Abstract. This Correspondence gives a simple proof of the fundamental coding theorem which states that the capacity of a binary symmetric memoryless channel can be approached using a linear code.

1 Introduction

The pioneering role of Shannon’s coding theorem [1] is obviously unquestionable. The coding theorem states that the capacity of a binary symmetric memoryless channel can be approached using a (possibly nonlinear) code. The significance of Shannon’s coding theorem also lies in its proof which is based on the use of random codes, which inspired numerous coding algorithms. Because linear codes are much easier for coding and decoding, an improved version of the Shannon’s coding theorem for linear codes is needed. This result was first established by Elais [2] which proved that the same claim by Shannon holds for linear codes. Elias’ proof is also based on the use of random codes and is somewhat lengthy.

Despite the monumental importance of the coding theorem, its proof is usually considered to be relatively difficult, and thus avoided in most introductory textbooks. The purpose of this Correspondence is to offer a simple proof of the fundamental coding theorem for linear codes.
2 Coding Theorem and Its Proof

In this section, we provide a simple proof for the fundamental coding theorem for linear codes. The statement of the result is the same as in [2]. But our proof is not based on the use of random codes.

Theorem 1 Consider a binary symmetric memoryless channel with error probability \(0 < p < 1/2\). Given any \(\delta > 0\) and \(R < C(p)\), where

\[
C(p) = 1 - H(p) = 1 + p \log_2 p + (1 - p) \log_2(1 - p)
\]

is the channel capacity, there exists a linear code \(C(n, k)\) with the coding ratio \(k/n > R\) and the probability of minimum-distance decoding error less than \(\delta\).

Proof: Any linear code \(C(n, k)\) can be represented as a binary linear mapping \(c = Tm\), where \(c \in B^{n \times 1}\), \(m \in B^{k \times 1}\) and \(T \in B^{n \times k}\). Rewriting the above as \(c = Mt\), where \(t' = [t_1\ t_2\ \cdots\ t_n]'\); \(t_i\) is the \(i\)-th row of \(T\) and \(M \in B^{n \times nk}\) is given by

\[
M = \begin{bmatrix}
m & 0 & 0 & \cdots & 0 \\
0 & m & 0 & \cdots & 0 \\
\vdots & & & & \vdots \\
0 & 0 & \cdots & 0 & m
\end{bmatrix}
\]

Obviously, \(M\) has full row rank iff \(m \neq 0\).

Take any \(\epsilon > 0\) be such that \(C(p + \epsilon) > R\). By the weak law of large numbers, the probability that the number of transmission errors exceeds \((p + \epsilon)n\) is less than \(\delta\) when \(n\) is sufficiently large. Thus, it is sufficient to show that there exists \(t \in B^{nk \times 1}\) such that no non-zero \(c\) has weight less than \((p + \epsilon)n\). The number of non-zero elements in \(B^{n \times 1}\) with weight less than \((p + \epsilon)n\) is bounded by ([3], p. 115)

\[
\sum_{i=1}^{(p+\epsilon)n} \binom{n}{i} \leq 2^{H(p+\epsilon)n}
\]

For each of these elements and each non-zero \(m\), there are only \(2^{nk-n}\) solutions of \(t\). Note that there are \(2^k - 1\) non-zero \(m\). So all together, the number of \(t\) for which there exists some non-zero \(m\) such that the resulting \(c\) has weight less than \((p + \epsilon)n\) is bounded by

\[
2^{H(p+\epsilon)n}2^{nk-n}(2^k - 1) < 2^{nk+k-C(p+\epsilon)n} - 1 \leq 2^{nk} - 1
\]
Hence, there exists some $t \in B^{nk \times 1}$ (or equivalently, $T \in B^{n \times k}$) such that $c = Tm$ has weight at least $(p + \epsilon)n$ for every $m \in B^k$, i.e., the corresponding decoding error has probability less than $\delta$.

References

