Low Density Parity Check Codes over $GF(q)$

David J.C. MacKay

Cavendish Laboratory,
Madingley Road,
Cambridge, CB3 0HE.
United Kindom.
Email: mackay@mrao.cam.ac.uk

Abstract — **Gallager's low density parity check codes over $GF(2)$ have been shown to have near Shannon limit performance when decoded using a probabilistic decoding algorithm. In this paper we report the empirical performance of the analogous codes defined over $GF(q)$ for $q > 2$.**

I BACKGROUND

Codes defined in terms of a non-systematic low density parity check matrix [1, 2] are asymptotically good, and can be practically decoded with Gallager's belief propagation algorithm [3, 4, 5]. Our proof in [5] shows that they are asymptotically good codes for a wide class of channels, not just for the memoryless binary symmetric channel.

We expect the generalization of these codes to finite fields $GF(q)$ for $q > 2$ to be useful for the q-ary symmetric channel, and possibly for other channels such as the binary symmetric channel.

Definition 1 *The* weight *of a vector or matrix is the number of non-zero elements in it. We denote the weight of a vector* $\mathbf{x}$ *by* $w(\mathbf{x})$*. The* density *of a source of random elements is the expected fraction of non-zero bits. A source of elements drawn from $GF(q)$ is* sparse *if its density is less than $(q - 1)/q$. A vector* $\mathbf{v}$ *is* very sparse *if its density vanishes as its length increases, for example, if a constant number t of its elements are non-zero. The* overlap *between two vectors is the number of non-zero elements in common between them.*

II CONSTRUCTION

The code is defined in terms of a very sparse, non-systematic, random parity check matrix $\mathbf{A}$. A transmitted block length N and a source block length K are selected. We define $M = N - K$ to be the number of parity checks. We select a *column weight t*, which is an integer greater than or equal to 3. We create a rectangular $M \times N$ matrix [M rows and N columns] $\mathbf{A}$ at random with exactly weight t per column and a weight per row as uniform as possible, and with the overlap between any two columns being either zero or one. If N/M is chosen to be an appropriate ratio of integers then the number per row can be constrained to be exactly tN/M. The non-zero elements are either drawn randomly from the non-zero elements of $GF(q)$ or according to a carefully chosen distribution. We then use Gaussian elimination and reordering of columns to derive an equivalent parity check matrix in systematic form $[\mathbf{P}\,\mathbf{I}_M]$, from which the generator matrix of the code can be obtained. There is a possibility that the rows of $\mathbf{A}$ are not independent (though for odd t, this has small probability); in this case, $\mathbf{A}$ is a parity check matrix for a code with the same N and with smaller M, that is, a code with *greater* rate than assumed in the following.

III VARIATIONS FOR BINARY SYMMETRIC CHANNELS

The issue of the choice of the non-zero elements in each row of the matrix $\mathbf{A}$ can be explored theoretically by computing bounds on the entropy of the parity check vector given by $\mathbf{z} = \mathbf{Ax}$, where $\mathbf{x}$ is a sample from the assumed channel noise model. The larger the entropy of $\mathbf{z}$, the closer the code might be able to get to capacity [5]. In the case of the q-ary symmetric channel, the entropy of one bit of $\mathbf{z}$ is independent of the choice of the elements in the corresponding row of $\mathbf{A}$. But in the case where the noise is that of a binary symmetric channel (assuming $q = 2^p$), some choices of the elements in a row of $\mathbf{A}$ are superior to others. We have found optimal selections for $GF(4)$, $GF(8)$ and $GF(16)$ by exhaustive search.

IV DECODING

The decoding algorithm is an appropriate generalization of the belief propagation algorithm used by Gallager [1] and MacKay and Neal [3, 4, 5]. The complexity of decoding scales as Ntq^2.

V RESULTS

We expect in early 1997 to have empirical results for codes over $GF(4)$, $GF(8)$ and $GF(16)$, applied to the q-ary symmetric channel, the binary symmetric channel, and the binary-input Gaussian channel.

REFERENCES

[1] R. G. Gallager, "Low density parity check codes", *IRE Trans. Info. Theory*, vol. IT-8, pp. 21–28, Jan 1962.

[2] R. G. Gallager, *Low Density Parity Check Codes*, Number 21 in Research monograph series. MIT Press, Cambridge, Mass., 1963.

[3] D. J. C. MacKay and R. M. Neal, "Good codes based on very sparse matrices", in *Cryptography and Coding. 5th IMA Conference*, Colin Boyd, Ed., number 1025 in Lecture Notes in Computer Science, pp. 100–111. Springer, Berlin, 1995.

[4] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes", *Electronics Letters*, vol. 32, no. 18, pp. 1645–1646, August 1996.

[5] D. J. C. MacKay, "Good error correcting codes based on very sparse matrices", To be submitted to IEEE transactions on Information Theory. Available from http://wol.ra.phy.cam.ac.uk/, 1996.

Submitted to ISIT-97   September 24, 1996