Theorem 4.6 (Entropy Bound) Let $\mathcal{C}$ be a D-ary uniquely decodable code for a source random variable X with entropy $H_D(X)$. Then the expected length of $\mathcal{C}$ is lower bounded by $H_D(X)$, i.e.,

$$L \geq H_D(X). \tag{1}$$

This lower bound is tight if and only if $l_i = -\log_D p_i$ for all i.

Proof

1. Since $\mathcal{C}$ is uniquely decodable, the lengths of its codewords satisfy the Kraft inequality. Write

$$L = \sum_i p_i l_i = \sum_i p_i \log_D D^{l_i}$$

and recall that

$$H_D(X) = -\sum_i p_i \log_D p_i.$$

2. Then

$$\begin{aligned}
L - H_D(X) &= \sum_i p_i (\log_D p_i + \log_D D^{l_i}) \\
&= \sum_i p_i \log_D (p_i D^{l_i}) \\
&= (\ln D)^{-1} \sum_i p_i \ln(p_i D^{l_i}) \\
&\geq (\ln D)^{-1} \sum_i p_i \left(1 - \frac{1}{p_i D^{l_i}}\right) \qquad (2) \\
&= (\ln D)^{-1} \sum_i (p_i - D^{-l_i}) \\
&= (\ln D)^{-1} \left[\sum_i p_i - \sum_i D^{-l_i}\right] \\
&= (\ln D)^{-1} \left[1 - \sum_i D^{-l_i}\right] \\
&\geq (\ln D)^{-1} (1 - 1) \qquad (3) \\
&= 0.
\end{aligned}$$

This proves the entropy bound in (1).

3. In order for this bound to be tight, both (2) and (3) have to be tight simultaneously. Now (2) is tight if and only if $p_i D^{l_i} = 1$, or $l_i = -\log_D p_i$ for all i. If this holds, we have

$$\sum_i D^{-l_i} = \sum_i D^{\log_D p_i} = \sum_i p_i = 1,$$

i.e., (3) is also tight. This completes the proof of the theorem.