Lösungen zu Übungsblatt 9

9.1

Wir unterscheiden folgende Fälle entsprechend dem Wert von a:

$a = 0$: In diesem Fall ist $X = Y$ und $C = \max I(X;Y) = \max H(X) = 1$.

$a \in \{\pm 1\}$: Im Fall $a = 1$ kann Y die drei verschiedenen Werte $0, 1, 2$ annehmen und die bedingten Verteilungen sind gegeben durch $P(Y = 0|X = 0) = P(Y = 1|X = 0) = \frac{1}{2}$ und $P(Y = 1|X = 1) = P(Y = 2|X = 1) = \frac{1}{2}$. Der Kanal ist damit ein binärer Auslöschungskanal mit Auslöschungswahrscheinlichkeit $\varepsilon = \frac{1}{2}$, dessen Kapazität $C = 1 - \varepsilon = \frac{1}{2}$ ist (siehe Vorlesung). Im Fall $a = -1$ erhält man analog $C = \frac{1}{2}$.

$a \notin \{0, \pm 1\}$: In diesem Fall kann Y die vier verschiedenen Werte $0, 1, a, 1 + a$ annehmen. Da X bei gegebenem Y eindeutig bestimmt ist, gilt $H(X|Y) = 0$ und damit $C = \max I(X;Y) = \max H(X) = 1$.

9.2

Wir bezeichnen mit A die Eingabe und mit B die Ausgabe.

a) Gemäss Bayes' Regel ist

$$P(A = a|B = b) = \frac{P(A = a)P(B = b|A = a)}{P(B = b)}.$$

Es gilt $P(A = a_0) = \frac{1}{2}$, $P(B = b|A = a_0) = (1 - \varepsilon)^i \varepsilon^{3-i}$ und $P(B = b) = \frac{1}{2}(1 - \varepsilon)^i \varepsilon^{3-i} + \frac{1}{2}(1 - \varepsilon)^{3-i} \varepsilon^i$, wobei wir für gegebenes b mit $i \in \{0, 1, 2, 3\}$ die Anzahl der Nullen in b bezeichnen. Damit folgt

$$P(A = a_0|B = b) = \frac{\frac{1}{2}(1 - \varepsilon)^i \varepsilon^{3-i}}{\frac{1}{2}(1 - \varepsilon)^i \varepsilon^{3-i} + \frac{1}{2}(1 - \varepsilon)^{3-i} \varepsilon^i}.$$

b) Wegen $\varepsilon < \frac{1}{2}$ gilt für den beschriebenen Decodierer $\hat{A}(b) = \arg\max_a P(b|a)$ für jedes b (d.h. $\hat{A}$ ist ein *Maximum-Likelihood-Decodierer*); da die Eingangssymbole hier gleichverteilt sind, ist $\arg\max_a P(b|a) = \arg\max_a P(a|b)$ für jedes b, und damit ist $\hat{A}$ auch ein Maximum-A-Posteriori-Decodierer (siehe Aufgabe 3) und minimiert somit die Fehlerwahrscheinlichkeit. (Alternativ kann man mit Hilfe der in (a) berechneten Wahrscheinlichkeiten argumentieren.)

c) Die Fehlerwahrscheinlichkeit ist gleich der Wahrscheinlichkeit, dass zwei oder drei Bits eines Dreierblocks falsch übertragen werden, d.h.

$$P(\hat{A} \neq A) = \binom{3}{2}\varepsilon^2(1 - \varepsilon) + \varepsilon^3 = 3\varepsilon^2(1 - \varepsilon) + \varepsilon^3.$$

d) Der optimale Decodierer ist (wie immer) der Maximum-A-Posteriori-Decodierer, der in unserem Fall ein gegebenes $b \in \{0,1\}^{2n+1}$ als a_0 decodiert, wenn in b mindestens $n+1$ Nullen vorkommen, und als a_1 sonst. Die Fehlerwahrscheinlichkeit ist gleich der Wahrscheinlichkeit, dass die Anzahl falsch übertragener bits $> n+1$ ist, d.h.

$$P(\hat{A} \neq A) = \sum_{i=n+1}^{2n+1} \binom{2n+1}{i} \varepsilon^i (1-\varepsilon)^{2n+1-i}$$

Die erwartete Anzahl falsch übertragener bits ist dagegen $(2n+1)\varepsilon < n+1$; da die Wahrscheinlichkeit einer Abweichung von diesem Erwartungswert wegen des Gesetzes der grossen Zahlen für $n \to \infty$ gegen 0 strebt, strebt auch die Fehlerwahrscheinlichkeit für $n \to \infty$ gegen 0.

9.3

a) Wir können die Fehlerwahrscheinlichkeit für einen Codierer $\hat{X} = g(Y)$ schreiben als

$$P_e(g) = \sum_y P(y) \sum_{x \neq g(y)} P(x|y) = \sum_y P(y)(1 - P(g(y)|y)),$$

und damit gilt

$$P_e(g) = \sum_y P(y)(1 - P(g(y)|y)) \geq \sum_y P(y)(1 - P(g_{MAP}(y)|y)) = P_e(g_{MAP}),$$

denn nach Definition von g_{MAP} ist $P(g_{MAP}(y)|y) \geq P(x|y)$ für alle x, y.

b) Es gilt

$$H(Z) = -\sum_z P_Z(z) \log P_Z(z) \geq -\sum_z P_Z(z) \log P_Z(z^*) = -\log P_Z(z^*),$$

da die Funktion $p \mapsto -\log p$ monoton fällt. Durch Anwendung auf die bedingte Verteilung $P_{X|Y=y}$ und Verwendung der Definition von g_{MAP} folgt

$$H(X|Y=y) \geq -\log P(\hat{X} = X|Y=y).$$

c) Für $\hat{X} = g_{MAP}(Y)$ gilt

$$\begin{aligned}
P(\hat{X} \neq X) &= \sum_y P(y)(1 - P(g_{MAP}(y)|y)) \\
&= 1 - \sum_y P(y) P(g_{MAP}(y)|y) \\
&= 1 - \sum_y P(y) 2^{\log P(g_{MAP}(y)|y)} \\
&\leq 1 - \sum_y P(y) 2^{-H(X|Y=y)} \\
&\leq 1 - 2^{-\sum_y P(y) H(X|Y=y)} \\
&= 1 - 2^{H(X|Y)},
\end{aligned}$$

wobei die erste Ungleichung aus dem Ergebnis in (b) folgt und die zweite aus der in der Aufgabenstellung angegebenen Anwendung der Jensen-Ungleichung.