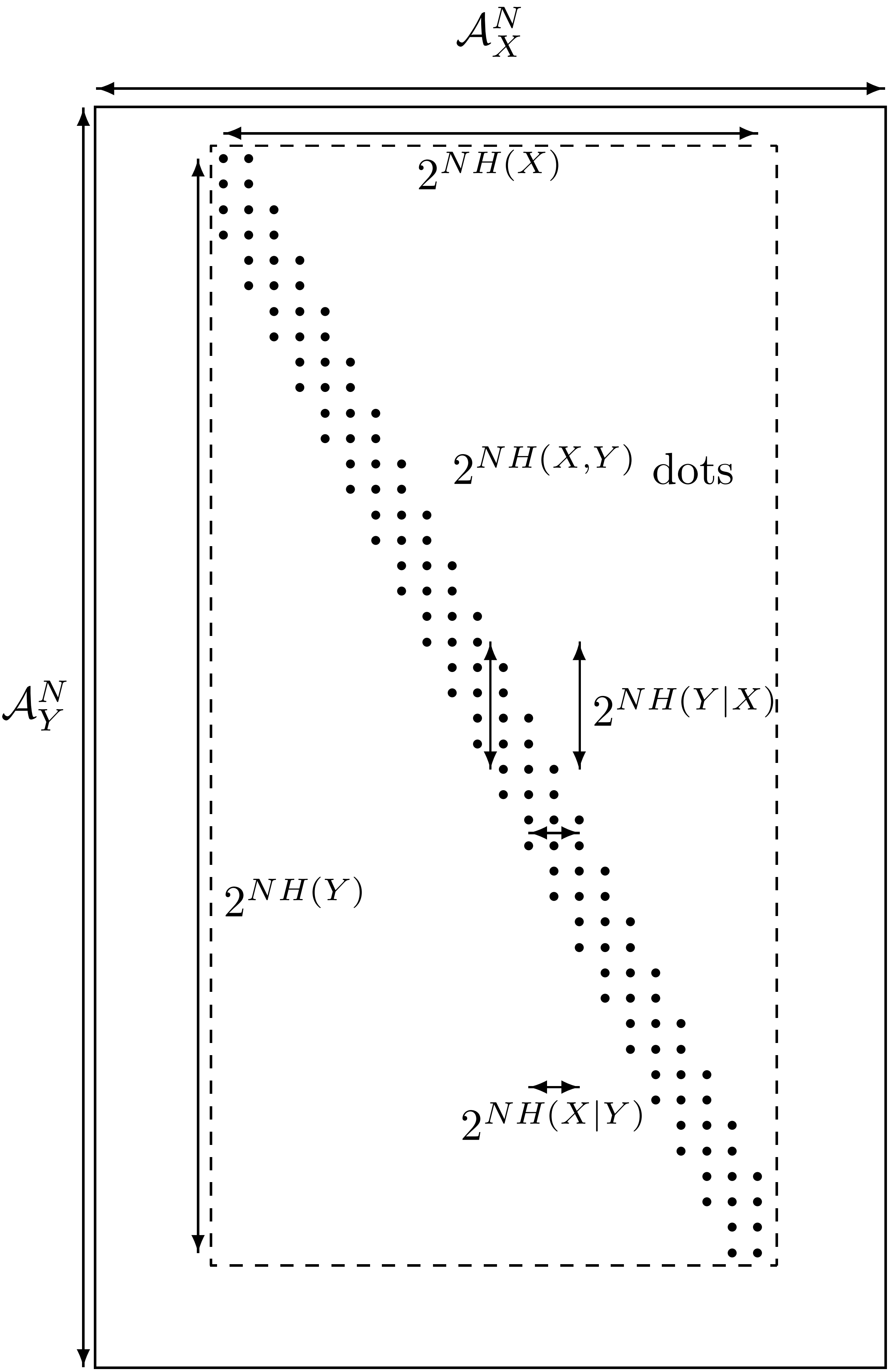
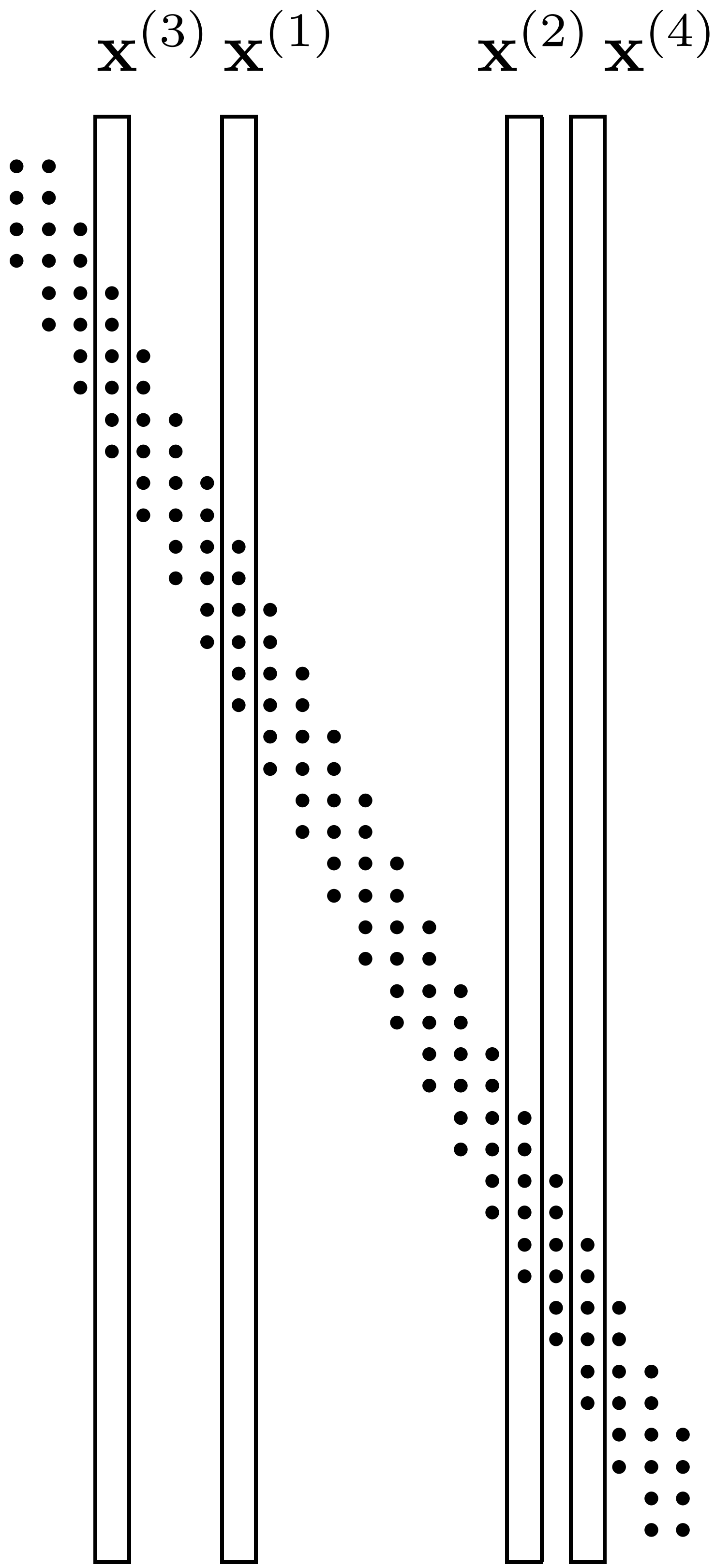


Figure 10.2. The jointly-typical set. The horizontal direction represents $\mathcal{A}_X^N$, the set of all input strings of length N. The vertical direction represents $\mathcal{A}_Y^N$, the set of all output strings of length N. The outer box contains all conceivable input–output pairs. Each dot represents a jointly-typical pair of sequences $(\mathbf{x}, \mathbf{y})$. The total number of jointly-typical sequences is about $2^{NH(X,Y)}$.

so the probability of hitting a jointly-typical pair is roughly

$$2^{NH(X,Y)}/2^{NH(X)+NH(Y)} = 2^{-NI(X;Y)}.$$

Book by David MacKay

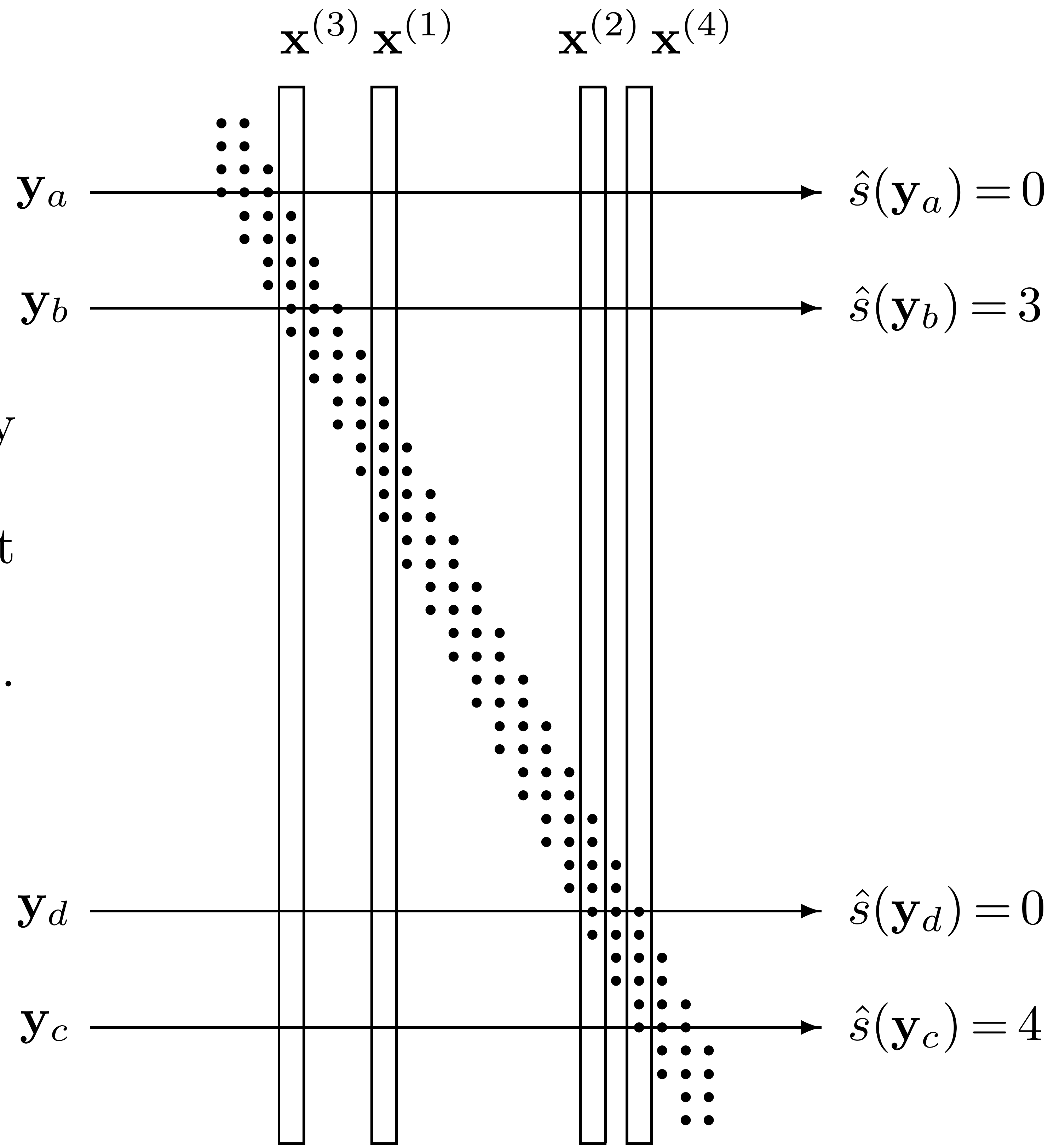


(a)

Figure 10.4. (a) A random code.

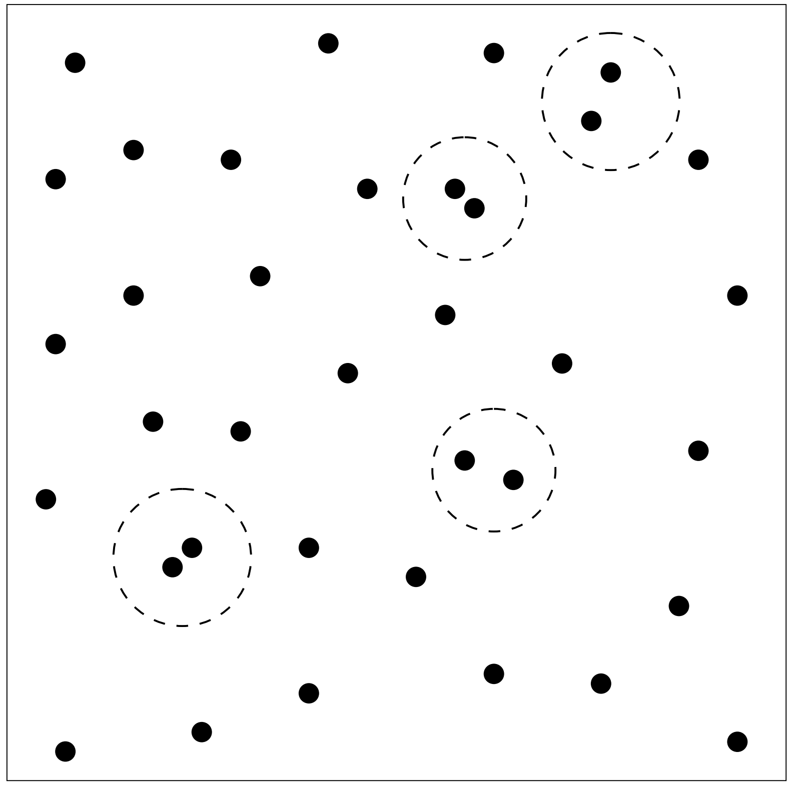
Book by David MacKay

(b) Example decodings by the typical set decoder. A sequence that is not jointly typical with any of the codewords, such as $\mathbf{y}_a$, is decoded as $\hat{s} = 0$. A sequence that is jointly typical with codeword $\mathbf{x}^{(3)}$ alone, $\mathbf{y}_b$, is decoded as $\hat{s} = 3$. Similarly, $\mathbf{y}_c$ is decoded as $\hat{s} = 4$. A sequence that is jointly typical with more than one codeword, such as $\mathbf{y}_d$, is decoded as $\hat{s} = 0$.



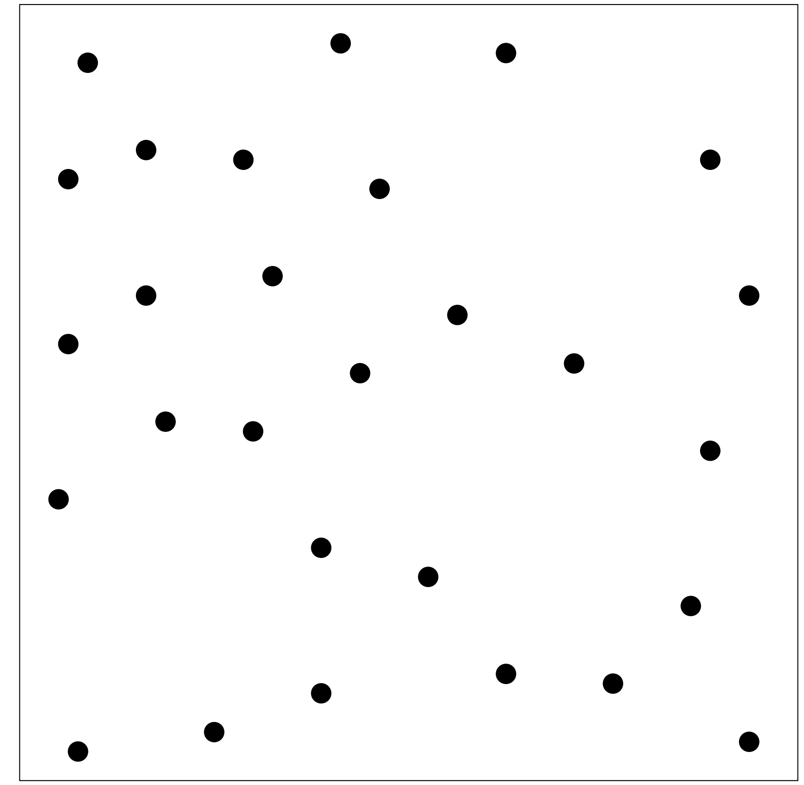
(b)

Book by David MacKay



(a) A random code . . .

⇒



(b) expurgated

Figure 10.5. How expurgation works. (a) In a typical random code, a small fraction of the codewords are involved in collisions – pairs of codewords are sufficiently close to each other that the probability of error when either codeword is transmitted is not tiny. We obtain a new code from a random code by deleting all these confusable codewords. (b) The resulting code has slightly fewer codewords, so has a slightly lower rate, and its maximal probability of error is greatly reduced.

Book by David MacKay

Analogy

Imagine that we wish to prove that there is a baby in a class of one hundred babies who weighs less than 10 kg. Individual babies are difficult to catch and weigh. Shannon's method of solving the task is to scoop up all the babies and weigh them all at once on a big weighing machine. If we find that their *average* weight is smaller than 10 kg, there must exist *at least one* baby who weighs less than 10 kg – indeed there must be many! Shannon's method isn't guaranteed to reveal the existence of an underweight child, since it relies on there being a tiny number of elephants in the class. But if we use his method and get a total weight smaller than 1000 kg then our task is solved.

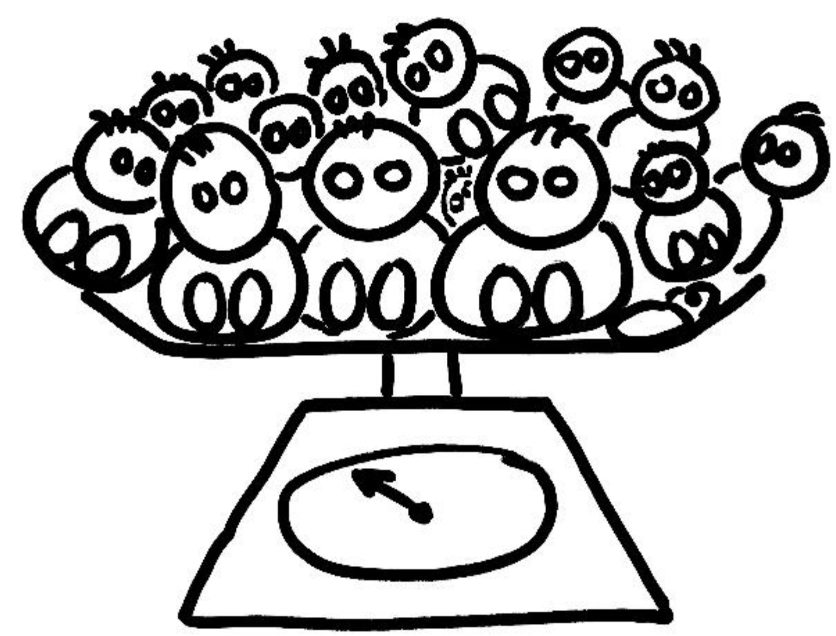


Figure 10.3. Shannon's method for proving one baby weighs less than 10 kg.

Book by David MacKay