Lecture Notes 7: Analysis of BP for the BEC via density evolution

Lecturer: Rüdiger Urbanke *Scribe:Andrei Giurgiu*

1 Analysis of Belief Propagation for the BEC via Density Evolution

We consider a binary erasure channel with probability of erasure ϵ and an (l, r)-LDPC code. Let $P_b^{BP}(l, r, \epsilon, n, L)$ denote the bit error probability of BP decoding an ensemble of size n, using L iterations (a computation tree of depth L). We know from the last lecture that the limit of this quantity as n goes to infinity is given by $F(\epsilon, x_{L-1})$, where

$$F(\epsilon, x) = \epsilon(1 - (1 - x)^{r-1})^l.$$

The quantity x corresponds to the probability of erasure at each iteration, and we can assume $x_0 = 1$. It evolves after each iterations according to the recurrence $x_i = f(\epsilon, x_{x-1})$, where

$$f(\epsilon, x) = \epsilon(1 - (1 - x)^{r-1})^l.$$

We analyze the sequence $\{x_i\}$ and ask whether it converges to 0 or not. In case it does, the decoding is successful, otherwise it is not. Note that convergence depends on ϵ, l, and r.

Remark 1. *The function $f(\epsilon, x)$ is increasing in ϵ and x for $x, \epsilon \in [0, 1]$.*

Lemma 1. *We have that*

- *the x_i are decreasing in i,*
- *$x_i(\epsilon) \leq x_i(\epsilon')$ if $\epsilon \leq \epsilon'$.*

Proof.

- By induction. The first two elements of the sequence are $x_0 = 1$ and $x_1 = f(\epsilon, x_0) = \epsilon$. We assume $x_{i-1} \leq x_{i-2}$ as the induction hypothesis. Since $f(\epsilon, \cdot)$ is increasing, we obtain $f(\epsilon, x_{i-1}) \leq f(\epsilon, x_{i-2})$. The left hand side is equal to x_i, and the right hand side to x_{i-1}, and we deduce that $x_i \leq x_{i-1}$.
- By induction. We first have $x_1(\epsilon) = \epsilon \leq \epsilon' = x_1(\epsilon)$. The general statement is deduced as follows:

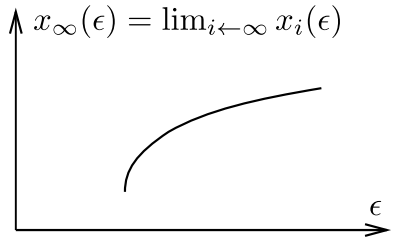
$$x_i(\epsilon) = f(\epsilon, x_{i-1}(\epsilon)) \overset{(a)}{\leq} f(\epsilon', x_{i-1}(\epsilon)) \overset{(b)}{\leq} f(\epsilon', x_{i-1}(\epsilon')) = x_i(\epsilon'),$$

where inequality (a) follows from the fact that f is increasing in ϵ, and inequality (b) follows from it being increasing in x, together with the induction hypothesis.

□

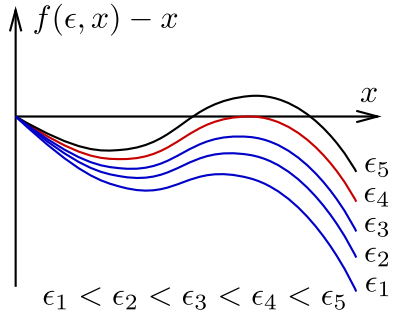
From the first part of the previous lemma, it follows that $x_i(\epsilon)$ converges in $[0,1]$. From the second part, it follows that if $x_i(\epsilon) \to 0$ for some ϵ, then $x_i(\epsilon') \to 0$ for all $\epsilon' < \epsilon$.

We denote by $x_\infty(\epsilon)$ the limit $\lim_{i\to\infty} x_i(\epsilon)$. Then x_∞ is increasing in ϵ.



Hence we can define the quantity ϵ^{BP} as $\sup \epsilon : x_\infty(\epsilon) = 0$; this is called *the threshold*.

There is a graphical way to characterize this threshold. Note that x_∞ is a fixpoint of $f(\epsilon, \cdot)$, i.e. $f(\epsilon, x_\infty) = x_\infty$. Thus, if $f(\epsilon, x) - x < 0$ for all $x < 0$, then $x_\infty = 0$, and as soon as there is a fixpoint $f(\epsilon, x) = x$ in the interval $(0,1)$, we have that $x_\infty > 0$.



Example. For a (3,6)-LDPC code, we have that $\epsilon^{BP} \sim 0.4294$. The transmission rate is $R = 1 - \frac{l}{r} = \frac{1}{2}$, and we know there exists some decoding procedure for erasure probabilities as high as $\epsilon^{Shannon} = 1 - R = \frac{1}{2}$.

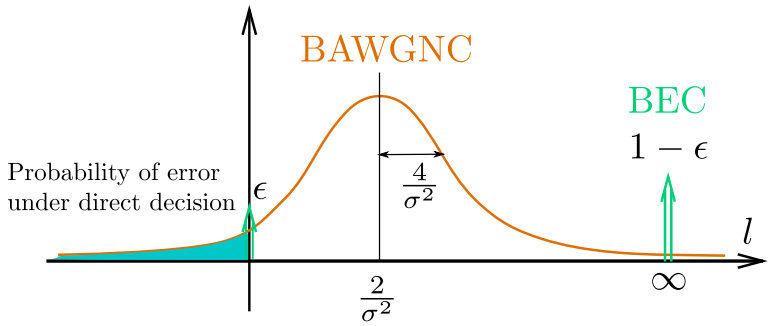
A result by Luby, Mitzenmacher, Shokrollahi, Spielman and Stemann [LMSSS 97] states that there exist LDPC codes so that $\epsilon^{BP} = \epsilon^{Shannon}$. Note that if we stick to regular codes, like $(3,6)$- or $(4,8)$-LDPC, we cannot achieve capacity.

We define Λ_i as the fraction of variable nodes of degree i in the ensemble; in particular, we have that $\Lambda_i \geq 0$ and $\sum_i \Lambda_i = 1$. Likewise we define R_i as the fraction of check nodes of degree i. The question is how to choose Λ_i and R_i such that we achieve capacity. In the cited paper it is shown that there must be nodes of arbitrarily large degree, otherwise capacity will not be attained.

We could also define ϵ^{MAP}, corresponding to a MAP decoder. Note that analyzing MAP is harder than analyzing BP. All three thresholds (ϵ^{BP}, ϵ^{MAP}, and $\epsilon^{Shannon}$) can be arbitrarily close.

2 Generalization to BMS Channels (in particular BAWGN Channel)

We consider a BAWGNC channel, where the output Y is given in terms of the input X by $Y = X + Z$, where $Z \sim \mathcal{N}(0, \sigma^2)$ is the noise (independent of the input).



If we fix the input to $X = 1$, then $Y \sim \mathcal{N}(1, \sigma^2)$. In this case, the log-likelihood is

$$l = \log \frac{p(y|x=1)}{p(y|x=-1)} = \log \frac{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-1)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+1)^2}{2\sigma^2}}} = \frac{1}{2\sigma^2}[(y+1)^2 - (y-1)^2] = \frac{2y}{\sigma^2}$$

We then define the random variable L, whose distribution is $\mathcal{N}\left(\frac{2}{\sigma^2}, \sigma^2 \left(\frac{2}{\sigma^2}\right)^2\right) \sim \mathcal{N}\left(\frac{2}{\sigma^2}, \frac{4}{\sigma^2}\right)$. Let $a(l)$ be the probability density function of L.

Direct decision based on L is performed as follows: if L is positive, decide $\hat{X} = 1$, if L is negative, decide $\hat{X} = -1$, and for $L = 0$, decide $\hat{X} = 1$ or $\hat{X} = 0$, each with probability $\frac{1}{2}$. The bit error probability in this case is given by

$$P_b = \int_{-\infty}^{0} a(l)dl = Q\left(\frac{\mu_L}{\sigma_L}\right) == Q\left(\frac{1}{\sigma}\right) \sim e^{-\frac{1}{2\sigma^2}}.$$

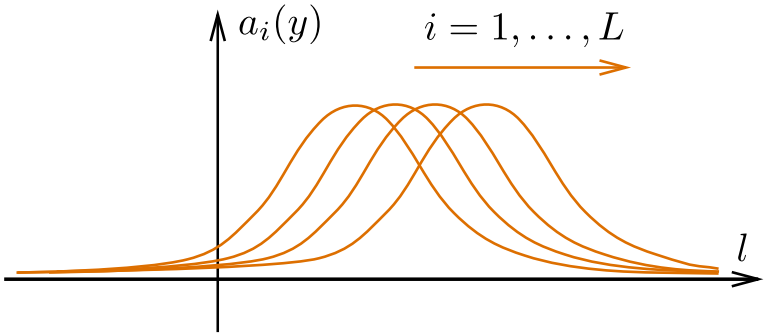
Note that P_b tends to 0 as σ approaches 0.

We now look at the message passing rules in the generalized setting.

- At variable nodes, we assume the incoming messages are $L_1, \ldots, L_{l-1}$, with (i.i.d.) distributions $b_i(y)$. The outgoing message is $L = \sum_i^{l-1} L_i$, with L having the distribution $a_{i+1}(y)$.
- At check nodes, we assume the incoming messages are $L_1, \ldots, L_{l-1}$, with (i.i.d.) distributions $a_i(y)$. The outgoing message is $L = 2\tanh^{-1}\left(\prod_{i=1}^{r-1} \tanh\left(\frac{L_i}{2}\right)\right)$, with L having the distribution $b_{i+1}(y)$.

We consider the distributions a_i after each iteration. By population dynamics approach or numerical calculations we observe that the mass of the distribution migrates toward $+\infty$, and

$$\lim_{n\to\infty} P_b^{BP}(l, r, \sigma, n, L) = \int_{-\infty}^{0} a_L(y)dy.$$



Channel degradation

In order to imitate the proofs that we have obtained for BEC, we need to define an order on the distributions. For two distributions $a(y)$ and $b(y)$, the intuition behind $a(y) \prec b(y)$ is that $a(y)$ is "better" than $b(y)$, and the sequence $a_i(y)$ decreases, as is the case for the BEC.

To define the order relation, we associate to each probability distribution $a(y)$ a binary symmetric channel, with $p(y|x = 1) = a(y)$, and (due to symmetry) $p(y|x = -1) = a(-y)$. Note that the log-likelihood distribution of $p(y|x)$ is exactly $a(y)$, and thus $a(y)$ can be represented by a channel.

For two BMS channels $p(y|x)$ and $q(z|x)$, they are ordered by degradation as $p(y|x) \prec q(z|x)$ if there exists a BMS channel $r(z|y)$, i.e. $q(z|x)$ is the composition of $p(y|x)$ and $r(z|y)$. p

Claim 1. *The $a_i(y)$ are monotonically decreasing under degradation, i.e. $a_1(y) \succ a_2(y) \succ \ldots \succ a_i(y) \succ \ldots$. This is similar in the case of BEC, where we had $x_1 > x_2 > \ldots > x_i > \ldots$.*

Claim 2. *Given two channels AWGN(σ) and AWGN(σ'), we have that $a_i(y; \sigma) \prec a_i(y, \sigma')$.*