STAT 206A: Gibbs Measures Elchanan Mossel

Lecture 11

Lecture date: Oct 3 *Scribe: Alex Dimakis*

In this lecture we are interested in determining the basic properties of the weight enumerator function, which is defined as

$$\mathcal{W}(w) = \{\# \text{ of codewords with weight w}\}. \tag{1}$$

In particular we will be estimating the *expected* weight enumerator function

$$\overline{\mathcal{W}}(w) = \mathbb{E}\mathcal{W}(w), \tag{2}$$

where the expectation is taken over the ensemble of LDPC codes. We would like to show that codewords will have a minimum weight that grows linearly in the blocklength N, and we will be analyzing weights that are constant fractions of N. So define ω as the linear fraction $w = N\omega$ and $\varphi(\omega)$ to be the exponential growth rate of $\overline{\mathcal{W}}$, i.e.

$$\overline{\mathcal{W}}(w = N\omega) = e^{N\varphi(\omega)}. \tag{3}$$

As in the previous lecture, assume that the variables which are assigned one, are colored red, and the others blue. To have a valid codeword, we need every check to be adjacent to an even number of red edges. To simply the combinatorics, we look at the regular case first.

I. Regular codes

Assume all the variables have degree l, all checks k, there are N variables and M checks and therefore

$$Nl = Mk = F = \text{number of edges}. \tag{4}$$

The corresponding degree distribution is $\Lambda = x^l, P = x^k$. Now we need to count how many codewords have weight w, equivalently how many colorings exist with each check adjacent to an even number of red edges. Let

$$m_r = \{\#\text{of checks adjacent to r red edges}\} \tag{5}$$

and observe that

$$\sum_r r m_r = lw = \{\ \#\text{ of red edges }\}. \tag{6}$$

The number of ways one can color the edges from the check node side, is

$$C(k, m) = \sum_{m_0, m_2, m_4, \cdots} \binom{M}{m_0, m_2, m_4, \cdots} \binom{k}{0}^{m_0} \binom{k}{2}^{m_2} \cdots, \tag{7}$$

since we need to assign edges to zero-red degree checks, two-red degree checks etc, and then count all the ways they can be mapped to the k incoming edges of each check. We now proceed with a claim:
Claim:

$$\mathbf{P}(\text{red edges from right match edges from left}) = \frac{(lw)!(F-lw)!}{F!} = \binom{F}{lw}^{-1}. \tag{8}$$

Proof: From the check size, the number of ways to color the edges is $C(k,m)$. From the variable side, there are

$$(lw)!(F-lw)! \tag{9}$$

ways to order the lw red and the remaining blue edges, and a total $F!$ permutations. □

Putting everything together, we find that

$$\overline{\mathcal{W}}(w) = \binom{N}{w}\frac{(lw)!(F-lw)!}{F!}C(k,M). \tag{10}$$

Now to obtain the asymptotic behavior, use $\omega = w/N$, $m_r = x_r M$ and Stirling's formula,

$$\varphi(\omega) = max_{x_r}\left[(1-l)H(\omega) + \frac{l}{k}\sum_r(-x_r\log x_r + x_r\log\binom{k}{r}\right], \tag{11}$$

where the maximum is taken over all choices of of fractions of checks $x_0, x_2, x_4, \ldots$, subject to the constraints

$$x_r = \frac{m_r}{M}, \tag{12}$$

$$\sum_r x_r = 1, \tag{13}$$

$$\sum_r r x_r = l\omega. \tag{14}$$

which simply constraint x_r to be proper fractions of checks.

Exercise: Show by imposing two lagrange multipliers that the maximum is obtained by setting $x_r = cz^r\binom{k}{r}$ for proper constants c, z which are implicitly determined by the constraints

$$c = \frac{2}{(1+z)^k + (1-z)^k}, \tag{15}$$

and

$$\omega = z\frac{(1+z)^{k-1} - (1-z)^{k-1}}{(1+z)^k + (1-z)^k}. \tag{16}$$

By plugging back the optimizing x_r into the expression for $\varphi(\omega)$ we obtain

$$\varphi(\omega) = (1-l)H(\omega) + \frac{l}{k}\log\frac{(1+z)^k + (1-z)^k}{2} - \omega l\log z. \tag{17}$$

One can now use this expression to optimize the degrees so that $\varphi(\omega)$ is negative for small weights. This will ensure that the code ensemble has a large distance.

II. General Degree distributions

We now want to compute the behavior of the exponent of of the (expected) weight enumerating function for a general LDPC(Λ, P) code. It is essentially the same technique, but the computation is heavier because we need to count the ways the edges can be adjacent to the variables and the checks, maintaining their degree distributions. The main problem is that the number of red edges E is no longer determined by w.

Claim: The number of ways of choosing X variables, with $|X| = w$, that is adjacent to E red edges is

$$\text{coeff}[\prod_l (1 + xy^l)^{N\Lambda_l}, x^w y^E] \tag{18}$$

This denotes the coefficient of $x^w y^E$ in the formal power series expansion of $\prod_l (1+xy^l)^{N\Lambda_l}$, This is because, selecting each varible of degree Λ_l contributes l edges (the variable counting edges in the generating function is y) and one red node. There are $\Lambda_l N$ variables of degree l and the cofficient of $x^w y^E$ corresponds exactly to the number of ways that E edges can be mapped to w variables.

Similarly we have to count the number of ways E red edges can be adjacent to the check nodes. The constraint now is that each check must have an even number of adjacent red edges:

Claim: The number of ways of mapping E edges to M check nodes in such a way that every check has even degree is

$$\text{coeff}[\prod_k q_k(z)^{MP_k}, z^E], \tag{19}$$

where

$$q_k(z) = \frac{1}{2}(1+z)^k + \frac{1}{2}(1-z)^k. \tag{20}$$

The number of ways one can match the variables and the checks is (as in the regular case) $E!(F-E)!$, where

$$F = N\Lambda'(1) = MP'(1) \tag{21}$$

is the total number of edges.

Therefore, putting the pieces together, we obtain

$$\overline{\mathcal{W}}(w) = \sum_{E=0}^{F} \frac{E!(F-E)!}{F!} \text{coeff}[\prod_{l=1}^{l_{max}} (1 + xy^l)^{N\Lambda_l}, x^w y^E] \text{coeff}[\prod_{k=2}^{k_{max}} q_k(z)^{MP_k}, z^E]. \tag{22}$$

We want to estimate the leading exponential behavior for large N, when $w = N\omega$. Set $E = F\xi = N\Lambda'(1)\xi$. The asymptotic behavior of the generating function coefficient terms can be estimated using the saddle point method. The main idea that we will use is that by the Cauchy theorem

$$\mathrm{coeff}[\prod_k q_k(z)^{MP_k}, z^E] = \oint \frac{1}{z^{N\lambda'(1)\xi+1}} \prod_{k=2}^{k_{max}} q_k(z)^{MP_k} \frac{dz}{2\pi i} = \oint \frac{F(z)^N}{z} \frac{dz}{2\pi i}. \tag{23}$$

where the integral is over any path that contains the origin. In the next lecture we will obtain asymptotic approximations for these integrals.