Examples of Noisy Channels

- analogue telephone line over which two modems communicate digital information
- a teacher mumbling at the board
- radio communication link between “curiosity” on Mars and earth
- reproducing cells, where daughter cells contain DNA from the parents’ cell
- a disk drive
- …
Discrete Channels

Def: A discrete channel is denoted by \((\mathcal{X}, P_{Y|X}, \mathcal{Y})\) where \(\mathcal{X}\) is a finite input set, \(\mathcal{Y}\) is a finite output set and \(P_{Y|X}\) is a conditional probability distribution, d.h.

\[
\forall x \in \mathcal{X} \quad \forall y \in \mathcal{Y} : P_{Y|X}(y|x) \geq 0
\]

\[
\forall x \in \mathcal{X} : \sum_{y} P_{Y|X}(y|x) = 1
\]

\(P_{Y|X}(y|x) = \text{the probability that the channel outputs } y\) when given \(x\) as input
Figure 1.5. A binary data sequence of length 10 000 transmitted over a binary symmetric channel with noise level $f = 0.1$. [Dilbert image Copyright © 1997 United Feature Syndicate, Inc., used with permission.]
| Received sequence $r$ | Likelihood ratio $\frac{P(r|s=1)}{P(r|s=0)}$ | Decoded sequence $\hat{s}$ |
|----------------------|---------------------------------|-----------------|
| 000                  | $\gamma^{-3}$                   | 0               |
| 001                  | $\gamma^{-1}$                   | 0               |
| 010                  | $\gamma^{-1}$                   | 0               |
| 100                  | $\gamma^{-1}$                   | 0               |
| 101                  | $\gamma^{1}$                    | 1               |
| 110                  | $\gamma^{1}$                    | 1               |
| 011                  | $\gamma^{1}$                    | 1               |
| 111                  | $\gamma^{3}$                    | 1               |

**Algorithm 1.9.** Majority-vote decoding algorithm for $R_3$. Also shown are the likelihood ratios (1.23), assuming the channel is a binary symmetric channel; $\gamma \equiv (1 - f)/f$. 

*Book by David MacKay*
<table>
<thead>
<tr>
<th>s</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>000</td>
<td>000</td>
<td>111</td>
<td>000</td>
<td>111</td>
<td>111</td>
<td>000</td>
</tr>
<tr>
<td>n</td>
<td>000</td>
<td>001</td>
<td>000</td>
<td>000</td>
<td>101</td>
<td>000</td>
<td>000</td>
</tr>
<tr>
<td>r</td>
<td>000</td>
<td>001</td>
<td>111</td>
<td>000</td>
<td>010</td>
<td>111</td>
<td>000</td>
</tr>
<tr>
<td>(\hat{s})</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Corrected errors: *  
Undetected errors: *
Exercise 1.2: 

Figure 1.11. Transmitting 10,000 source bits over a binary symmetric channel with $f = 10\%$ using a repetition code and the majority vote decoding algorithm. The probability of decoded bit error has fallen to about 3%; the rate has fallen to 1/3.
Def: A \((M,n)\)-code for the channel \((\mathcal{X}, P_{Y|X}, \mathcal{Y})\) consists of

1. message set: \([M] = \{1, 2, \ldots, M\}\)
2. encoding function: \(e : [M] \to \mathcal{X}^n\)
   - codebook: \(\{e(1), e(2), \ldots, e(M)\}\)
3. deterministic decoding function assigning a guess to each possible received vector \(d : \mathcal{Y}^n \to [M]\)

The rate of a \((M,n)\)-code denotes the transmitted bits per channel use

\[ R := \frac{\log M}{n} \]
Figure 1.12. Error probability $p_b$ versus rate for repetition codes over a binary symmetric channel with $f = 0.1$. The right-hand figure shows $p_b$ on a logarithmic scale. We would like the rate to be large and $p_b$ to be small.
Table 1.14. The sixteen codewords \( \{ t \} \) of the \((7, 4)\) Hamming code. Any pair of codewords differ from each other in at least three bits.
Figure 1.17. Transmitting 10000 source bits over a binary symmetric channel with $f = 10\%$ using a (7, 4) Hamming code. The probability of decoded bit error is about 7\%.
Figure 1.18. Error probability $p_b$ versus rate $R$ for repetition codes, the $(7, 4)$ Hamming code and BCH codes with blocklengths up to 1023 over a binary symmetric channel with $f = 0.1$. The righthand figure shows $p_b$ on a logarithmic scale.
Figure 1.19. Shannon’s noisy-channel coding theorem. The solid curve shows the Shannon limit on achievable values of \((R, p_b)\) for the binary symmetric channel with \(f = 0.1\). Rates up to \(R = C\) are achievable with arbitrarily small \(p_b\). The points show the performance of some textbook codes, as in figure 1.18.

The equation defining the Shannon limit (the solid curve) is \(R = C/(1 - H_2(p_b))\), where \(C\) and \(H_2\) are defined in equation (1.35).
$C \approx 0.53$. Let us consider what this means in terms of noisy disk drives. The repetition code $R_3$ could communicate over this channel with $p_b = 0.03$ at a rate $R = 1/3$. Thus we know how to build a single gigabyte disk drive with $p_b = 0.03$ from three noisy gigabyte disk drives. We also know how to make a single gigabyte disk drive with $p_b \approx 10^{-15}$ from sixty noisy one-gigabyte drives (exercise 1.3, p.8). And now Shannon passes by, notices us juggling with disk drives and codes and says:

‘What performance are you trying to achieve? $10^{-15}$? You don’t need sixty disk drives – you can get that performance with just two disk drives (since $1/2$ is less than 0.53). And if you want $p_b = 10^{-18}$ or $10^{-24}$ or anything, you can get there with two disk drives too!’
\( C \sim 0.53 \). Let us consider what this means in terms of noisy disk drives. The repetition code \( R_3 \) could communicate over this channel with \( p_b = 0.03 \) at a rate \( R = 1/3 \). Thus we know how to build a single gigabyte disk drive with \( p_b = 0.03 \) from three noisy gigabyte disk drives. We also know how to make a single gigabyte disk drive with \( p_b \sim 10^{-15} \) from sixty noisy one-gigabyte drives (exercise 1.3, p.8). And now Shannon passes by, notices us juggling with disk drives and codes and says:

‘What performance are you trying to achieve? 10\(^{-15}\)? You don’t need sixty disk drives – you can get that performance with just two disk drives (since 1/2 is less than 0.53). And if you want \( p_b = 10^{-18} \) or \( 10^{-24} \) or anything, you can get there with two disk drives too!’

[Strictly, the above statements might not be quite right, since, as we shall see, Shannon proved his noisy-channel coding theorem by studying sequences of block codes with ever-increasing blocklengths, and the required blocklength might be bigger than a gigabyte (the size of our disk drive), in which case, Shannon might say ‘well, you can’t do it with those tiny disk drives, but if you had two noisy terabyte drives, you could make a single high-quality terabyte drive from them’.]