Modern Channel Coding

Ingmar Land & Jossy Sayir

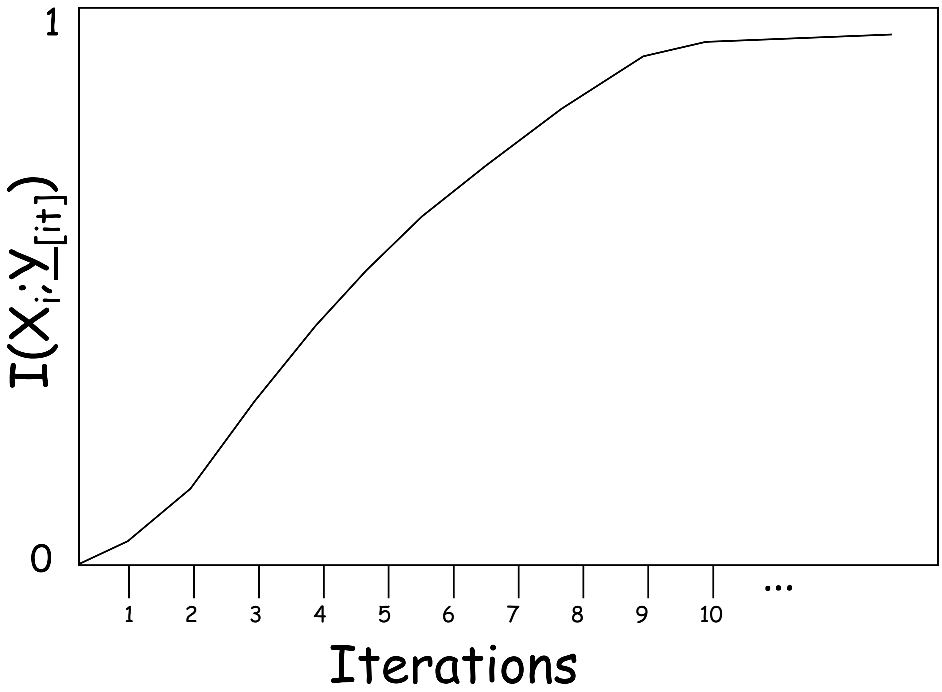
Lecture 4: EXIT Charts

ACoRN Summer School 2007

Iterative Decoding

- How does the mutual information evolve in an iterative decoding algorithm?
- We have learned that it is possible to optimize LDPC codes so as to maximize their threshold
- We will see that we can design capacity-achieving, iteratively decodable families of LDPC codes!!
  (i.e., threshold → capacity)
- What is the implication in terms of mutual information?

Mutual Information Trajectory



Mutual Information Trajectory

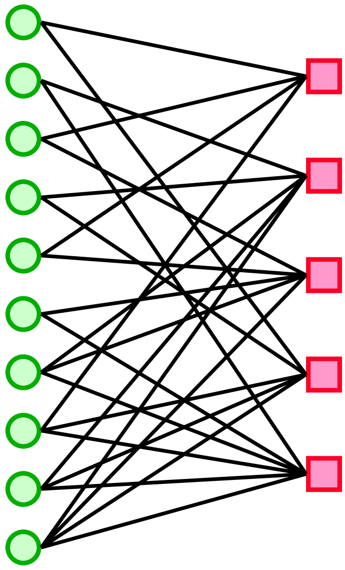
- The L-values calculated in the tree are optimal in the sense of a MAP-calculator, i.e., $L(X_i|\underline{Y}_{[it]})$ is a sufficient statistic for $\underline{Y}_{[it]}$:

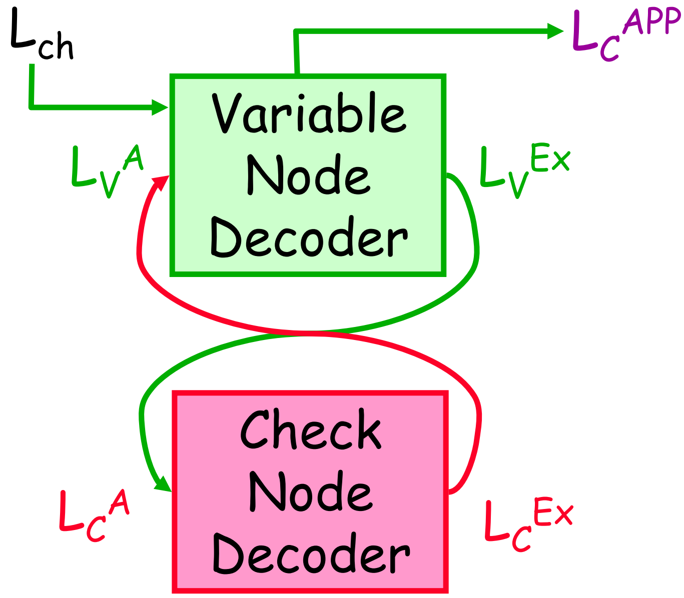
$$I(X_i\,;\,L(X_i|\underline{Y}_{[it]})) = I(X_i\,;\,\underline{Y}_{[it]})$$

- We can also draw the trajectory at half-iterations (after variable nodes & after check nodes)

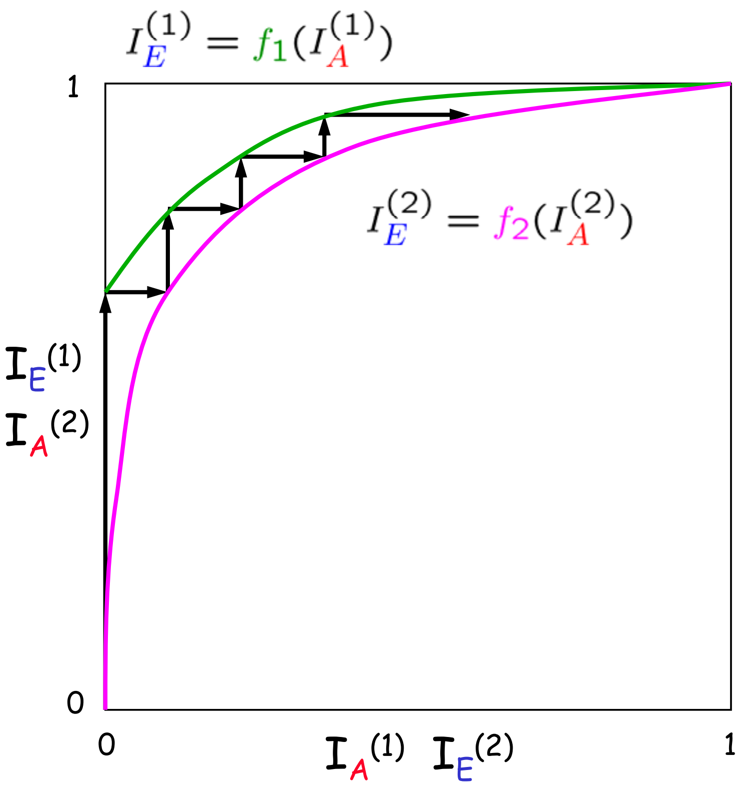
- But: the output messages of variable nodes and check nodes are extrinsic L-values, whereas the mutual information trajectory we consider now is for a-posteriori L-values

Message Passing





Tracking of Messages



This assumes that the decoder depends only on mutual information!

Problem:
How to compute the "transfer functions" f_1 and f_2?

Tracking of Messages

Tracking of messages would mean tracking of pdfs

(→ Density Evolution)

Instead of tracking the pdfs we reduce the problem to tracking of mutual information between the messages and the codeword which are scalar quantities

$$I_A = \frac{1}{N}\sum_{n=1}^{N} I(X_i; A_i)$$

$$I_E = \frac{1}{N}\sum_{n=1}^{N} I(X_i; E_i)$$

$$I_A^{(1)} = 0$$

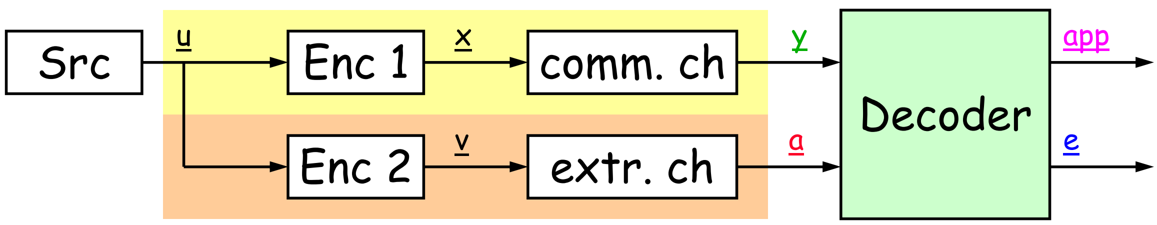
$$\Rightarrow I_E^{(1)} \rightarrow I_A^{(2)}$$

$$\Rightarrow I_E^{(2)} \rightarrow I_A^{(1)}$$

$$\cdots$$

I_A, I_E ..... average symbolwise mutual information

Extrinsic Channel Model



A-priori messages are modeled as independent noisy observations of the encoded source.

Assumptions:

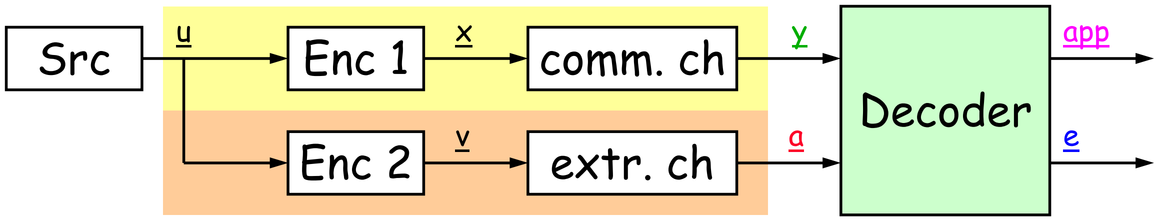
- independent observations
- model for extrinsic channel

$$I_A = I(X_i; A_i) = I(V_i; A_i)$$

$$I_E = I(X_i; E_i) \leq I(X_i; \underline{Y}\,\underline{A}_{\backslash i})$$

with equality if the decoder is optimal

Transfer Functions

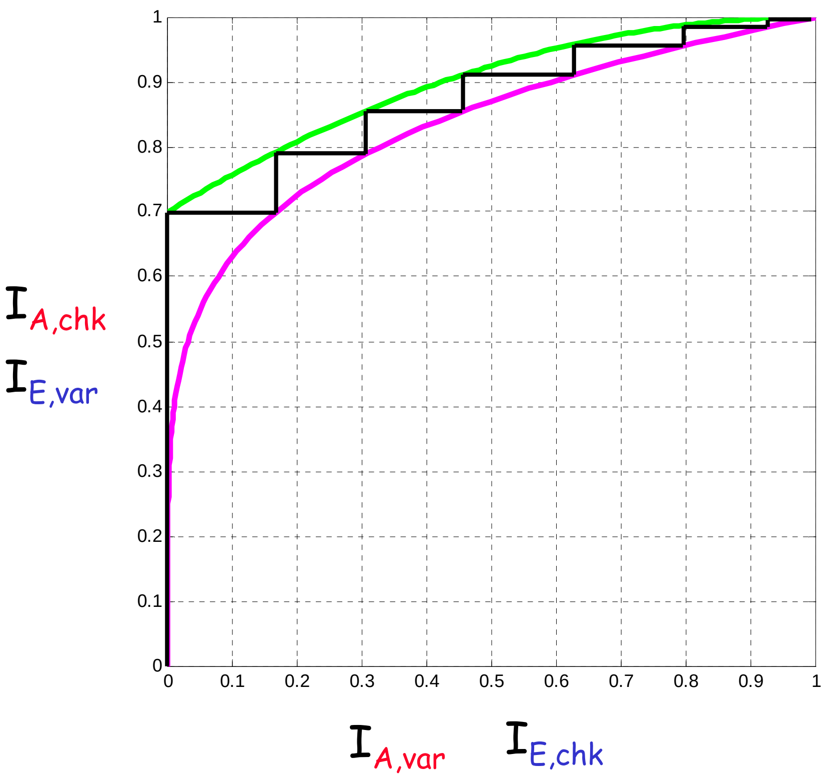


Assuming a model for the extrinsic channel we can vary I_A by varying the channel parameter.

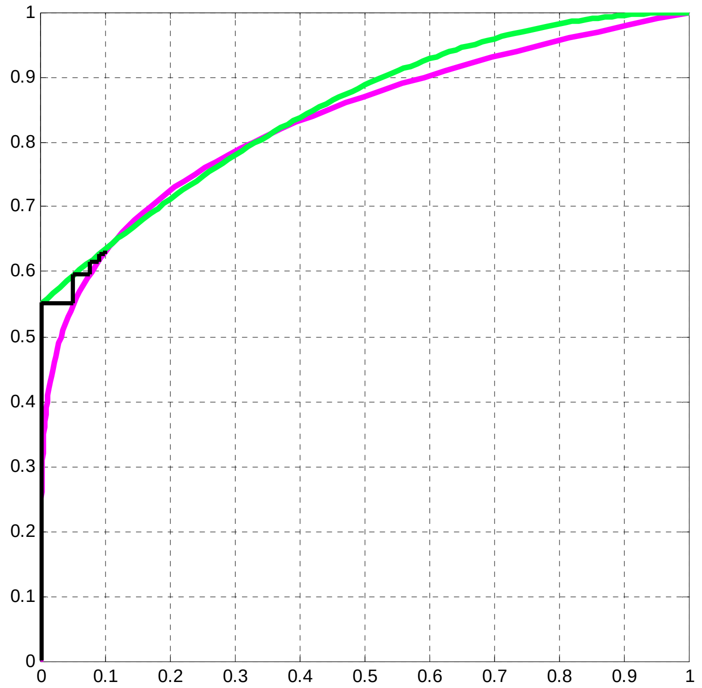
At the output of the decoder we can measure/calculate I_E $\Rightarrow I_E = f(I_A)$

This is only exact if the model of the extrinsic channel is correct!

EXIT Chart of LDPC Code



Intersecting Curves



$I_{A,chk}$

$I_{E,var}$

$I_{A,var}$ $I_{E,chk}$

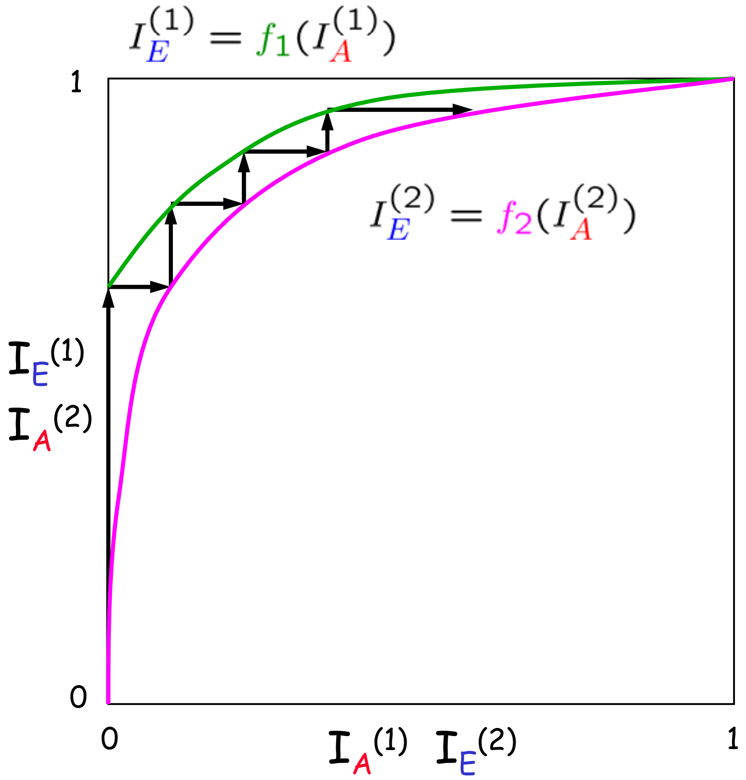
Extrinsic Information Transfer Charts (Stephan ten Brink)

Photo by Jossy Sayir

(Stephan is the guy on the right, not the clown on the left)

Stephan did his PhD at the U of Stuttgart, then worked for Bell Labs in the U.K., then in New Jersey. He is currently with RealTek. He is a regular visitor of ftw. and TU Wien.

Tracking of Messages



This assumes that the decoder depends only on mutual information!

Problem:
How to compute the "transfer functions" f_1 and f_2?

Tracking of Messages

Tracking of messages would mean tracking of pdfs.

Instead of tracking the pdfs we reduce the problem to tracking of mutual information between the messages and the codeword which are scalar quantities.

$$I_A = \frac{1}{N}\sum_{n=1}^{N} I(X_i; A_i)$$

$$I_E = \frac{1}{N}\sum_{n=1}^{N} I(X_i; E_i)$$

$$I_A^{(1)} = 0$$

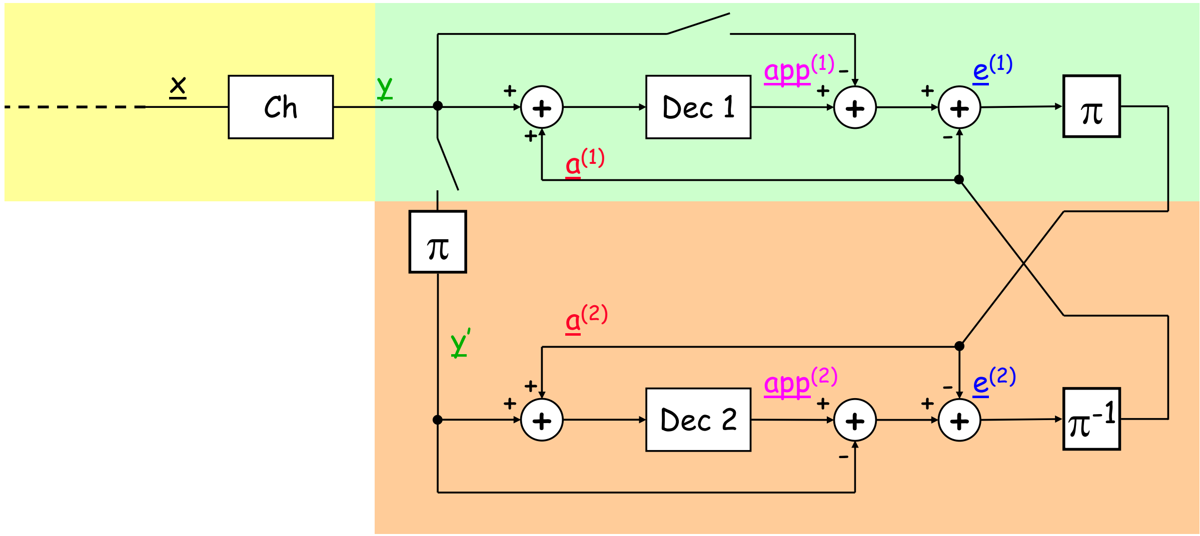
$$\Rightarrow I_E^{(1)} \rightarrow I_A^{(2)}$$

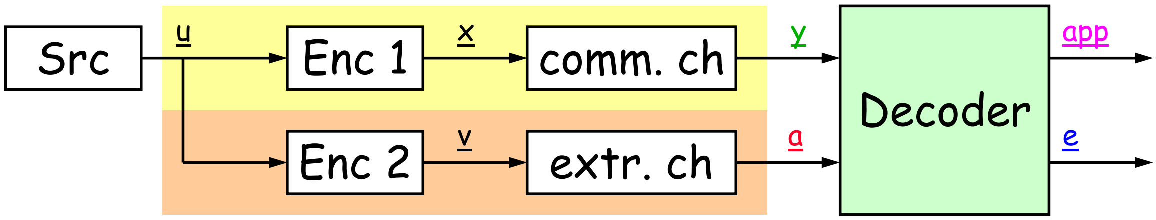
$$\Rightarrow I_E^{(2)} \rightarrow I_A^{(1)}$$

$$\ldots$$

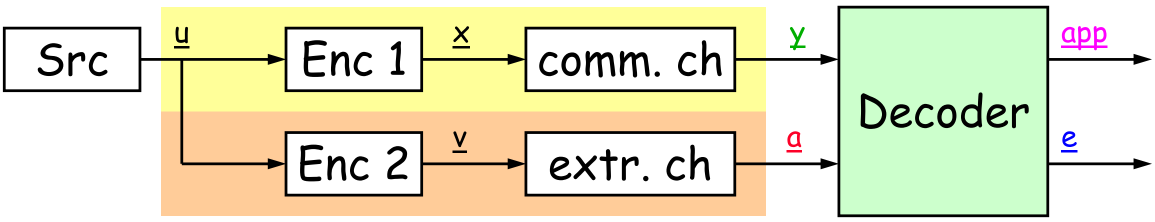
I_A, I_E ..... average symbolwise mutual information

Extrinsic Channel Model





Extrinsic Channel Model



A-priori messages are modeled as independent noisy observations of the encoded source.

Assumptions:

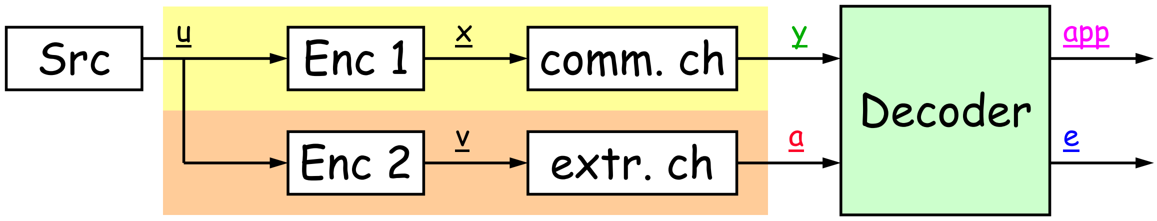
- independent observations
- model for extrinsic channel

$$I_A = I(X_i; A_i) = I(V_i; A_i)$$

$$I_E = I(X_i; E_i) \leq I(X_i; \underline{Y}\,\underline{A}_{\setminus i})$$

with equality if the decoder is optimal

Transfer Functions

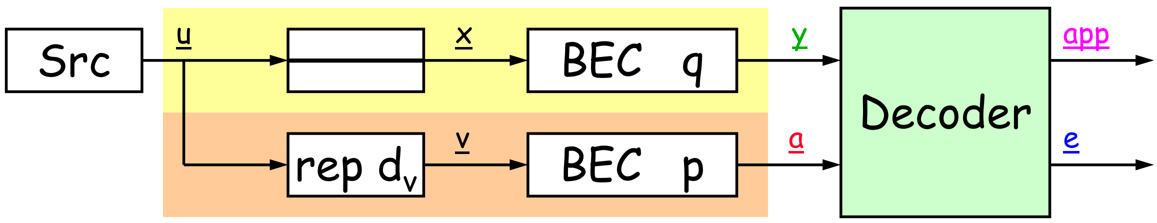


Assuming a model for the extrinsic channel we can vary I_A by varying the channel parameter.

At the output of the decoder we can measure/calculate I_E $\Rightarrow I_E = f(I_A)$

This is only exact if the model of the extrinsic channel is correct!

Variable Nodes and BEC

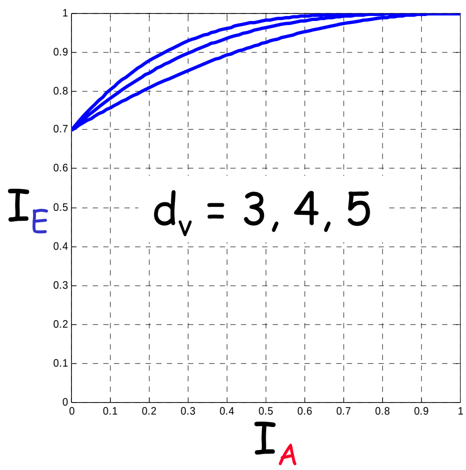


Extrinsic channel is modeled as BEC (exact).

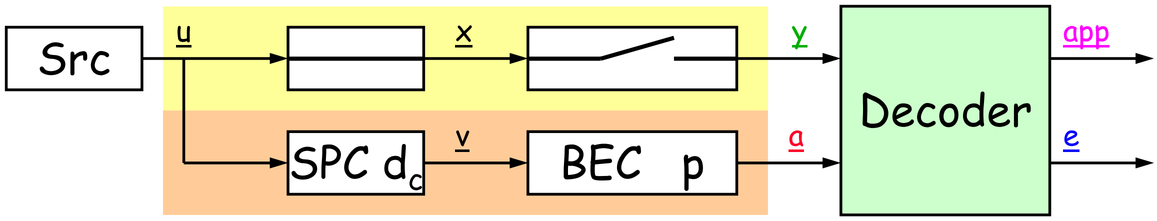
$$I_A = I(X_i; A_i) = I(V_i; A_i) = 1 - p$$

$$I_E = I(X_i; YA_{\backslash i}) = 1 - qp^{d_v - 1}$$

$$I_E = 1 - q(1 - I_A)^{d_v - 1}$$



Check Nodes and BEC

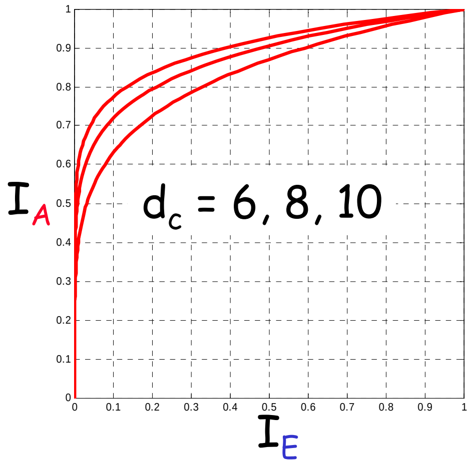


SPC ... single parity check

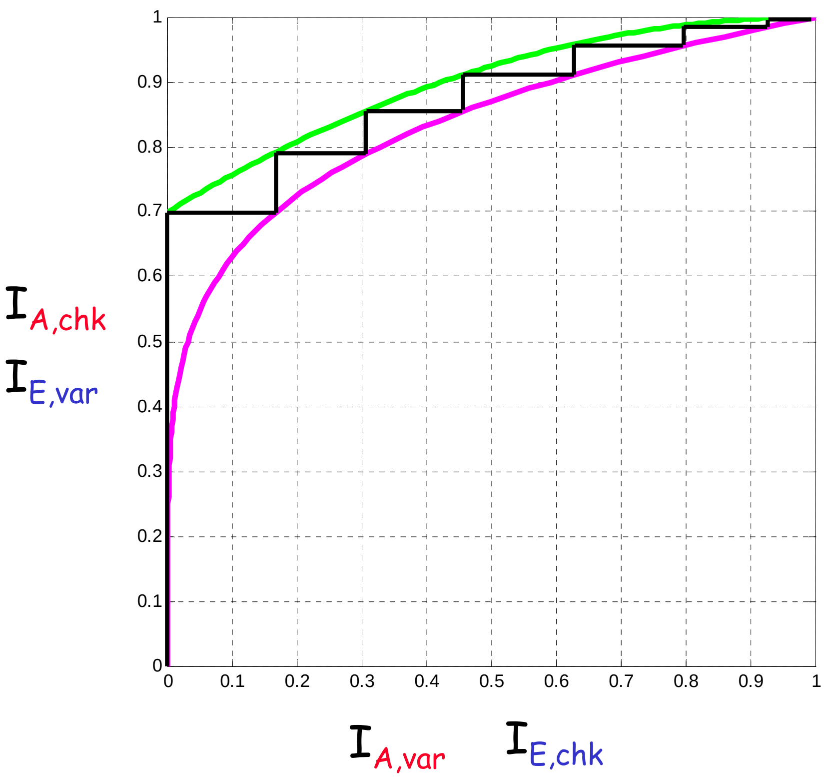
$$I_A = I(X_i; A_i) = I(V_i; A_i) = 1 - p$$

$$I_E = I(X_i; \underline{Y A_{\setminus i}}) = (1-p)^{d_c - 1}$$

$$I_E = (I_A)^{d_c - 1}$$



EXIT Chart of LDPC Code



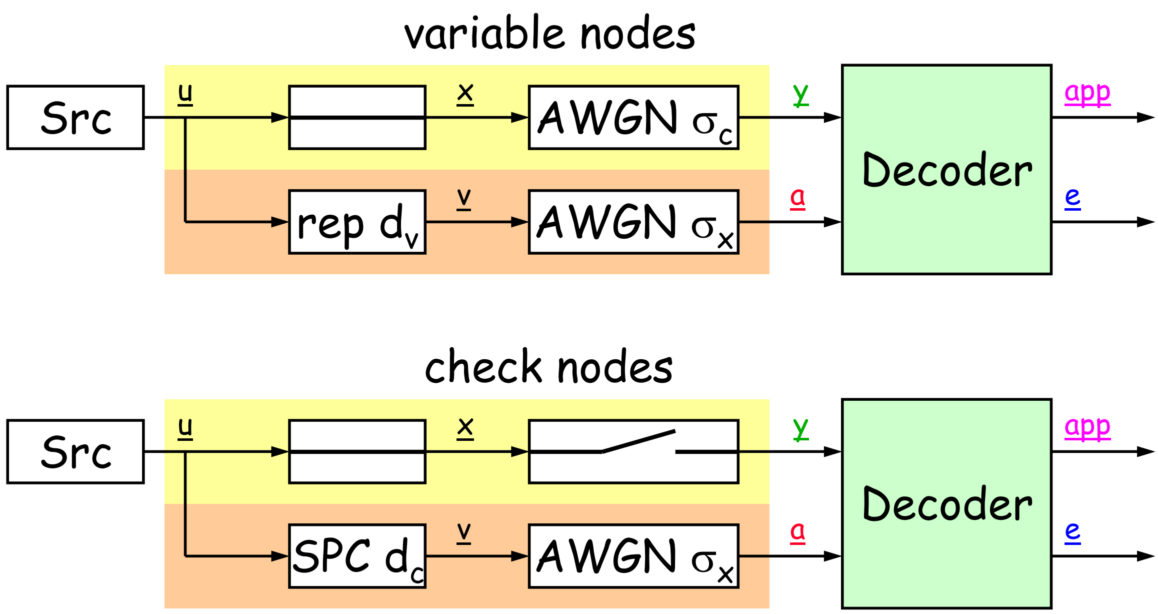
Other Channels

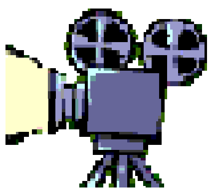
Modeling the extrinsic channel as a BEC is exact if the communication channel is a BEC.

For other communication channels, the assumption of the extrinsic channel is in general an approximation.

If the communication channel is an AWGN channel, we model the extrinsic channel also as an AWGN, but this is only an approximation!

AWGN Channel





Convolutional Codes

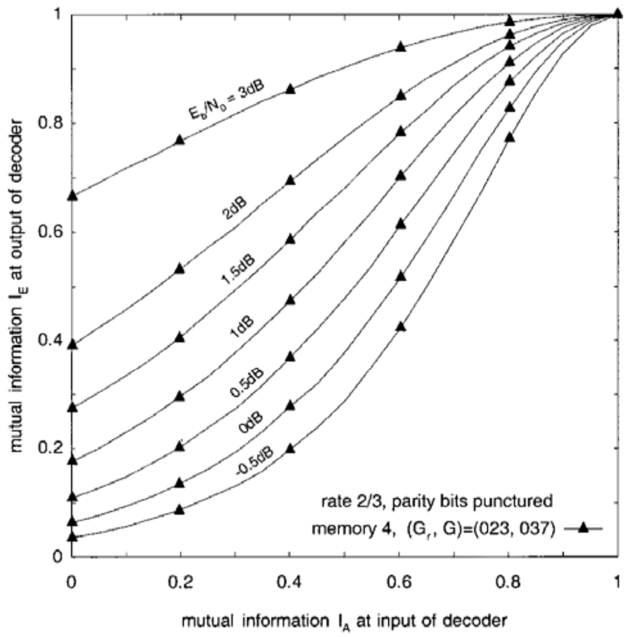
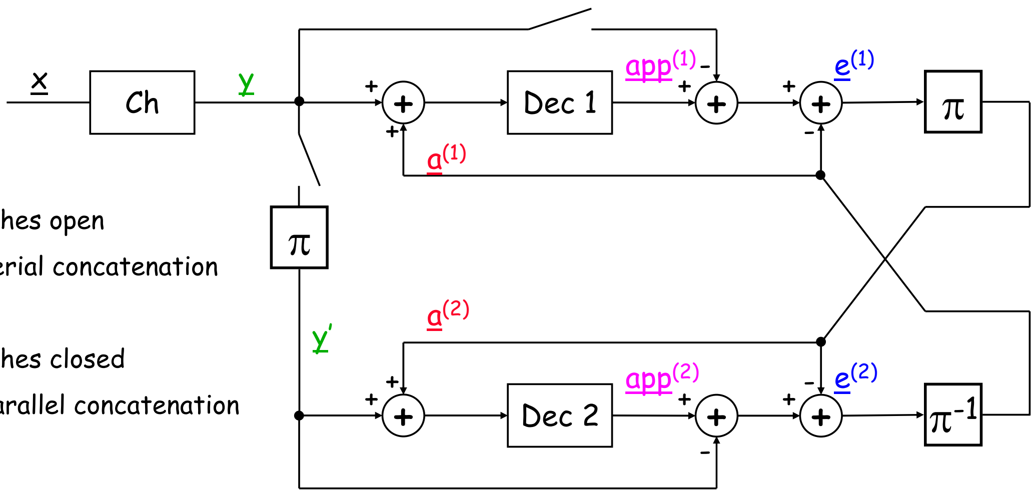


Fig. 2. Extrinsic information transfer characteristics of soft in/soft out decoder for rate 2/3 convolutional code; E_b/N_0 of channel observations serves as parameter to curves.

Stephan ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes", IEEE Trans. Comm. October 2001

Serial / Parallel Concatenation



switches open

→ serial concatenation

switches closed

→ parallel concatenation

Serial concatenation:

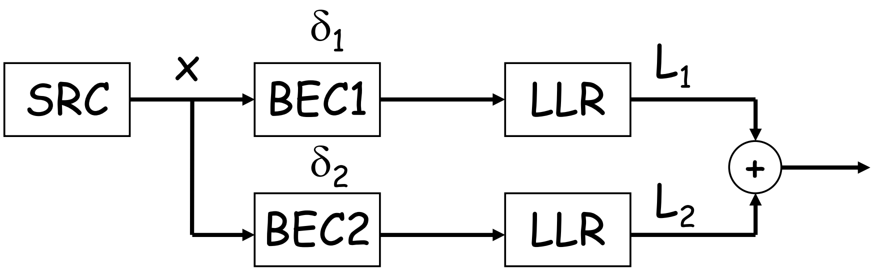
e = app - a

Parallel concatenation:

e = app - a - y

Information Combining BEC

What is the effect on mutual information when we add L-values?



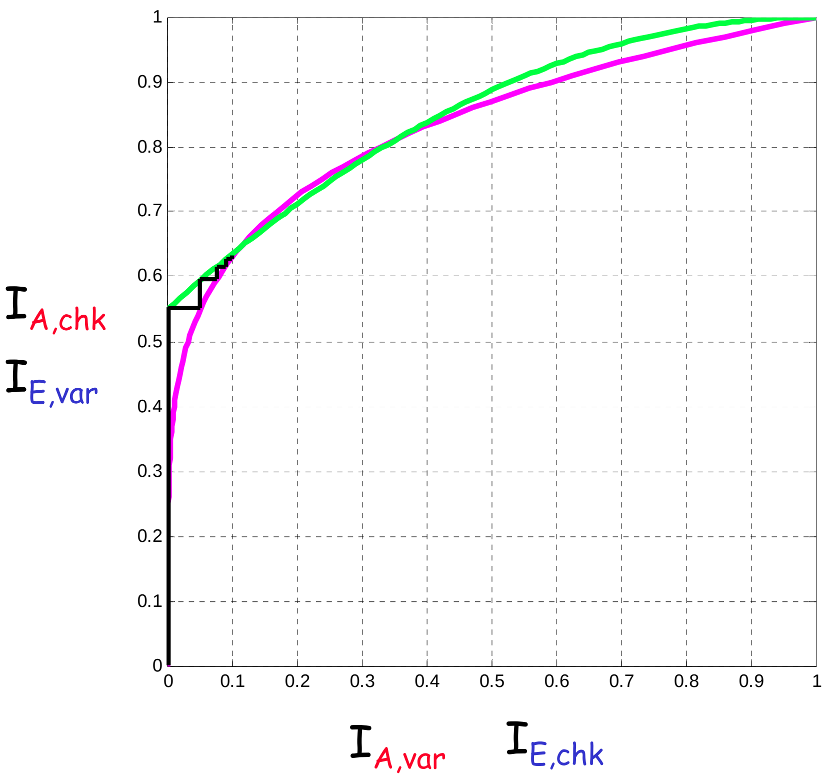
$I_1 = I(X;L_1) = 1 - \delta_1$

$I_2 = I(X;L_2) = 1 - \delta_2$

$I(X;L_1L_2) = 1 - \delta_1\delta_2$

$= 1 - (1-I_1)(1-I_2)$

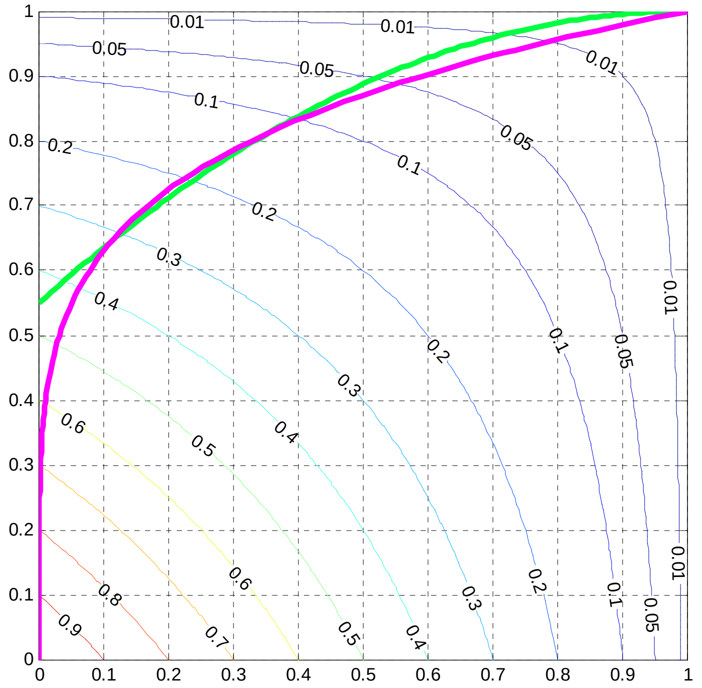
Intersecting Curves



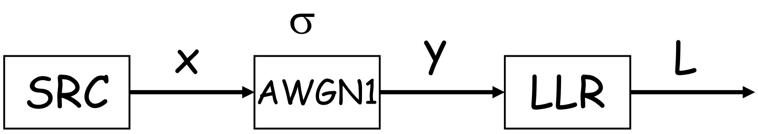
BER from EXIT Chart (BEC)

app = a + e

I(X;APP) = 1 – P_b = 1 – (1-I_A)(1-I_E)



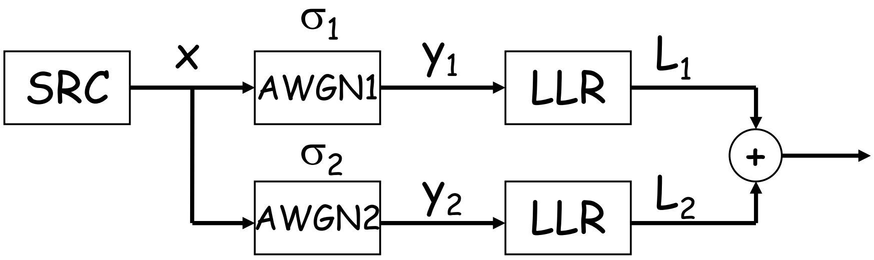
Information Combining AWGN



$$L = \frac{2y}{\sigma^2} \qquad \Rightarrow \qquad L \sim \mathcal{N}\left(\frac{2}{\sigma^2}, \frac{4}{\sigma^2}\right) = \mathcal{N}\left(\frac{\sigma_L^2}{2}, \sigma_L^2\right)$$

$$I(X;L) = \underbrace{1 - \int_{-\infty}^{\infty} \frac{e^{-\frac{(\zeta-\sigma_L/2)^2}{2\sigma_L{}^r}}}{\sqrt{2\pi}\sigma_L} \log_2\left(1 + e^{-\zeta}\right) d\zeta}_{J(\sigma_L)}$$

Information Combining AWGN



$$I_1 = I(X;L_1) = J(\sigma_{L1})$$

$$I_2 = I(X;L_2) = J(\sigma_{L2})$$

$$I(X;L_1L_2) = J\left(\sqrt{\sigma_{L_1}^2 + \sigma_{L_2}^2}\right)$$

$$= J\left(\sqrt{J^{-1}(I_1)^2 + J^{-1}(I_2)^2}\right)$$

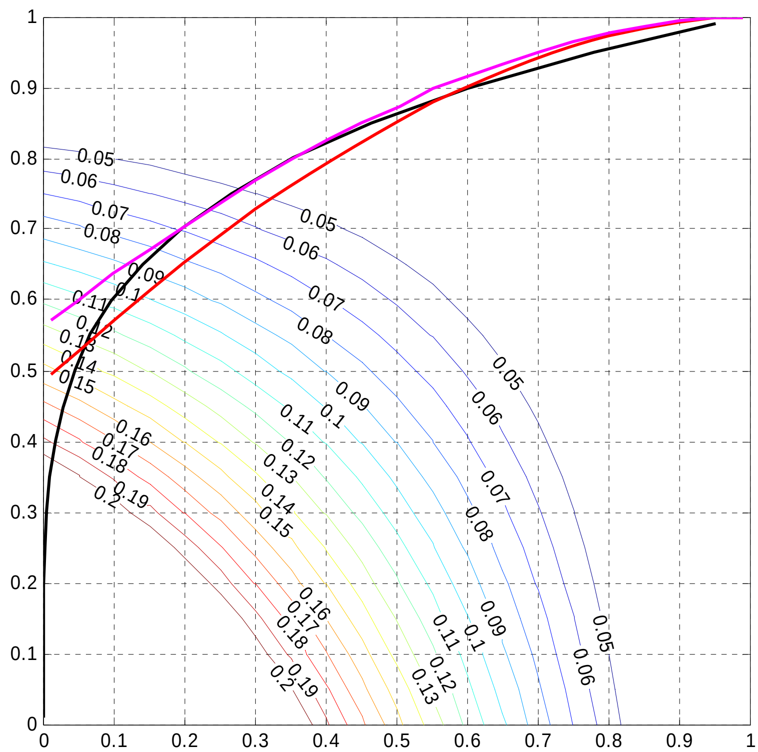
BER from EXIT Chart (AWGN)

E_b/N_0 = 0.0dB

P_b = 0.13

E_b/N_0 = 1.0dB

P_b = 0.07



Independent Observations

Messages received from the extrinsic channel are independent observations, which is only fulfilled if $N \to \infty$

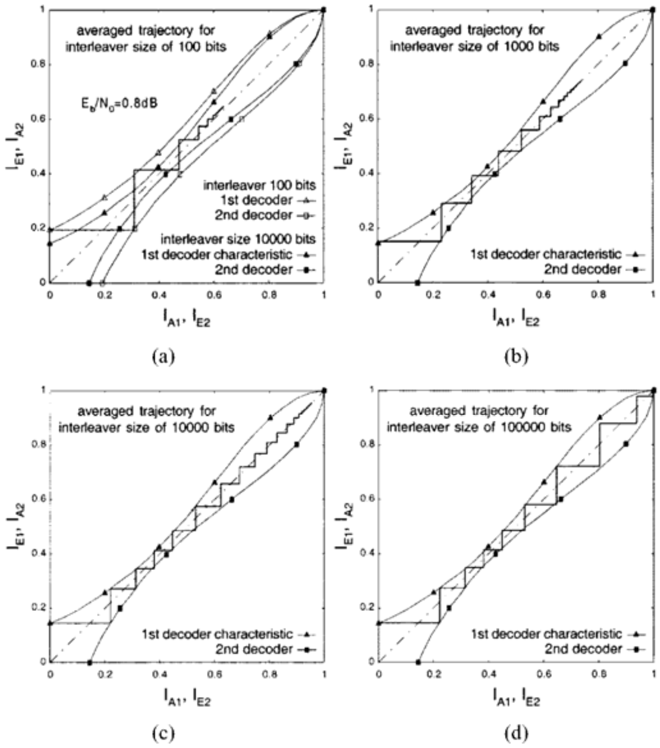
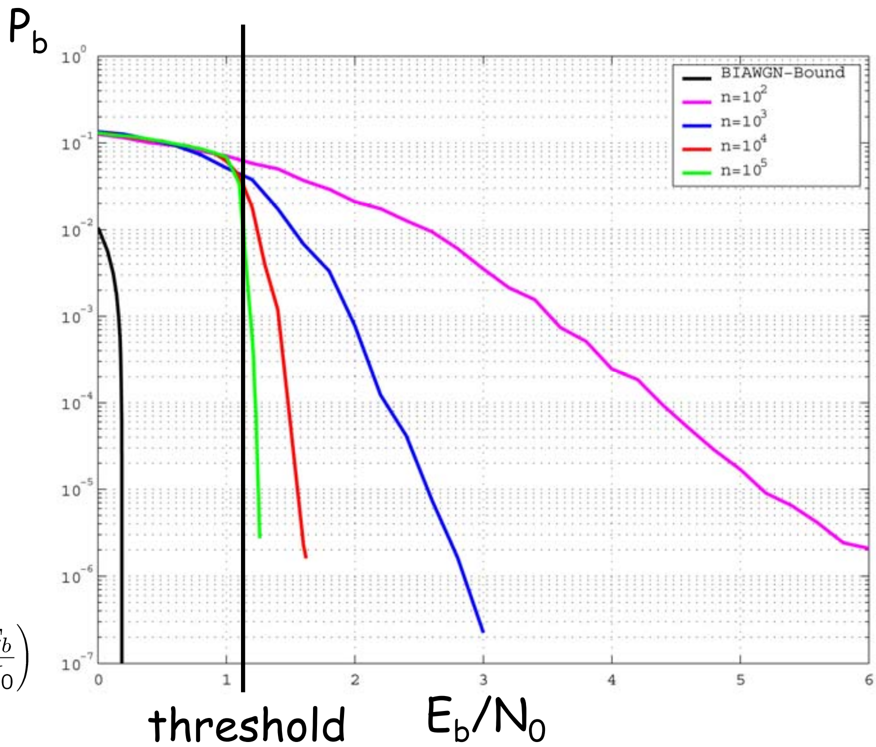


Fig. 8. Averaged decoding trajectories for different interleaver lengths; PCC rate 1/2 memory 4, $(G_r, G) = (023, 037)$; averaged over 10^8 information bits.

Statistics

We use statistical quantities, which are only correct if $N \to \infty$



$$E_s = R \cdot E_b \qquad \sigma^2 = \frac{N_0}{2}$$

$$E_s := 1$$

$$\frac{E_b}{N_0} = \frac{E_s}{R2\sigma^2} \Rightarrow \sigma^2 = \frac{1}{2R} \cdot \left(\frac{E_b}{N_0}\right)$$

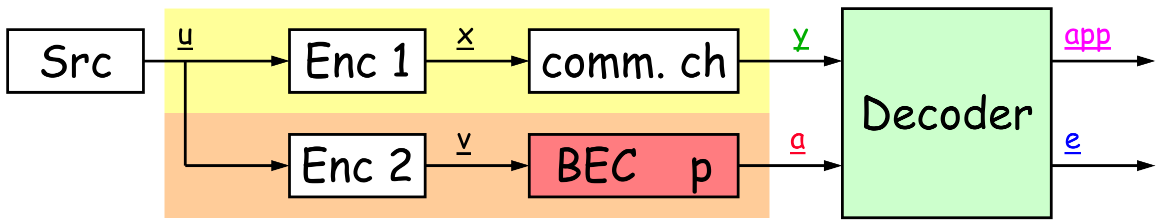
Summary of Assumptions

- Messages received from the extrinsic channel are independent observations, which is only fulfilled if $N \to \infty$

- We use statistical quantities, which are only correct if $N \to \infty$

- We model extrinsic messages with an extrinsic channel. This can only be done exact for the BEC. The Gaussian assumption is an approximation.

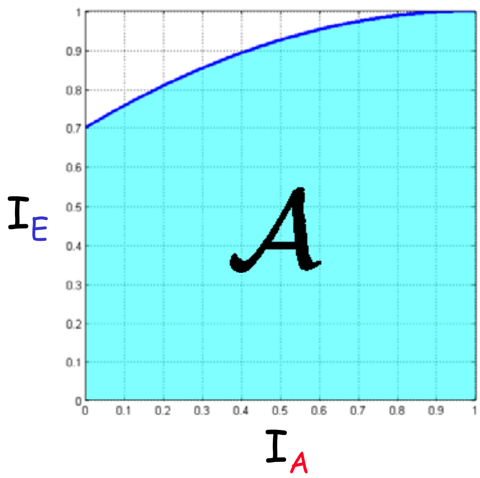
Area Property



$$m = |\underline{v}|$$

$$I_A = \frac{1}{m}\sum_{i=1}^{m} I(V_i; A_i) = 1 - p$$

$$\mathcal{A} = \int_0^1 I_E(I_A)dI_A = \int_0^1 I_E(p)dp$$



Derivation of Area Property 1

$$I_E = \frac{1}{m}\sum_{i=1}^{m} \underbrace{I(V_i; \underline{Y}\underline{A}_{\setminus i})}$$

$$H(V_i) - \underbrace{H(V_i|\underline{Y}\underline{A}_{\setminus i})}$$

1 → $H(V_i)$

$$\underbrace{\sum_{\underline{a}_{\setminus i}} P(\underline{a}_{\setminus i}) \cdot \underbrace{H(V_i|\underline{Y}, \underline{A}_{\setminus i} = \underline{a}_{\setminus i})}_{H(V_i|\underline{Y}, \underline{V}_{\mathcal{S}} = \underline{v}_{\mathcal{S}})}} \qquad \mathcal{S} = \{b = 1\dots m | b \neq i, A_b \neq \Delta\}$$

$$0 \leq j = |\mathcal{S}| \leq m-1$$

$$\sum_{j=0}^{m-1} (1-p)^j \cdot p^{m-1-j} \sum_{|\mathcal{S}|=j, i\notin\mathcal{S}} H(V_i|\underline{Y}\underline{V}_{\mathcal{S}})$$

$$I_E = \frac{1}{m}\sum_{i=1}^{m}\left[1 - \sum_{j=0}^{m-1} (1-p)^j \cdot p^{m-1-j} \sum_{|\mathcal{S}|=j} H(V_i|\underline{Y}\underline{V}_{\mathcal{S}})\right]$$

$$I_E = 1 - \frac{1}{m}\sum_{i=1}^{m}\sum_{j=0}^{m-1} (1-p)^j \cdot p^{m-1-j} \sum_{|\mathcal{S}|=j} H(V_i|\underline{Y}\underline{V}_{\mathcal{S}})$$

Derivation of Area Property 2

$$I_E = 1 - \frac{1}{m}\sum_{i=1}^{m}\sum_{j=0}^{m-1} (1-p)^j \cdot p^{m-1-j} \sum_{|\mathcal{S}|=j} H(V_i|\underline{Y}\underline{V}_{\mathcal{S}})$$

$$\mathcal{A} = \int_0^1 1 - \frac{1}{m}\sum_{i=1}^{m}\sum_{j=0}^{m-1} (1-p)^j \cdot p^{m-1-j} \sum_{|\mathcal{S}|=j} H(V_i|\underline{Y}\underline{V}_{\mathcal{S}})dp$$

$$\mathcal{A} = 1 - \frac{1}{m}\sum_{i=1}^{m}\sum_{j=0}^{m-1} \int_0^1 (1-p)^j \cdot p^{m-1-j} dp \sum_{|\mathcal{S}|=j} H(V_i|\underline{Y}\underline{V}_{\mathcal{S}})$$

$$\mathcal{A} = 1 - \frac{1}{m}\underbrace{\sum_{i=1}^{m}\sum_{j=0}^{m-1} \frac{1}{m\binom{m-1}{j}} \sum_{|\mathcal{S}|=j} H(V_i|\underline{Y}\underline{V}_{\mathcal{S}})}_{H(\underline{V}|\underline{Y})}$$

$$\boxed{\mathcal{A} = 1 - \frac{1}{m}H(\underline{V}|\underline{Y}) = 1 - \frac{1}{m}H(\underline{X}|\underline{Y})}$$

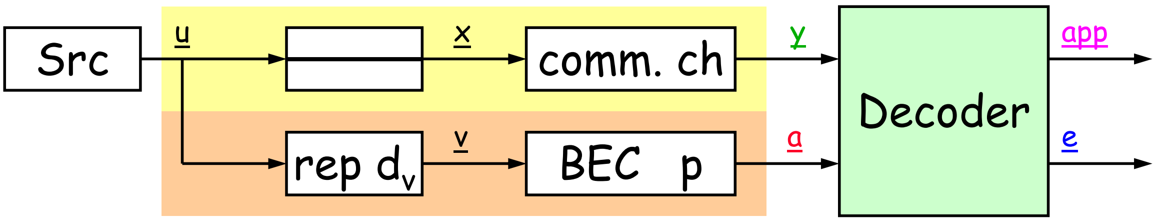
Derivation of H(V|Y)

$$\begin{aligned} H(\underline{V}) &= H(V_1) + H(V_2|V_1) + H(V_3|V_1V_2) \\ &= H(V_1) + H(V_3|V_1) + H(V_2|V_1V_3) \\ &= H(V_2) + H(V_1|V_2) + H(V_3|V_1V_2) \\ &= H(V_2) + H(V_3|V_2) + H(V_1|V_2V_3) \\ &= H(V_3) + H(V_1|V_3) + H(V_2|V_1V_3) \\ &= H(V_3) + H(V_2|V_3) + H(V_1|V_2V_3) \end{aligned}$$

$$\begin{aligned} 3! \cdot H(\underline{V}) \;&=\; 2 \cdot (H(V_1) + H(V_2) + H(V_3)) + \\ &+\; H(V_1|V_2) + H(V_1|V_3) + H(V_2|V_1) + H(V_2|V_3) + H(V_3|V_1) + H(V_3|V_2) \\ &+\; 2 \cdot (H(V_1|V_2V_3) + H(V_2|V_1V_3) + H(V_3|V_1V_2)) \end{aligned}$$

$$\sum_{i=1}^{m} \sum_{j=0}^{m-1} \frac{1}{m\binom{m-1}{j}} \sum_{|\mathcal{S}|=j} H(V_i|\underline{Y}\underline{V}_{\mathcal{S}})$$

Variable Nodes

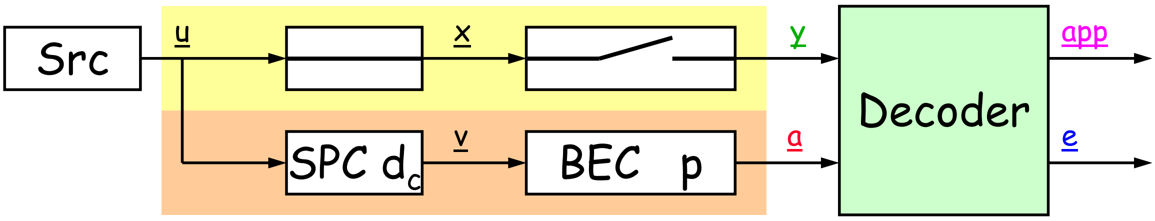


$$|\underline{u}| = k \qquad m = |\underline{v}| = k \cdot d_v$$
$$|\underline{x}| = k$$

$$\mathcal{A}_v = 1 - \frac{1}{m} H(\underline{X}|\underline{Y}) = 1 - \frac{H(\underline{X}) - I(\underline{X};\underline{Y})}{k \cdot d_v}$$

$$= 1 - \frac{k - k \cdot I(X_1;Y_1)}{k \cdot d_v} = 1 - \frac{1-C}{d_v}$$

Check Nodes



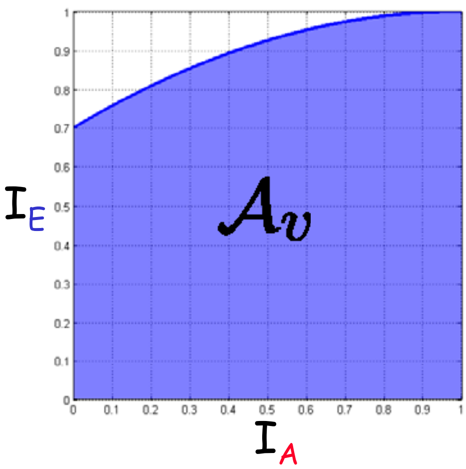
$$|\underline{u}| = k \qquad m = |\underline{v}| = k \cdot \frac{d_c}{d_c - 1}$$
$$|\underline{x}| = k$$

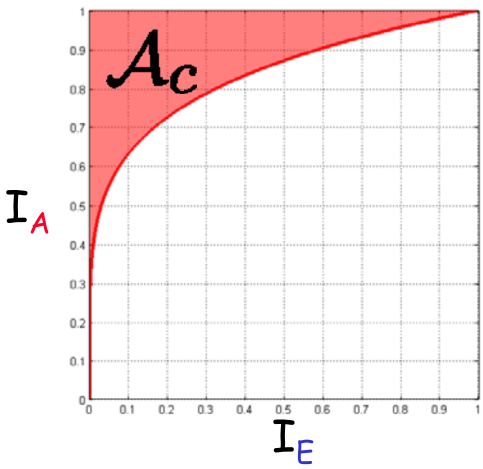
$$\mathcal{A}_c = 1 - \frac{1}{m} H(\underline{V}|\underline{Y}) = 1 - \frac{H(\underline{V}) - I(\underline{V};\underline{Y})}{m}$$

$$= 1 - \frac{k \cdot (d_c - 1)}{k \cdot d_c} = \frac{1}{d_c}$$

Area of LDPC Component Codes

$$\mathcal{A}_v = 1 - \frac{1 - C}{d_v} \qquad \mathcal{A}_c = \frac{1}{d_c}$$





Necessary condition for successful decoding:

$$1 - \mathcal{A}_v < \mathcal{A}_c$$

Consequences of Area Property

$$1 - \mathcal{A}_v < \mathcal{A}_c$$

$$1 - 1 + \frac{1-C}{d_v} < \frac{1}{d_c}$$

$$1 - C < \frac{d_v}{d_c}$$

$$C > 1 - \frac{d_v}{d_c} = R$$

"Surprising" result:

The area property tells us that the decoder can only converge if the rate is smaller than capacity!

More Consequences...

Suppose the condition for convergence is fulfilled

$$1 - \mathcal{A}_v = \gamma \cdot \mathcal{A}_c < \mathcal{A}_c \qquad 0 \le \gamma < 1$$

$$\mathcal{A}_c = \frac{1}{d_c}$$

$$1 - \mathcal{A}_v = \gamma \cdot \mathcal{A}_c = \frac{1-C}{d_v}$$

$$R = 1 - \frac{d_v}{d_c} = 1 - \frac{1-C}{\gamma} = \frac{C-(1-\gamma)}{\gamma} < C$$

What is the result of this inequality?

Area and Rate Loss

$$R = 1 - \frac{d_v}{d_c} = 1 - \frac{1-C}{\gamma} = \frac{C-(1-\gamma)}{\gamma} < C$$

If $\gamma \to 1$ we can transmit at rates that approach capacity.
If $\gamma < 1$ we are bounded from capacity.

$\gamma \to 1$ means that $1 - A_v = A_c$

Furthermore, the curves must not intersect.

$\Rightarrow$ The curves have to be matched.

Code design reduces to curve fitting!

Curve Fitting – Code Mixture

We only considered regular codes, where every symbol has the same properties. Therefore, averaging over all symbols is equivalent to the mutual information of an arbitrarily symbol.

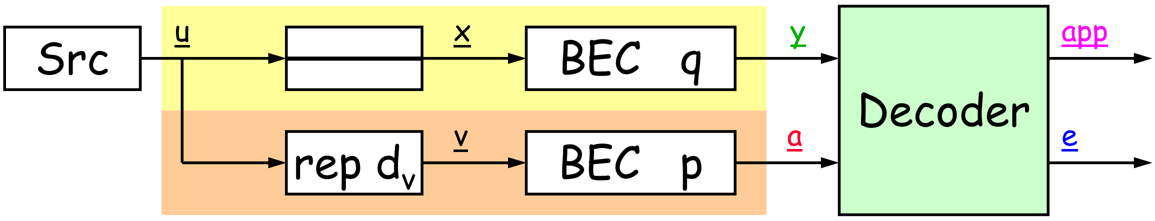
$$I_E = \frac{1}{m}\sum_{i=1}^{m} I(V_i; E_i) = I(V_1; E_1)$$

If we partition m into n_u groups $j=1 \ldots n_u$ each with length l_j, we can write I_E as

$$I_E = \sum_{j=1}^{n_u} \frac{l_j}{m}\left[\frac{1}{l_j}\sum_{i=1}^{l_j} I(V_{ji}; E_{ji})\right] = \sum_{j=1}^{n_u} \gamma_j I_{E_j} \qquad \gamma_j = \frac{l_j}{m} = \frac{l_j}{\sum_{j=1}^{n_u} l_j}$$

The resulting EXIT function is the weighted average of the EXIT functions of the groups.

Example – Variable Mixture



70% of the variable nodes have d_v=2
30% of the variable nodes have d_v=5

$$\gamma_1 = \frac{0.7 \cdot k \cdot 2}{0.7 \cdot k \cdot 2 + 0.3 \cdot k \cdot 5} = 0.48 \qquad \gamma_2 = \frac{0.3 \cdot k \cdot 5}{0.7 \cdot k \cdot 2 + 0.3 \cdot k \cdot 5} = 0.52$$

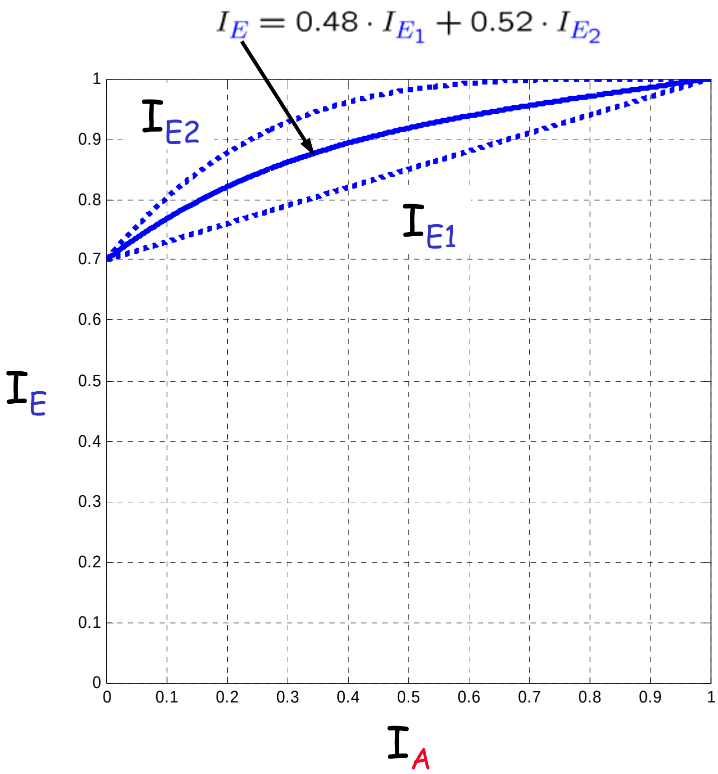
$$I_{E_j} = 1 - qp^{d_{vj}-1} \qquad I_E = \gamma_1 \cdot \left[1 - qp^{d_{v1}-1}\right] + \gamma_2 \cdot \left[1 - qp^{d_{v2}-1}\right]$$

$$I_E(p) = 1 - q \cdot \sum_{j=1}^{n_u} \gamma_j \cdot p^{d_{vj}-1}$$

This is a polynomial in p

Note that $\sum \gamma_j = 1$

Example – Variable Mixture



Curve Fitting

Lets fix the EXIT function of the check node decoder.

$$I_{Ec} = (I_{Ac})^{d_c - 1}$$

For curve fitting, we can exchange the following quantities

$$I_{Ec} = I_{Av} \qquad I_{Ev} = I_{Ac}$$

Therefore, we can write the EXIT function of the variable node decoder as the inverse EXIT function of the check node decoder.

$$I_{Av} = (I_{Ev})^{d_c - 1}$$

$$I_{Ev} = (I_{Av})^{\frac{1}{d_c - 1}} = (1 - p)^{\frac{1}{d_c - 1}}$$

Taylor Series Expansion

$$I_{Ev} = (I_{Av})^{\frac{1}{d_c - 1}} = (1 - p)^{\frac{1}{d_c - 1}}$$

Assuming for example d_c=5 we can expand I_{Ev} as a Taylor series

$$I_{Ev} = 1 - \left[\frac{1}{4}p + \frac{3}{32}p^2 + \frac{7}{128}p^3 + \cdots\right]$$

Truncating the Taylor series and normalizing the coefficients to 1 results in

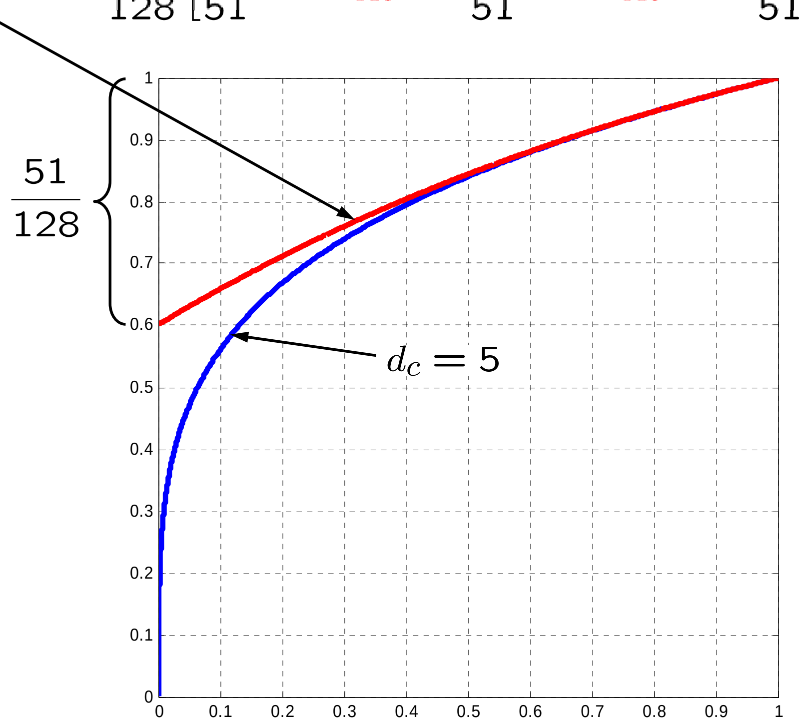
$$I_{Ev} = 1 - \frac{51}{128}\left[\frac{32}{51}p + \frac{12}{51}p^2 + \frac{7}{51}p^3\right]$$

Compare this with the transfer function of the mixture of variable nodes...

$$I_E(p) = 1 - q \cdot \sum_{j=1}^{n_u} \gamma_j \cdot p^{d_{vj} - 1}$$

Curve Fitting

$$I_{Ev} = 1 - \frac{51}{128}\left[\frac{32}{51}(1 - I_{Av}) + \frac{12}{51}(1 - I_{Av})^2 + \frac{7}{51}(1 - I_{Av})^3\right]$$



Even more Consequences...

Using the same model as for the variable and check node decoder, it can be shown that the areas for a serial concatenated code with an outer code $R_{out}=k_{out}/n_{out}$ and an inner code $R_{in}=k_{in}/n_{in}$ are given by

$$\mathcal{A}_{out} = 1 - R_{out} \qquad \mathcal{A}_{in} = \frac{I(\underline{X};\underline{Y})}{n_{in} \cdot R_{in}}$$

The same necessary condition $1-\mathcal{A}_{out} < \mathcal{A}_{in}$ leads to

$$R_{out} \cdot R_{in} < \frac{I(\underline{X};\underline{Y})}{n_{in}} \leq C$$

If the inner code has rate < 1, i.e. $I(\underline{X};\underline{Y})/n_{in} < C$ then we can not achieve capacity with serial concatenated codes!