Tutorial on Iterative Detection and Decoding

S. ten Brink

Institute of Telecommunications (INÜ), University of Stuttgart, Germany

24.04.2015

Zandvoort aan je
School of Inf. Theory
Outline

1. Introduction
2. Soft Output Decoding
3. Serially Concatenated Codes (SCC)
4. Parallel Concatenated Codes (PCC)
5. Low-Density Parity Check (LDPC) Codes
6. Iterative Detection
7. Future Trends
Outline

1. Introduction
2. Soft Output Decoding
3. Serially Concatenated Codes (SCC)
4. Parallel Concatenated Codes (PCC)
5. Low-Density Parity Check (LDPC) Codes
6. Iterative Detection
7. Future Trends
Intro - Channel Coding

- channel code, channel interface, channel

\[\text{MIMO} \quad \text{QAM} \quad \text{AWGN}\]

\[
\begin{array}{c}
\text{CH. mod.} \\
\text{CH. int.} \\
\text{channel} \\
\text{CH. dec.} \\
\text{it. dec.}
\end{array}
\]

\[\text{will code at this}\]
Intro - Simple Channels: BEC

- Binary Erasure Channel (BEC)
- Channel quality parameter: Erasure Probability $P_{\text{erasure}}$
- Capacity $C_{\text{BEC}} = 1 - P_{\text{erasure}}$

Lo used to model prior knowledge (late)
Intro - Simple Channels: BSC

- Binary Symmetric Channel (BSC)
- Channel quality parameter: Error probability $P_{err}$
- Capacity $C_{BSC}(P_{err}) = 1 - H_b(P_{err})$
  with $H_b(P_{err}) = -P_{err} \cdot \log P_{err} - (1 - P_{err}) \cdot \log (1 - P_{err})$
Intro - Simple Channels: AWGN

- Additive White Gaussian Noise (AWGN) channel
- $n$ Gaussian distributed, mean zero, variance $\sigma^2$
- Channel quality parameter: $SNR = E_s/2\sigma^2 = E_s/N_0$
- Capacity $C_{AWGN} = \log_2 (1 + SNR)$
Intro - Simple Codes: Repetition

- \((N, K)\) block code of rate \(R = K/N\)
- repetition codes \(R = 1/N\), \(N\)-fold repetition \(c = (u, \ldots, u)\)
- on BEC: can correct \(N - 1\) erasures
- on BSC: can correct \(\lfloor(N - 1)/2\rfloor\) errors
- before Shannon: we have to sacrifice rate to achieve reliability
- Shannon (1948): arbitrarily good reliability achievable as long as the code rate \(R\) is below the channel capacity \(C\)
- important for coding guys: SNR normalized to energy used per information bit

\[
\frac{E_b}{N_0} = \frac{1}{M_b R} \left( \frac{E_s}{N_0} \right)
\]
Intro - Simple Codes: Single Parity Check

- $(N,N-1)$ block code of rate $R = \frac{N-1}{N}$
- $c = (u_1, u_2, ..., u_{N-1}, u_1 \oplus u_2 \oplus ... \oplus u_{N-1})$
- on BEC: can correct one erasure
- on BSC: can not correct errors; only detect odd number of errors
Intro - Simple Codes: Multiple Parity Checks

- e.g., (7,4) Hamming code, rate $R = 4/7$
- $c = (u_1, u_2, u_3, u_4, u_1 \oplus u_2 \oplus u_3, u_1 \oplus u_2 \oplus u_4, u_1 \oplus u_3 \oplus u_4)$
- on BEC: can correct two erasures
- on BSC: can correct one error; detect two errors
- extended Hamming code: additional parity bit $u_2 \oplus u_3 \oplus u_4$
Intro - Comparison of Simple Codes

- An Experiment: Comparison of four $R = 1/2$ codes
- code 1: $c = (u, u)$
- code 2: $c = (u_1, u_2, u_1 \oplus u_2, u_1)$
- code 3: $c = (u_1, u_2, u_3, u_1 \oplus u_2, u_1 \oplus u_3, u_2 \oplus u_3)$
- code 3: $c = (u_1, u_2, u_3, u_4, u_1 \oplus u_2 \oplus u_3, u_1 \oplus u_2 \oplus u_4, u_1 \oplus u_3 \oplus u_4, u_2 \oplus u_3 \oplus u_4)$
Intro - Simple Codes, BER Chart

- code 1 (repetition) has no coding gain
- code 4 has highest coding gain
---

**Intro - Mutual Information Chart vs $E_s/N_0$**

- AWGN-channel: repetition code has some gain...?
Intro - Mutual Information Chart vs $E_b/N_0$

- Now versus $E_b/N_0$: repetition code has no coding gain!
Intro - Need for Coding

- long, random-like codes (with structure) suffice
- various coding schemes emerged over the past 50 years
- and: channel interface should make channel “look nice” (Gaussian-like) for channel code
- soft-input decoding seems important
- soft-output decoding...?
Note: this is an attempt of a coding timeline with progression on it, decoding, it, detection as incomplete, biased, ... send an email with improvement.

S. ten Brink Iterative Detection and Decoding 24.04.2015

tenbrink@ieee.org
Intro - Channel Coding Timeline, 1940-1960

1940's
- Information Theory (Shannon)
- Hamming Code
- Golay Code

1950's
- Reed-Muller Code
- Convolutional Code, Product Code (Elias)
- Cyclic Codes (NN)
- BCH Codes
- Reed-Solomon Codes
Intro - Channel Coding Timeline, 1960-1980

1960's

Sequential Decoding

LDPC Codes (Gallager)

Concatenated Codes (Forney)

Viterbi Decoder

1970's

"CODING IS DEAD"

Chase Decoder

BCJR Decoder

Trellis Coded Modulation (Ungerböck)

Pioneer 9
- R=1/2 conv. code,
- memory 20, seq. dec.

Mariner '69 (32,6) RM,
- corr. dec.

Pioneer 10, 11 (2,1,32)
- conv. memory 31, seq.

Voyager 1,2 (2,1,7) conv.
- Viterbi Decoder

Really?... Soft in loud
Intro - Channel Coding Timeline, 1980-2000

1980's
- Tanner Code
- Pondération des symboles... (Battail)
- SOVA (Hagenauer, Höher)
- Separable MAP filters (Lodge, Hagenauer)
- Turbo Codes (Berrou, Glavieux, Thitimajshima)
- CCSDS Telemetry Std concatenation (255,223) RS, (2,1,7) conv. code
- Galileo, concatenation (255,223) RS, (4,1,15) conv.

1990's
- Irregular LDPC Codes (MacKay)
- Iterative Demapping
- It. dec. product codes (Pyndiah et al.)
- Repeat-Acc. Codes (Divsalar el al.)
- Density Evolution Richardson, Urbanke
- Iterative MUD (Moher)
- Iterative EQ (Bauch, Franz, Hagenauer)
- Belief Propagation (McEliece, MacKay, Cheng)
- Convolutional LDPCCC (Felström, Zigangirov)
- CCSDS Telemetry Std, Turbo Codes added
# Intro - Coding History Timeline, 2000-2020

## 2000's
- Belief Prop., Factor Graphs (Kschischang, Frey, Löliger)
- Polar Codes (Arikan)
- List Sphere Detection (Hochwald, tB)
- LDPC modulation (tB, Kramer, Ashikhmin)
- DVB-S2 (LDPCC)
- UMTS (turbo)

## 2010's
- "THE PHY LAYER IS DEAD"
- SpatCC, Threshold Saturation (Kudekar, Richardson, Urbanke)
- SpatCC, Universality Property (Kudekar, Richardson, Urbanke)
- Staircase Codes (Kschischang)
- Spatial Coupling
- Optical long haul (Staircase; SpatCC)
- LTE (turbo)
- 802.11ac/ad (LDPCC)
- WLAN 802.11n (LDPCC)
Outline

1 Introduction
2 Soft Output Decoding
3 Serially Concatenated Codes (SCC)
4 Parallel Concatenated Codes (PCC)
5 Low-Density Parity Check (LDPC) Codes
6 Iterative Detection
7 Future Trends
Soft Decoding - Log-Likelihood Ratio Values

- Discrete-time channel model for binary, antipodal signaling
  \[ y = x + n; \ x \in \{ \pm \sqrt{E_s} \} \]

- realizations of \( n \), e.g., i.i.d. Gaussian; in the following \( E_s = 1 \)

- A posteriori L-value ("soft channel output") given by the ratio of the a posteriori probabilities \( P(x = \pm 1 \mid y) \)
  \[ L(x \mid y) = \ln \frac{P(x = +1 \mid y)}{P(x = -1 \mid y)}. \]

- Applying Bayes’ rule yields
  \[ P(x = +1 \mid y) = \frac{p(y \mid x = +1)}{p(y)} \cdot P(x = +1). \]

  - a priori probability \( P(x = +1) \) that \( x = +1 \) was transmitted,
  - channel output PDF \( p(y \mid x = +1) \) conditioned on the transmitted symbol \( x = +1 \)
  - and \( p(y) = P(x = -1) \cdot p(y \mid x = -1) + P(x = +1) \cdot p(y \mid x = +1) \)
Soft Decoding - Log-Likelihood Ratio Values

- Finally, we obtain the a posteriori L-value as

\[ L(x|y) = \ln \frac{P(x = +1)}{P(x = -1)} + \ln \frac{p(y|x = +1)}{p(y|x = -1)} \]

\[ = L_A(x) + L_{ch}(x|y) \]

- Sign of L-value: hard decision
- Absolute value \(|L(x|y)|\): reliability of the decision
- A priori L-value \(L_A(x)\): accounts for available prior knowledge on \(x\)
- Channel L-value \(L_{ch}(x|y)\): knowledge on \(x\) based on \(y\)
- Additions/subtractions rather than multiplications/divisions
• Connection between a priori probability $P(x = 1)$ and corresponding L-value $L_A(x) = \ln \frac{P(x=+1)}{P(x=-1)} = \ln \frac{P(x=+1)}{1-P(x=+1)}$
Soft Decoding - Log-Likelihood Ratio Values

- For the AWGN channel we have

\[ p(y|x) = \frac{1}{\sigma \sqrt{2\pi}} \cdot \exp \left[ -\frac{(y-x)^2}{2\sigma^2} \right] \]

and obtain

\[ L_{ch}(x|y) = \ln \frac{\exp \left[ -\frac{(y-1)^2}{2\sigma^2} \right]}{\exp \left[ -\frac{(y+1)^2}{2\sigma^2} \right]} = \frac{2}{\sigma^2} \cdot y \]

- Channel L-values are simply weighted versions of the channel observations \( y = x + n \)
- Can be easily included into a metric for soft input decoding (e.g. Viterbi algorithm)
Soft Decoding - Log-Likelihood Ratio Values

- Channel L-values $L_{ch}(x|y)$ for transmitted symbol $x = 1$

- It can be shown that the mean value $\mu_{L_{ch}}$ and the variance $\sigma_{L_{ch}}^2$ satisfy $\mu_{L_{ch}} = \frac{\sigma_{L_{ch}}^2}{2}$
Soft Decoding - Difference of ML versus APP Decoding

- Encoder: maps $K$ information bits $u$ into the sequence of $N$ coded bits $c = \text{map}(u)$
  (or antipodal sequence $x(u \in \{\pm 1\})$ respectively)
- Mapping is one-to-one and thus invertible $u = \text{map}^{-1}(c)$
- Received signal from the channel: $y = x + n$.
- Realizations of $n$ are i.i.d. Gaussian, mean zero, variance $\sigma^2$
- Channel decoder: obtain estimate $\hat{u}$ of transmitted bit sequence $u$
Soft Decoding - ML and APP Decoding

- Two decoding rules based on different optimization criteria:
  - Minimizing the sequence error probability \( P_{seq} \)
  - Minimizing the bit error probability \( P_b \)
  - Minimizing \( P_{seq} \): maximum likelihood sequence estimation (MLSE, or ML)
  - Minimizing \( P_b \): take hard decision on a posteriori probabilities (APP) of information bits (APP, or MAP decoding)

Viterbi decoding
Soft Decoding - ML-Decoding

- ML decoding: Receiver generates all possible $2^K$ transmitted codeword hypotheses $c$ (message hypotheses $u = \text{map}^{-1}(c)$) and selects that one which maximizes the sequence a posteriori probability

$$\max_{\forall u} P(u \mid y)$$

- Assuming that all transmitted message vectors $u$ are equally likely, we obtain a maximization of the likelihood function

$$\max_{\forall u} p(y \mid u)$$

- Complexity of direct approach (testing all $2^K$ message hypotheses) grows exponentially in the message length $K$

- Viterbi algorithm: Exploits trellis structure of convolutional codes; complexity grows only linearly in $K$
Soft Decoding - ML-Decoding

- For the AWGN channel, the likelihood computation reduces to

$$\tilde{p}(y|x(u)) = \sum_{i=0}^{N-1} x_i \cdot y_i = x \cdot y^T$$

- A simple correlation is sufficient to find the best matching hypothesis $x$ (codeword $c$, message $u$ respectively)
Soft Decoding - APP Decoding

- Maximum a posteriori probability criterion
  - optimizes the bit (symbol) error probability rather than the sequence error probability
- Provides soft output values, expressed in terms of a posteriori log-likelihood ratio values (L-values)
  - based on a posteriori probabilities with respect to the bits (not the sequence, as opposed to MLSE)
- Maximum A Posteriori Probability decoding is abbreviated as “APP decoding”, or “MAP decoding”
Soft Decoding - APP Decoding

- Example: Compute a posteriori L-values for an arbitrary rate $1/2$ block code with $K = 2$ and $N = 4$
- E.g., the a posteriori L-value of bit $u_0$ conditioned on the received vector $y = (y_0, y_1, y_2, y_3)$ is

$$L_D(u_0|y) = \ln \frac{P(u_0 = 0|y)}{P(u_0 = 1|y)}$$

The probability

$$P(u_0 = 0|y) = P(u_0 = 0, u_1 = 0|y) + P(u_0 = 0, u_1 = 1|y)$$

is the a posteriori probability that the transmitted information bit was $u_0 = 0$
Soft Decoding - APP Decoding

- Next we apply Bayes’ rule

\[
P(u_0, u_1 | y) = \frac{p(y | u_0, u_1)}{p(y)} \cdot P(u_0, u_1)
\]

- The information bits \( u_0, u_1 \) (e. g. output of a memoryless source) can be assumed to be independent, and we have

\[
P(u_0, u_1) = P(u_0) \cdot P(u_1)
\]
Soft Decoding - APP Decoding

- We obtain

\[ L_D(u_0 | y) = \ln \frac{p(y|u_0=0,u_1=0) \cdot P(u_0=0) \cdot P(u_1=0) + p(y|u_0=0,u_1=1) \cdot P(u_0=0) \cdot P(u_1=1)}{p(y|u_0=1,u_1=0) \cdot P(u_0=1) \cdot P(u_1=0) + p(y|u_0=1,u_1=1) \cdot P(u_0=1) \cdot P(u_1=1)} \]

\[ = \ln \frac{P(u_0=0)}{P(u_0=1)} + \ln \frac{p(y|u_0=0,u_1=0) \cdot \frac{P(u_1=0)}{P(u_1=1)} + p(y|u_0=0,u_1=1)}{p(y|u_0=1,u_1=0) \cdot \frac{P(u_1=0)}{P(u_1=1)} + p(y|u_0=1,u_1=1)} \]

\[ = L_A(u_0) + \ln \frac{p(y|u_0=0,u_1=0) \cdot \exp L_A(u_1) + p(y|u_0=0,u_1=1)}{p(y|u_0=1,u_1=0) \cdot \exp L_A(u_1) + p(y|u_0=1,u_1=1)} \]

\[ L_E'(u_0 | y) \]

- A posteriori L-value \( L_D(u_0 | y) \) is composed of its a priori L-value \( L_A(u_0) \) and the channel-and-“extrinsic” L-value \( L_E'(u_0 | y) \)
- For systematic codes: can separate \( L_E'(u_0 | y) \) into channel observation \( L_{ch}(u_0 | y_0) \) and “pure” extrinsic L-value \( L_E(u_0 | y_0) \)
- The extrinsic L-value: captures all information we learn about bit \( u_0 \) based on code redundancy and observation of bit \( u_1 \)
Suppose that the first \( K = 2 \) bits of the codeword \( c \) are the systematic bits, \( c = (c_0 = u_0, c_1 = u_1, c_2, c_3) \).

Then we can factorize the PDF for the systematic bits \( u_i, 0 \leq i < 2 \), into

\[
p(y|u) = p(y|c = \text{map}(u)) = p(y_i|c_i) \cdot p(y_i[c[i]) = p(y_i|c_i) \cdot p(y_i|u)
\]

and find for \( u_0 \)

\[
L_D(u_0|y) = L_A(u_0) + \underbrace{\ln \frac{p(y_0|u_0 = 0)}{p(y_0|u_0 = 1)}}_{L_{ch}(u_0|y_0)} + \ln \frac{p(y_{[0]}|u_0 = 0, u_1 = 0) \cdot \exp L_A(u_1) + p(y_{[0]}|u_0 = 0, u_1 = 1)}{p(y_{[0]}|u_0 = 1, u_1 = 0) \cdot \exp L_A(u_1) + p(y_{[0]}|u_0 = 1, u_1 = 1)}
\]

\[
= L_E(u_0|y_{[0]})
\]
Soft Decoding - APP Decoding

- Notation: \( y_{[i]} \) is the vector \( y \) where the \( i \)th element is omitted, i.e.
  \[ y_{[i]} = (y_0, y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_{N-1}) \]
- For a general number of \( K \) information bits we find

\[
L_D (u_i | \underline{y}) = L_A (u_i) + \ln \frac{\sum_{u \in U_{i,0}} p (\underline{y} | u) \cdot \exp \left( \sum_{j \in J_{i,u}} L_A (u_j) \right)}{\sum_{u \in U_{i,1}} p (\underline{y} | u) \cdot \exp \left( \sum_{j \in J_{i,u}} L_A (u_j) \right)} L_E'(u_i | \underline{y})
\]

with \( U_{i,0} \) being the set of \( 2^{K-1} \) bit vectors \( u \) having \( u_i = 0 \)
\[
U_{i,0} = \{ u | u_i = 0 \}, \quad U_{i,1} = \{ u | u_i = 1 \}
\]
and \( J_{i,u} \) being the set of indices \( j \) with
\[
J_{i,u} = \{ j | 0 \leq j < K, j \neq i \wedge u_j = 0 \}
\]
Soft Decoding - APP Decoding

- For systematic codes with \( \mathbf{c} = (c_0 = u_0, \ldots, c_{K-1} = u_{K-1}, c_K, \ldots, c_{N-1}) \), we can extract the "pure" extrinsic information for \( u_i, 0 \leq i < K \):

\[
L_D(u_i | y) = L_A(u_i) + L_{ch}(u_i) + \ln \left( \sum_{u \in U_{i,0}} p(y_{[i]} | u) \cdot \exp \left( \sum_{j \in J_{i,u}} L_A(u_j) \right) \right) - \ln \left( \sum_{u \in U_{i,1}} p(y_{[i]} | u) \cdot \exp \left( \sum_{j \in J_{i,u}} L_A(u_j) \right) \right)
\]
Outline

1. Introduction
2. Soft Output Decoding
3. Serially Concatenated Codes (SCC)
4. Parallel Concatenated Codes (PCC)
5. Low-Density Parity Check (LDPC) Codes
6. Iterative Detection
7. Future Trends
SCC - Encoder and Iterative Decoder

- Index “1”: Elements belonging to inner encoding/decoding
- Index “2”: Elements belonging to the outer encoding/decoding
Serially concatenated code consisting of outer rate 1/2 memory 2 recursive systematic convolutional code \((G_r, G) = (07, 05)\) and inner rate 1 memory 1 (differential) code \((G_r, G) = (03, 02)\)
Bit error rate curves for serially concatenated code example; interleaver size $4 \cdot 10^5$ coded bits
SCC - BER Chart

We identify three typical regions of the BER chart:

- The region of low $E_b/N_0 < E_b/N_0|_{\text{cliff}}$ with negligible iterative BER reduction
- The turbo cliff region at about $E_b/N_0 \approx E_b/N_0|_{\text{cliff}}$ with persistent iterative BER reduction over many iterations
- The BER floor region for moderate to high $E_b/N_0$-values in which a rather low BER can be reached after just a few number of iterations

Property of the particular concatenation used
SCC - Inner Transfer Characteristics

- Experiment: Rate 1 mappings
- Compact definition of $M$-bit rate 1 mappings: regard vector realizations $u$ and $c$ as integer values $v$, $0 \leq v < 2^M$, with $v_u = \sum_{m=0}^{M-1} u_m \cdot 2^m$, and $v_c$ correspondingly
- The value $v_u$ serves as an index to the elements of a vector $V$ which contains the definition of the mapping

$$V = (v_c (v_u = 0), v_c (v_u = 1), \ldots, v_c (v_u = 2^M - 1))$$

- Example: The most simple mapping is the identity mapping $c = u$, given by the vector $V_{id} = (0, 1, 2, \ldots, 2^M - 1)$; an arbitrary 2-bit mapping is defined by $V_2 = (0, 3, 1, 2)$, an arbitrary 3-bit mapping by $V_3 = (0, 4, 1, 2, 3, 5, 7, 6)$
- There are $(2^M)!$ different $M$-bit mappings possible; but many are equivalent
Notes

\[ J(u; y) = J(c; y) = J(x; y) \]

\[ \text{id. mapping: } J(y; y) = H \cdot C_0 \]

\[ \text{mapping is information preserving (reconstructs)} \]
SCC - Inner Transfer Characteristics

- Conditional Mutual Information and Discrete A Priori Knowledge:
- Assumption: AWGN channel, BPSK
- (Average) mutual information between transmitted information bit vector \( u \) and the noise-corrupted channel output vector \( y = x + n \):

\[
I(U;Y) = \sum_{\forall u} P(U = u) \cdot \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} p(\xi | U = u) \cdot \frac{p(\xi | U = u)}{p(\xi)} \cdot \frac{d\xi_0 \ldots d\xi_{M-1}}{M-\text{fold integration}}.
\]

- \( p(\xi | U = u) \) is the PDF of the AWGN channel
- The bits \( u_m \) are independent, \( P(U_m = 0) = P(U_m = 1) = 1/2 \)
- Consequently, the a priori probability of a vector realization \( u \) is \( P(U = u) = 1/2^M \)
SCC - Inner Transfer Characteristics

- Chain rule of mutual information: write $I(U;Y)$ as a sum of $M$ bitwise conditional mutual informations $I_L$

$$I(U;Y) = \sum_{L=0}^{M-1} I_L = M \cdot C_G \left( \frac{E_b}{N_0} \right) \leq M$$

- $I_L$ is a short-hand notation of

$$I_L = I(U_m; Y \mid \text{other bits known})$$

$$0 \leq I_L \leq 1$$

- The bar indicates that $I_L$ is averaged
  - over bitwise mutual information with respect to all $M$ bits
  - all possible $\binom{M-1}{L}$ combinations to choose $L$ known bits out of the total of $M-1$ other bits
  - and over all $2^L$ bit vector realizations thereof
SCC - Inner Transfer Characteristics

- Vector channel $0 \leq I(U;Y) \leq M$ can be viewed as being composed of $M$ parallel sub-channels with mutual information $0 \leq I_L \leq 1$ each.
- Allows interesting interpretation:
  - The mapping only influences the partitioning of the total amount of mutual information $M \cdot C_G$ among the different conditional sub-channels $I_L$,
  - whereas the sum $\sum I_L$ always adds up to the constant value $M \cdot C_G$, independently of the applied mapping
- Quantities $I_L$ are well suited for characterizing different mappings
- Partitioning of mutual information among the sub-channels $I_L$ has a strong impact on the behavior of the particular mapping in an iterative decoding scheme
SCC - Inner Transfer Characteristics

- Example: 5-bit mappings
- Two randomly chosen examples \((E_b/N_0 = 1\text{dB assuming } R = 1/2)\)

<table>
<thead>
<tr>
<th>5-bit mappings</th>
<th>(M = 5) parallel sub-channels (I_L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(V_{5,id})</td>
<td>(I_0)</td>
</tr>
<tr>
<td>0.562</td>
<td>0.562</td>
</tr>
<tr>
<td>0.416</td>
<td>0.487</td>
</tr>
<tr>
<td>0.249</td>
<td>0.410</td>
</tr>
</tbody>
</table>

- The mappings are defined as

\[
V_{5,id} = (0, 1, 2, ..., 30, 31) \quad \checkmark
\]

\[
V_{5,1} = (17, 16, 8, 0, 4, 20, 12, 28, 18, 26, 2, 7, 6, 22, 14, 30, 1, 9, 25, 24, 5, 21, 13, 29, 3, 19, 11, 27, 23, 15, 31, 10)
\]

\[
V_{5,2} = (29, 10, 31, 27, 21, 17, 8, 22, 30, 6, 9, 12, 1, 13, 14, 26, 19, 24, 5, 16, 28, 2, 7, 15, 25, 3, 20, 23, 18, 4, 0, 11)
\]

- For the identity mapping we find \(I_L = C_G\)
- For other mappings: \(I_L\) increases with \(L\), \(I_{L-1} < I_L\), \(1 \leq L < M\)
- Chain rule: the \(I_L\) sum up to constant \(M \cdot C_G\) for all mappings
SCC - Inner Transfer Characteristics

channel \( I_{E',L} \) and \( I_{E',BEC} \) at output of demapper

(a priori information \( I_{A,L} \) and \( I_{A,BEC} \) at input of demapper)

discrete points \((I_{A,L}, I_{E',L})\) for mapping \( V_{5,id} \)

\((I_{A,L}, I_{E',L})\) for mapping \( V_{5,1} \)

\((I_{A,L}, I_{E',L})\) for mapping \( V_{5,2} \)

continuous transfer characteristics \( I_{E',BEC} = T(I_{A,BEC}) \) for mappings \( V_{5,id}, V_{5,1}, \) and \( V_{5,2} \)

One of \( 5 \)\^{th} \( (a/4, 0/4) \)

\( (1/4, I_{1}) \)

\( (1/2, I_{2}) \)

\( (3/4, I_{3}) \)

\( (1, I_{4}) \)

\( \frac{1}{4} = \frac{1}{4} \)
Gaussian (Continuous) A Priori Knowledge:

- From simulations of actual decoder: Extrinsic L-values $E_2$ tend to be Gaussian-like distributed.

\[
E_2 \sim N(0, \sigma_A^2)
\]

- Model a priori input $A$ as an independent Gaussian random variable $n_A$ with variance $\sigma_A^2$ and mean zero.
- In conjunction with the known transmitted inner information bits $\tilde{u} = 1 - 2u$, $\tilde{u} \in \{\pm 1\}$, we write

\[
A = \frac{2}{\tilde{\sigma}_A^2} \cdot (\tilde{u} + \tilde{n}_A) = \mu_A \cdot \tilde{u} + n_A
\]

with $\mu_A = 2/\tilde{\sigma}_A^2$ and $\sigma_A^2 = 4/\tilde{\sigma}_A^2$.
SCC - Inner Transfer Characteristics

- Transfer characteristics for Gaussian distributed a priori knowledge
- Note: $I_{E'}(0)$ and $I_{E'}(1)$ independent of a priori distribution
**SCC - Inner Transfer Characteristics**

- Extrinsic transfer characteristics of some inner rate 1 codes
- Note: Non-recursive codes do not go up to $I_{E_1}(1) \approx 1$
SCC - Outer Transfer Characteristic

- In the same way as for inner codes, we can derive transfer characteristics of the outer code.
- We consider the a priori information $I_{A_2} = I(C_2; A_2)$, and the extrinsic information $I_{E_2} = I(C_2; E_2)$ of the decoder output, to get the transfer characteristic:

$$I_{E_2} = T_2(I_{A_2})$$

- For the computation we assume $A_2$ to be Gaussian distributed, and measure histograms $p_{E_2}(\xi | C_2 = 0), p_{E_2}(\xi | C_2 = 1)$.
- The outer transfer characteristics are independent of the $E_b/N_0$-value.
- Note that the axes are swapped: Input $I_{A_2}$ is on ordinate, output $I_{E_2}$ on abscissa.
- This is in preparation of the design tool where we connect both inner and outer transfer characteristic in a single diagram (EXIT chart).
Extrinsic transfer characteristics of some outer rate 1/2 codes

Note: Rate 1/2 repetition code is just diagonal line $I_{E_2} = I_{A_2}$
Notes

A first EXIT chart:
SCC - Extrinsic Information Transfer Chart

- To account for the iterative nature of the sub-optimal decoding algorithm, both decoder characteristics are plotted into a single diagram; for second decoder, axes are swapped.

- This diagram is referred to as extrinsic information transfer chart (EXIT chart) since the exchange of extrinsic information can be visualized as a decoding trajectory.

- Provided that independence (large interleaver) and Gaussian assumptions hold for modelling extrinsic information (a priori information respectively), the decoding trajectory that can be graphically obtained by simply drawing a zigzag-path into the EXIT chart (bounded by the decoder transfer characteristics) should match with the trajectory computed by simulations.

- For simulations (next), interleaver size $4 \cdot 10^5$ coded bits used.
- $E_b/N_0 = 0.6\text{dB}$, trajectory gets stuck
- $E_b/N_0 = 1.1$dB, convergence to low BER through narrow tunnel
• $E_b/N_0 = 2.5\text{dB}$, convergence tunnel wide open
For $E_b/N_0 = 0.6$ dB the trajectory gets stuck

For $E_b/N_0 = 1.1$ dB the inner transfer characteristic has been raised just high enough to open a narrow tunnel (“bottleneck”) for the trajectory to “sneak through” and to converge towards low BER ($\approx 10^{-6}$, depending on interleaver size)

At $E_b/N_0 = 2.5$ dB, less iterations are needed to get down to low BER

For short interleavers the trajectory tends to diverge from the characteristics towards smaller extrinsic output after a few iterations, owing to increasing correlation of extrinsic information

Main advantage of EXIT chart:

- Only simulations of individual component decoders are required
- Transfer characteristics can be used in any combination
SCC - BER from EXIT Chart

- Note: BER contour lines are independent of the $E_b/N_0$-value
SCC - Code Design Examples

**Pinch-off 0.41dB**
- Outer repetition codes and systematic doping yields early turbo cliffs

**Pinch-off 0.27dB**
Outline

1 Introduction
2 Soft Output Decoding
3 Serially Concatenated Codes (SCC)
4 Parallel Concatenated Codes (PCC)
5 Low-Density Parity Check (LDPC) Codes
6 Iterative Detection
7 Future Trends
PCC - Encoder and Iterative Decoder

1st encoder
1st APP decoder
1st encoder

2nd encoder
2nd APP decoder
2nd encoder

AWGN
n

systematic (information) bits

u

binary source

systematic bits and 1st parity bits

systematic bits and 2nd parity bits

systematic (information) bits

u

AWGN
n

systematic bits and 1st parity bits

systematic bits and 2nd parity bits

systematic bits and 1st parity bits

systematic bits and 2nd parity bits

hard decision

binary sink

(uncoded) information bits
coded bits

with respect to APP processing block

S. ten Brink Iterative Detection and Decoding 24.04.2015 70/118
Iterative “turbo” decoding was introduced by C. Berrou, A. Glavieux in 1993.

First scheme to make use of **extrinsic** information.

Typically, **PCC are systematic**.

The **classic turbo codes** of 1993 consist of memory 4 constituent codes with polynomials $(G_r, G) = (037, 021)$. 
- Rate 2/3 constituent codes (for rate 1/2 PCC)
• Different code memory (left), different code polynomials for memory fixed to 4 (right); $E_b/N_0 = 0.8$dB
PCC - EXIT Chart

- PCC rate $1/2$, memory 4, $(G_r, G) = (023, 037)$; interleaver $10^6$ bits
PCC - Code Design Examples

- Decoding trajectories for rate 1/2 PCC with turbo cliff below 0.5dB
PCC - Webdemo

- for your edutainment, turbo code webdemo
  - [http://webdemo.inue.uni-stuttgart.de/webdemos/03_theses/turboCodingM/](http://webdemo.inue.uni-stuttgart.de/webdemos/03_theses/turboCodingM/)
Outline

1. Introduction
2. Soft Output Decoding
3. Serially Concatenated Codes (SCC)
4. Parallel Concatenated Codes (PCC)
5. Low-Density Parity Check (LDPC) Codes
6. Iterative Detection
7. Future Trends
LDPCC - Definitions

- Invented in 60’s, regained popularity in 1997 (irregular LDPCC)
- A \((d_v, d_c)\)-regular LDPC code is a binary linear block code that has a parity-check matrix \(H\)
  - with \(d_v\) ones in each column
  - with \(d_c\) ones in each row
- Example: \((d_v = 2, d_c = 4)\)-regular LDPC code
- Described by \((N - K) \times N\) matrix \(H\) (parity-check matrix) \(Hc^T = 0\)

\[
H = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}
\]

- \(H\) is sparse
LDPCC - Irregular Code

- Example of an irregular code; parity check matrix

\[
H = \begin{pmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 1
\end{pmatrix}
\]
LDPCC - Iterative Decoder

- Variable node decoder: viewed as inner code, typically irregular
- Check node decoder: viewed as outer code, typically regular
LDPCC - Component Decoders

rate 1/4 repetition code (degree 4)

\[ L_i = L_{ch} + \sum_{j \neq i} L_j \]

Add Rep

rate 3/4 single parity check code (degree 4)

\[ L_i = \sum_{j \neq i} \oplus L_j \approx \prod_{j \neq i} sgn L_j \cdot \min \left( \left| L_j \right| \right) \]

weakest link dominates

"soft majority vote"
LDPCC - VND EXIT curves

- repetition codes of rate $1/d_v$ (here: $E_b/N_0 = 1$dB at $R = 1/2$)
LDPCC - CND EXIT curves

- single parity check codes of rate \( \frac{(d_c - 1)}{d_c} \)
LDPCC - Superposition of EXIT Curves

- Code design by mixing nodes of different degree (irregular LDPCC)
- Transfer characteristic of resulting curve is a linear combination of individual curves of respective degrees
- E.g., variable node curve (mixture of \(D\) different degrees)

\[
I_{E,VND}(I_A) = \sum_{i=1}^{D} b_i I_{E,VND}(I_A, d_{v,i})
\]

with edge fractions \(b_i\),

\[
\sum_{i=1}^{D} b_i = 1
\]

- LDPC code design boils down to curve fitting in EXIT chart!
Example of a regular LDPC code over AWGN channel

- Convergence to low bit error rate at 1.3dB
- Example of an irregular LDPC code over AWGN channel
- Convergence to low bit error rate at 0.6dB (issue: degree 2 variables...)
LDPCC - Webdemo

- for your edutainment, LDPC code webdemo

- [http://webdemo.inue.uni-stuttgart.de/webdemos/03_theses/ldpcExit/]
Notes

Random linear $\mathbf{A} \sim O(N^2)$ encoding complexity

$H = \mathbb{R}^n$ in practice

Coding/Decoding complexity $\sim \Theta(n)$

Repeat accumulate cod
Outline

1. Introduction
2. Soft Output Decoding
3. Serially Concatenated Codes (SCC)
4. Parallel Concatenated Codes (PCC)
5. Low-Density Parity Check (LDPC) Codes
6. Iterative Detection
7. Future Trends
Iterative Detection - Basic Structures

(a) MIMO map \( s \) \( x_1 \) \( \Pi \) \( y_2 \) ENC \( x_2 \) source

(b) MIMO map \( s \) \( x_1 \) \( \Pi \) ENC1 \( \Pi_1 \) ENC2

(c) MIMO map \( s \) \( x_1 \) DEC1 \( \Pi_1 \) DEC2

AWGN

\( H \)

\( n \)

S. ten Brink Iterative Detection and Decoding 24.04.2015 94/118
Iterative Detection - Inner MIMO Detector

\[ E_b/N_0 = 3 \text{dB}, \text{Gray mapping, QPSK} \]

- 1x1
- 2x1
- 4x1

\[ E_b/N_0 = 3 \text{dB}, \text{Gray mapping, QPSK} \]

- 1x1
- 2x2
- 4x4

- 2x2, 64QAM
- 2x2, 16QAM

\[ E_b/N_0 = 3 \text{dB}, \text{Gray mapping, QPSK} \]

- 4x8, 3dB
- 4x8, 4dB
- 4x8, 5dB

\[ E_b/N_0 = 3 \text{dB}, \text{Gray mapping, QPSK} \]

- 4x2, 3dB
- 4x4, 3dB
- 4x8, 3dB

\[ E_b/N_0 = 3 \text{dB}, \text{Gray mapping, QPSK} \]
Iterative Detection - EXIT Chart

- 4x4, QPSK MIMO detector $E_b/N_0 = 2\text{dB}$
- 4x1, QPSK MIMO detector $E_b/N_0 = 9\text{dB}$

outer rate 1/2 memory 2 PCC (8 internal iterations)
outer rate 1/2 memory 2 convolutional code
Iterative Detection - Outer LDPC Codes

- irregular vs regular LDPC codes, $4 \times 2$ MIMO, $E_b/N_0 = 3.3$dB
Current fitting: find best matching mixture of RFP code

Channel input

RFP

WD

SPC

CMD

New matching

Iterative detection & LDPC decoding
Outline

1. Introduction
2. Soft Output Decoding
3. Serially Concatenated Codes (SCC)
4. Parallel Concatenated Codes (PCC)
5. Low-Density Parity Check (LDPC) Codes
6. Iterative Detection
7. Future Trends
Future Trends

- recently, it was shown that spatially coupled codes (e.g. convolutional LDPC codes) are “universal”
  - can universally achieve capacity over binary-input memoryless symmetric-output channels
  - good for various detection front-ends...
- specific degree distribution design becomes obsolete
  - 1960s: Gallager, LDPC codes, regular
  - 1990s: MacKay, Richardson, Urbanke et al: make them irregular to get close to capacity
  - 2010s: Kudekar, Richardson, Urbanke, universality of spatial coupled (LDPC) codes: regular codes suffice

→ back to regular... (?)
Future Trends - Spatial Coupling

- Degree profile optimization for dedicated channel detector, e.g.
  - MIMO detector (multiple antennas)
  - equalizer (multipath channel)
  - QAM mapping (QPSK, 16QAM…)
  - differential coding

- For each detector (or each channel), different degree profile needed!

- Thus, with spatially coupled codes as “universally good codes”...
  - should be no matching to channel interface needed anymore!
Future Trends - Spatial Coupling

- Pariy check matrix a terminated convolutional LDPC code,

\[ H_L = \begin{pmatrix} H_0 & H_0 \\ H_1 & H_1 \\ \vdots & \vdots \\ H_m & H_m \end{pmatrix} \]

\[ (L+m)M' \times LN' \]

- slight irregularities (in check node degrees) kick off decoding wave below BP threshold (but above MAP threshold)
- Regular (3,6) code, extended into an conv. LDPC code $L = 100$;
  $\frac{E_b}{N_0} = 0.85\text{dB}$
Future Trends - Spatial Coupling

- Tunneling through the pinch-off... against conventional wisdom; better than BP threshold

["Modulation/Detection with Spatially Coupled Codes", L. Schmalen, tB, IEEE/ITG Conf. on. SCC, Jan. 2013]
Future Trends - Spatial Coupling, Detection Experiment

- experiment: spat. coupled code with BICM-detector
- three different labelings (16-QAM)
  - Gray labeling (most flat detector EXIT curve)
  - set partitioning (SP)
Future Trends - Spatial Coupling, Detection Experiment

- density evolution on protograph; replication factor up to $L = 100$
- capacity of 16-QAM with rate $R \rightarrow 3/4$-code at $E_b/N_0 \approx 4.53$dB

["Modulation/Detection with Spatially Coupled Codes", L. Schmalen, tB, IEEE/ITG Conf. on. SCC, Jan. 2013]
Future Trends - Spatial Coupling, Detection Experiment

DE thresholds for the spatially coupled and the corresponding regular ensembles and optimized irregular ensembles for BICM-ID; Capacity limit $E_{b}/N_{0 \min} = 4.528$ dB

<table>
<thead>
<tr>
<th>Code</th>
<th>Gray</th>
<th>SP</th>
<th>M16a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatially coupled code</td>
<td>4.712</td>
<td>4.577</td>
<td>4.745</td>
</tr>
<tr>
<td>Opt. irregular (for Gray)</td>
<td>4.600</td>
<td>6.217</td>
<td>6.344</td>
</tr>
<tr>
<td>Opt. irregular (for SP)</td>
<td>&gt; 7 dB</td>
<td>4.739</td>
<td>4.983</td>
</tr>
<tr>
<td>Opt. irregular (for M16a)</td>
<td>&gt; 7 dB</td>
<td>4.729</td>
<td>4.959</td>
</tr>
</tbody>
</table>

- simulation results: $L = 50$, codeword length $N = 200000$ bits, 16-QAM with rate $R \rightarrow 3/4$-code
- spatially coupled codes (regular) can be universally good
  - avoids degree 2 variable nodes, ...

["Modulation/Detection with Spatially Coupled Codes", L. Schmalen, tB, IEEE/ITG Conf. on. SCC, Jan. 2013]
Summary

- Iterative decoding: to approach capacity
- Iterative detection and decoding: include channel interface in iterative decoding loop
- LDPC codes: Degree profile matching
- Spatially Coupled Codes: Can be universally good, regular codes suffice
- Most communication problems in practice can benefit from iterative detection and decoding

Thank you! It was fun!
Some References (biased, incomplete) - 1

Some References (biased, incomplete) - 2

- L. Schmalen, S. ten Brink, "Combining Spatially Coupled LDPC Codes with Modulation and Detection", 9th International ITG Conference on Systems, Communications and Coding (SCC), 2013, pp. 1 - 6