STAT 206A: Gibbs Measures 10

Lecture 10

Lecture date: Sept 28 *Scribe: Aleksandr Simma*

1 Introduction to Linear Codes

For the purposes of coding, we will be working with linear algebra on finite fields. We will only need to work with the field F_2^n, composed of the elements $\{0,1\}^n$ and the operations $\oplus$ and $\cdot$.

Definition 1 *A linear code C is a linear subspace of* F_2^n

Claim 2 *For every linear code C, there exist the parity check matrix H and the generator matrix G which satisfy* $C = \{x \in F_2^n : Hx = 0\}$ *and* $C = imageG = \{Gy : y \in \{0,1\}^m\}$

These matrices H and G allow us to characterize and study codes. However, for a particular code, H and G are not necessarily unique. The parity check matrix has a clear interpretation as an adjacency matrix of the code's factor graph: each row represents a factor node, and the value of each column in the row indicates whether the corresponding variable node is connected to the factor node or not.

Definition 3 *The factor graph of a parity check matrix H has n variable nodes* $x_1, \ldots, x_n$ *and m factor nodes. Each factor node corresponds to a row of H and is connected to all the x nodes which have non-zero coefficients in the row.*

Each factor node (and row of H) prevents some values of x from being code words. Thus, any x which is not restricted by any of the factors is a valid code word, and so the uniform distribution factorizes into a product of indicators, one for each factor nodes. $P_H(x) = \frac{1}{Z}\prod_{a=0}^{m} 1(H_a X = 0)$

Claim 4 *Observe that for any linear subspace, the addition of a constraint linearly independent from the previous constraints eliminates half the elements in the subspace, so it follows that* $|C| = 2^{N-Rank(H)} \geq 2^{N-M}$

2 Ensembles of factor graphs with a given degree distribution.

Recall that $\Lambda(x) = \sum_{n=0}^{\infty} \Lambda_n x^n$ where Λ_n is the fraction of variable nodes of degree N. Analogously, $P(x) = \sum_{n=0}^{\infty} \mathbb{P}_n x^n$ where $\mathbb{P}_n$ is the fraction of factor nodes of degree n.

One way of defining a uniform distribution over factor nodes is by stating that

Definition 5 $\mathbb{D}_N(\Lambda, P)$ *is the uniform distribution over a graph containing N variable nodes, and the degree profile given by* (Λ, P)

However, this definition raises several questions. First, does a single graph satisfying these constraints exist? How can we sample from this distribution? It turns out that it is very difficult to sample from the distribution coming from this definition, even on regular graphs. Thus, this definition is not very useful.

An alternative definition is the "Configuration Model," which results in a different, but close, distribution over graphs.

Definition 6 *Define a uniform distribution over random graphs with* $N\Lambda_i$ *verteces of degree* i*, and* MP_i *factor nodes of degree* i*.*

We can sample this model with the following algorithm:

1. For each of the $N\Lambda_i$ variable nodes, create i variable sub-nodes.
2. For each of the MP_i factor nodes, create i factor sub-nodes.
3. Randomly, connect each variable sub-node to a factor sub-node to produce a graph.

The a diagram of the generating process is in Figure 1.

Although this is not the same model as the prior, intractable definition, the models are closely related when the degrees of the nodes are not too high. For purposes of coding, when there are multiple (m) edges between a particular variable and factor node, they are replaced by $m \mod 2$. This produces the same code, as double-xor is always 0, but eliminates the problem of duplicate edges.

Exercise 7 *Suppose* $\Lambda(1), P(1) < 0$ *and let* m *be the number of parallel edges. Show that* $E(m) = O(1)$

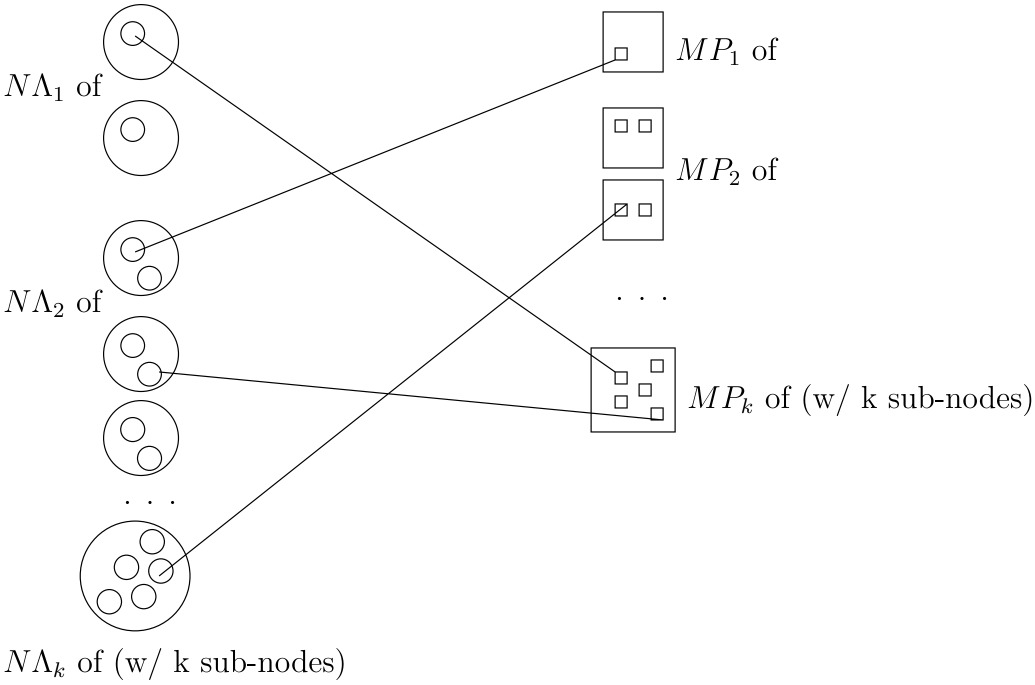

Figure 1: Generating a random factor graph. The edges in the middle are random, but every small circle on the left has to be paired with a small square on the right.

Exercise 8 *Let $\bar{\Lambda}$ be the degree distribution of the variable nodes before removing the parallel edges, and Λ be the degree distribution after. Similarly, let $\bar{P}$ be the degree distribution of factor nodes before parallel edge removal, and P be the degree distribution after.*

Show that $\mathbf{E}(\sum_l |\Lambda_l - \bar{\Lambda}_l| + \sum_l |P_l - \bar{P}_l|) = O(\frac{1}{n})$

Definition 9 *A low density parity check code (LDPCC) is an ensemble of codes of* $D_N(\Lambda, P)$

Underlying this definition is the assumption that we truncate Λ, P at $n(N)$ such that $n << n(N) << N$

The design rate of the code is $\frac{N-M}{N} = 1 - \frac{\Lambda'(1)}{P'(1)}$

Since the code is linear, the distribution of code words around a particular code word X is the same for all X. Thus, in order to study the properties of the code, it is sufficient to consider $X = 0$.

One desired property of a code is that the code words be far apart, so that drift from errors in transmission would be lower than the spacing between code words.

Definition 10 *For a linear code C, the weight-enumerator function $W_N(\omega)$ is the number of words of weight ω, where $w(x) = d_H(x, 0) = \{i : x_i = 1\}$*

For $LDPCC_N(\Lambda, P)$, let $\bar{W}(\omega) = \mathbf{E}(W(\omega))$. We write $\bar{W}(N\omega)$ for $\bar{W}(N\omega) \approx \exp(N\varphi(\omega))$ meaning $\varphi(\omega) = \frac{\log \bar{W}(N\omega)}{N}$ plus lower-order terms.

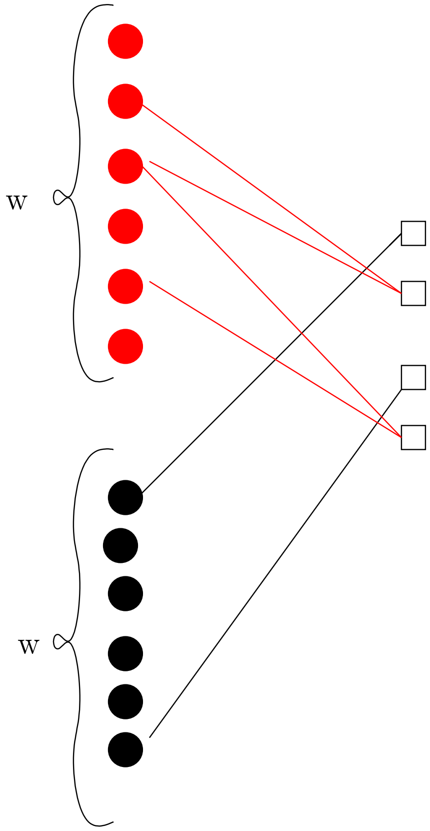

Figure 2: The bipartite graph used for counting LDPCCs for estimating $\bar{W}$

It is not easy to analyze $\bar{W}$ but analyzing W is much harder. Computing $\bar{W}$ involves looking at an x such that $W(x) = w$ and counting $LDPCC_N(\Lambda, P)$ for which x is a code word. In order to do that, we construct a bipartite graph with variable nodes on the left, and factors on the right. We represent edges from $x_i = 1$ to a factor node with red edges, and edges from $x_i = 0$ by black edges. In that case, it is sufficient to count the number of factor graphs where every factor node is adjacent to an even number of red edges. This construction can be seen in Figure 2.