Student Number: [redacted] Name: Bryan Hoang

3. (15 points)

(a) **Answer:**

Let $\mathbb{Q}$ be the transition matrix of the Markov source. The transition probabilities making up the entries of $\mathbb{Q}$ are as follows:

$$\begin{aligned} p_{0,0} &= \frac{T-R+\Delta}{T+\Delta}\cdot\frac{T^{-1}}{T^{-1}} = \frac{1-\rho+\delta}{1+\delta} \\ p_{0,1} &= \frac{R}{T+\Delta}\cdot\frac{T^{-1}}{T^{-1}} = \frac{\rho}{1+\delta} \\ p_{1,0} &= \frac{T-R}{T+\Delta}\cdot\frac{T^{-1}}{T^{-1}} = \frac{1-\rho}{1+\delta} \\ p_{1,1} &= \frac{R+\Delta}{T+\Delta}\cdot\frac{T^{-1}}{T^{-1}} = \frac{\rho+\delta}{1+\delta} \\ &\Rightarrow \boxed{\mathbb{Q} = \begin{bmatrix} \frac{1-\rho+\delta}{1+\delta} & \frac{\rho}{1+\delta} \\ \frac{1-\rho}{1+\delta} & \frac{\rho+\delta}{1+\delta} \end{bmatrix}} \end{aligned}$$

Let $\pi = (\pi_0, \pi_1)$ be the stationary distribution of the Markov source. It satisfies the property of remaining unchanged by the operation of transition matrix on it, which gives

$$\begin{aligned} \pi &= \pi\mathbb{Q} \\ (\pi_0, \pi_1) &= (\pi_0, \pi_1)\begin{bmatrix} \frac{1-\rho+\delta}{1+\delta} & \frac{\rho}{1+\delta} \\ \frac{1-\rho}{1+\delta} & \frac{\rho+\delta}{1+\delta} \end{bmatrix} \\ \Rightarrow &\begin{cases} \pi_0 = \frac{1-\rho+\delta}{1+\delta}\pi_0 + \frac{1-\rho}{1+\delta}\pi_1 \\ \pi_1 = \frac{\rho}{1+\delta}\pi_0 + \frac{\rho+\delta}{1+\delta}\pi_1 \end{cases} \\ \Rightarrow &\begin{cases} \cancel{\pi_0} + \cancel{\delta\pi_0} = \cancel{\pi_0} - \rho\pi_0 + \cancel{\delta\pi_0} + \pi_1 - \rho\pi_1 \\ \pi_1 + \cancel{\delta\pi_1} = \rho\pi_0 + \rho\pi_1 + \cancel{\delta\pi_1} \end{cases} \\ \Rightarrow &\begin{cases} \pi_0 = \frac{1-\rho}{\rho}\pi_1 & (1) \\ \pi_1 = \frac{\rho}{1-\rho}\pi_0 & (2) \end{cases} \\ (1) + (2) &\Rightarrow \pi_0 + \pi_1 = \frac{1-\rho}{\rho}\pi_1 + \frac{\rho}{1-\rho}\pi_0 \\ &\Rightarrow 1 = \frac{1-\rho}{\rho}\pi_1 + \frac{\rho}{1-\rho}\pi_0 \qquad (3) \end{aligned}$$

Student Number: [redacted] Name: Bryan Hoang

since π is a probability distribution. Then

$$
\begin{aligned}
(1) \to (3) \Rightarrow 1 &= \frac{1-\rho}{\rho}\pi_1 + \cancel{\frac{\rho}{1-\rho}}\left(\cancel{\frac{1-\rho}{\rho}}\pi_1\right) \\
\rho &= \pi_1 - \rho\pi_1 + \rho\pi_1 \\
\pi_1 &= \rho \qquad (4) \\
(4) \to (1) \Rightarrow \pi_0 &= \frac{1-\rho}{\cancel{\rho}}\cancel{\rho} \\
\pi_0 &= 1 - \rho
\end{aligned}
$$

Therefore, the stationary distribution of the Markov source is $\boxed{\pi = (1-\rho, \rho)}$.

According to Example 3.17 in the textbook and by the fact that the time-invariant Markov source has its initial distribution p_{Z_1} given by the stationary distribution π (i.e., $p_{Z_1} = \pi$), **the Markov source is indeed a stationary process.**

(b) **Answer:**

Since the Markov source is a stationary process, it is also identically distributed. Then by the chain rule for mutual entropy,

$$
\begin{aligned}
I(Z_2; Z_3) &= I(Z_3; Z_2) \\
&= H(Z_3) - H(Z_3|Z_2) \qquad (5) \\
&= H(Z_1) - H(Z_2|Z_1) \quad \because \text{the Markov source is ID and TI} \\
&= -\sum_{a\in\mathcal{X}} p_{Z_1}(a)\log_2 p_{Z_1}(a) + \sum_{a\in\mathcal{X}}\sum_{b\in\mathcal{X}} p_{Z_2,Z_1}(b,a)\log_2 p_{Z_2|Z_1}(b|a) \\
&= -\sum_{a\in\mathcal{X}} p_{Z_1}(a)\log_2 p_{Z_1}(a) + \sum_{a\in\mathcal{X}} p_{Z_1}(a)\sum_{b\in\mathcal{X}} p_{Z_2|Z_1}(b|a)\log_2 p_{Z_2|Z_1}(b|a) \\
&= [-(1-\rho)\log_2(1-\rho) - \rho\log_2\rho] \\
&\quad + \Bigg[(1-\rho)\bigg[\left(1-\frac{\rho}{1+\delta}\right)\log_2\left(1-\frac{\rho}{1+\delta}\right) + \frac{\rho}{1+\delta}\log_2\frac{\rho}{1+\delta}\bigg] \\
&\quad + \rho\bigg[\frac{1-\rho}{1+\delta}\log_2\frac{1-\rho}{1+\delta} + \frac{\rho+\delta}{1+\delta}\log_2\frac{\rho+\delta}{1+\delta}\bigg]\Bigg] \qquad (6)
\end{aligned}
$$

Let

$$
h(p) = -(1-p)\log_2(1-p) - p\log_2 p \qquad (7)
$$

be the binary entropy function. Then the expression for $I(Z_2; Z_3)$ in (6) can be re-written as

$$
\boxed{I(Z_2; Z_3) = h(\rho) - (1-\rho)h\left(\frac{\rho}{1+\delta}\right) - \rho h\left(\frac{\rho+\delta}{1+\delta}\right)}
$$

Student Number: [redacted] Name: Bryan Hoang

Applying the chain rule to the conditional mutual entropy gives

$$\begin{aligned}
I(Z_2; Z_3|Z_1) &= I(Z_3; Z_2|Z_1) \\
&= H(Z_3|Z_1) - H(Z_3|Z_1, Z_2) \\
&= H(Z_3|Z_1) - H(Z_3|Z_2) && \because \text{the Markov source has memory 1} \quad (8) \\
&= H(Z_3|Z_1) - H(Z_2|Z_1) && \because \text{the Markov source is ID and TI} \quad (9)
\end{aligned}$$

$H(Z_2|Z_1)$ has been calculated as the second top level term in (6). To calculate $H(Z_3|Z_1)$, we will need $[p_{Z_3|Z_1}(b|a)] = \mathbb{Q}^2 \ \forall\, a, b \in \mathcal{X}$.

$$\begin{aligned}
\mathbb{Q}^2 &= \begin{bmatrix} 1 - \frac{\rho}{1+\delta} & \frac{\rho}{1+\delta} \\ \frac{1-\rho}{1+\delta} & \frac{\rho+\delta}{1+\delta} \end{bmatrix} \cdot \begin{bmatrix} 1 - \frac{\rho}{1+\delta} & \frac{\rho}{1+\delta} \\ \frac{1-\rho}{1+\delta} & \frac{\rho+\delta}{1+\delta} \end{bmatrix} \\
&= \begin{bmatrix} \left(1 - \frac{\rho}{1+\delta}\right)^2 + \frac{\rho(1-\rho)}{(1+\delta)^2} & \left(1 - \frac{\rho}{1+\delta}\right)\left(\frac{\rho}{1+\delta}\right) + \frac{\rho(\rho+\delta)}{(1+\delta)^2} \\ \left(1 - \frac{\rho}{1+\delta}\right)\left(\frac{1-\rho}{1+\delta}\right) + \frac{(\rho+\delta)(1-\rho)}{(1+\delta)^2} & \frac{\rho(1-\rho)}{(1+\delta)^2} + \left(\frac{\rho+\delta}{1+\delta}\right)^2 \end{bmatrix} \\
&= \begin{bmatrix} 1 - \frac{\rho(2\delta+1)}{(1+\delta)^2} & \frac{\rho(2\delta+1)}{(1+\delta)^2} \\ 1 - \frac{\rho(2\delta+1)+\delta^2}{(1+\delta)^2} & \frac{\rho(2\delta+1)+\delta^2}{(1+\delta)^2} \end{bmatrix} \\
&= \begin{bmatrix} p_{Z_3|Z_1}(0|0) & p_{Z_3|Z_1}(1|0) \\ p_{Z_3|Z_1}(0|1) & p_{Z_3|Z_1}(1|1) \end{bmatrix} \qquad (10)
\end{aligned}$$

Then

$$\begin{aligned}
&H(Z_3|Z_1) \\
&= -\sum_{a\in\mathcal{X}} \sum_{b\in\mathcal{X}} p_{Z_3,Z_1}(b,a) \log_2 p_{Z_3|Z_1}(b|a) \\
&= -\sum_{a\in\mathcal{X}} p_{Z_1}(a) \sum_{b\in\mathcal{X}} p_{Z_3|Z_1}(b|a) \log_2 p_{Z_3|Z_1}(b|a) \\
&= (1-\rho)\Big[-p_{Z_3|Z_1}(0|0) \log_2 p_{Z_3|Z_1}(0|0) - p_{Z_3|Z_1}(1|0) \log_2 p_{Z_3|Z_1}(1|0)\Big] \\
&\quad + \rho\Big[-p_{Z_3|Z_1}(0|1) \log_2 p_{Z_3|Z_1}(0|1) - p_{Z_3|Z_1}(1|1) \log_2 p_{Z_3|Z_1}(1|1)\Big] \\
&= (1-\rho)h\left(\frac{\rho(2\delta+1)}{(1+\delta)^2}\right) + \rho h\left(\frac{\rho(2\delta+1)+\delta^2}{(1+\delta)^2}\right) \qquad \text{by (7) and (10)} \quad (11)
\end{aligned}$$

Thus, combining the results from (9), (11), (6), and (7) gives

$$\begin{aligned}
&= I(Z_2; Z_3|Z_1) \\
&= H(Z_3|Z_1) - H(Z_2|Z_1) \\
&= \left[(1-\rho)h\left(\frac{\rho(2\delta+1)}{(1+\delta)^2}\right) + \rho h\left(\frac{\rho(2\delta+1)+\delta^2}{(1+\delta)^2}\right)\right] - \left[(1-\rho)h\left(\frac{\rho}{1+\delta}\right) + \rho h\left(\frac{\rho+\delta}{1+\delta}\right)\right] \\
&\boxed{= (1-\rho)\left[h\left(\frac{\rho(2\delta+1)}{(1+\delta)^2}\right) - h\left(\frac{\rho}{1+\delta}\right)\right] + \rho\left[h\left(\frac{\rho(2\delta+1)+\delta^2}{(1+\delta)^2}\right) - h\left(\frac{\rho+\delta}{1+\delta}\right)\right]}
\end{aligned}$$

Student Number: [redacted] Name: Bryan Hoang

(c) **Answer:**

Proof. From (5) and (8), we have

$$\begin{aligned} I(Z_2; Z_3) - I(Z_2; Z_3|Z_1) &= H(Z_3) - H(Z_3|Z_2) - H(Z_3|Z_1) + H(Z_3|Z_2) \\ &= H(Z_3) - H(Z_3|Z_1) \\ &\geq 0 \end{aligned}$$

since conditiong reduces entropy, and Z_3 is not independent of Z_1. □