Correction to Exercise 01:

Recall that : $P_X(x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = -1 \end{cases}$

and $P_W(w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{w^2}{2\sigma^2}\right).$

Let $Y = X + W.$

Question 1:

- X is a discrete random variable (only two possible values)
- W is a continuous random variable ($P_W(w)$ is defined on the whole set of real numbers).
- Hence, Y is a continuous random variable

$\Rightarrow$ X is defined by a pmf while Y and W are defined by their pdf.

Question 2:

Let $y \in \mathbb{R}$ and $x \in \{-1, 1\}$.

We have that

$$\begin{aligned} P_{Y|X}(y|x) &= \mathbb{P}(Y = y \mid X = x). \\ &= \mathbb{P}(W = y - x \mid X = x) \\ &= P_{W|X}(y - x \mid x). \\ &\overset{(a)}{=} P_W(y - x) \end{aligned}$$

(a) is due to the fact : W and X are independent

Question 3:

Let $y \in \mathbb{R}$, we have that :

$$P_Y(y) = \sum_{x \in \{-1,1\}} P_{YX}(y,x)$$ $\longrightarrow$ marginal from joint laws.

$$= \sum_{x \in \{-1,1\}} P_X(x)\, P_{Y|X}(y|x)$$

$$= P_X(-1)\, P_{Y|X}(y|-1) + P_X(1)\, P_{Y|X}(y|1)$$

$$= (1-p)\, P_W(y+1) + p\, P_W(y-1).$$

which proves the identity.

Question 4:

One can see that

$$P_Y(y) = p\, \underbrace{P_W(y-1)}_{\text{Gaussian centered around } 1} + (1-p)\, \underbrace{P_W(y+1)}_{\text{Gaussian centered around } -1}$$

mixture

Hence, the probability distribution of y is a mixture of two Gaussian distribution centered around $+1$ and -1 respectively.

Question 5:

a) Mean of Y:

$$\mathbb{E}(Y) = \mathbb{E}(X) + \mathbb{E}(W).$$

$\rightarrow$ by definition, W is a centered Gaussian noise

$$\mathbb{E}(W) = 0$$

$\longrightarrow$ As for the expectation of X, we have.

$$\mathbb{E}(X) = \sum_{x \in \{-1,1\}} x\, P_X(x) = 1 \cdot p - 1 \cdot (1-p)$$

$$= 2p - 1.$$

$\Rightarrow$ Hence, the expectation of Y is given by

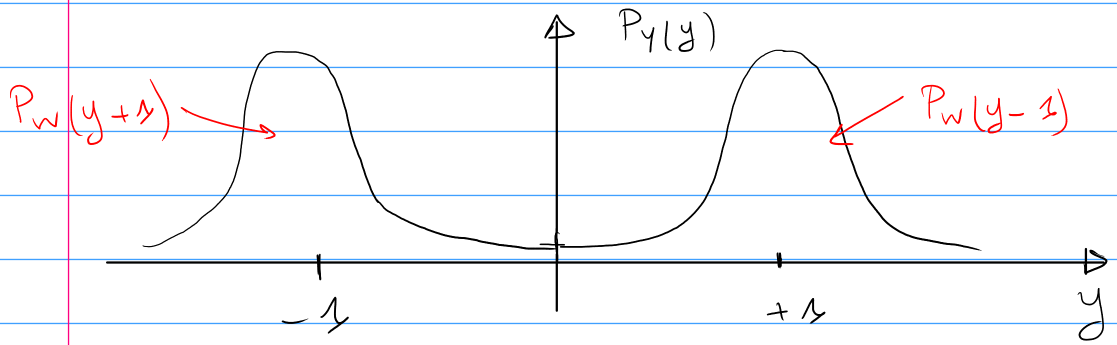
$$\mathbb{E}(Y) = \mathbb{E}(X) = 2p - 1.$$

Question 6:

Graphical representation: Assume $p = 1-p = 1/2$

then, $\mathbb{E}(Y) = 0$ and

$$P_Y(y) = \frac{1}{2} P_W(y-1) + \frac{1}{2} P_W(y+1).$$



Question 7:

Let $p = 1/2$ $\quad (P_X(1) = P_X(-1) = 0{,}5)$.

We have that:

$$\mathcal{I}(X;Y) = H(X) - H(X|Y) \qquad \text{(discrete entropy).}$$

Let us compute:

$$\rightarrow H(X) = \sum_{x \in \{-1,1\}} - P_X(x) \log_2 (P_X(x)).$$

$$= -\frac{1}{2} \cdot \log_2(1/2) - \frac{1}{2} \log_2(1/2)$$

$$= 1/2 + 1/2 = 1.$$

$$\rightarrow H(X|Y) = \int_{y \in \mathbb{R}} P_Y(y) \, H(X|Y=y) \, dy$$

where $H(X|Y=y)$ is the uncertainty on X knowing that we have already observed a value of $Y=y$.

By definition, we have that:

$$H(X|Y=y) = -\sum_{x \in \{-1,1\}} P_{X|Y}(x|y) \log_2 (P_{X|Y}(x|y))$$

$$H(X|Y=y) = -P_{X|Y}(-1|y)\log_2\left(P_{X|Y}(-1|y)\right)$$
$$- P_{X|Y}(1|y)\log_2\left(P_{X|Y}(1|y)\right).$$

Note at this point that

$$P_{X|Y}(-1|y) + P_{X|Y}(1|y) = \sum_{x\in\{-1,1\}} P_{X|Y}(x|y) = 1$$

by definition of a conditional law.

Hence : $P_{X|Y}(-1|y) = 1 - P_{X|Y}(1|y)$.

Thus : $$H(X|Y=y) = -\left[1-P_{X|Y}(1|y)\right]\log_2\left(1-P_{X|Y}(1|y)\right)$$
$$- P_{X|Y}(1|y)\log_2\left(P_{X|Y}(1|y)\right)$$
$$= H_2\left(P_{X|Y}(1|y)\right).$$

Let us go back to the original mutual information.

$$I(X;Y) = H(X) - H(X|Y)$$
$$= 1 - \int_y P_Y(y)\, H_2\left(P_X(1|y)\right)dy.$$

Question 8:

A mutual information has no physical unit, however it is measured in bits/s/Hz (if log is in base 2) and nats (if log is in base e).

It models the amount of information that is shared between two random variables. These random variables could be the input/output of a given channel, in which case the mutual information measures the maximal rate that could be transmitted over this channel.

Question 9:

Let us show that $I(X;Y) \leq 1$.

From the previous questions we have that

Method 1 $$I(X;Y) = 1 - \underbrace{\int_y \underbrace{P_Y(y)}_{\geq 0} \underbrace{H_2\left(P_{X|Y}(1|y)\right)}_{\geq 0}}_{\geq 0} dy.$$

Method 2: The term $H(X|Y)$ being a discrete entropy, it is positive by definition.

Hence $$I(X;Y) = H(X) - H(X|Y)$$
$$\leq H(X) = 1.$$

The fact that the mutual information is smaller than 1 implies that, no matter how good the channel $P_{Y|X}$ is, the max information rate which can transmit through the channel is no larger than the entropy of the source which is 1 bit.

Question 10:

At high $\frac{E_b}{N_0}$, σ^2 is very small, which means very little noise. In this case $Y \approx X$, and hence
$$I(X;Y) = H(X) - \underbrace{H(X|Y)}_{=0},$$
which yields a mutual information at its maximum value 1.

Question 11 :

At low $\frac{E_b}{N_0}$, σ^2 is very high, and thus, the noise is very strong. As a consequence the channel output Y seems almost independent from X, and hence

$$I(X;Y) \gtrsim 0.$$

Note that the mutual information being positive, its minimum value is 0.

Question 12 :

As $\frac{E_b}{N_0}$ increases, σ^2 decreases, hence the channel becomes of better quality and thus, its mutual information increases.

Question 14 :

If X was a 4-valued random variable, then its entropy would be bounded by

$$H(X) \leq \log_2(4) = 2 \text{ bits}$$

and equal to 2 if X was uniformly distributed.

Hence, $I(X;Y) \leq H(X) \leq 2$ bits,

the maximum possibly achievable information rate would be 2 bits.