EXIT Functions
Ruediger Urbanke, EPFL
February 17th, 2010
~ 2 hours

For the most recent version of these slides visit

Modern Coding Theory
Cambridge University Press

Thursday, February 18, 2010
Density Evolution for BEC: Recall

\[ x_\ell = \epsilon \lambda (1 - \rho (1 - x_{\ell-1})) . \]

\[ f(\epsilon, x) = \epsilon \lambda (1 - \rho(1 - x)) \]

The threshold \( \epsilon^{BP} \) is the largest \( \epsilon \) such that the graph of \( f(\epsilon, x) - x \) is negative.
Component Functions

\[ c(x) = 1 - \rho(1 - x) \]

\[ \nu_\epsilon(x) = \epsilon \lambda(x) \]

\[ f(\epsilon, x) = \epsilon \lambda(1 - \rho(1 - x)) \]

\[ f(\epsilon, x) = \nu_\epsilon(c(x)) \]

\[ f(\epsilon, x) - x \leq 0 \]

\[ c(x) < \nu_\epsilon^{-1}(x), \quad x \in (0, 1). \]
Visualization
c(x) < \nu^{-1}_\epsilon(x), \quad x \in (0, 1).
Visualization

\[ c(x) < \nu_\epsilon^{-1}(x), \quad x \in (0, 1). \]
Matching Condition

\[ \epsilon = 0.35 \]

\[ \nu_\epsilon^{-1}(x) \]

\[ c(0.35) \approx 0.884 \]

\[ c(x) \]

\[ x \]
Matching Condition
Matching Condition
Matching Condition

\[(\epsilon \int \lambda) + (1 - \int \rho) \leq 1\]
Matching Condition

\[(\epsilon \int \lambda) + (1 - \int \rho) \leq 1\]

\[1 - C_{Sha} = \epsilon \leq \frac{\int \rho}{\int \lambda} = 1 - r(\lambda, \rho).\]
Capacity Achieving Degree Distributions for BEC
Capacity Achieving Degree Distributions for BEC

\[ \epsilon \lambda (1 - \rho (1 - x)) = x \]

matching condition
Capacity Achieving Degree Distributions for BEC

\[ \epsilon \lambda (1 - \rho(1 - x)) = x \]  

\[ \hat{\lambda}_\alpha(x) = 1 - (1 - x)^\alpha = \sum_{i=1}^{\infty} \binom{\alpha}{i} (-1)^{i-1} x^i, \quad \rho_\alpha(x) = x^{\frac{1}{\alpha}} \]  

matching condition
Capacity Achieving Degree Distributions for BEC

$$\epsilon \lambda (1 - \rho (1 - x)) = x$$  \hspace{1cm} \text{matching condition}

$$\hat{\lambda}_\alpha(x) = 1 - (1 - x)^{\alpha} = \sum_{i=1}^{\infty} \binom{\alpha}{i} (-1)^{i-1} x^i$$,

$$\rho_\alpha(x) = x^{\frac{1}{\alpha}}$$

$$\binom{\alpha}{i} (-1)^{i-1} = \frac{\alpha(\alpha - 1) \cdots (\alpha - i + 1)}{i!} (-1)^{i-1} = \frac{\alpha}{i} \frac{\alpha}{i-1} \cdots (1 - \alpha)$$  \hspace{1cm} \text{pos. coeff?}
Capacity Achieving Degree Distributions for BEC

\[ e\lambda(1 - \rho(1 - x)) = x \]

matching condition

\[ \hat{\lambda}_\alpha(x) = 1 - (1 - x)^\alpha = \sum_{i=1}^{\infty} \binom{\alpha}{i} (-1)^{i-1} x^i, \quad \rho_\alpha(x) = x^{\frac{1}{\alpha}} \]

pos. coeff?

\[ \binom{\alpha}{i} (-1)^{i-1} = \frac{\alpha(\alpha - 1) \cdots (\alpha - i + 1)}{i!} (-1)^{i-1} = \frac{\alpha}{i} \frac{\alpha - \alpha}{i - 1} \cdots (1 - \alpha). \]

\[ \hat{\lambda}_\alpha^{(N)}(x) \quad \lambda_\alpha^{(N)}(x) = \frac{\hat{\lambda}_\alpha^{(N)}(x)}{\hat{\lambda}_\alpha^{(N)}(1)} \]
we have two parameters, namely $N$ and $\alpha$, and two conditions, namely rate and threshold; can we choose the parameters so that we get a capacity-achieving pair?
Capacity Achieving Degree Distributions for BEC

\[
\int_0^1 \lambda^{(N)}_\alpha(x)dx = \sum_{i=1}^{N-1} \binom{\alpha}{i} \frac{(-1)^i}{i+1} = \frac{\alpha - \binom{\alpha}{N}(-1)^{N-1}}{\alpha + 1},
\]

\[
\lambda^{(N)}_\alpha(1) = \sum_{i=1}^{N-1} \binom{\alpha}{i}(-1)^{i-1} = 1 - \frac{N}{\alpha} \binom{\alpha}{N}(-1)^{N-1},
\]

\[
\int_0^1 \rho_\alpha(x)dx = \frac{\alpha}{1 + \alpha}.
\]

\[
x = \hat{\lambda}_\alpha(1 - \rho_\alpha(1 - x)) \geq \hat{\lambda}^{(N)}_\alpha(1 - \rho_\alpha(1 - x)) = \hat{\lambda}^{(N)}_\alpha(1) \hat{\lambda}^{(N)}_\alpha(1 - \rho_\alpha(1 - x)).
\]

\[
e^{\text{BP}}(\alpha, N) \geq \hat{\lambda}^{(N)}_\alpha(1).
\]

\[
r(\alpha, N) = 1 - \frac{\int_0^1 \rho_\alpha(x)dx}{\int_0^1 \lambda^{(N)}_\alpha(x)dx} = 1 - \hat{\lambda}^{(N)}_\alpha(1) \frac{\int_0^1 \rho_\alpha(x)dx}{\int_0^1 \hat{\lambda}^{(N)}_\alpha(x)dx} = \frac{N}{\alpha} \binom{\alpha}{N}(-1)^{N-1}(1 - 1/N)
\]

\[
\leq \frac{1 - N \binom{\alpha}{N}(-1)^{N-1}}{N - N \binom{\alpha}{N}(-1)^{N-1}}.
\]

\[
r(N) = 1 - \epsilon + O(1/N), \quad \delta(N) \leq \frac{1 - (1-\epsilon)}{N - (1-\epsilon)} = O(1/N).
\]
EXIT Functions For General Channels

\[ c(x) = 1 - (1 - x)^5 \]

\[ v_{\varepsilon}^{-1}(x) = (x/\varepsilon)^{1/2} \]

generalization?
**Entropy Functional**

**Definition 4.40 (Entropy Functional).** The entropy $H(a)$ associated with a symmetric $L$-density $a$ is $H$

$$H(a) = \int_{-\infty}^{+\infty} a(y) \log_2(1 + e^{-y}) dy$$
EXIT Functions For General Channels
EXIT Functions For General Channels

Example 4.133 ([n, n−1, 2] Parity-Check Code − {BSC(h)}). Again by symmetry, \( h(h) = h_i(h), \) \( i \in [n], \) and

\[
    h(h) = H(a_{\text{BSC}(\varepsilon)}^{(n-1)}) = H(a_{\text{BSC}(\frac{1-(1-2\varepsilon)^{n-1}}{2})}) = h_2 \left( \frac{1 - (1 - 2\varepsilon)^{n-1}}{2} \right),
\]

where \( \varepsilon = h_2^{-1}(h). \)
EXIT Functions For General Channels

**Example 4.133 ([n, n - 1, 2] Parity-Check Code – {BSC(h)})**. Again by symmetry, \( h(h) = h_i(h), \ i \in [n], \) and

\[
h(h) = H(a_{BSC(\varepsilon)}^{(n-1)}) = H(a_{BSC}(\frac{1-(1-2\varepsilon)^{n-1}}{2})) = h_2 \left( \frac{1 - (1 - 2\varepsilon)^{n-1}}{2} \right),
\]

where \( \varepsilon = h_2^{-1}(h) \).

**Example 4.132 ([n, 1, n] Repetition Code)**. Assume that transmission takes place over a family of BMS channels characterized by their L-densities \( \{a_{BMS(h)}\} \). By symmetry, \( h(h) = h_i(h), \ i \in [n], \) and

\[
h(h) = H(a_{\oplus BMS(h)}^{(n-1)}).
\]

If we specialize to the family \( \{\text{BSC}(h)\} \) we get

\[
h(h) = \sum_{i=0}^{n-1} \binom{n-1}{i} \varepsilon^i \varepsilon^{n-1-i} h_2 \left( \frac{\varepsilon^{n-1-2i}}{\varepsilon^{n-1-2i} + \varepsilon^{n-1-2i}} \right),
\]
EXIT Functions For General Channels
Theorem 4.141 (Extremes of Information Combining). Let $a$ and $b$ represent two BMS channels and fix $H(b) = h$, $0 \leq h \leq 1$. Then

- Repetition code:
  \[ H(a)_h \leq H(a \bigotimes b_{\text{BEC}(h)}) \leq H(a \bigotimes b) \leq H(a \bigotimes b_{\text{BSC}(h)}), \]

- Parity-check code:
  \[ H(a \bigoplus b_{\text{BSC}(h)}) \leq H(a \bigoplus b) \leq H(a \bigoplus b_{\text{BEC}(h)}) \cdot \frac{1-(1-H(a))(1-h)}{1-(1-H(a))(1-h)}. \]
Extremes of Information Combining

**Lemma 4.41 (Duality Rule for Entropy).** Let $a$ and $b$ denote two symmetric $L$-densities. Then

$$H(a \otimes b) + H(a \boxtimes b) = H(a) + H(b).$$
Extremes of Information Combining

**Lemma 4.41 (Duality Rule for Entropy).** Let $a$ and $b$ denote two symmetric $L$-densities. Then

$$H(a \otimes b) + H(a \boxtimes b) = H(a) + H(b).$$

$$H(c) = \int c(z) \log_2(1 + e^{-z})dz$$

$$= \int \int a(x)b(y) \log_2(1 + e^{-2 \tanh^{-1}(\tanh(x/2) \tanh(y/2)))}dx \ dy$$

$$= \int \int a(x)b(y) \log_2\left(\frac{(1 + e^{-x})(1 + e^{-y})}{1 + e^{-x-y}}\right)dx \ dy$$

$$= H(a) + H(b) - H(a \circ b).$$
Proof
Proof

\[ a(y) = \int_0^1 w_a(h) a_{\text{BSC}(h_2^{-1}(h))}(y) \, dh \]  
\[ b(y) = \int_0^1 w_b(h) a_{\text{BSC}(h_2^{-1}(h))}(y) \, dh. \]
Proof

\[ a(y) = \int_{0}^{1} w_a(h) a_{\text{BSC}(h^{-1}(h))}(y) dh \]
\[ b(y) = \int_{0}^{1} w_b(h) a_{\text{BSC}(h^{-1}(h))}(y) dh. \]

\[ H(a \ast b) = \int_{0}^{1} \int_{0}^{1} w_a(h_a) w_b(h_b) h_2(\varepsilon_a \bar{\varepsilon}_b + \bar{\varepsilon}_a \varepsilon_b) dh_a dh_b \]
\[ = \int_{0}^{1} w_a(h_a) \left( \int_{0}^{1} w_b(h_b) h_2(\varepsilon_b(1 - 2\varepsilon_a) + \varepsilon_a) dh_b \right) dh_a \]
Proof

\[ a(y) = \int_0^1 w_a(h) a_{\text{BSC}(h^{-1}\{h\})}(y) \, dh \]
\[ b(y) = \int_0^1 w_b(h) a_{\text{BSC}(h^{-1}\{h\})}(y) \, dh. \]

\[ H(a \Box b) = \int_0^1 \int_0^1 w_a(h_a) w_b(h_b) h_2(\epsilon_a \tilde{\epsilon}_b + \tilde{\epsilon}_a \epsilon_b) \, dh_a \, dh_b \]
\[ = \int_0^1 w_a(h_a) \left( \int_0^1 w_b(h_b) h_2(\epsilon_b (1 - 2\epsilon_a) + \epsilon_a) \, dh_b \right) \, dh_a \]

\[ \epsilon_{a/b} = h_2^{-1}(h_{a/b}) \]
Proof

\[ a(y) = \int_0^1 w_a(h) a_{BSC(h_2^{-1}(h))}(y) \, dh \quad b(y) = \int_0^1 w_b(h) a_{BSC(h_2^{-1}(h))}(y) \, dh. \]

\[ H(a \boxtimes b) = \int_0^1 \int_0^1 w_a(h_a) w_b(h_b) h_2(\epsilon_a \bar{\epsilon}_b + \bar{\epsilon}_a \epsilon_b) \, dh_a \, dh_b \]
\[ = \int_0^1 w_a(h_a) \left( \int_0^1 w_b(h_b) h_2(\epsilon_b(1 - 2\epsilon_a) + \epsilon_a) \, dh_b \right) \, dh_a \]

\[ \epsilon_{a/b} = h_2^{-1}(h_{a/b}) \quad h_2(h_2^{-1}(u)(1 - 2v) + v) \]
Proof

$v$ fixed; as a function of $u$

$h_2\left(h_2^{-1}(u)(1-2v)+v\right)$

$v$ fixed; as a function of $u$

non-decreasing and convex in $u$, $u \in [0, 1]$
Proof -- Lower Bound

\[ \int w_b(h_b) h_2(e_b(1-2\varepsilon_a) + \varepsilon_a) dh_b \geq h_2(h_2^{-1}(h)(1-2\varepsilon_a) + \varepsilon_a) \]

\[ H(a \boxtimes b) \geq H(a \boxtimes b_{\text{BSC}(h_2^{-1}(h))}) \]
Proof -- Upper Bound On Entropy

\[ h_2(\epsilon_b(1 - 2\epsilon_a) + \epsilon_a) \leq h_a(1 - h_b) + h_b = 1 - (1 - h_a)(1 - h_b) \]

\[ H(a \boxtimes b) \leq 1 - (1 - H(a))(1 - h) = H(a \boxtimes b_{BEC(h)}) . \]
Bounds On Threshold via Extremes of Information Combining

universal lower bound on threshold
Bounds On Threshold via Extremes of Information Combining

\[ H(a_{BSC(h)} \otimes a_{BSC(\hat{h})}) = (e^2 \hat{e} + e^2 \hat{e}) h_2 \left( \frac{e^2 \hat{e}}{e^2 \hat{e} + e^2 \hat{e}} \right) + 2e\hat{e} h_2 (\hat{e}) + (e^2 \hat{e} + e^2 \hat{e}) h_2 \left( \frac{e^2 \hat{e}}{e^2 \hat{e} + e^2 \hat{e}} \right) \]

universal lower bound on threshold
Bounds On Threshold via Extremes of Information Combining

\[ H(a_{BSC} \otimes a_{BSC}) = (e^2 \hat{e} + e^2 \hat{e}) h_2 \left( \frac{e^2 \hat{e}}{e^2 \hat{e} + e^2 \hat{e}} \right) + 2e\hat{e} h_2 (e) + (e^2 \hat{e} + e^2 \hat{e}) h_2 \left( \frac{e^2 \hat{e}}{e^2 \hat{e} + e^2 \hat{e}} \right) \]

\[ 1 - (1 - h)^5 \]

universal lower bound on threshold
Bounds On Threshold via Extremes of Information Combining

\[ H(a_{BSC(h)}^2 \otimes a_{BSC(\hat{h})}) = (e^2 \hat{e} + e^2 \hat{e}) h_2 \left( \frac{e^2 \hat{e}}{e^2 \hat{e} + e^2 \hat{e}} \right) + 2e\hat{e} h_2(\hat{e}) + (e^2 \hat{e} + e^2 \hat{e}) h_2 \left( \frac{e^2 \hat{e}}{e^2 \hat{e} + e^2 \hat{e}} \right) \]

universal lower bound on threshold

\[ 1 - (1 - h)^5 \]
(G)EXIT Functions and the Area Theorem
(G)EXIT Functions and the Area Theorem

All bits are transmitted through a family of BMS channels parameterized by the entropy $h$. 
(G)EXIT Functions and the Area Theorem

All bits are transmitted through a family of BMS channels parameterized by the entropy $h$.

**Area Theorem = Definition of GEXIT Function**

$$g(h) \triangleq \frac{dH(X | Y(h))}{dh}$$  \hspace{1cm} (GEXIT Function)

$$n \int_{h_a}^{h_b} g(h) \, dh = H(X | Y(h_b)) - H(X | h_a).$$
(G)EXIT Functions and the Area Theorem
Since the channel is memoryless

\[ H(X|Y) = H(X_i|Y) + H(X_{\sim i}|Y, X_i) \]
\[ = H(X_i|Y) + H(X_{\sim i}|X_i, Y_{\sim i}). \]
(G)EXIT Functions and the Area Theorem

Since the channel is memoryless

\[ H(X|Y) = H(X_i|Y) + H(X_{\sim i}|Y, X_i) \]
\[ = H(X_i|Y) + H(X_{\sim i}|X_i, Y_{\sim i}). \]

Therefore, \( \frac{\partial H(X|Y)}{\partial h_i} = \frac{\partial H(X_i|Y)}{\partial h_i} + 0, \) so that

\[ g(h) = \frac{dH(X|Y)}{ndh} = \frac{1}{n} \sum_{i=1}^{n} \left. \frac{\partial H(X_i|Y)}{\partial h_i} \right|_{h_i=h} \triangleq \frac{1}{n} \sum_{i=1}^{n} g_i(h) \]
GEXIT versus EXIT
GEXIT versus EXIT

For the BEC, $H(X_i|Y) = h_i \cdot H(X_i|Y_{\sim i})$ so that

$$g(h) = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial H(X_i|Y)}{\partial h_i} = \frac{1}{n} \sum_{i=1}^{n} H(X_i|Y_{\sim i}) = h(h).$$

For the BEC, GEXIT = EXIT!
For the BEC, \( H(X_i|Y) = h_i \cdot H(X_i|Y_{\sim i}) \) so that

\[
g(h) = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial H(X_i|Y)}{\partial h_i} = \frac{1}{n} \sum_{i=1}^{n} H(X_i|Y_{\sim i}) = h(h).
\]

For the BEC, GEXIT = EXIT!

**Definition (EXIT function)**

\[
h_i(h) \triangleq H(X_i|Y_{\sim i}(h))
\]

\[
h(h) = \frac{1}{n} \sum_i h_i(h)
\]
Example 3.71 (Parity-Check Code). Consider the binary parity-check code with parameters $C[n, n-1, 2]$ and a uniform prior on the set of codewords. By symmetry, $h_i(e) = h(e)$ for all $i \in [n]$ and

$$h(e) = 1 - (1 - e)^{n-1}.$$ 

Example 3.72 (Repetition Code). Consider the binary repetition code $C[n, 1, n]$ with a uniform prior on the set of codewords. By symmetry, $h_i(e) = h(e)$ for all $i \in [n]$ and

$$h(e) = e^{n-1}.$$
(G)EXIT Functions and the Area Theorem

**Example 3.73 ([7, 4, 3] Hamming Code).** The parity-check matrix $H$ of the [7, 4, 3] binary Hamming code is stated explicitly in Example 1.24. Assuming a uniform prior on the set of codewords, a tedious calculation reveals (see Problem 3.30) that

$$h_{\text{Ham}}(\epsilon) = 3\epsilon^2 + 4\epsilon^3 - 15\epsilon^4 + 12\epsilon^5 - 3\epsilon^6.$$  

**Example 3.83 (Area Theorem Applied to [7, 4, 3] Hamming Code).**

$$\int_0^1 h_{\text{Ham}}(\epsilon) \, d\epsilon = \int_0^1 (3\epsilon^2 + 4\epsilon^3 - 15\epsilon^4 + 12\epsilon^5 - 3\epsilon^6) \, d\epsilon = \frac{4}{7}.$$
MAP versus BP

Example: LDPC \( \left( \frac{x+4x^3}{5}, \frac{x^3+4x^7}{5} \right) \)

- BP threshold: \( h^{BP} \approx 0.427264 \)
- Shannon threshold: \( h^{Sh} = 0.5 \)
MAP versus BP

Example: LDPC($\frac{x+4x^3}{5}, \frac{x^3+4x^7}{5}$)

$$\int_0^1 h^{MAP}(h) dh = r$$
MAP versus BP

Example: LDPC\(\left(\frac{x+4x^3}{5}, \frac{x^3+4x^7}{5}\right)\)

Construct \(\bar{h}\) such that

\[
\int_{0}^{1} h^{BP}(h) \, dh = r
\]

But MAP EXIT \(\leq\) BP EXIT, then

\[
h^{MAP} \leq \bar{h}.
\]

\[
0.4273 \lesssim \bar{h} \approx 0.4820 \lesssim 0.5000
\]
Theorem 3.120 (EXIT Function for Regular Degree Distributions). Consider the $(1, r)$-regular degree distribution, let $P(x)$ denote the associated trial entropy, and define $\epsilon(x) = x/\lambda(1 - \rho(1 - x))$. Let $x^{\text{MAP}}$ be the unique positive solution of $P(x) = 0$ and define $\epsilon^{\text{MAP}} = \epsilon(x^{\text{MAP}})$. Then

$$h(\epsilon) = \begin{cases} 0, & \epsilon \in [0, \epsilon^{\text{MAP}}), \\ h^{\text{BP}}(\epsilon), & \epsilon \in (\epsilon^{\text{MAP}}, 1], \end{cases}$$

and for $\epsilon^{\text{MAP}} \leq \epsilon \leq 1$

$$\lim_{n \to \infty} \mathbb{E}_G[H_G(X | Y(\epsilon))/n] = \int_0^\epsilon h(\epsilon')d\epsilon' = P(x(\epsilon)),$$

where $x(\epsilon)$ is the largest solution of the equation $\epsilon(x) = \epsilon$. 
MAP versus BP

Example: LDPC\(\left(\frac{x+4x^3}{5}, \frac{x^3+4x^7}{5}\right)\)

\[
\left(\frac{x}{\lambda(1 - \rho(1 - x))}, \Lambda(1 - \rho(1 - x))\right)
\]

\text{AREA} = \text{rate}
(G)EXIT Functions and the Area Theorem
(G)EXIT Functions and the Area Theorem
Van der Waals and Maxwell
Van der Waals and Maxwell

Van der Waals isotherm (canonical at $T = 0.85$)

$P + \frac{2}{V^2} (3V - 1) = 8T$

compression of gas

$P_c$

$V$

$0 \frac{1}{3}$

$8$
Questions

Maxwell Construction
In the original Maxwell construction the areas represent work. Do we have a similar interpretation in the setting of iterative decoding?
Questions

✦ Optimization?
✦ Can we achieve capacity in general?
✦ How much loss of BP versus MAP?


