ, and hence for all allowed choices of γ. The right hand side (RHS) can be optimized to obtain a tighter upper bound as

$$\epsilon \leq \min_{0 \leq \gamma \leq 1} \exp\left(\frac{-n(\gamma(I(X;Y) - R))^2 Z_p^2}{2V_p}\right) + \exp\left(\frac{-n\bar{\gamma}(I(X;Y) - R)}{\log_2 e}\right).$$

Note: One could extend the optimization interval and replace the strict inequalities with weak inequalities in order to obtain a compact region and attain the minimum. This does not change the optimal value of RHS as the minimum is clearly not attained at 0 or 1.

Problem 3 (Strong Converse).

a) As the channel is BSC(p), we have the probability

$$W^n\big(\mathcal{Y}(u) \cap B_{q,n}(u) | U = u\big) = \sum_{i=0}^{\lfloor nq \rfloor} \binom{n}{i} p^i (1-p)^{n-i}.$$

Therefore by the result of 1) e) we have

$$W^n\big(\mathcal{Y}(u) \cap B_{q,n}(u) | U = u\big) \leq 2^{-nD_2(q||p)}$$

b) As $y^n \neq B_{q,n}(u)$, we also know that $d_H(y^n, \text{enc}(u)) > nq$. This implies that

$$W^n\big(y^n | U = u\big) = p^{d_H(y^n, \text{enc}(u))}(1-p)^{n - d_H(y^n, \text{enc}(u))} \leq p^{nq}(1-p)^{n(1-q)}. \qquad \forall y^n \notin B_{q,n}(u)$$

Where the last inequality is due to the fact that $p \leq (1-p)$.

c) We have the correct decoding probability equals to:

$$\begin{aligned}
&= P(\text{dec}(Y^n) = u | U = u) \\
&= P(Y^n \in \mathcal{Y}(u) | U = u) \\
&= P(Y^n \in \mathcal{Y}(u) \cap Y^n \in B_{q,n}(u) | U = u) + P(Y^n \in \mathcal{Y}(u) \cap Y^n \notin B_{q,n}(u) | U = u) \\
&\leq P(Y^n \in B_{q,n}(u) | U = u) + |\mathcal{Y}(u)| p^{nq}(1-p)^{n(1-q)} \\
&\leq 2^{-nD_2(q||p)} + \frac{|\mathcal{Y}(u)|}{2^n} 2^n 2^{nq \log p + n(1-q)\log(1-p)} \\
&= 2^{-nD_2(q||p)} + \frac{|\mathcal{Y}(u)|}{2^n} 2^{n(1 - h_2(q) - D_2(q||p))}
\end{aligned}$$

d) From the result in (c), we have the following inequality

$$\begin{aligned}
P(\text{dec}(Y^n) = U) &= \sum_{u \in \mathcal{U}} P(\text{dec}(Y^n) = u | U = u) P(U = u) \\
&\leq 2^{-nR} \sum_{u \in \mathcal{U}} 2^{-nD_2(q||p)} + \frac{|\mathcal{Y}(u)|}{2^n} 2^{n(1 - h_2(q) - D_2(q||p))} \\
&= 2^{-nR} \left(2^{nR} 2^{-nD_2(q||p)} + 2^{n(1 - h_2(q) - D_2(q||p))} \sum_{u \in \mathcal{U}} \frac{|\mathcal{Y}(u)|}{2^n} \right) \\
&= 2^{-nD_2(q||p)} + 2^{-n(R - 1 + h_2(q) + D_2(q||p))}
\end{aligned}$$

e) Observe that the function $f(h) = 1 + h\log_2 p + (1-h)\log_2(1-p)$ is continuous. Hence if $f(a) = k_a$, $f(b) = k_b$ and $k^* \in [k_a, k_b]$, there is a value $c \in [a, b]$ such that $f(c) = k^*$. We note that $f(p) = C$ and $f(0) = 1$. Hence if we take $k^* = \min(1, \frac{R+C}{2})$, there is always a $q \in [0, p]$ such that:

$$C < f(k^*) = \min\left(1, \frac{R+C}{2}\right) < R.$$

f) If we define

$$f(h) = 1 + h\log_2 p + (1-h)\log_2(1-p),$$

then the result of d. is equal to:

$$P(\text{dec}(Y^n) = U) \leq 2^{-nD_2(q||p)} + 2^{-n(R-f(q))} \qquad \forall q \leq p$$

By the result of e., we know that there exists $q^* < p$ such that $f(q^*) < R$ if $R > C$. Plugging this $q*$ to our inequality gives us

$$P(\text{dec}(Y^n) = U) \leq 2^{-nD_2(q^*||p)} + 2^{-nc)}$$

where $c = R - f(q^*) > 0$. Because $D_2(q^*||p) > 0$ as $q^* \neq p$, then we have

$$\lim_{n\to\infty} P(\text{dec}(Y^n) = U) = 0.$$

This implies that

$$\lim_{n\to\infty} P(\text{dec}(Y^n) \neq U) = 1 - \lim_{n\to\infty} P(\text{dec}(Y^n) = U) = 1.$$