Code description

Let $x_1, \ldots, x_k \in \mathbb{F}$ be input symbols that we wish to encode and transmit over a channel with erasure parameter ϵ. Fix an integer r and let $\delta > 0$. Let $\Omega(x) = \sum_{d=1}^{D} \Omega_D x^D$ be the generating function of a distribution on the integers $\{1, \ldots, D\}$. To each integer d in this range, we associate a $[d+r, d]_{\mathbb{F}}$-MDS code C_d. We define the generalized LT code with degree distribution $\Omega(x)$ as follows.

- **Encoding**: Pick d according to the distribution Ω. Pick d input symbols uniformly at random. Consider these symbols as message symbols and using the code C_d, generate r parity symbols from them. For sending the symbols over the channel, we have two choices:
  a. Pack the r symbols together and send them over the channel. This assumes that the channel has input alphabet $\mathbb{F}^r$.
  b. Send each of the r symbols independently over the channel.
- **Decoding**: Start decoding when you have collected k *parity* symbols. Use belief propagation to recover the k input symbols. Let the edges of the decoding graph carry messages "erasure" and "non-erasure" during belief propagation.

We will first analyze scenario a. Here we consider one output symbol to consist of r packed parity symbols generated from some subset of d input symbols according to C_d. The r symbols are either all lost or all received, which makes decoding easier to analyze.

Density evolution

Fix a decoding round. For a random edge of the decoding graph, let x be the probability that this edge carries a non-erasure message from the input symbol to the output symbol side. Let q be the probability that a random edge carries a non-erasure message from the output symbol to the input symbol size, and let q_d be the probability that a random edge connected to an output symbol of degree d carries a non-erasure message from the output symbol to the input symbol side. This happens when less than r of the

remaining $d-1$ edges of the output symbol carry erasures, so that

$$q_d = \sum_{j=0}^{r-1} \binom{d-1}{j} (1-x)^j x^{d-1-j}.$$

Averaging over all output symbols degrees, we get

$$q = \sum_d \frac{d\Omega_d}{\Omega'(1)} q_d.$$

To compute the new non-erasure probability from the input symbol to the output symbol side x_{new}, note that a random edge connected to an input symbol of degree d carries an erasure (from the input to the output side) if and only if the other $d-1$ edges carry erasures (from the output to the input side). This happens with probability $(1-q)^{d-1}$. Note that $d-1$ (as well as d) is sampled from a Poisson distribution with mean $\lambda = \frac{\Omega'(1)}{r}$, so that x_{new} is given by

$$x_{\text{new}} = 1 - e^{-\frac{\sum_d d\Omega_d q_d}{r}}.$$

The density evolution constraint for correct decoding is $x_{\text{new}} > x$, so that we are looking for a degree distribution that satisfies

$$\sum_d d\Omega_d \sum_{j=0}^{r-1} \binom{d-1}{j} (1-x)^j x^{d-1-j} = -r \ln(1-x). \tag{1}$$

Theorem 1. *The degree distribution $\Omega(x)$ given by*

$$\begin{aligned} \Omega_1 &= \cdots = \Omega_r = 0 \\ \Omega_d &= \frac{r}{d(d-1)}, \quad d > r \end{aligned} \tag{2}$$

ensures correct decoding of k input symbols from k parity symbols, in expectation.

Proof. It suffices to show that this distribution satisfies (1). We thus need to prove that

$$\sum_{\substack{d \geq r \\ 0 \leq i \leq j \leq r-1}} \frac{1}{d} \binom{d}{j} \binom{j}{i} (-1)^{j-i} x^{d-i} = \sum_{\ell \geq 1} \frac{x^\ell}{\ell}.$$

Extracting the ℓth coefficient of both series, we need to prove the equality

$$\sum_{\substack{d\geq r\\ 0\leq j\leq r-1}} \frac{1}{d}\binom{d}{j}\binom{j}{d-\ell}(-1)^{j-d+\ell} = \begin{cases} 0, & \ell = 0\\ \frac{1}{\ell}, & \ell \geq 1. \end{cases}$$

For $\ell = 0$, the fact that $j < d$ ensures that the binomial coefficient $\binom{j}{d-\ell}$ is equal to 0, so that the desired equality holds. For $\ell \geq 1$, we prove the equality by induction over r.
Let $r = 1$. Then the only term that survives in the summation

$$\sum_{d\geq 1} \frac{1}{d}\binom{d}{0}\binom{0}{d-\ell}(-1)^{-d+\ell}$$

is that for which $d = \ell$, so that the summation is indeed equal to $\frac{1}{\ell}$. Now we suppose

$$\sum_{\substack{d\geq r\\ 0\leq j\leq r-1}} \frac{1}{d}\binom{d}{j}\binom{j}{d-\ell}(-1)^{j-d+\ell} = \frac{1}{\ell}, \quad \ell \geq 1$$

and wish to prove that

$$\sum_{\substack{d\geq r+1\\ 0\leq j\leq r}} \frac{1}{d}\binom{d}{j}\binom{j}{d-\ell}(-1)^{j-d+\ell} = \frac{1}{\ell}, \quad \ell \geq 1. \tag{3}$$

Consider the difference of the left-hand sides of the two equations above

$$\begin{aligned} &\sum_{\substack{d\geq r\\ 0\leq j\leq r-1}} \frac{1}{d}\binom{d}{j}\binom{j}{d-\ell}(-1)^{j-d+\ell} - \sum_{\substack{d\geq r+1\\ 0\leq j\leq r}} \frac{1}{d}\binom{d}{j}\binom{j}{d-\ell}(-1)^{j-d+\ell}\\ &= \sum_{0\leq j\leq r-1} \frac{1}{r}\binom{r}{j}\binom{j}{r-\ell}(-1)^{j-r+\ell} - \sum_{d\geq r+1} \frac{1}{d}\binom{d}{r}\binom{r}{d-\ell}(-1)^{r-d+\ell}. \end{aligned} \tag{4}$$

In what follows, we will make use of three standard binomial coefficient identities (for proofs, see [1]): *symmetry*

$$\binom{n}{k} = \binom{n}{n-k}, \quad n \in \mathbb{N}, k \in \mathbb{Z}, \tag{5}$$

upper negation

$$\binom{r}{k} = (-1)^k \binom{k-r-1}{k}, \quad k \in \mathbb{Z}, \tag{6}$$

and *Vandermonde convolution*

$$\sum_k \binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}, \quad n \in \mathbb{Z}. \tag{7}$$

On one hand,

$$\begin{aligned}
& \sum_{0\le j\le r-1} \frac{1}{r}\binom{r}{j}\binom{j}{r-\ell}(-1)^{j-r+\ell} \\
= \; & \frac{1}{r}\sum_{0\le j\le r-1} \binom{r}{r-j}\binom{j}{j-r+\ell}(-1)^{j-r+\ell} && (8)\\
= \; & \frac{1}{r}\sum_{0\le j\le r-1} \binom{r}{r-j}\binom{\ell-r-1}{j-r+\ell} && (9)\\
= \; & \frac{1}{r}\left(\binom{\ell-1}{\ell} - \binom{r}{0}\binom{\ell-r-1}{\ell}\right) && (10)\\
= \; & 0,
\end{aligned}$$

where (8) follows from symmetry, (9) from upper negation, and (10) from Vandermonde convolution.

On the other hand,

$$\begin{aligned}
& \sum_{d\geq r+1} \frac{1}{d}\binom{d}{r}\binom{r}{d-\ell}(-1)^{r-d+\ell} \\
= & \sum_{d\geq r+1} \frac{1}{r}\binom{d-1}{r-1}\binom{r}{d-\ell}(-1)^{r-d+\ell} \\
= & \sum_{d\geq r+1} \frac{1}{r}\binom{d-1}{d-r}\binom{r}{r-d+\ell}(-1)^{r-d+\ell} \\
= & \frac{1}{r}(-1)^{\ell} \sum_{d\geq r+1} \binom{-r}{d-r}\binom{r}{r-d+\ell} && (11) \\
= & \frac{1}{r}(-1)^{\ell} \left(\binom{0}{\ell} - \binom{-r}{0}\binom{r}{\ell}\right) && (12) \\
= & \frac{1}{r}\binom{\ell-r-1}{\ell} && (13) \\
= & 0,
\end{aligned}$$

where (11) and (13) follow from upper negation and (12) from convolution. This proves that the difference in (4) is indeed zero, which gives us the desired equality in (3). □

References

[1] Concrete Mathematics: A Foundation for Computer Science. Graham, R., Knuth, D., and Patashnik, O. Second edition, Addison-Wesley 1989.