Chapter 20
Error-Correcting Codes and Dirty Paper Coding

20.1 Introduction and Background

In the following we are concerned with impressing information on an independent signal, such as an image or an audio stream with the aim of the additional energy used consistent with reliable detection of the information. Information can even be impressed on background noise with no apparent signal present. A secondary aim is that in impressing the information, the independent signal should suffer a minimal amount of degradation or distortion to the point that in some circumstances the difference is virtually undetectable.

20.2 Description of the System

The following, for simplicity, is first described in terms of using binary codes and binary information. It is shown later that the method may be generalised to non-binary codes and non-binary information. The independent signal or noise is denoted by the waveform $v(t)$ and the information carrying signal to be impressed on the waveform $v(t)$ is denoted by $s(t)$. The resulting waveform $w(t)$ is simply given by the sum:

$$w(t) = v(t) + s(t) \tag{20.1}$$

The decoder which is used to determine $s(t)$ from the received waveform will usually be faced with additional noise, interference and sometimes distortion due to the receiving equipment or the transmission. With no distortion, the input to the decoder is denoted by $r(t)$ and given by:

$$r(t) = v(t) + s(t) + n(t) \tag{20.2}$$

M. Tomlinson et al., *Error-Correction Coding and Decoding*,
Signals and Communication Technology,
DOI 10.1007/978-3-319-51103-0_20

In its simplest form $s(t)$ carries only one bit of information and

$$s(t) = k_0 s_0(t) - k_1 s_1(t) \tag{20.3}$$

to convey data 0, and

$$s(t) = k_0 s_1(t) - k_1 s_0(t) \tag{20.4}$$

to convey data 1.

The multiplicative constants, k_0 and k_1 are chosen to adjust the energy of the information carrying signal and k_1 is used to reduce the correlation of the alternative information carrying signal that could cause an error in the decoder. The multiplicative constants, k_0 and k_1 are normally chosen as a function of $v(t)$, the main component of interference in the decoder, which is attempting to decode $r(t)$.

In conventional communications, $s_0(t)$ (or $s_1(t)$) is transmitted or stored and $s_0(t)$ (or $s_1(t)$) is decoded despite the presence of interference or noise. $s(t)$ is added to $v(t)$ and $s_0(t)$ (or $s_1(t)$) is decoded from the composite waveform $v(t) + s(t)$ despite the presence of additional interference or noise.

Noting that the transmitter has no control over the independent signal or noise $v(t)$, a good strategy is to choose $s_0(t)$ and $s_1(t)$ from a large number of possible waveforms in order to produce waveforms which have a large correlation with respect to $v(t)$. Each possible waveform is constrained to be a codeword from an (n, k, d_{min}) error-correcting code. In one approach using binary codes, the 2^k codewords are partitioned into two disjoint classes, codewords having even parity and codewords having odd parity. The codeword $s_0(t)$ is the even parity codeword with highest correlation out of all even parity codewords and the codeword $s_1(t)$ is the odd parity codeword with highest correlation out of all odd parity codewords. The idea is that $w(t)$ should have maximum correlation with $s_0(t)$ if the information data is 0 compared to any of the other $2^k - 1$ codewords. Conversely if the information data is 1, $w(t)$ should have maximum correlation with $s_1(t)$ compared to any of the other $2^k - 1$ codewords. As there is a minimum Hamming distance of d_{min}, between codewords, this prevents small levels of additional noise or interference causing an error in detecting the data in the decoder.

As an example, consider a typical sequence of 47 Gaussian noise samples $v(t)$ as shown in Fig. 20.1. A binary quadratic residue [4] code, described in Chap. 4, the (47, 24, 11) code is used and the highest correlation codeword having even parity is determined using a near maximum likelihood decoder, the modified Dorsch decoder described in Chap. 15. The waveform of Fig. 20.1 is input to the decoder. The highest correlation codeword, which has a correlation value of 20.96 is the codeword:

{01001001100110010010000100001000010000100 10000000000}

The highest correlation, odd parity codeword, is then determined. This codeword, which has a correlation value of 22.65, is the codeword:

{ 11101001011000011001100100010000010000 00100000000 }

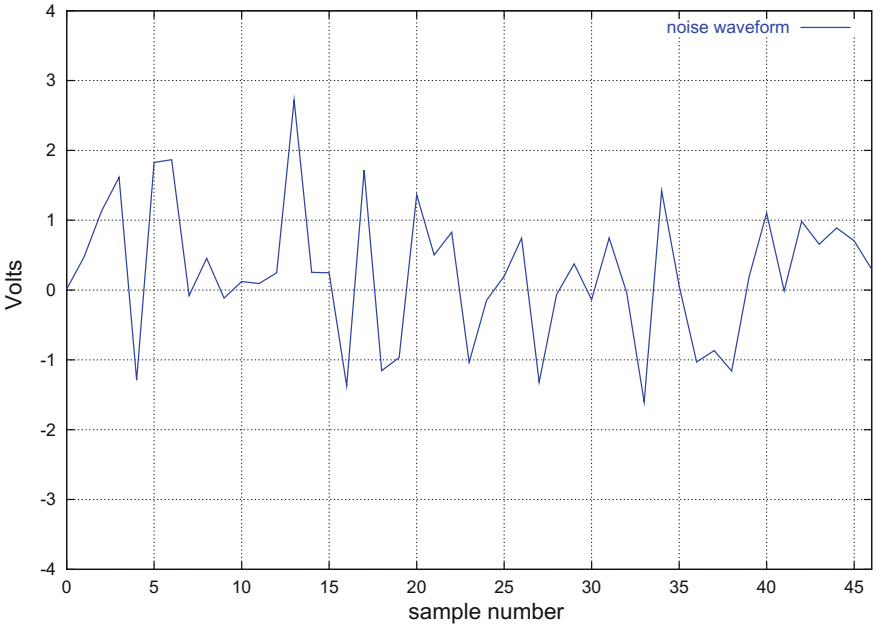

Fig. 20.1 Noise waveform to be impressed with data

It should be noted that in carrying out the correlations, codeword 1's are mapped to -1's and codeword 0's are mapped to $+1$'s.

The information bit to be impressed on the noise waveform is say, data 0, in which case the watermarked waveform $w(t)$ needs to produce a maximum correlation with an even parity codeword. Correspondingly, the value given to k_0 is 0.156 and to k_1 is 0.02 in order to make sure that the codeword which produces the maximum correlation with the marked waveform is the previously found even parity codeword:

{0 1 0 0 1 0 0 1 1 0 0 1 1 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0}

The marked waveform $w(t)$ is as shown in Fig. 20.2. It may be observed that the difference between the marked wavefordm and the original waveform is small. In the decoder it is found that the codeword with highest correlation with the marked waveform $w(t)$ is indeed the even parity codeword:

{0 1 0 0 1 0 0 1 1 0 0 1 1 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 }

and this codeword has a correlation of 28.31.

One advantage of this watermarking system over conventional communications is that the watermarked waveform may be tested using the decoder. If there is insufficient margin, adjustments may be made to the variables k_0 and k_1 and a new

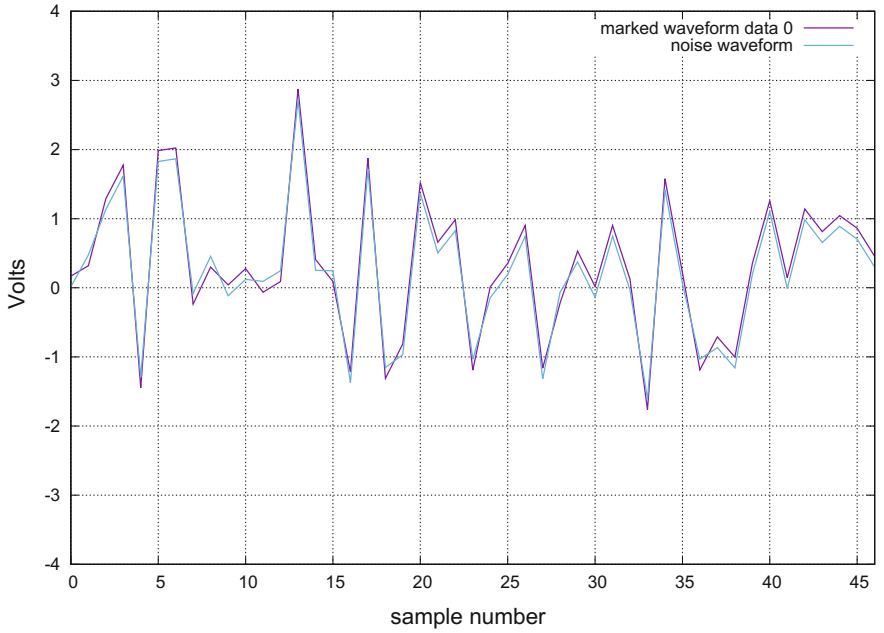

Fig. 20.2 Noise waveform impressed with data 0

watermarked waveform produced. Conversely, if there is more than adequate margin, adjustments may be made to the variables k_0 and k_1, so that there is less degradation to the original waveform $v(t)$.

The highest correlation, odd parity codeword with correlation 25.31 is the codeword:

{ 1 1 1 0 1 0 0 1 0 1 1 0 0 0 0 1 1 0 0 1 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 }

It should be noted that this odd parity codeword is the same odd parity codeword as determined in the encoder, but this is not always the case depending upon the choice of values for k_0 and k_1.

For the case where the information bit is a 1, the marked waveform $w(t)$ needs to produce a maximum correlation with an odd parity codeword. In this case, the value of k_0 is 0.043 and the value of k_1 is 0.02 and $s(t) = k_0 s_1(t) - k_1 s_0(t)$. The marked waveform $w(t)$ is as shown in Fig. 20.3. This time in the decoder it is found that the codeword with highest correlation with $w(t)$ is indeed the odd parity codeword:

{ 1 1 1, 0 1 0 0 1 0 1 1 0 0 0 0 1 1 0 0 1 1 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 }

and this codeword has a correlation of 24.70. The highest correlation,even parity, codeword has a correlation of 22.02.

In the encoding and decoding procedure above, the maximum correlation codeword needs to be determined. For short codes a maximum likelihood decoder [6]

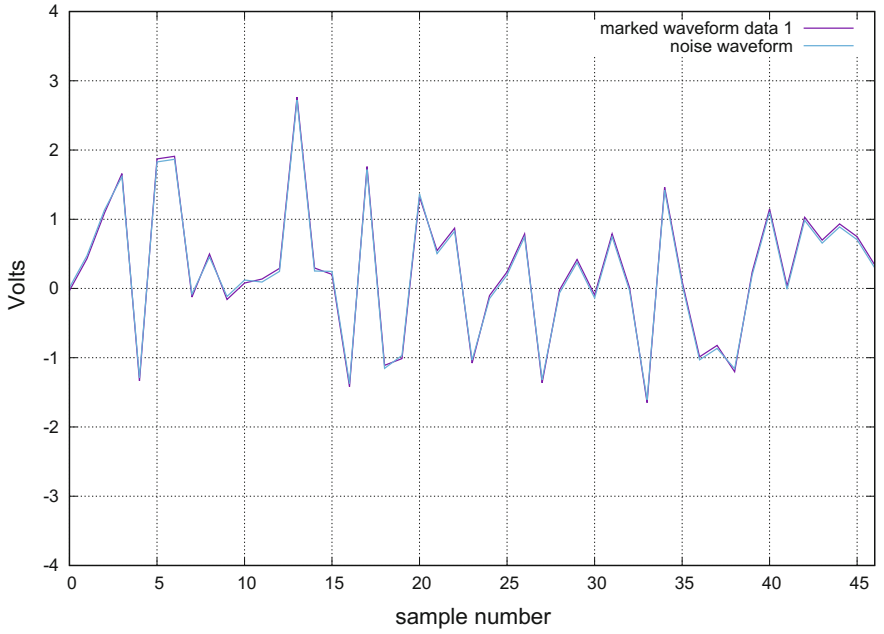

Fig. 20.3 Noise waveform impressed with data 1

may be used. For medium length codes, up to 200 bits long, the near maximum likelihood decoder, the modified Dorsch decoder of Chap. 15 is the best choice. For longer codes, decoders such as an LDPC decoder [3], turbo code decoder [1], or turbo product code decoder [7] may be used in conjunction with the appropriate iterative decoder. An example of a decoder for LDPC codes is given in by Chen [2].

Once the maximum correlation codeword has been found, codewords with similar, high correlation values, may be found from the set of codewords having small Hamming distance from the highest correlation codeword. Linear codes are the most useful codes because the codewords with high correlations with the target waveform are given by the sum of the highest correlation codeword and the low-weight codewords of the code, modulo q, (where $GF(q)$ is the base field [4] of the code). The low-weight codewords of the code are fixed and may be derived directly as described in Chaps. 9 and 13, or determined from the weight distribution of the dual code [4].

For practical implementations, the most straightforward approach is to restrict the codes to binary codes less than 200 bits long and determine the high correlation codewords by means of the modified Dorsch decoder. This conveniently, can output a ranked list of the high cross correlation codewords is together with their correlation values. It is straightforward to modify the decoder so as to provide the output codewords in odd and even parity classes, with the maximum correlation codeword for each class. The results for the example above were determined in this way.

Additional information may be impressed upon the independent signal or noise by partitioning the 2^k codewords into more disjoint classes (other than binary). For example four disjoint classes may be obtained by partitioning the codewords according to odd and even parity for the odd numbered codeword bits and odd and even parity for the even numbered codeword bits. Namely, if the codewords are represented as:

$$c(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + \cdots + c_{k-1} x^{k-1} \tag{20.5}$$

then the codewords are partitioned according to the values of p_0 and p_1 given by

$$p_0 = c_0 + c_2 + c_4 + c_6 \cdots + c_{k-1} \text{ modulo } 2 = 0$$
$$p_1 = c_1 + c_3 + c_5 + c_7 \cdots + c_{k-2} \text{ modulo } 2 = 0$$

or with the result

$$p_0 = c_0 + c_2 + c_4 + c_6 \cdots + c_{k-1} \text{ modulo } 2 = 1$$
$$p_1 = c_1 + c_3 + c_5 + c_7 \cdots + c_{k-2} \text{ modulo } 2 = 1$$

Clearly the procedure may be extended to m parity bits by partitioning the 2^k codewords into 2^m disjoint classes. In this case, following encoding, m bits of information will be conveyed by the marked waveform and determined from the codeword which has the highest correlation with the marked waveform. This is by virtue of which of the 2^m classes this codeword resides.

An alternative to this procedure is to use non-binary codes [5] with a base field of $GF(q)$ as described in Chap. 7. For convenience a base field of $GF(2^m)$ may be used so that each symbol of a codeword is represented by m bits. In this case codewords are partitioned into 2^m classes according to the value of the overall parity sum:

$$p_0 = c_0 + c_1 + c_2 + c_3 + c_4 + \cdots + c_{k-1} \text{ modulo } 2^m \tag{20.6}$$

The n non-binary symbols of each codeword may be mapped into n Pulse Amplitude Modulation (PAM) symbols [6] or into $n.m$ binary symbols or a similar hybrid combination before correlation with $v(t)$.

Rather than maximum correlation with the waveform to be marked, codewords may be chosen that have near zero correlation with the waveform to be marked. Information is conveyed by the watermarked marked waveform by the addition of a codeword to $v(t)$, which is orthogonal or near orthogonal to the codeword which has maximum correlation to the independent signal or noise waveform $v(t)$. In this case, the codeword with maximum correlation to $v(t)$ is denoted as $s_{max}(t)$. Codewords that are orthogonal or near orthogonal to $s_{max}(t)$ are denoted as $s_{max,i}(t)$ for $i = 1$ to 2^m. The signal impressed upon $v(t)$ is:

$$s(t) = k_0 s_{max}(t) + k_1 s_{max,\eta}(t) \tag{20.7}$$

where η determines which one of the 2^m orthogonal codewords is impressed on the waveform to convey the m bits of information data. The addition of the maximum correlation codeword $k_0 s_{max}(t)$ to $v(t)$ is to make sure that $s_{max}(t)$ is still the codeword with maximum correlation after the waveform has been marked. Although the codewords $s_{max,i}(t)$ for $i = 1$ to 2^m are orthogonal to $k_0 s_{max}(t)$ they are not necessarily orthogonal to $v(t)$. In this case, the signal impressed upon $v(t)$ needs to be:

$$s(t) = k_0 s_{max}(t) + \sum_{i=1}^{2^m} k_i s_{max,i}(t) \tag{20.8}$$

The coefficients k_i will usually be small in order to produce near zero correlation of the codewords $s_{max,i}$ with $w(t)$ except for the coefficient k_j in order to produce a strong correlation with the codeword $s_{max,j}$.

The choice of the multiplicative constants, k_0 and k_1 or the multiplicative constants k_i for the general case (these adjust the energy of the components of the information signal), depends upon the expected levels of additional noise or interference and acceptable levels of decoder error probability. If the marked signal to noise ratio is represented as SNR_z, the marked signal energy as E_z, and the difference in highest correlation to next highest correlation of the codewords is Δ_c, then the probability of decoder error $p(e)$ is lower bounded by:

$$p(e) \leqslant \frac{1}{2} erfc\left(\frac{\Delta_c^2 . SNR_z}{8 . E_z}\right)^{0.5} \tag{20.9}$$

This is under the assumption that there is only one codeword close in Euclidean distance to the maximum correlation codeword.

The multiplicative constants may be selected "open loop" or "closed loop". In "closed loop", which is a further variation of the system, the encoding is followed by a testing phase. After encoding, the information is decoded from the marked waveform and the margin for error determined. Different levels of noise or interference may be artificially added to the marked waveform, prior to decoding, in order to assist in determining the margin for error. If the margin for error is found to be insufficient, then the multiplicative constants may be adjusted and a new marked waveform $w(t)$ produced and tested.

In the decoder, once the maximum correlation codeword has been detected from the marked signal or noise waveform, candidate orthogonal, or near orthogonal codewords, are generated from the maximum correlation codeword and these codewords are cross correlated with the marked signal or noise waveform in order to determine which weighted orthogonal, or near orthogonal, codewords have been added to the marked signal or noise waveform. In turn the detected orthogonal, or near orthogonal, codewords from the cross correlation coefficients are used to determine the additional information which was impressed on the marked signal or noise waveform.

In order to clarify the description, Fig. 20.4 shows a block diagram of the encoder for the example of a system conveying two information bits. The independent signal

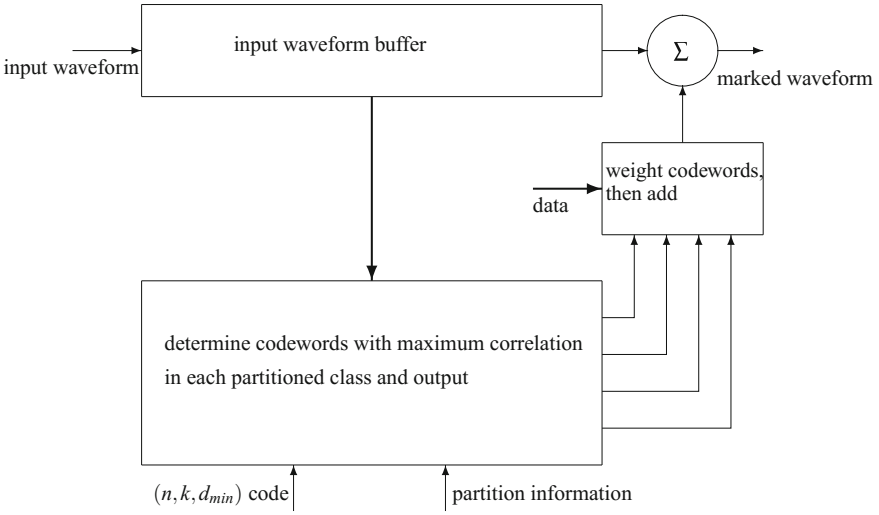

Fig. 20.4 Encoder for two information bits using near orthogonal codewords

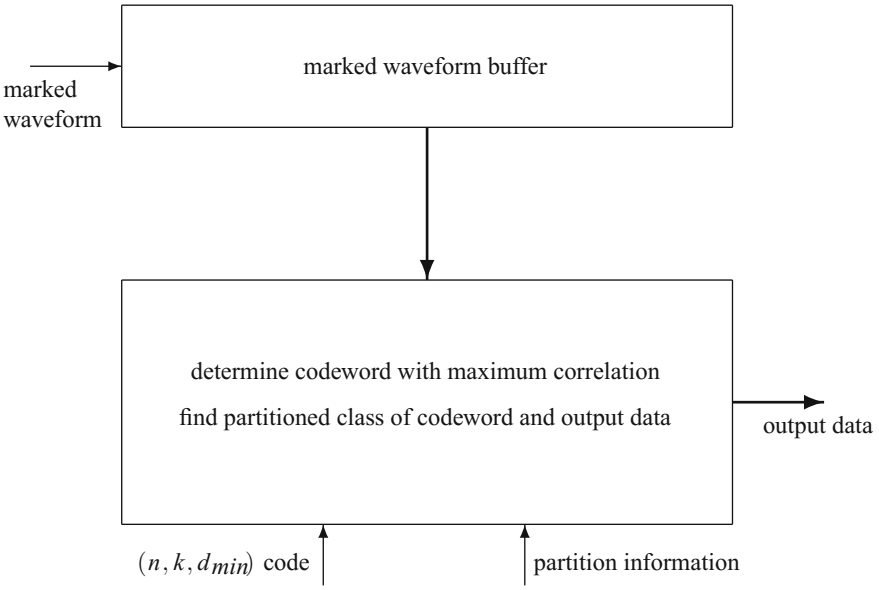

Fig. 20.5 Decoder for marked waveform containing orthogonal codewords

or noise is input to a buffer memory which feeds a maximum correlation decoder, which usually will be a modified Dorsch decoder. The maximum correlation decoder has as input the error-correcting code parameters (n, k, d_{min}) and the code partition information. In this case the partition information is used to partition the codewords

into four classes. The codewords, in each class, having highest correlation, and their correlation values are output as shown in Fig. 20.4. From the input data and these correlation values, the multiplicative constants are determined. The coefficients of each codeword are weighted by these constants, and added to the stored independent signal or noise to produce the marked waveform, which is output from the encoder.

Figure 20.5 shows a block diagram of the corresponding decoder. The marked waveform is input to the buffer memory which feeds a maximum correlation decoder. The error-correcting code parameters of the same (n, k, d_{min}) code and the code partition information are also input to the maximum correlation decoder. The codeword with the highest correlation is determined. The class in which the codeword resides is found and the two bits of data identifying this class are output from the decoder.

In a further approach, additional information may be conveyed by adding weighted codewords to the marked signal or noise waveform such that these codewords are orthogonal, or near orthogonal, to the codeword having maximum correlation with the marked signal or noise waveform.

20.3 Summary

This chapter has described how error-correcting codes can be used to impress additional information onto waveforms with a minimal level of distortion. Applications include watermarking and steganography. A method has been described in which the modified Dorsch decoder of Chap. 15 is used to find codewords from partitioned classes of codewords, whose waveforms may be used as a watermark which is almost invisible, and still be reliably detected.

References

1. Berrou, C., Thitimajshima, P., Glavieux, A.: Near Shannon limit error correcting coding and decoding: turbo codes. In: Proceedings of IEEE International Conference on Communications, pp. 1064-1070. Geneva, Switzerland (1993)
2. Chen, J., Fossorier, M.P.C.: Near optimum universal belief propagation based decoding of low-density parity check codes. IEEE Trans. Comm **50**(3), 406–414 (2002). March
3. Gallager, R.G.: Low-Density Parity Check Codes. M.I.T. Press, Cambridge (1963)
4. MacWilliams, F.J., Sloane, N.J.A.: The Theory of Error-Correcting Codes. North-Holland (1977)
5. Peterson, W., Weldon, E.J,Jr: Error-Correcting-Codes, 2nd edn. MIT Press, Cambridge (1972)
6. Proakis, J.G.: Digital Communications. McGraw-Hill, New York (1995)
7. Pyndiah, R.M.: Near-optimum decoding of product codes: block Turbo codes. IEEE Trans. Commun. **46**, 1003–1010 (1998). Aug