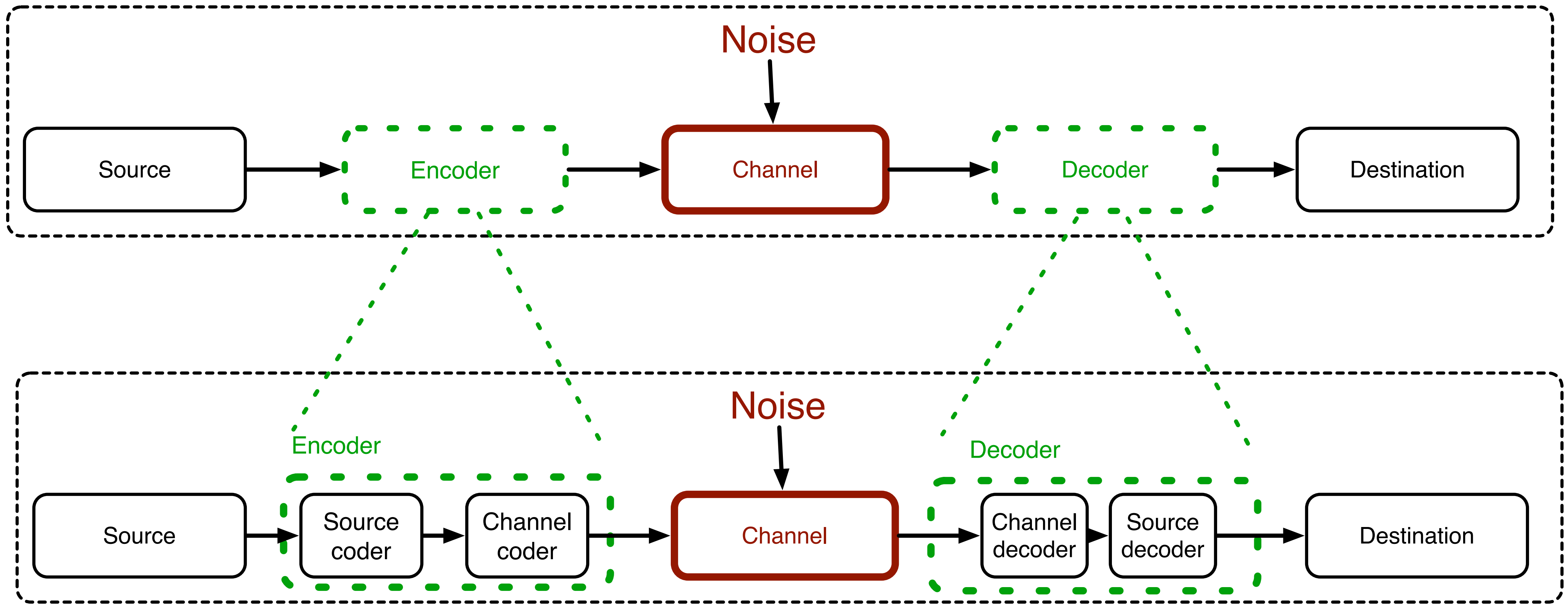
Source-channel separation



- For (time-varying) DMC we can design the source encoder and channel coder separately and still get optimum performance
- Not true for:
  - Correlated Channel and Source
  - Multiple access with correlated sources
  - Broadcast channel

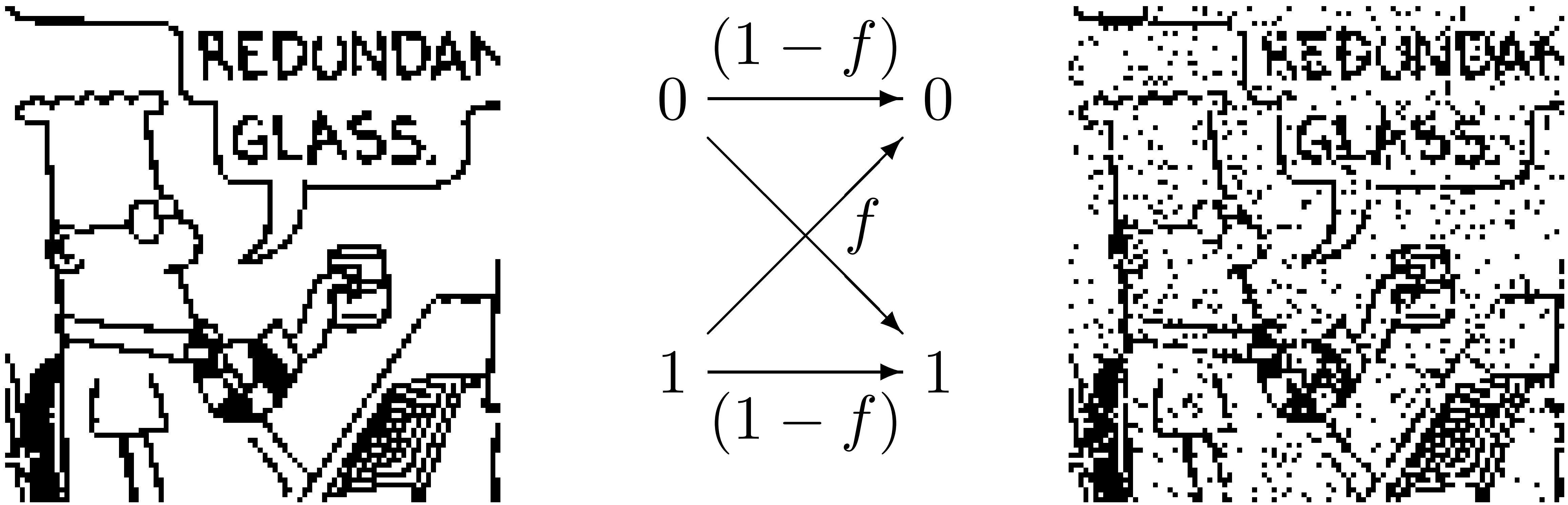


Figure 1.5. A binary data sequence of length 10 000 transmitted over a binary symmetric channel with noise level $f = 0.1$. [Dilbert image Copyright©1997 United Feature Syndicate, Inc., used with permission.]

Book by David MacKay

Received sequence $\mathbf{r}$	Likelihood ratio $\frac{P(\mathbf{r} \mid s=1)}{P(\mathbf{r} \mid s=0)}$	Decoded sequence $\hat{\mathbf{s}}$
000	γ^{-3}	0
001	γ^{-1}	0
010	γ^{-1}	0
100	γ^{-1}	0
101	γ^{1}	1
110	γ^{1}	1
011	γ^{1}	1
111	γ^{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

$$
\begin{array}{rccccccc}
\mathbf{s} & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
\mathbf{t} & \overbrace{0\,0\,0} & \overbrace{0\,0\,0} & \overbrace{1\,1\,1} & \overbrace{0\,0\,0} & \overbrace{1\,1\,1} & \overbrace{1\,1\,1} & \overbrace{0\,0\,0} \\
\mathbf{n} & 0\,0\,0 & 0\,0\,1 & 0\,0\,0 & 0\,0\,0 & 1\,0\,1 & 0\,0\,0 & 0\,0\,0 \\
\hline
\mathbf{r} & \underbrace{0\,0\,0} & \underbrace{0\,0\,1} & \underbrace{1\,1\,1} & \underbrace{0\,0\,0} & \underbrace{0\,1\,0} & \underbrace{1\,1\,1} & \underbrace{0\,0\,0} \\
\hat{\mathbf{s}} & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
\text{corrected errors} & & \star & & & & & \\
\text{undetected errors} & & & & & \star & &
\end{array}
$$

Book by David MacKay

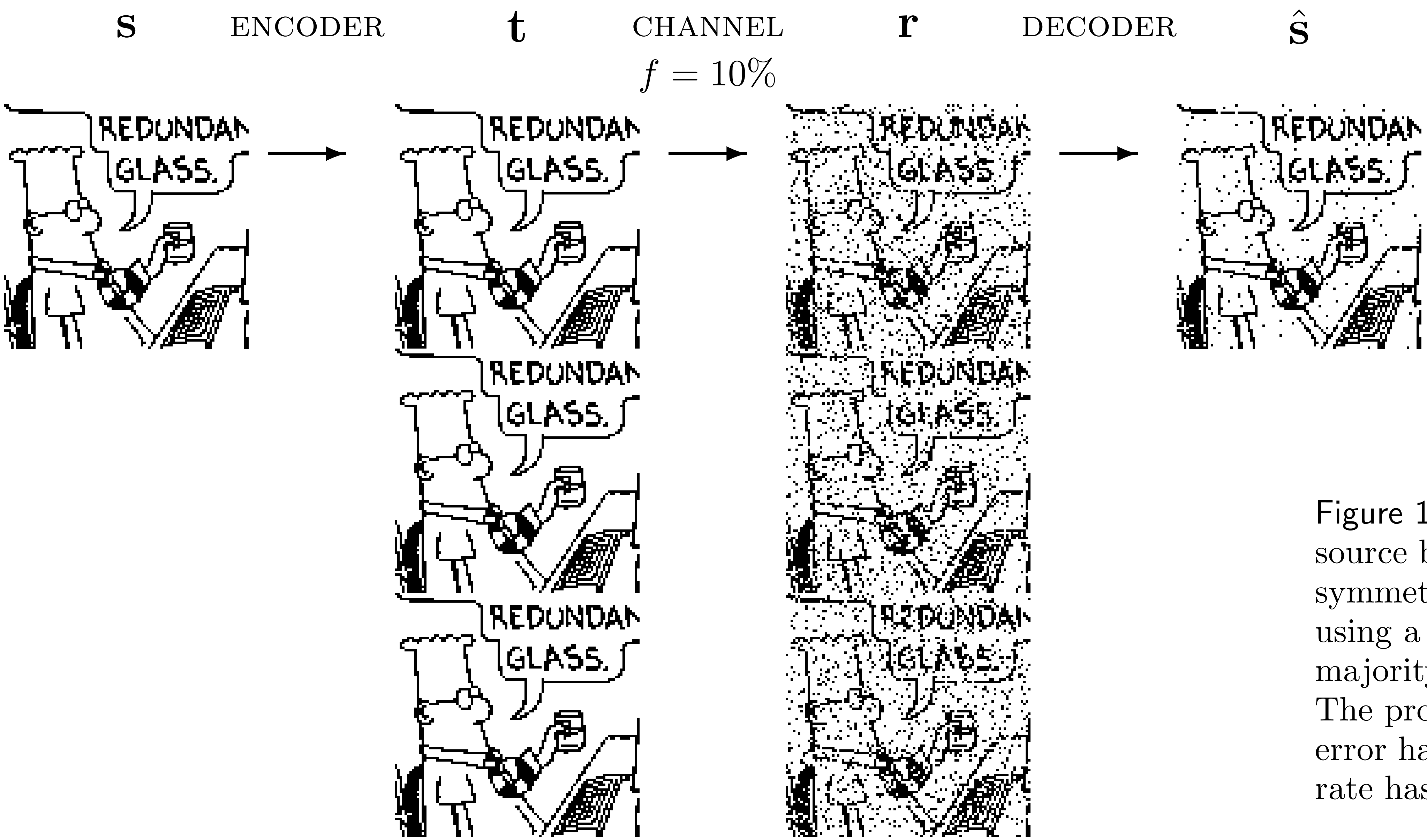


Figure 1.11. Transmitting 10 000 source bits over a binary symmetric channel with $f = 10\%$ using a repetition code and the majority vote decoding algorithm. The probability of decoded bit error has fallen to about 3%; the rate has fallen to 1/3.

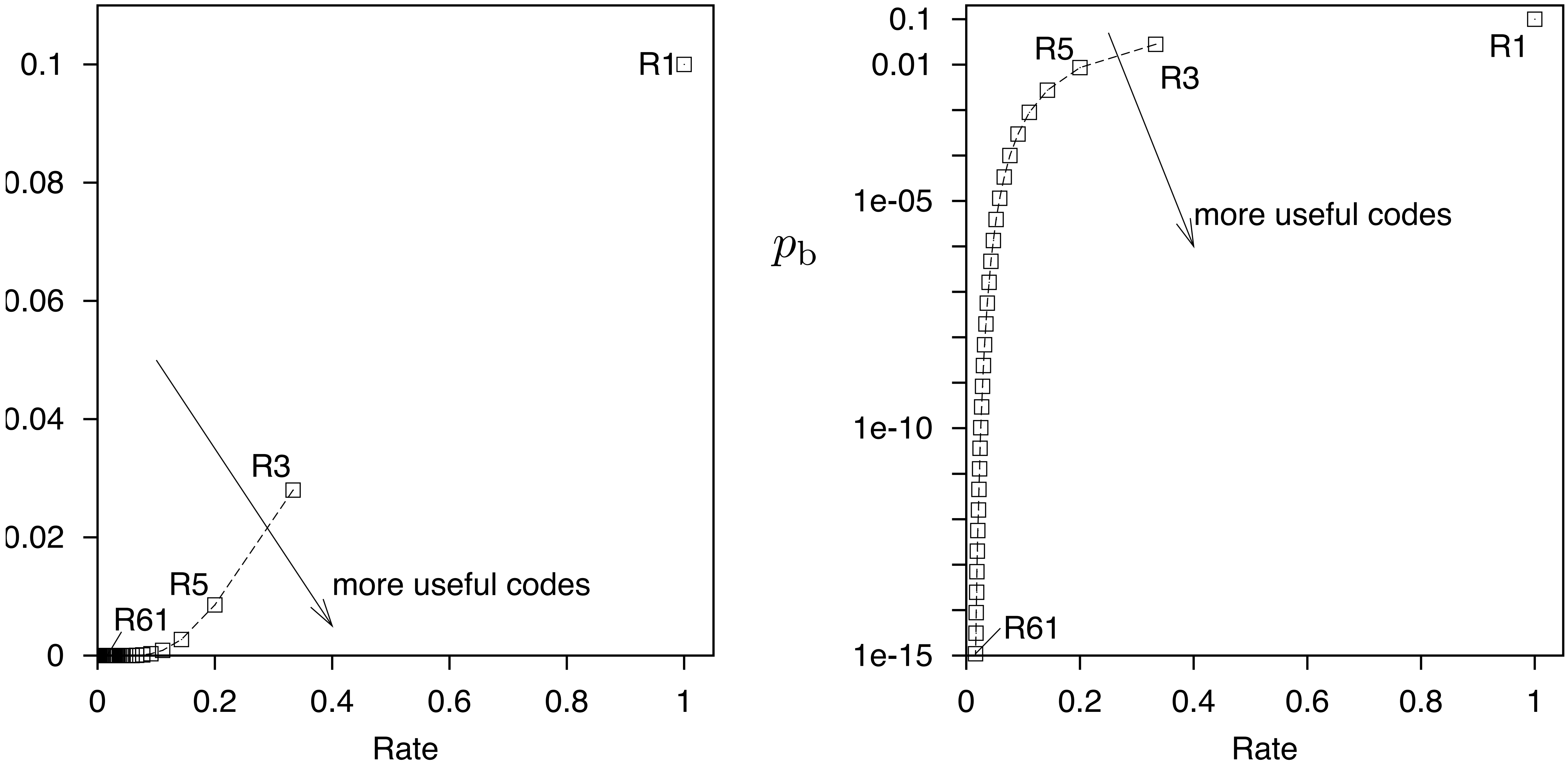


Figure 1.12. Error probability p_b versus rate for repetition codes over a binary symmetric channel with $f = 0.1$. The right-hand figure shows p_b on a logarithmic scale. We would like the rate to be large and p_b to be small.

Book by David MacKay

s	**t**	**s**	**t**	**s**	**t**	**s**	**t**
0000	0000000	0100	0100110	1000	1000101	1100	1100011
0001	0001011	0101	0101101	1001	1001110	1101	1101000
0010	0010111	0110	0110001	1010	1010010	1110	1110100
0011	0011100	0111	0111010	1011	1011001	1111	1111111

Table 1.14. The sixteen codewords $\{\mathbf{t}\}$ of the $(7, 4)$ Hamming code. Any pair of codewords differ from each other in at least three bits.

Book by David MacKay

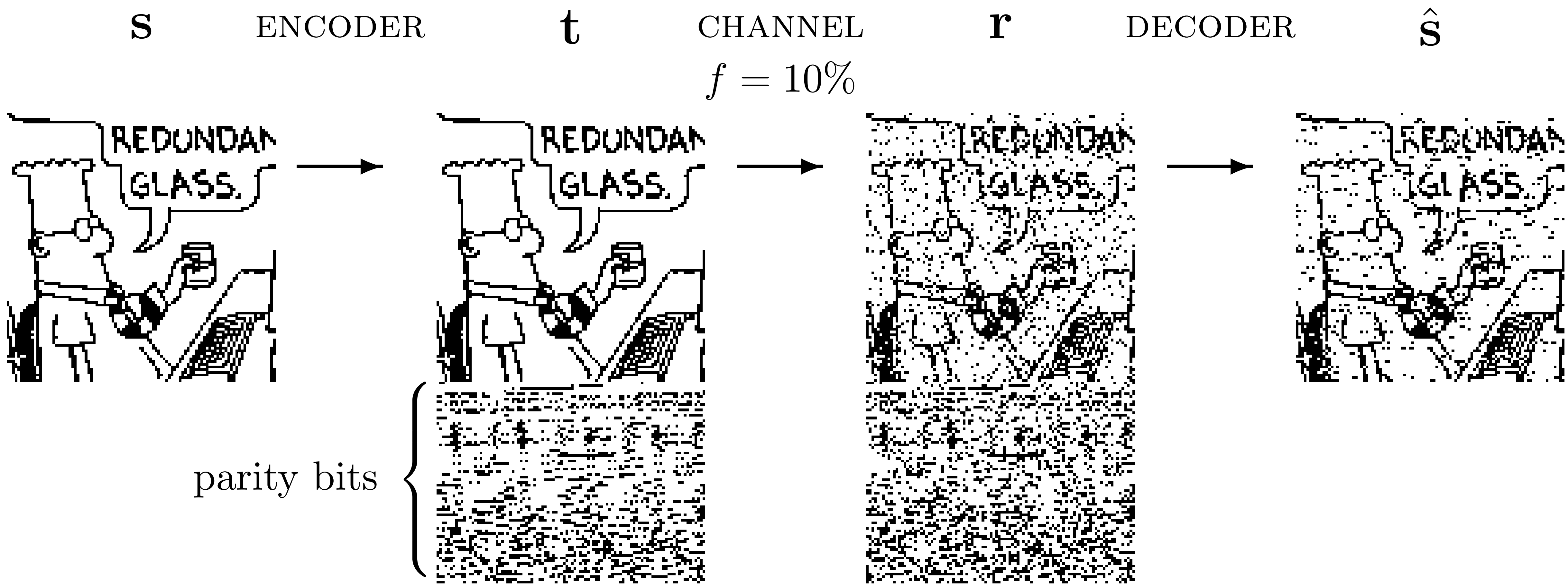


Figure 1.17. Transmitting 10 000 source bits over a binary symmetric channel with $f = 10\%$ using a $(7, 4)$ Hamming code. The probability of decoded bit error is about 7%.

Book by David MacKay

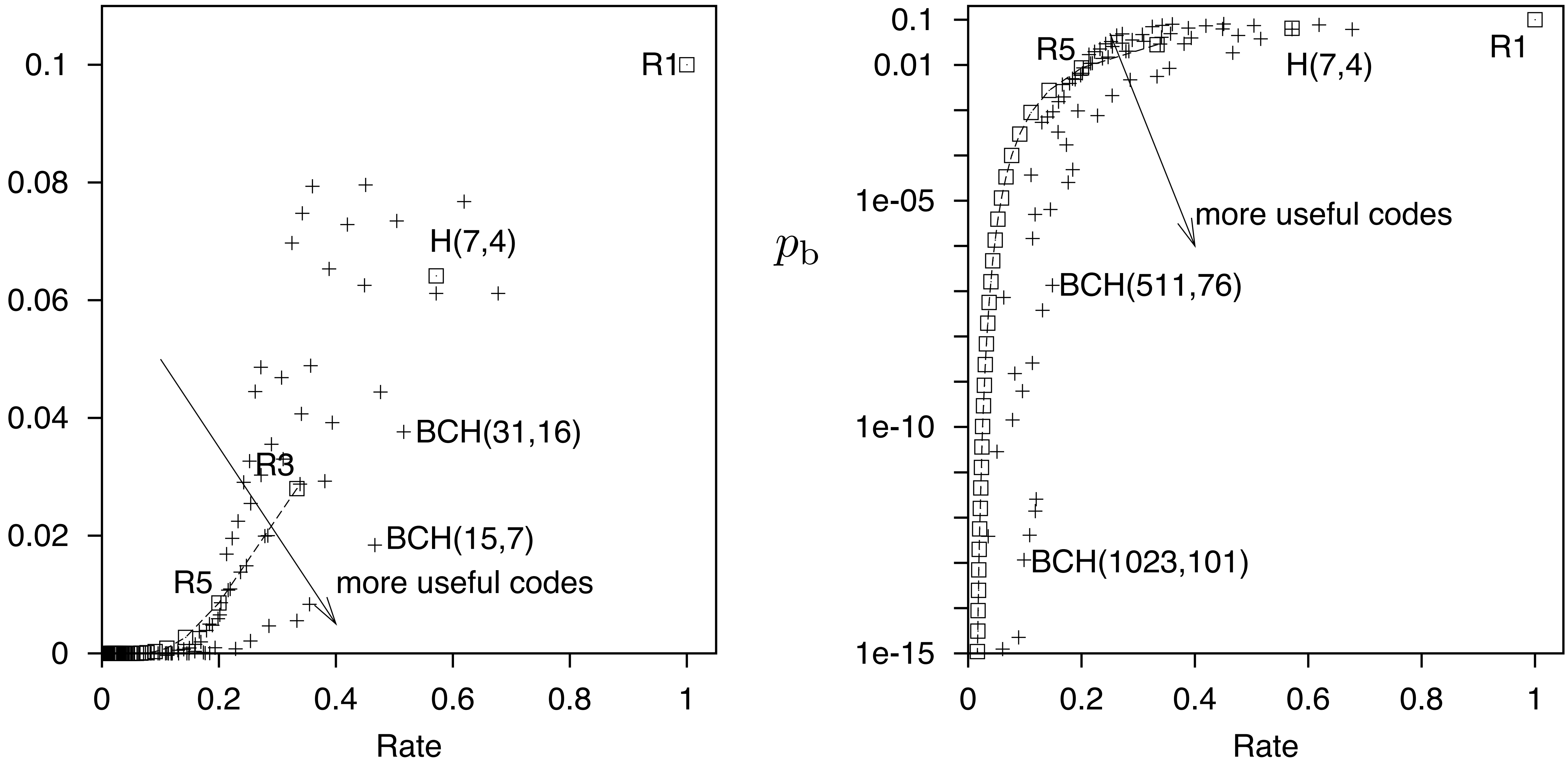
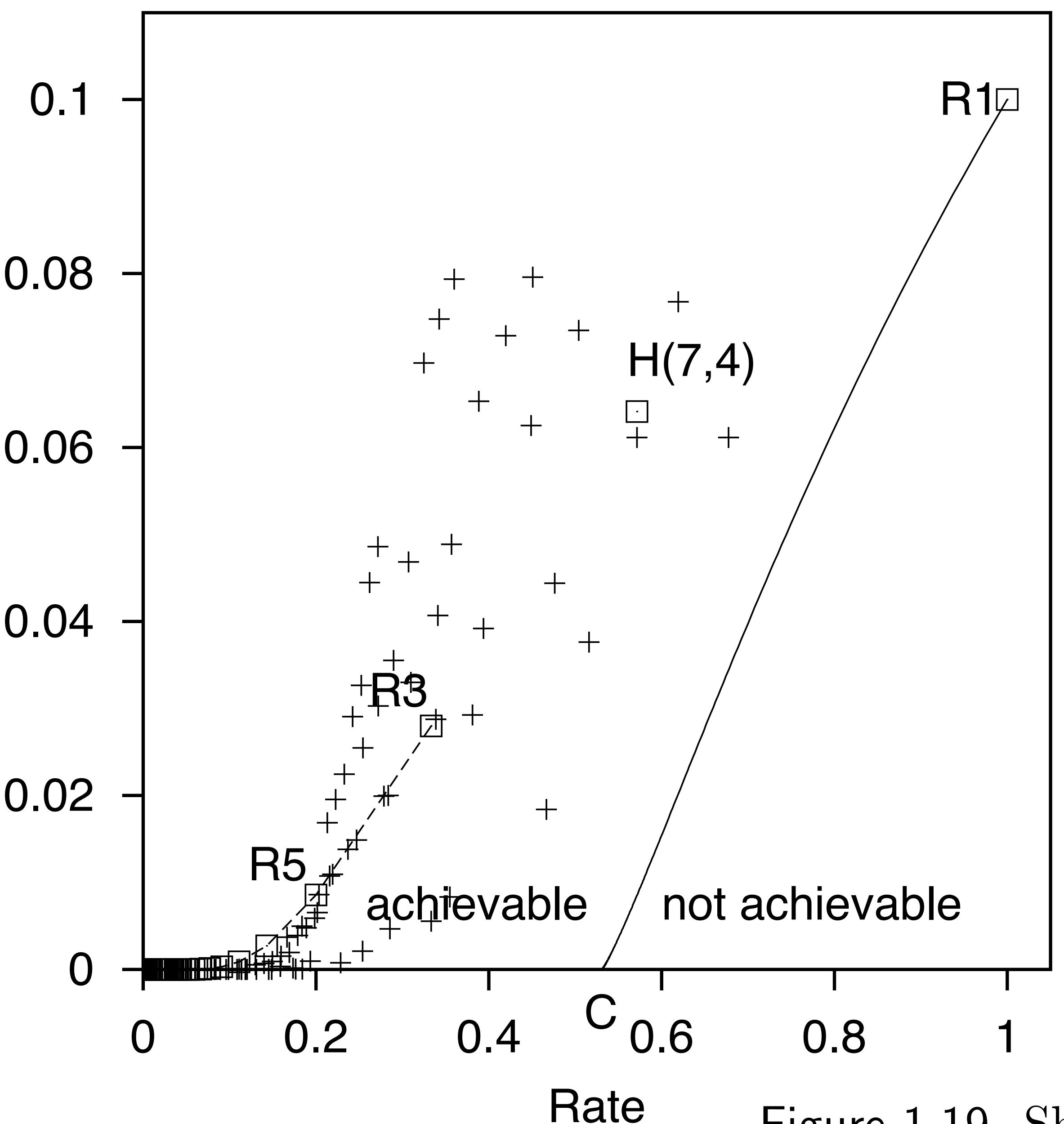


Figure 1.18. Error probability p_{b} versus rate R for repetition codes, the $(7, 4)$ Hamming code and BCH codes with blocklengths up to 1023 over a binary symmetric channel with $f = 0.1$. The righthand figure shows p_{b} on a logarithmic scale.

Book by David MacKay



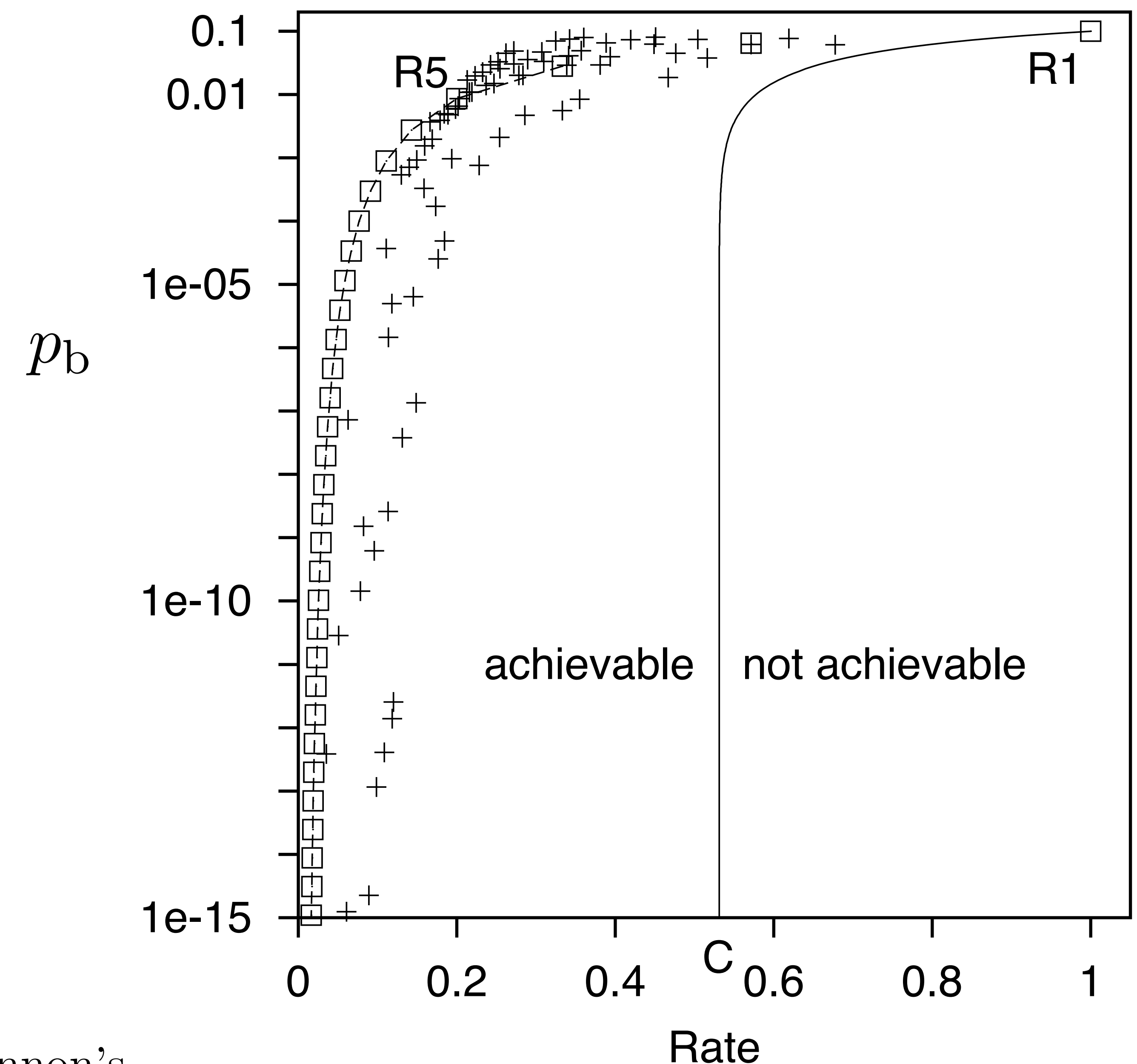


Figure 1.19. Shannon's noisy-channel coding theorem. The solid curve shows the Shannon limit on achievable values of (R, p_{b}) for the binary symmetric channel with $f = 0.1$. Rates up to $R = C$ are achievable with arbitrarily small p_{b}. The points show the performance of some textbook codes, as in figure 1.18.

The equation defining the Shannon limit (the solid curve) is $R = C/(1 - H_2(p_{\mathrm{b}}))$, where C and H_2 are defined in equation (1.35).

Book by David MacKay

$C \simeq 0.53$. Let us consider what this means in terms of noisy disk drives. The repetition code R_3 could communicate over this channel with $p_\mathrm{b} = 0.03$ at a rate $R = 1/3$. Thus we know how to build a single gigabyte disk drive with $p_\mathrm{b} = 0.03$ from three noisy gigabyte disk drives. We also know how to make a single gigabyte disk drive with $p_\mathrm{b} \simeq 10^{-15}$ 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 $p_\mathrm{b} = 10^{-18}$ or 10^{-24} or anything, you can get there with two disk drives too!'

Book by David MacKay

$C \simeq 0.53$. Let us consider what this means in terms of noisy disk drives. The repetition code R_3 could communicate over this channel with $p_{\mathrm{b}} = 0.03$ at a rate $R = 1/3$. Thus we know how to build a single gigabyte disk drive with $p_{\mathrm{b}} = 0.03$ from three noisy gigabyte disk drives. We also know how to make a single gigabyte disk drive with $p_{\mathrm{b}} \simeq 10^{-15}$ 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 $p_{\mathrm{b}} = 10^{-18}$ 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'.]

Book by David MacKay