The Turbo Principle in Wireless Communications

Joachim Hagenauer

Institute for Communications Engineering (LNT)
Munich University of Technology (TUM)

D-80290 München, Germany

Nordic Radio Symposium, Oulu, Finland, August 17th, 2004

LNT

Outline

- The turbo principle and its applications
- Log-Likelihood Ratios (LLR) and the APP Decoders
- Extrinsic Information Transfer (EXIT) charts
- Turbo applications: Coded equalization of multipath channels
- Turbo applications: Precoded QAM with irregular channel codes
- Turbo applications: Coded MIMO systems
- Conclusions

Introduction

History:

- 1948: Shannon's absolute limits in communications, e.g. 0.2 dB in E_b/N_0 for binary codes with rate 1/2 on AWGN channel
- 1962: Gallager's low density parity check codes with iterative decoding
- 1966: Forney: Concatenated codes
- before 1993: Concatenated codes (Viterbi plus RS codes) approach Shannon's limit by by 2.5 dB and with iterations by 1.5 dB.
- 1993: Berrou, Glavieux and Thitimajshima: Turbo decoding approaches Shannon's limit by 0.5 dB.
- 1995: Douillard, Glavieux, Berrou et al: Turbo equalization
- 1997: Turbo principle recognized as general method in communications systems
- 2001: Chung, Forney, Richardson, Urbanke :Iterative decoding of Irregular LDPC Codes within 0.0045 dB of Shannon limit

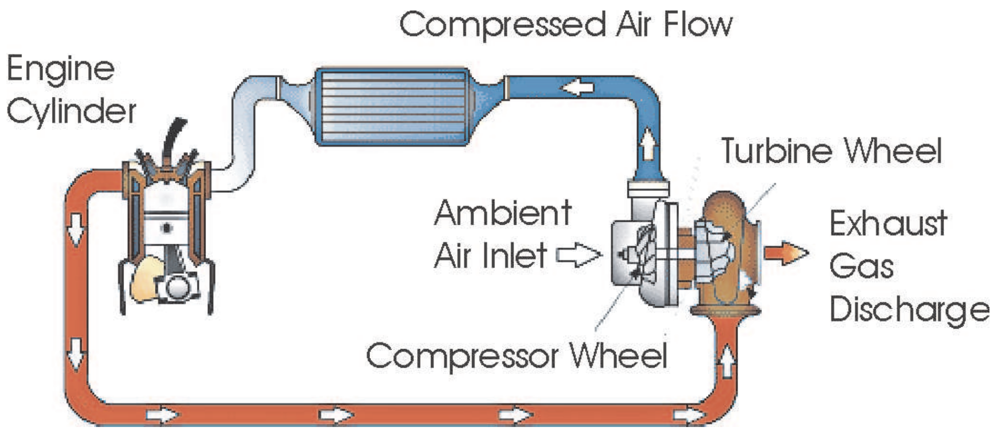
Introduction

The Turbo Principle comprises...

- ... a communication system with serial and/or parallel concatenations of components
- ... a posteriori probability (APP) symbol-by-symbol decoders/detectors
- ... soft-in/soft-out decoders/detectors
- ... interleavers between the components
- ... exchange of extrinsic information between components in the form of probabilities or log-likelihood ratios

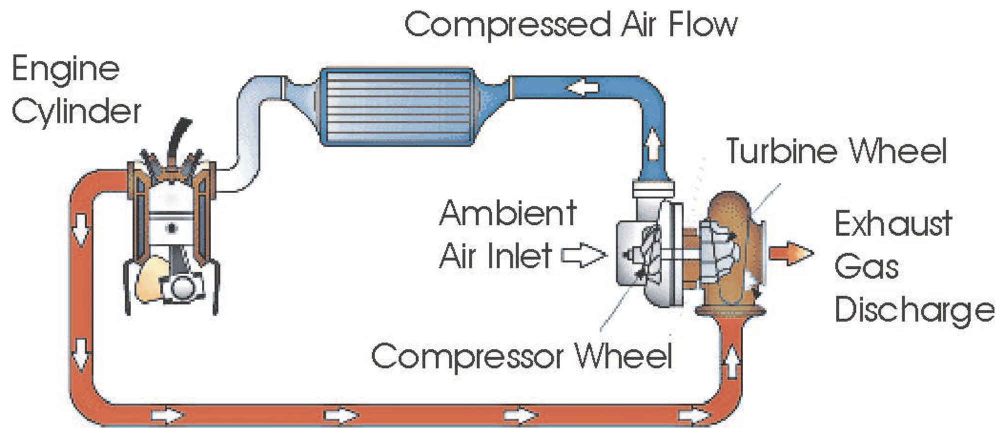
The Turbo Principle ...

... in mechanics

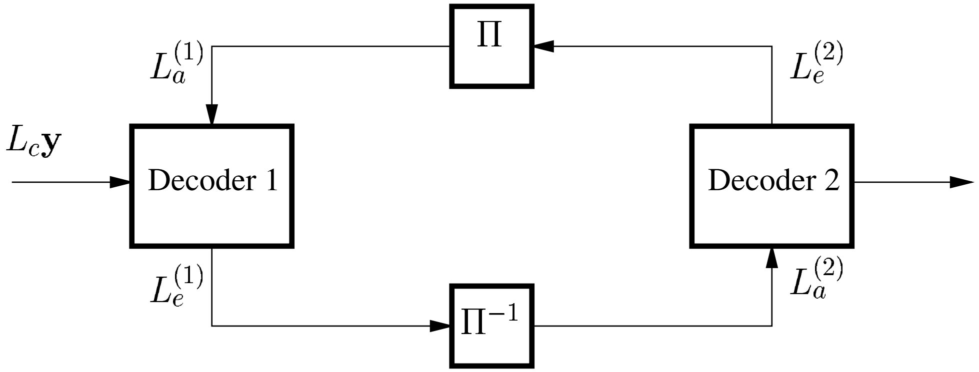


The Turbo Principle ...

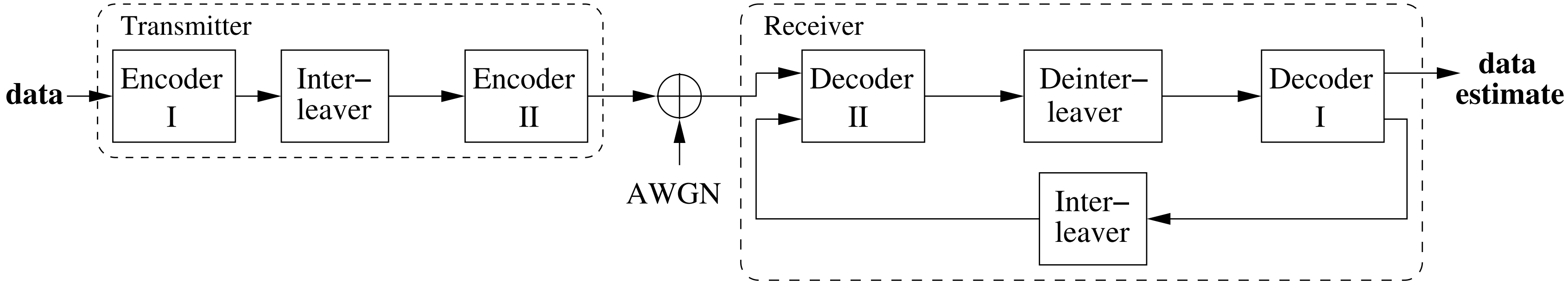
... in mechanics



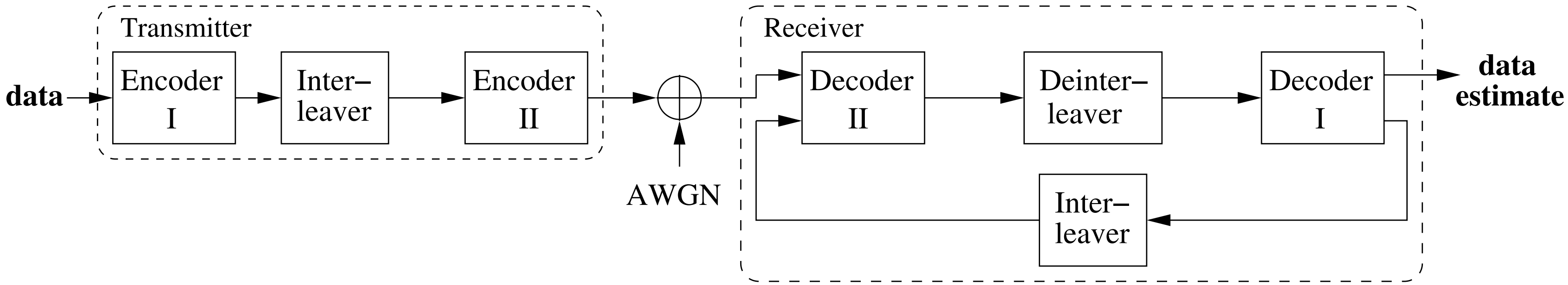
... in communications



Serial Concatenation



Examples for serial concatenation in communications systems



configuration	en-/decoder I (outer code)	en-/decoder II (inner code)
serial code concat.	FEC en-/decoder	FEC en-/decoder
turbo equalization	FEC en-/decoder	Multipath channel/detector
turbo BiCM	FEC en-/decoder	Mapper/demapper
turbo MIMO	FEC en-/decoder	Mapper & MIMO detector
turbo multiuser	FEC en-/decoder	SISO multiuser detector
turbo source-channel	source encoder	FEC en-/decoder
LDPC code/decoder	check nodes	variable nodes

Basics of Turbo Decoding

Log-Likelihood Ratios and the APP Decoders:

Let u be in GF(2) with the elements $\{+1, -1\}$, where $+1$ is the 'null' element under the $\oplus$ addition.
The log-likelihood ratio (LLR) or L-value of the binary variable is

$$L(u) = \ln \frac{P(u = +1)}{P(u = -1)} \tag{1}$$

with the inverse

$$P(u = \pm 1) = \frac{e^{\pm L(u)/2}}{e^{+L(u)/2} + e^{-L(u)/2}}. \tag{2}$$

Note: The sign of $L(u)$ is the hard decision and the magnitude $|L(u)|$ is the reliability of this decision.

The soft bit and the binary sum

The soft bit $\lambda(u)$ is

$$\begin{aligned}\lambda(u) = \mathrm{E}\{u\} &= (+1)\cdot P(u=+1) + (-1)\cdot P(u=-1)\\ &= \tanh(L(u)/2).\end{aligned}$$

GF(2) addition $u_1 \oplus u_2$ of two independent binary random variables:

$$\mathrm{E}\{\mathrm{u}_1 \cdot \mathrm{u}_2\} = \mathrm{E}\{\mathrm{u}_1\}\mathrm{E}\{\mathrm{u}_2\} = \lambda(u_1)\cdot\lambda(u_2).$$

L value of the sum :

$$L(u_1 \oplus u_2) = 2\tanh^{-1}(\tanh(L(u_1)/2)\cdot\tanh(L(u_2)/2)) = L(u_1) \boxplus L(u_2).$$

with the **boxplus** $\boxplus$ abbreviation.

Transmission and combining after fading/AWGN channels

The a posteriori probability (APP) in $y = ax + n$ is

$$P(x|y) = \frac{p(y|x)P(x)}{p(y)} \quad (3)$$

with the pdf

$$p(y|x) = \frac{1}{\sqrt{2\pi\sigma_c^2}} e^{\frac{-(y-ax)^2}{2\sigma_c^2}} \quad (4)$$

The complementary APP LLR equals

$$L_{CH} = L(x|y) = \ln \frac{P(x=+1|y)}{P(x=-1|y)} = L_c \cdot y + L(x). \quad (5)$$

$L(x)$ is the a priori LLR of x and L_c is the channel state information (CSI):

$$L_c = \frac{2a}{\sigma_c^2} = 4aE_s/N_0 \quad (6)$$

For statistically independent transmission

$$L(x|y_1, y_2) = L_{c_1}y_1 + L_{c_2}y_2 + L(x). \quad (7)$$

Practical Usefulness of Log-Likelihood Calculation

Did it rain in Helsinki at 8:00 am today?
A Yes (rain!) is binary coded as $+1$,transmitted over an unreliable link.
Two rain detection devices measured :

$$x_1 = +1$$

$$x_2 = +1$$

Additional a priori value is available: From Farmer's Almanac:
Probability of rain in Helsinki today is 75%

$$L(x) = \ln(0.75/0.25) = +1.1$$

	transmitted value x	received value y	channel state L_c	$L_c y$
link 1	+1.0	-1.5	2.0	-3.0
link 2	+1.0	+0.9	3.0	+2.7
a priori				+1.1

For statistically independent information

$L(x|y_1, y_2) = L_{c_1}y_1 + L_{c_2}y_2 + L(x) = +0.8$ with 31% error : rain in Helsinki !!.

The extrinsic information as a LLR

Assume a parity check equation of statistically independently transmitted bits x_j

$$\sum_{j=1}^{j=N}\!\oplus\, x_j = 0.(+1)$$

Then the extrinsic bit x_i equals

$$x_i = \sum_{j=1, j\neq i}^{j=N}\!\oplus\, x_j$$

and consequently the extrinsic LLR for this bit given the APP LLR's $L(x_j|y_j)$ of all the other bits equals

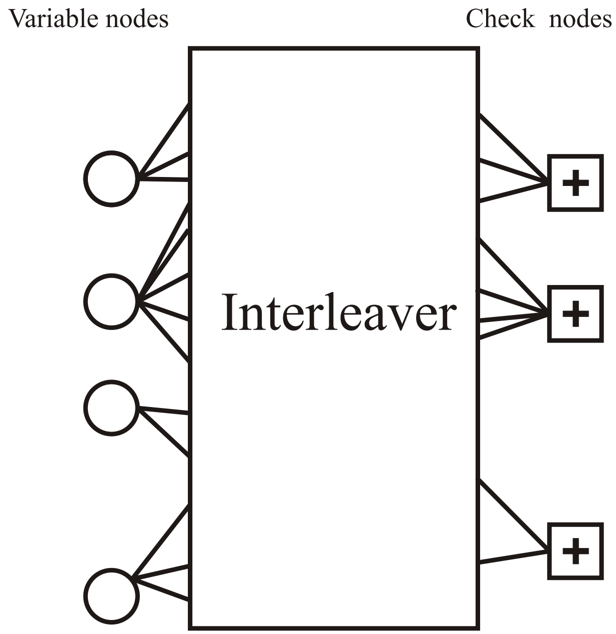
$$L_E(x_i) = \sum_{j=1, j\neq i}^{j=N}\!\boxplus\, L(x_j|y_j)$$

<u>Example</u>:
SPC code, $N = 3$, with $L(x_2|y_2) = -0.3,\ L(x_3|y_3) = -5.5.$
Then the extrinsic LLR for the first bit is

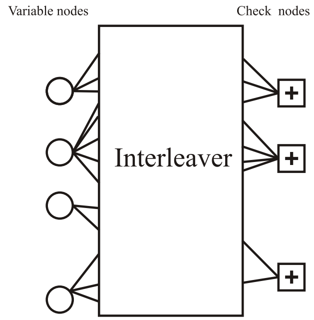
$$L_E(x_1) = -0.3 \boxplus -5.5 \approx +0.3.$$

Low Density Parity Check (LDPC) codes and their Turbo decoder

A low density parity check code of rate k/n can be described as a serial concatenation of n variable nodes as inner repetition codes with $n-k$ check nodes as outer single parity check nodes.



Irregular LDPC codes and their Turbo decoder (cnt')



i-th variable node (i-th code bit) with $d_{v,i}$ connections. $n-k$ check nodes where the i-th checks $d_{c,i}$ bits.
More than one extrinsic message

$$L_{i,j}^{(\mathrm{out})} = L_{c,i} \cdot y_i + \sum_{j=1, j\neq i}^{d_{v,i}} L_{i,j}^{(\mathrm{in})} \tag{8}$$

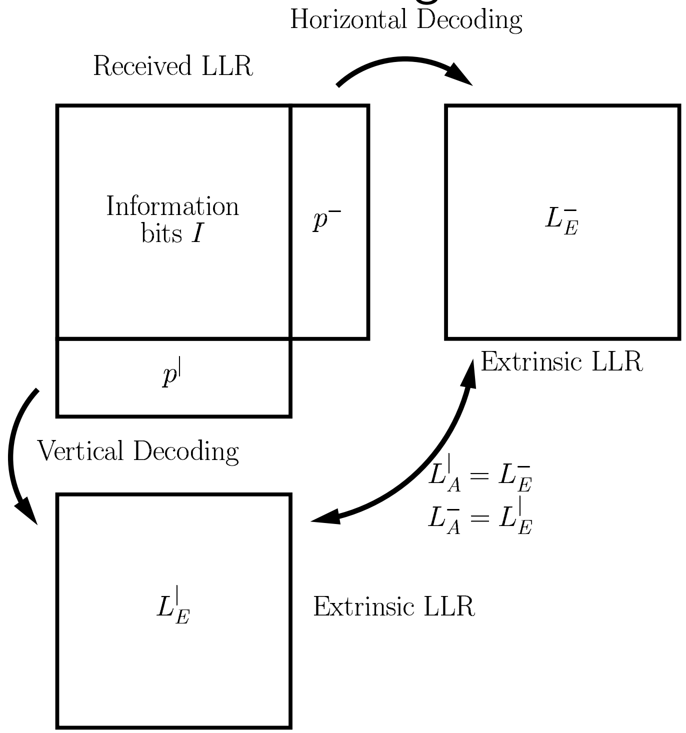
per code bit $x_i, i = 1 \ldots n$ is sent to the outer single parity check (SPC) decoders. The SCPC decoders return for $i = 1 \ldots n-k$

$$L_{i,j}^{(\mathrm{c,out})} = \sum_{j=1, j\neq i}^{d_{c,i}} \boxplus \quad L_{i,j}^{(\mathrm{c,in})} \tag{9}$$

The decoding result is the overall L value of the inner bits.

Principle of Turbo Decoding for a parallel concatenated scheme.

Exchange of extrinsic information between horizontal and vertical decoding



Showing the chaotic behavior of Turbo decoding in a demonstration

- The turbo decoder decodes a parallel concatenated rate $1/2$ code with memory 2, rate $2/3$ convolutional code as constituent codes
- Block interleaver: size 20 x 20 $=$ 400 information bits, 800 transmitted bits
- Decoder: SOVA algorithm with L-values
- Shown are the soft-output L-values of the information bits after each half iteration
- Display shows 20 x 20 interleaver matrix with

Red Circles for wrong bits
Green circles for correct bits

- Diameter of circles is the reliability (magnitude of L-values)
- Goal of decoding:

Big green circles !!!

Properties of LLR

The pdf's and the LLR's have the following properties

- The matched filter output y is Gaussian with

$$\mathcal{N}(m, \sigma_c^2) = \mathcal{N}(\pm a, \sigma_c^2)$$

.

- the APP LLR L_{CH} is also Gaussian with

$$\mathcal{N}(\pm\sigma_{CH}^2/2, \sigma_{CH}^2)$$

where $\sigma_{CH}^2 = 2aL_c$ and is determined by one parameter.

- the matched filter output has a symmetric pdf

$$p(-y|x=+1) = p(y|x=-1)$$

and is a LLR

- all LLR with symmetric distributions satisfies the consistency condition

$$p(-y|x) = e^{-L_c xy} p(y|x).$$

The Extrinsic Information Transfer (EXIT) chart

The mutual information between the equally likely X and the respective LLR's L for symmetric and consistent L-values simplifies to

$$I(L;X) = 1 - \int_{-\infty}^{+\infty} p(L|x=+1)\log_2(1+e^{-L})dL$$

$$I(L;X) = 1 - \mathrm{E}\{\log_2(1+e^{-L})\}$$

The expectation is over the one parameter distribution

$$p(L|x=+1) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(L-x\sigma^2/2)^2/2\sigma^2}$$

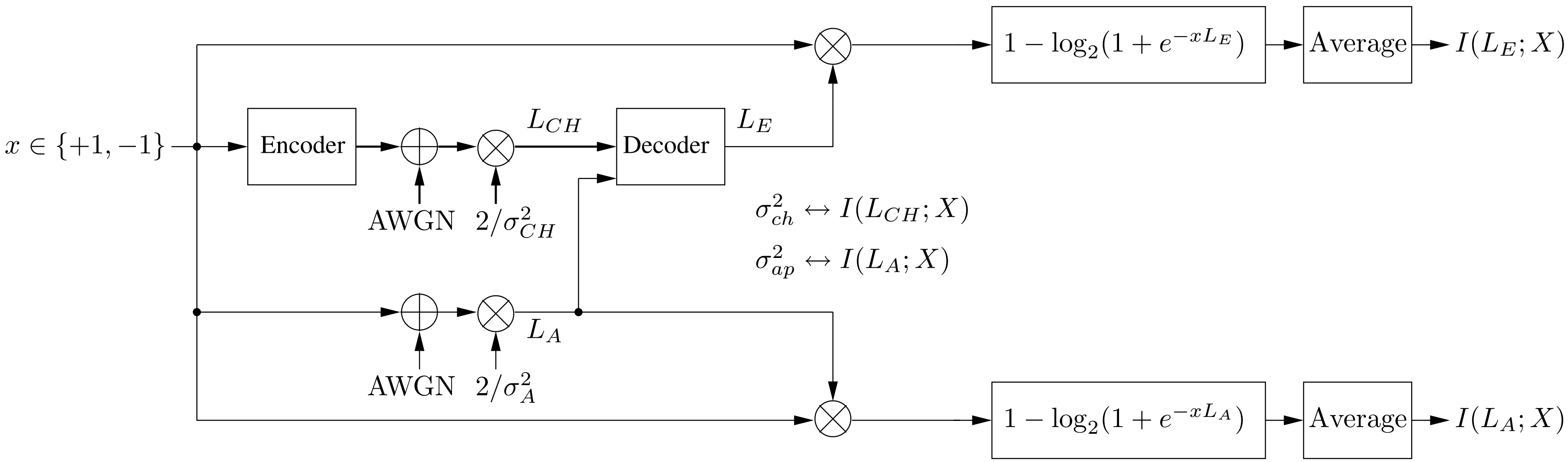
In case of L_E the expectation is over the measured generally non-Gaussian distribution:

$$1 - \mathrm{E}\{\log_2(1+e^{-L})\} \approx 1 - \frac{1}{N}\sum_{n=1}^{N}\log_2(1+e^{-x_n \cdot L_{E,n}})$$

Measurement of the EXIT chart

Experimental measurement from N samples $x_n \cdot L_{E,n}$ which are corrected for positive x by evoking the ergodic assumption

$$I(L;X) = 1 - \frac{1}{N} \sum_{n=1}^{N} \log_2(1 + e^{-x_n \cdot L_{E,n}})$$



The lower part of the diagram is the **a priori** channel channel.

Assumption for the EXIT charts:

Large interleavers to assure statistical independence.

Options for the a priori channel

We can model the a priori channel in several ways as

- an AWGN channel (see previous figure). Here we assume that the soft output of the other constituent decoder to be a consistent normally distributed LLR.
- a BEC with erasure probability ϵ. Here we assume the the other constituent decoder to be a genie aided decoder which converts errors into erasures. This genie aided decoder is the optimal (however unrealizable) SISO decoder. We replace the AWGN a priori channel by a multiplicative channel with an i.i.d. factor

$$a_n = \begin{cases} \infty & \text{with probability} \quad 1-\epsilon \\ 0 & \text{with probability} \quad \epsilon \end{cases}$$

  Then the the mutual information becomes

$$I(L;X) = 1 - \frac{1}{N} \sum_{n=1}^{N} \log_2(1+e^{-a_n}) = 1-\epsilon$$

  as it should be.

The Extrinsic Information Transfer (EXIT) Chart

Pioneered by Stephan ten Brink the information transfer for turbo decoding an the performance of turbo decoding in the fall-off region can be visualized by the EXIT Chart:

The information transfer function T is measured as

$$I(L_E; X) = T(I(L_A; X))$$

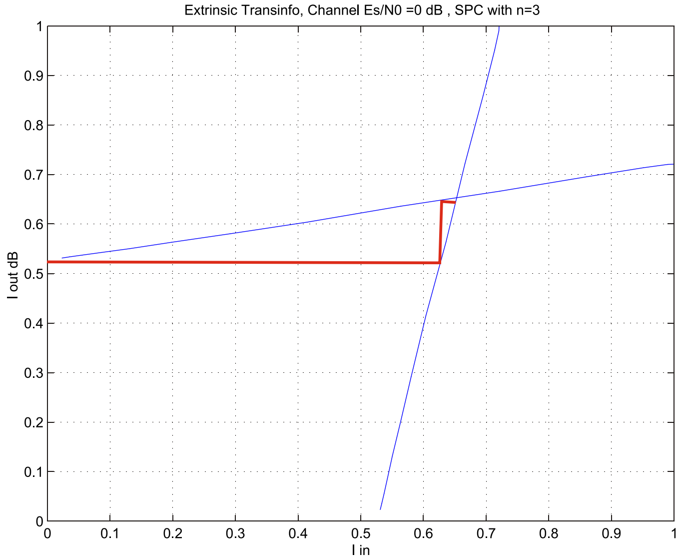
- The assumption is a large interleaver to assure statistical independance and Gaussian distribution for the input L_A with parameter $\sigma_a^2 \leftrightarrow I(L_A; X)$.
- For inner decoders in a serial concatenated scheme and for parallel concatenated schemes the additional parameter input is L_{CH} and σ_{CH}^2, the channel SNR or $I(L_{CH}; X)$ appears as parameter.
- For outer decoders in a serial concatenation only $L_A^{(o)}$ appears as input which is taken from the interleaved serial $L_E^{(i)}$.

Example of an EXIT chart

Parallel concatenated system with single parity check component codes SPC(n-1,n,2) on the AWGN channel: For bit $x_1 = x_2 \oplus x_3$ the extrinsic value is

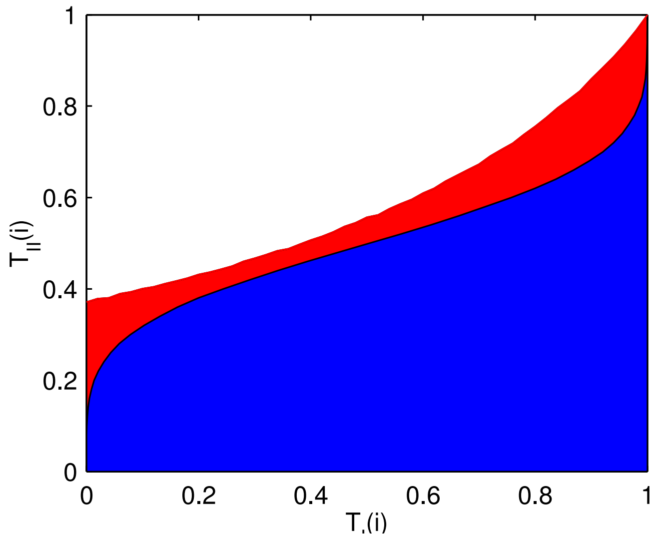
$$L_E = (L_{CH,2} + L_A) \boxplus L_{CH,3}.$$

For the parallel decoder the transfer function with axes are swapped. The iteration alternates between these curves.



The difference $I(L_E; X) - I(L_A; X)$ meaning the difference to the diagonal is the average information gain measured in bits per half iteration.

Observations with EXIT charts



Capacity property: if $R_{II} = 1$ then the area under the white area is

$$\int_0^1 T_{II}(i)\mathrm{d}i \approx C_{ui}$$

Rate property: if decoder is APP based then the blue area is

$$\int_0^1 T_I^{-1}(i)\mathrm{d}i \approx R_I$$

Red Area $= \int_0^1 (T_{II}(i) - T_I^{-1}(i))\,\mathrm{d}i > 0 \quad \Rightarrow \quad R_I < C_{ui}$

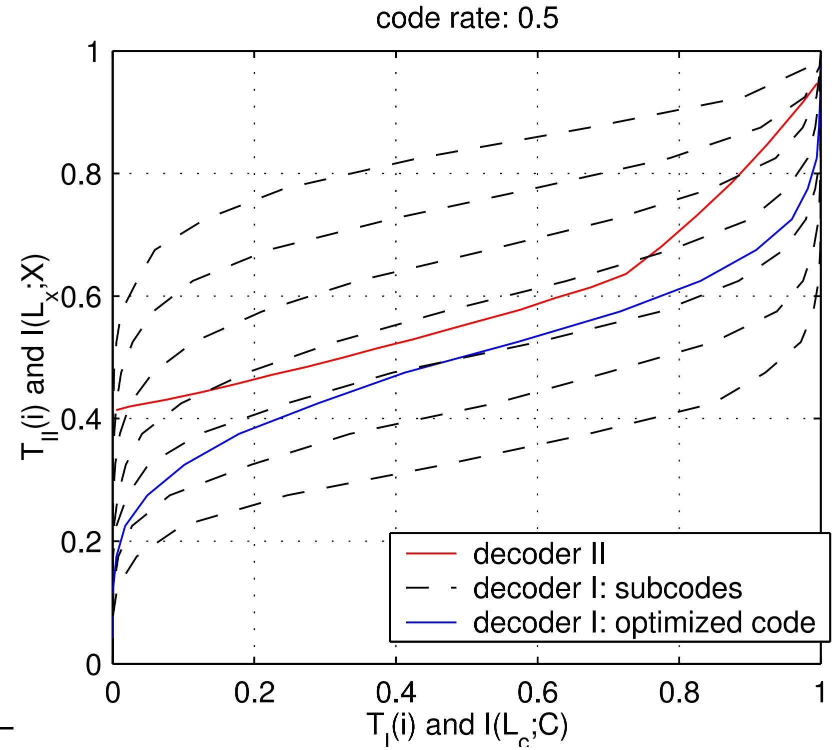
These properties were proven by Ashikhmin, Kramer, ten Brink for an a priori channel with a BEC model .

Turbo applications: Coded equalization of multipath channels

- Channel II: Static multipath, memory L= 4 with taps $\{h_l\} = \{0.32, 0.63, 0, 0.63, 0.32\}$, Modulation: BPSK, SNR of 4 dB
- Metric:

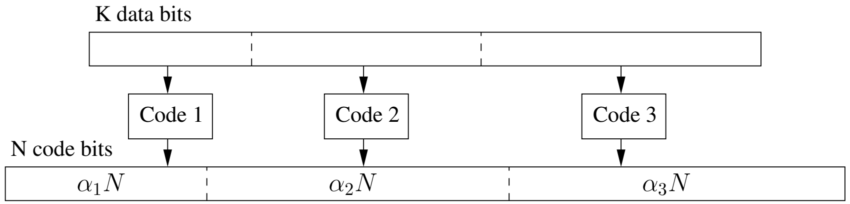
$$\ln p(\mathbf{y}_k \mid u_k) = B_{M_k} - \frac{1}{2\sigma^2}|y_k - \sum_{l=0}^{L} x_{k-l}h_l|^2,$$

EXIT Chart II: Behavior of the Channel



Outer irregular code matched to inner EXIT chart: MT,JH CISS 2002

Construction method:



Convolutional code , memory 4 (16 states), punctured to $L_R = 7$ rates (4/12, 5/12, 6/12, 7/12, 8/12, 9/12, 10/12)
Constraints on the α_k in this FEC code:

$$\sum_{k=1}^{L_R} \alpha_k = 1,$$

$$\sum_{k=1}^{L_R} \alpha_k R_k = 1/2, \quad \text{and} \quad \alpha_k \in [0,1], \ k = 1, \ldots, L_R.$$

Since the L-values of all subcode decoders are symmetric and consistent the overall transfer function I is

$$T_I(i) = \sum_{k=1}^{L_R} \alpha_k T_k(i).$$

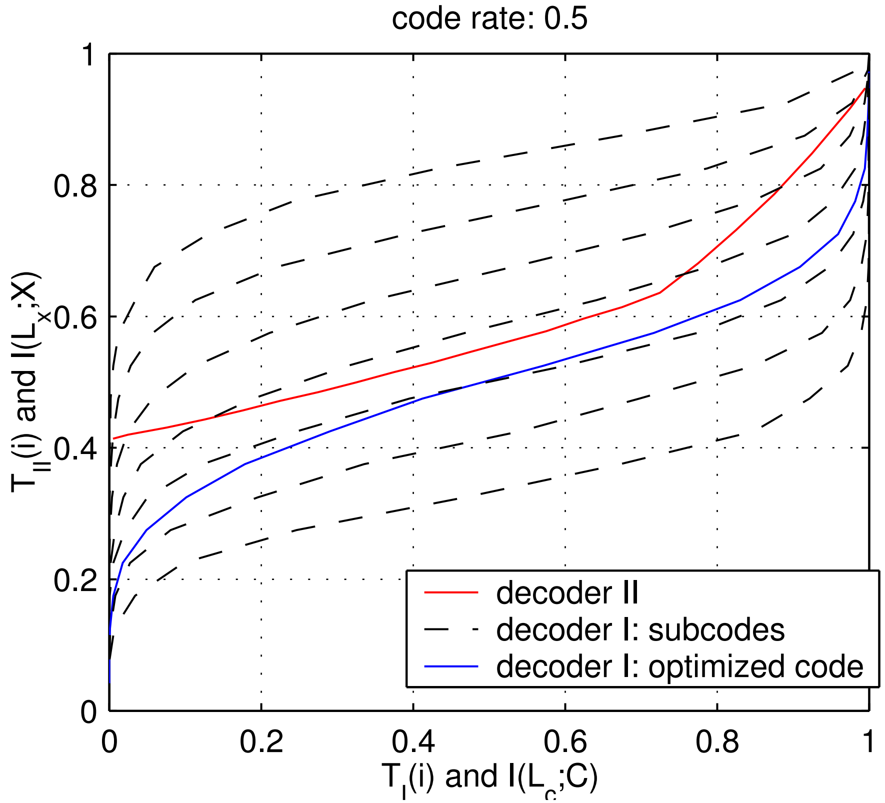
Outer Code matched to inner multipath EXIT chart

Optimization criterium:

Find the set $\{\alpha_i\}$ in such a way that the curves I and II keep a reasonable open tunnel for the iterations

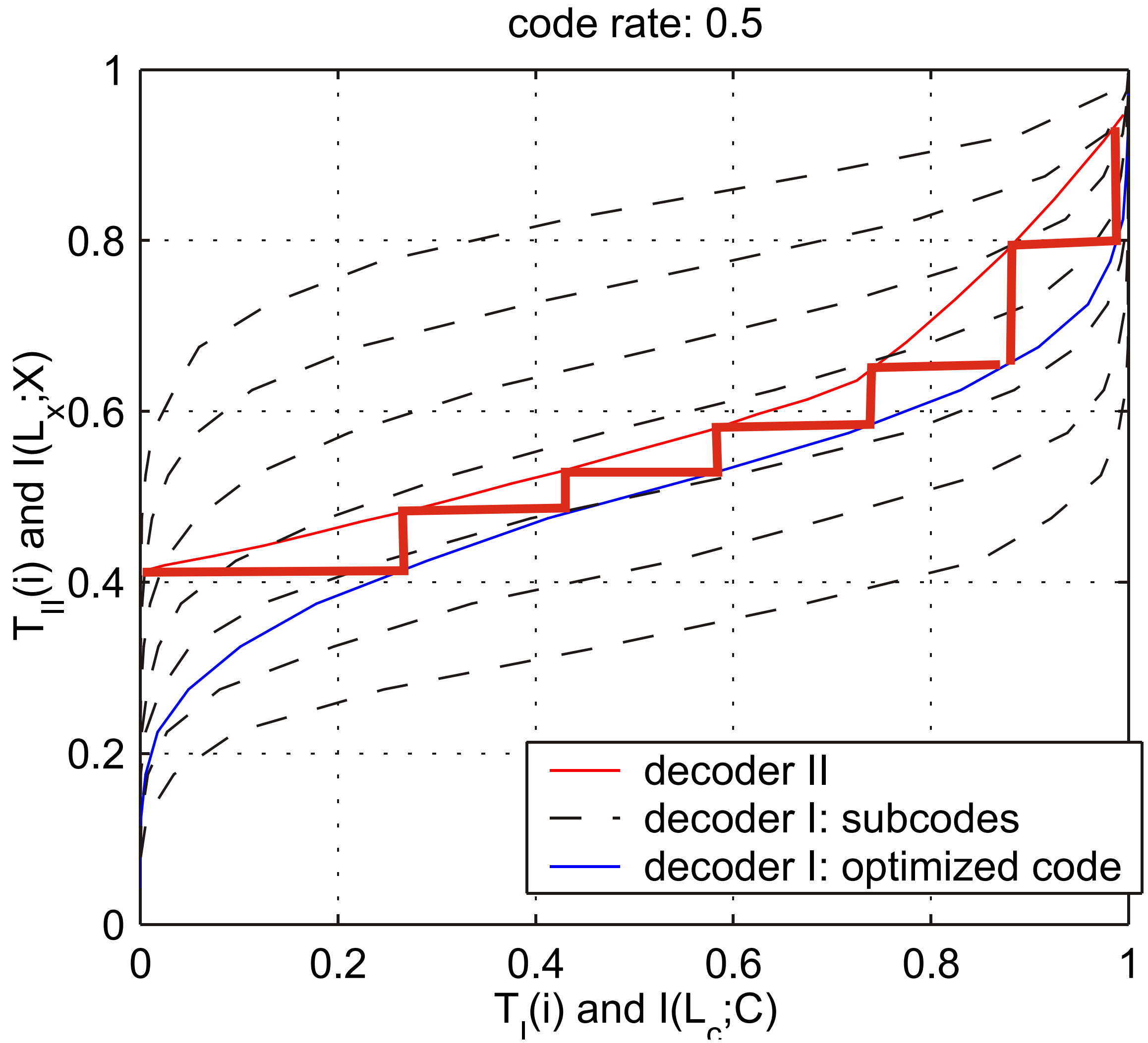
EXIT Chart II for multipath channel

EXIT Chart I for optimized irregular code

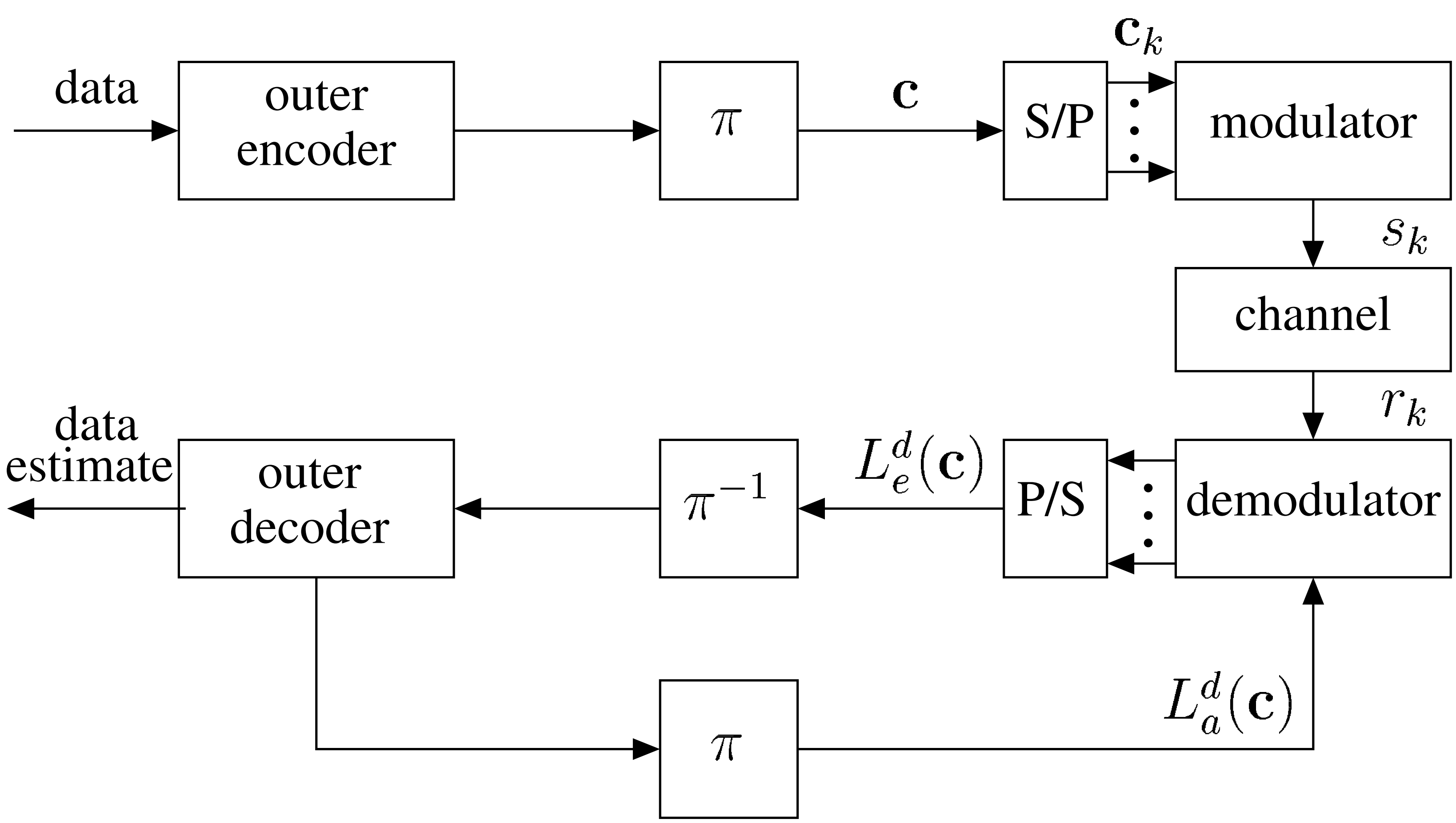


Outer Code matched to inner multipath channel

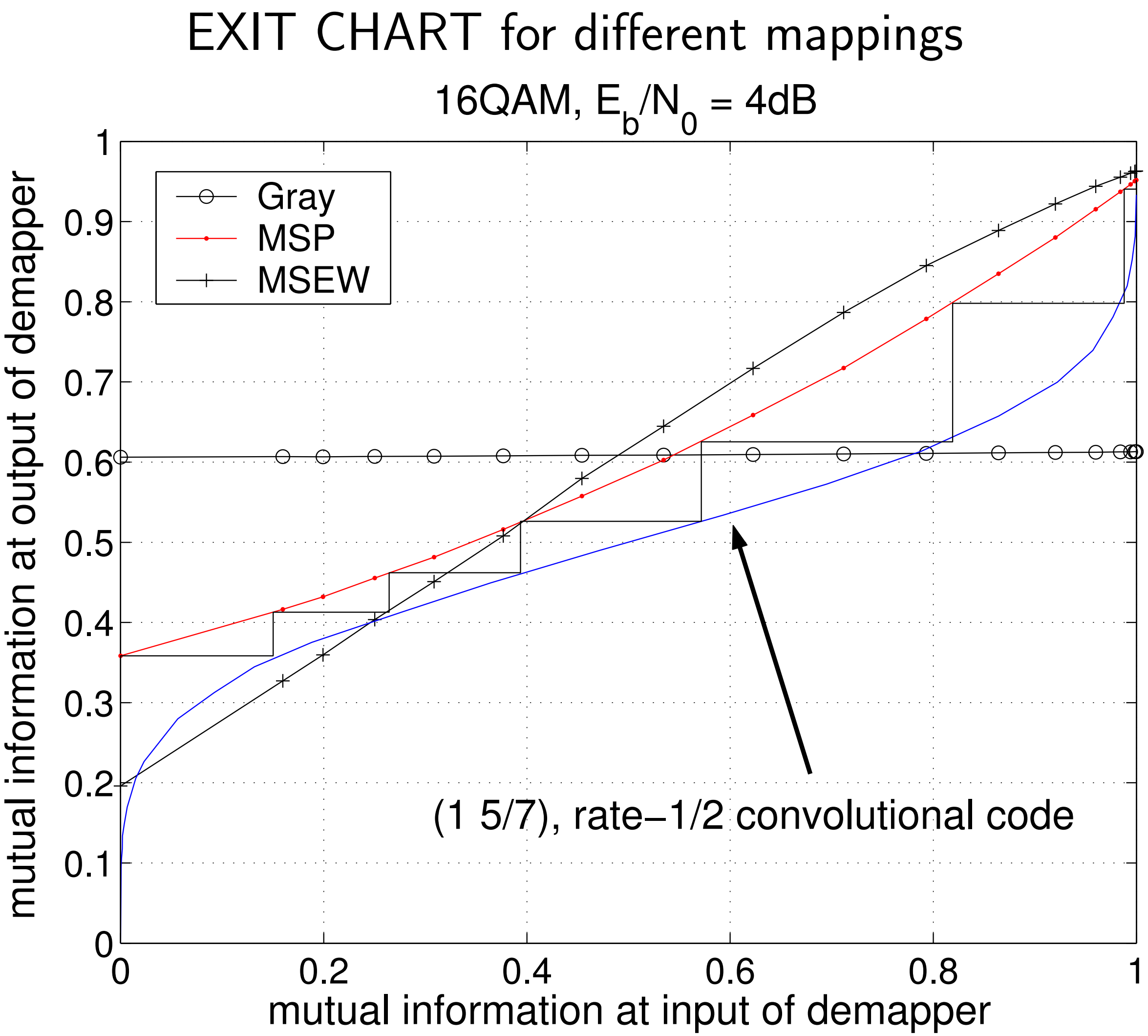
Iteration Trajectory Tunnel closes at rate 0.565



Turbo applications: QAM with channel codes



Turbo applications: 16-QAM with channel codes, rate 1/2, memory 2



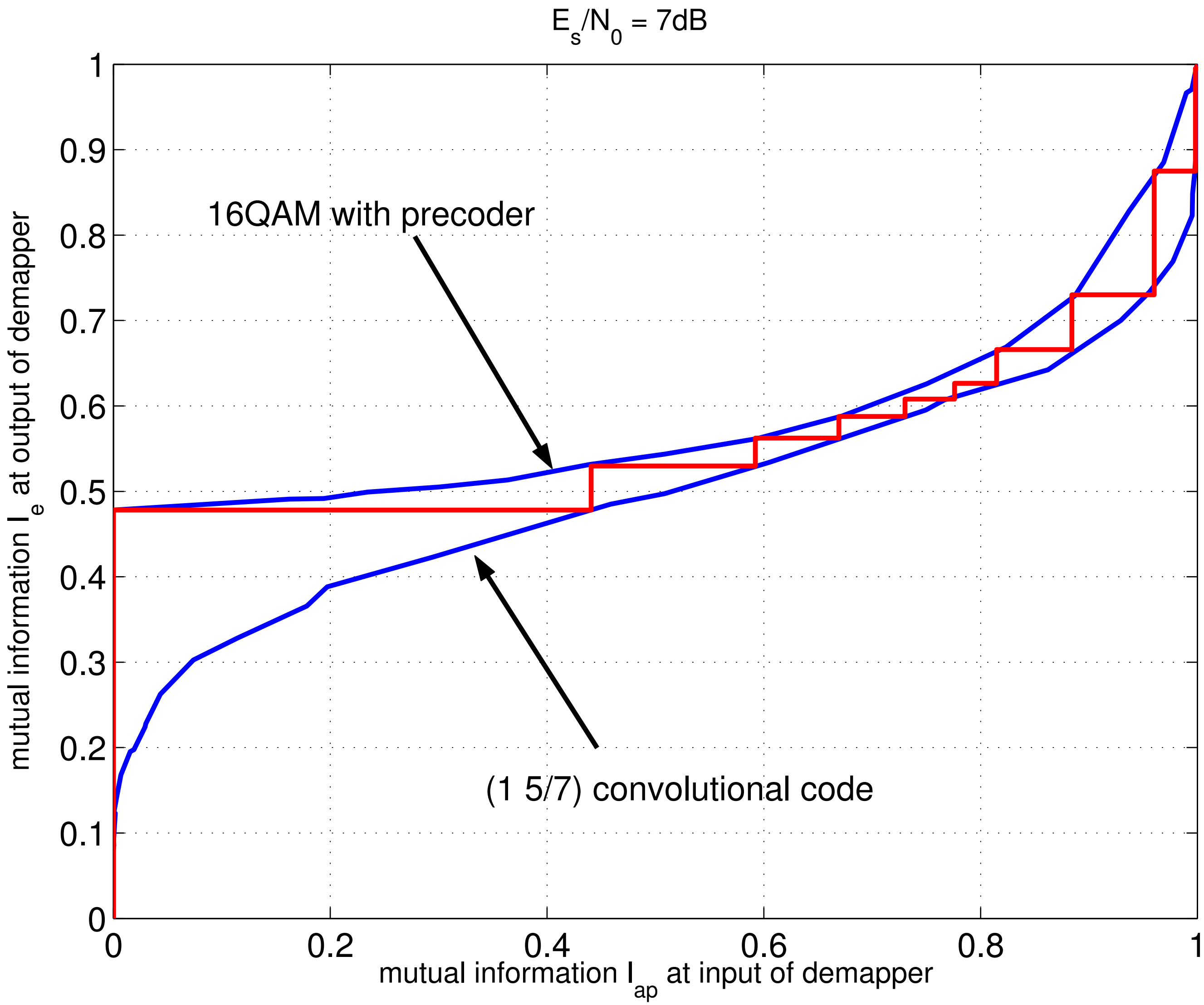
Turbo applications: Precoded QAM with channel codes

Problem: Iterations between inner QAM and outer channel code do not reach the error free point $(I_A, I_E) = (1.0, 1.0)$.

Solution: Differential Precoding:

The inner 'code' is a rate 1 (no redundancy added) differentially precoded QAM mapper

Iterations with precoded 16-QAM

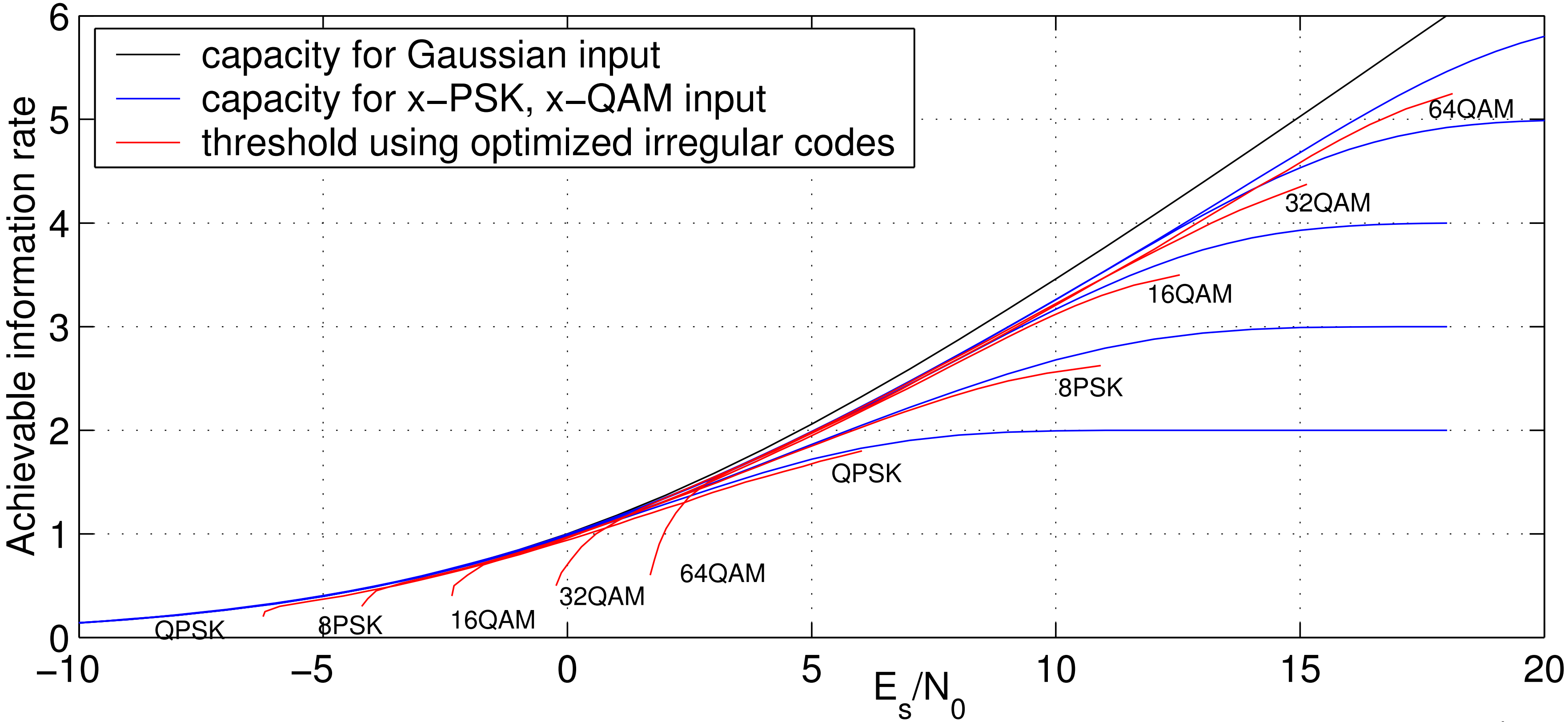


LNT

Summary of optimizations (c.f. Michael Tüchler)

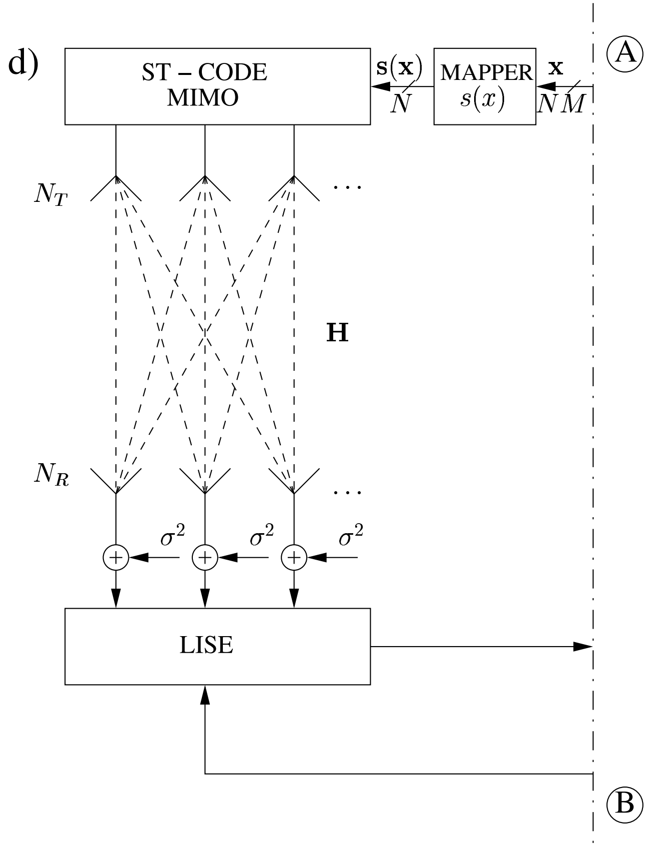
Inner channel:Precoded QAM

Outer channel: Irregular channel codes



Turbo applications: Coded MIMO systems: SB,JH, MW ICC 2003

Inner channel:
Multiple-Input/Multiple Output (MIMO) Antenna System with QAM



Outer channel: Irregular or standard channel codes

Turbo applications: Coded system: Inner MIMO channel

Channel matrix $\mathbf{H}$ contains the complex ergodic channel coefficients $h_k^{(i,j)}$.

Input to the turbo scheme: bit vector of size $M \cdot N \times 1$

$$\mathbf{x} = (\mathbf{x_1}, ...\mathbf{x_n}, ..., \mathbf{x_N})^{\mathrm{T}}$$

with the M-bit row vectors of size $1 \times N$

$$\mathbf{x_n} = (x_{1,n}, ...x_{m,n}, ..., x_{M,n}), \quad x_{m,n} \in \{+1, -1\}.$$

The constellation mapper maps M bits to one complex signal element $s_n(\mathbf{x_n})$ of the signal vector $\mathbf{s(x)}$

After transmission the received vector

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{s}(\mathbf{x}) + \mathbf{n_0}$$

Turbo applications: Coded system: Inner MIMO channel

Symbol APP detector soft-output all NM bits

$$L(\hat{x}_{m,n}) = \ln \frac{\sum_{x_{m,n}=+1} e^{-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{H}\mathbf{s}(\mathbf{x})\|^2+\frac{1}{2}L(\mathbf{x})^{\mathrm{T}}\mathbf{x}}}{\sum_{x_{m,n}=-1} e^{-\frac{1}{2\sigma^2}\|\mathbf{y}-\mathbf{H}\mathbf{s}(\mathbf{x})\|^2+\frac{1}{2}L(\mathbf{x})^{\mathrm{T}}\mathbf{x}}}.$$

$L(\mathbf{x})$ are the a priori values of all bits from the turbo feedback of the outer decoder.

Evaluation for all 2^{MN} possible data is prohibitively complex .

Example: 32-QAM with 4 antennas: 2^{20} values

Turbo applications: Coded system: Inner MIMO channel

SOLUTION: List sequential decoder LISS

Find a candidate list of transmit vectors $\mathbf{x}$ with the highest metrics and evaluate the soft-output only for those candidates.

Search for candidates:

- compute the center (zero forcing) solution

$$\hat{\mathbf{s}}(\mathbf{y}) = \mathbf{H}^{\mathrm{H}}(\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{y}$$

- Apply Cholesky factorization to obtain a lower triangular matrix $\mathcal{L}$

$$\mathcal{L}^{\mathrm{H}}\mathcal{L} = \mathbf{H}^{\mathrm{H}}\mathbf{H}.$$

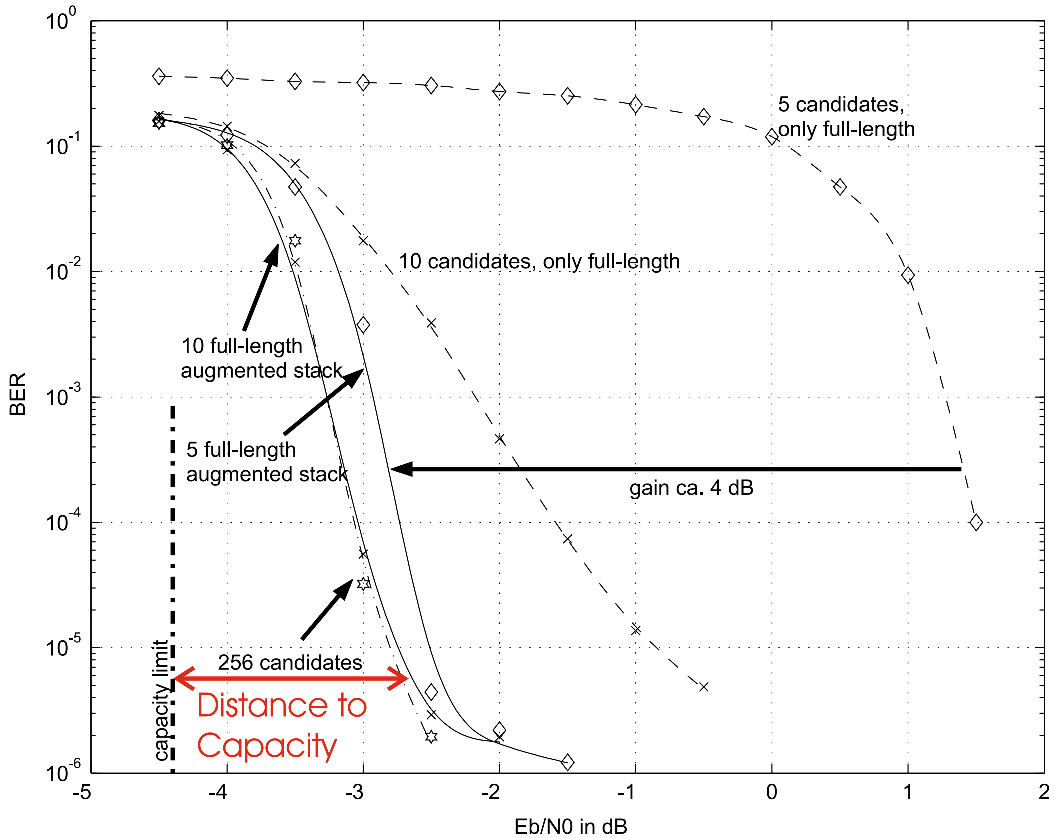
- Obtain an additive metric with increment

$$\lambda_n = -\left|\sum_{j=1}^{n} \frac{\ell_{n,j}}{2\sigma^2}(s(x_j) - \hat{s}_j)\right|^2 + \sum_{m=1}^{M} \frac{1}{2} x_{m,n} L(x_{m,n}).$$

- Use sequential search with the stack algorithm and augment pathes to full length

Turbo applications: Coded MIMO system:

Inner MIMO channel with LISt Sequential decoder (LISS)

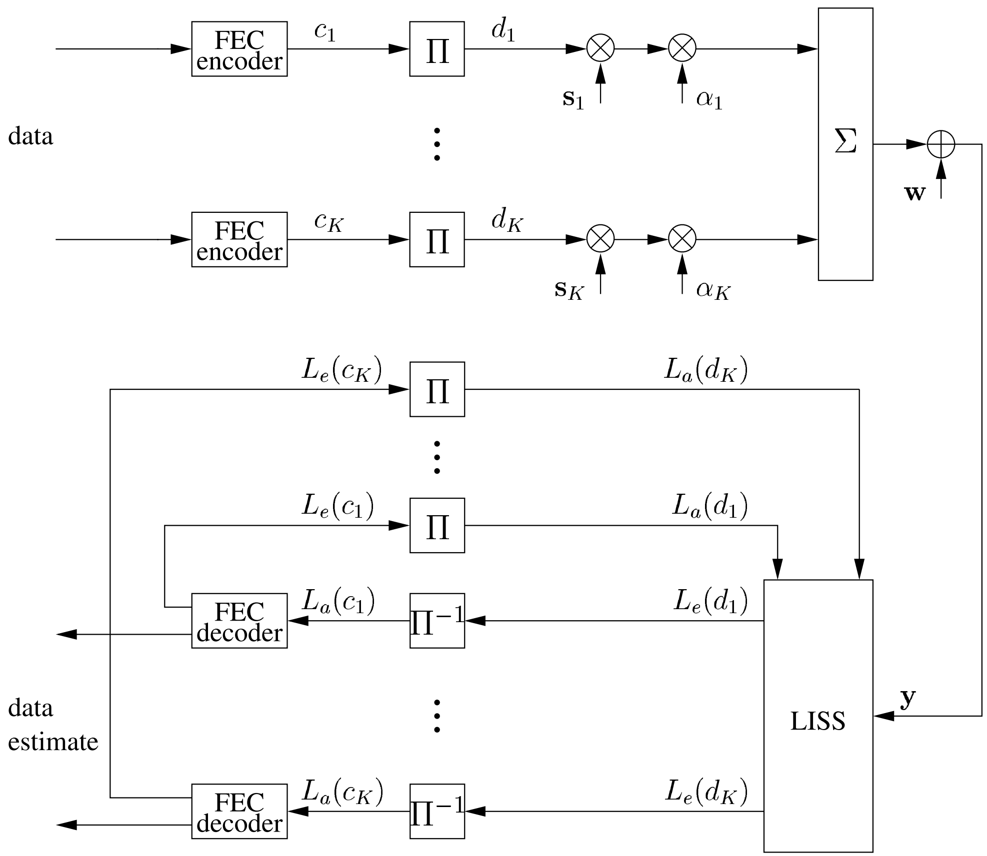


Simulation results: outer PCC code $r = 1/2$, QPSK, $n_T = n_R = 4$, $L_{\max} = 256$, number of candidates with and without stack augmentation

Coded Multiuser System with Turbo LISS Detector

Inner Rayleigh-Fading channel with LISS-detector

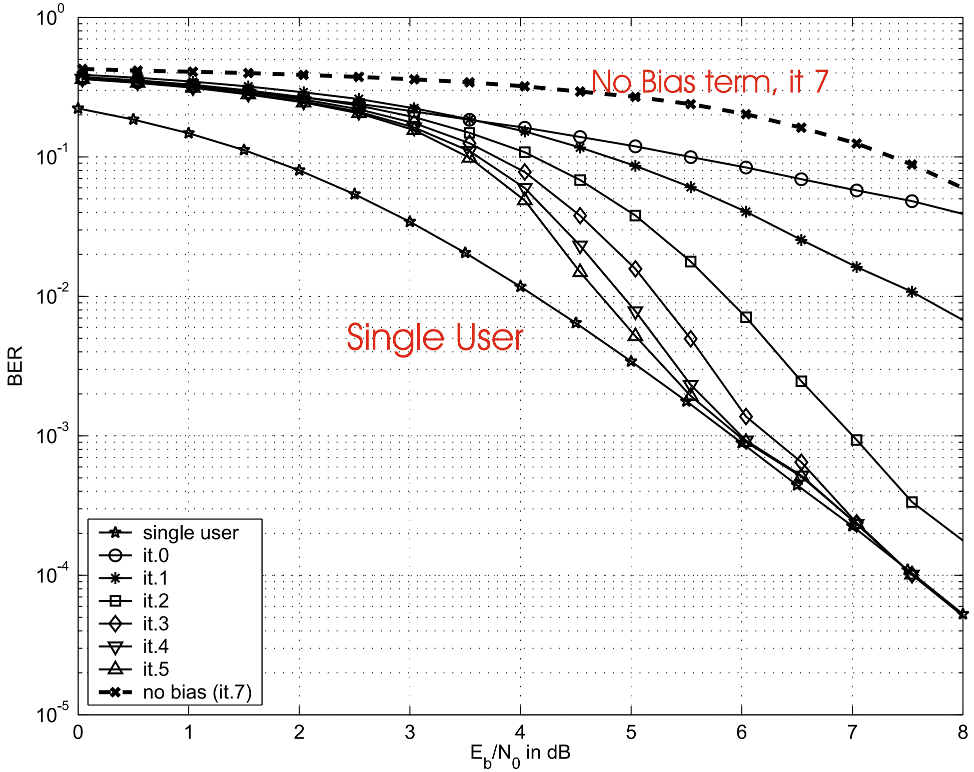
$K = 32$ users with random spreading codes of length $N = 32$
outer conv. code $r = 1/2$, BPSK.



Coded Multiuser System with Turbo LISS Detector

Simulation results: Inner Rayleigh-Fading channel with LISS-detector, stack size 256, aux. stack size 128

$K = 32$ users with random spreading codes of length $N = 32$
outer conv. code $r = 1/2$, BPSK.



Conclusions

- The Turbo Principle is a very general principle
- Most mobile communications schemes employ serial concatenation
- Extrinsic Information Transfer (EXIT) charts are very useful tools
- Approaching the capacity limits is possible
- Drawback: Large interleavers are necessary for optimum performance and analysis
- For systems with high number of states new sequential soft-in/soft-out decoders are presented
- There are many more applications to the Turbo Principle:
  - Turbo source compression
  - Joint source channel coding
  - Interference cancellation
  - Analog decoding