4. \( R \leq H(\frac{1}{2}) + \frac{\log n + 3}{n} \xrightarrow{n \to \infty} H(\frac{1}{2}) = 1 \text{ bit/source symbol} \), see notes.

b. The scheme is universal for the class of memoryless sources. As such, it does not leverage the memory present in \( X^n \) size window.

c. \( R = \lim_{n \to \infty} \frac{\log n + 2 \log \log n}{n} = \lim_{n \to \infty} \frac{\log n + 2 \log \log n - C}{n} \) see notes

\[ \text{size match}(n) = W = n \]

\[ \frac{W}{n} \]

d. LZ77 does not leverage the memory in the source.

\begin{align*}
2. & \quad \frac{1}{2} \quad 0 \quad \frac{1}{2} \quad 0 \quad \frac{1}{2} + \frac{1}{4} = \frac{3}{4} \\
& \quad \frac{1}{2} \quad \frac{1}{2} \quad 0 \quad \frac{1}{2} \quad 0 \quad \frac{1}{4} \quad \frac{1}{4}
\end{align*}

Maximum rate = \( I(X; Y) = H(Y) - H(Y|X) \)

\[ = H(\frac{1}{4}) - \frac{1}{2} H(Y|X = 1) \]

\[ = H(\frac{1}{4}) - \frac{1}{2} = 0.311 \text{ bits/source symbol} \]
3. a. $E = \{(0,0,0), (1,4,4)\}$ minimizes the prob. of error because the two codewords are at the maximum Hamming distance (=3)

b. $\Pr[\text{error}] = \Pr[\# \text{ bit flips} > 1] = 1 - \Pr[\# \text{ bit flips} \leq 1]
\quad = 1 - (0.2 \times 0.8^2 \times 3 + 0.8^3)
\quad = 0.104$

The above probability follows by considerig the optimal decoder that maps $y^n$ to the closest received codeword in $E$.

c. With 4 codewords, the Hamming dist between codewords $< 3$. The following codebook has Hamming distance 2 and is hence optimal

$E = \{(0,0,0), (0,1,4), (1,0,4), (1,1,0)\}$

d. Optimal decoder:

<table>
<thead>
<tr>
<th>$y^3$</th>
<th>Decoded message $\hat{M}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>000</td>
<td>1 (arbitrarily assigned)</td>
</tr>
<tr>
<td>001</td>
<td>3 (arbitrarily assigned)</td>
</tr>
<tr>
<td>010</td>
<td>2 (arbitrarily assigned)</td>
</tr>
<tr>
<td>011</td>
<td>2 (arbitrarily assigned)</td>
</tr>
<tr>
<td>100</td>
<td>4 (arbitrarily assigned)</td>
</tr>
<tr>
<td>101</td>
<td>3 (arbitrarily assigned)</td>
</tr>
<tr>
<td>110</td>
<td>4 (arbitrarily assigned)</td>
</tr>
<tr>
<td>111</td>
<td>3 (arbitrarily assigned)</td>
</tr>
</tbody>
</table>

$\Pr[\text{Error}] = \Pr[\text{Error} | M = \hat{M}]$
\quad = 1 - (\Pr[\text{no flips}] + \frac{1}{3} \Pr[1 \text{ flip}])
\quad = 1 - (0.8^3 + 0.8^2 \cdot 0.2)
\quad = 0.36$
4. \[ \Pr[\text{Error}] = \Pr[3 \text{ errors}] = \epsilon^3 = 0.2^3 = 0.008 \]

Smaller than for BSC.
We want to show that it is always possible to partition the vertices so that \( c(A \cup B) \geq \frac{m}{2} \).

To this end, we use the probabilistic method:

- Assign each vertex to \( A \) or \( B \) with equal probability and independently.

- Define
  \[
  X_i = \begin{cases} 
    1 & \text{if edge } i \text{ connects a vertex in } A \text{ and one in } B \\
    0 & \text{otherwise}
  \end{cases}
  \]

- Note that \( c(A \cup B) = \sum_{i=1}^{m} X_i \) and
  \[
  \mathbb{E}[X_i] = \Pr[X_i = 1] = \Pr[X_i = 0] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}
  \]

- So, we have
  \[
  \mathbb{E}[c(A \cup B)] = \sum_{i=1}^{m} \mathbb{E}[X_i] = \frac{m}{2}
  \]

\( \Rightarrow \) Since the average of \( c(A \cup B) \) is \( \frac{m}{2} \), there must be at least one partition such that \( c(A \cup B) = \frac{m}{2} \).
6. \( C = \infty \)

It is enough to choose \( X \) such that

\[
\sum_i \frac{1}{\mu_i} > \frac{1}{\mu}
\]

b. With cost constraint

\[
\begin{align*}
I(X;Y) &= h(Y) - h(Y|X) = h(Y) - h(Z) \\
&= h(Y) - \log_2(e\mu) \\
&\leq \log_2(e(\lambda + \mu)) - \log_2(e\mu) = \log_2(1 + \frac{\lambda}{\mu})
\end{align*}
\]

Max entropy lemma
\[
E[Y] = E[X] + E[Z] = \lambda + \mu
\]

7. \( I(X;Y,H) = I(X;H) + I(X;Y|H) \)

\[
= 0 + I(X;Y|H) \geq 0
\]

\[
\Rightarrow I(X;Y|H) \geq I(X;Y)
\]
8. We have

\[ H\left(\frac{1}{3}\right) \leq 10 \left(2^{-0.9}\right) \]

Note required bandwidth ratio
by source

\[ \Rightarrow 10 \geq \frac{H\left(\frac{1}{3}\right)}{0.1} = 10 \times 0.918 = 9.18 \]