Communications Research Centre Canada
An Agency of Industry Canada

Centre de recherches sur les communications Canada
Un organisme d'Industrie Canada

Information Theory, Shannon Limit and Recent Advances in Error Correction Codes for Terrestrial DTV Broadcasting

Yiyan Wu, Bo Rong, and Gilles Gagnon
Communications Research Centre Canada

ATSC 2nd Symposium on Next Generation Broadcasting TV
Rancho Las Palmas, Rancho Mirage, CA
February 15, 2011

Presentation Outline

- Information Theory and Shannon Limit
- Error correction code history
- Error correction code performance BER, S/N & E_b/N_o
- Convolutional code, block code, concatenated code
- Turbo code vs. LDPC code
- Iterative Decoding and Irreducible Error-Floor
- Space-Time Coding
- Joint Source and Channel Coding
- Cross-Layer Forward Error Correction Coding
- Data Rate vs. Robustness

Information Theory

History & Definition …

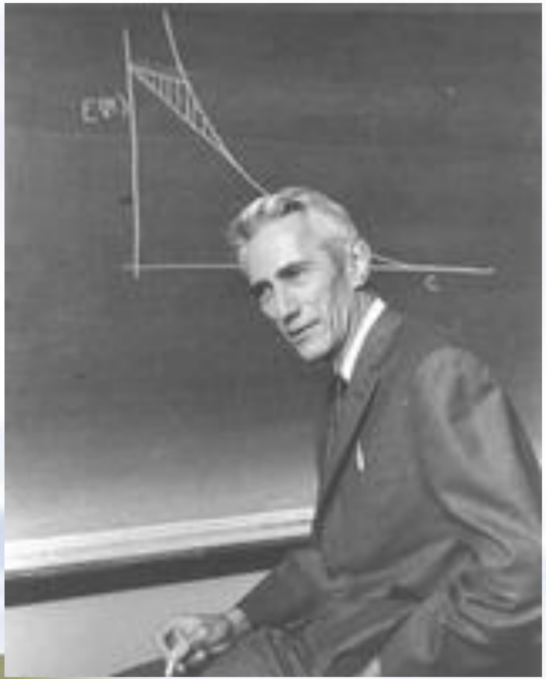
History: The publication of Claude Shannon's 1948 paper entitled "A Mathematical Theory of Communication".

Definition: Information theory is the **science of compression, storage, and transmission** of information. In other words, it was developed to find fundamental limits on signal processing operations such as data compression and on reliably storage and communicating data.

Shannon Limit

Who is Shannon

- Claude Elwood Shannon (April 30, 1916 – February 24, 2001) was an American mathematician, electronic engineer, and cryptographer known as "the father of information theory".
- He taught at MIT from 1956 until his retirement in 1978,



Shannon Limit

What is Shannon Limit

Shannon showed that any communications channel — a telephone line, a radio band, a fiber-optic cable — could be characterized by two factors: **bandwidth and noise**.

- Bandwidth is the range of electronic, optical or electromagnetic frequencies that can be used to transmit a signal;
- Noise is anything that can disturb that signal.

Given a channel with particular bandwidth and noise characteristics, Shannon showed how to calculate the maximum rate at which data can be sent without error. He called that rate the **channel capacity.** But people always call it — **Shannon limit**.

Shannon Limit

Channel Capacity: The maximum data rate at which the error-free communication over the channel is performed.

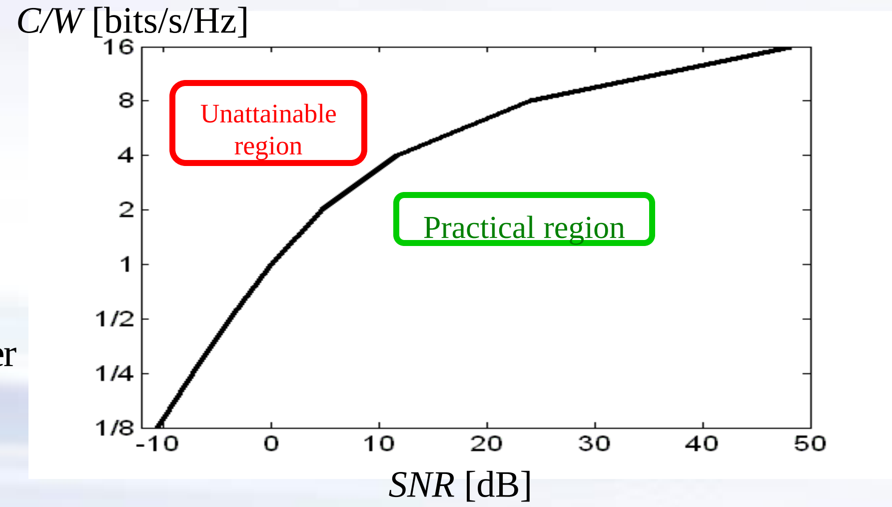
Channel capacity on AWGN channel (Shannon-Hartley capacity theorem):

$$C = W \log_2\left(1 + \frac{S}{N}\right) \quad \text{[bits/s]}$$

W [Hz]: Bandwidth

$S = E_b \cdot R_b$ [Watt]: Average received signal power

$N = N_0 \cdot BW$ [Watt]: Average noise power



Error Correction Schemes

There are two methods to correct the transmission errors:

1. Error detection and retransmission:
   – Good for point-to-point two-way communications.
   – Return channel and retransmission (Automatic Repeat reQuest-ARQ) needed;
2. Forward Error Correction (FEC)
   – Addition parity bits are transmitted, which can be used to correct the transmission errors in the receiving-end.

The broadcasting service are point-to-multi-point real time services.

Forward Error Correction is suited for broadcasting applications.

The rest of the presentation will focus on FEC codes.

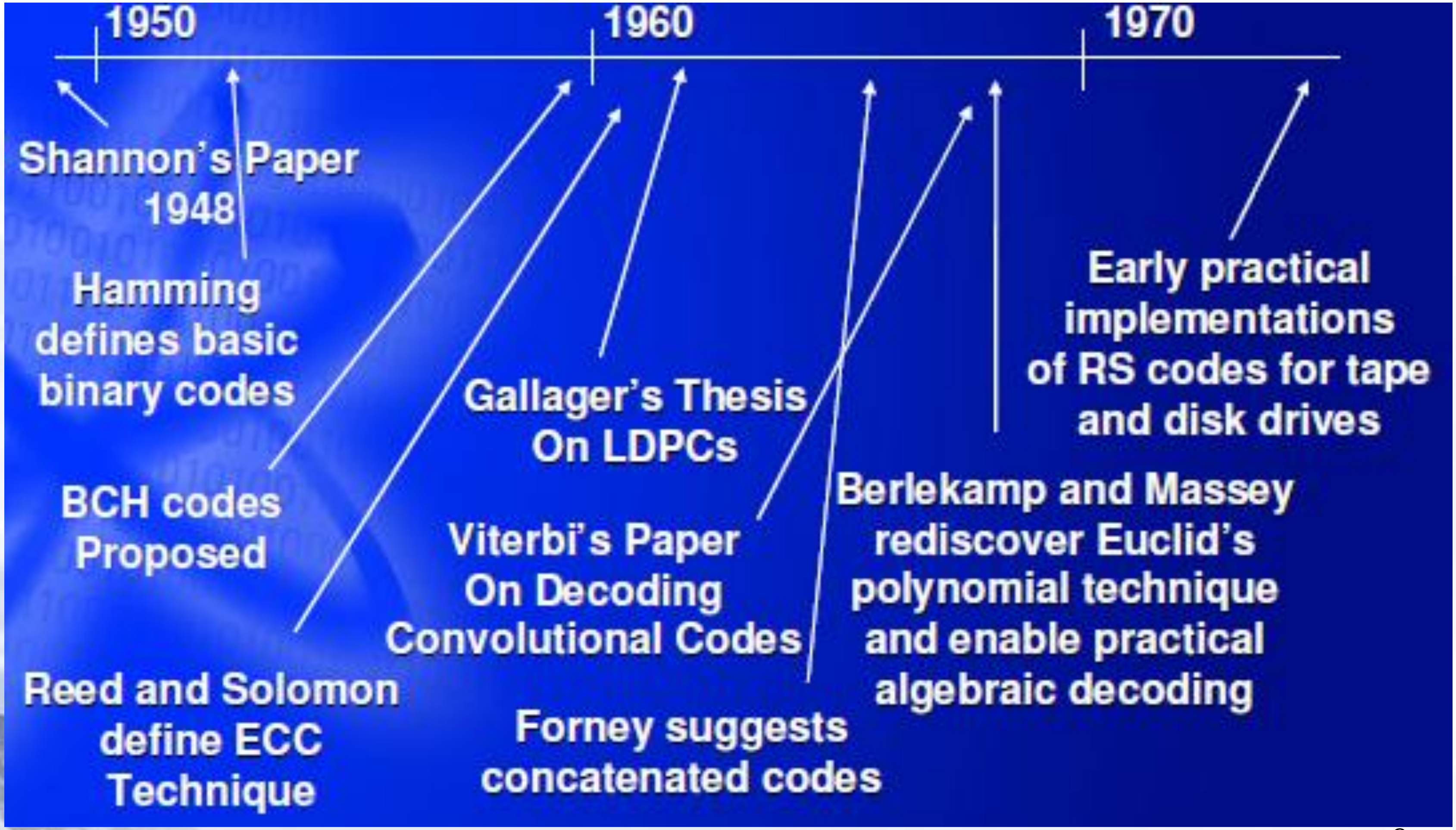
Error Correction Codes

There are two types of error correction codes:

1. Convolutional codes and Trellis Coded Modulation (TCM: combines convolutional code and modulation);

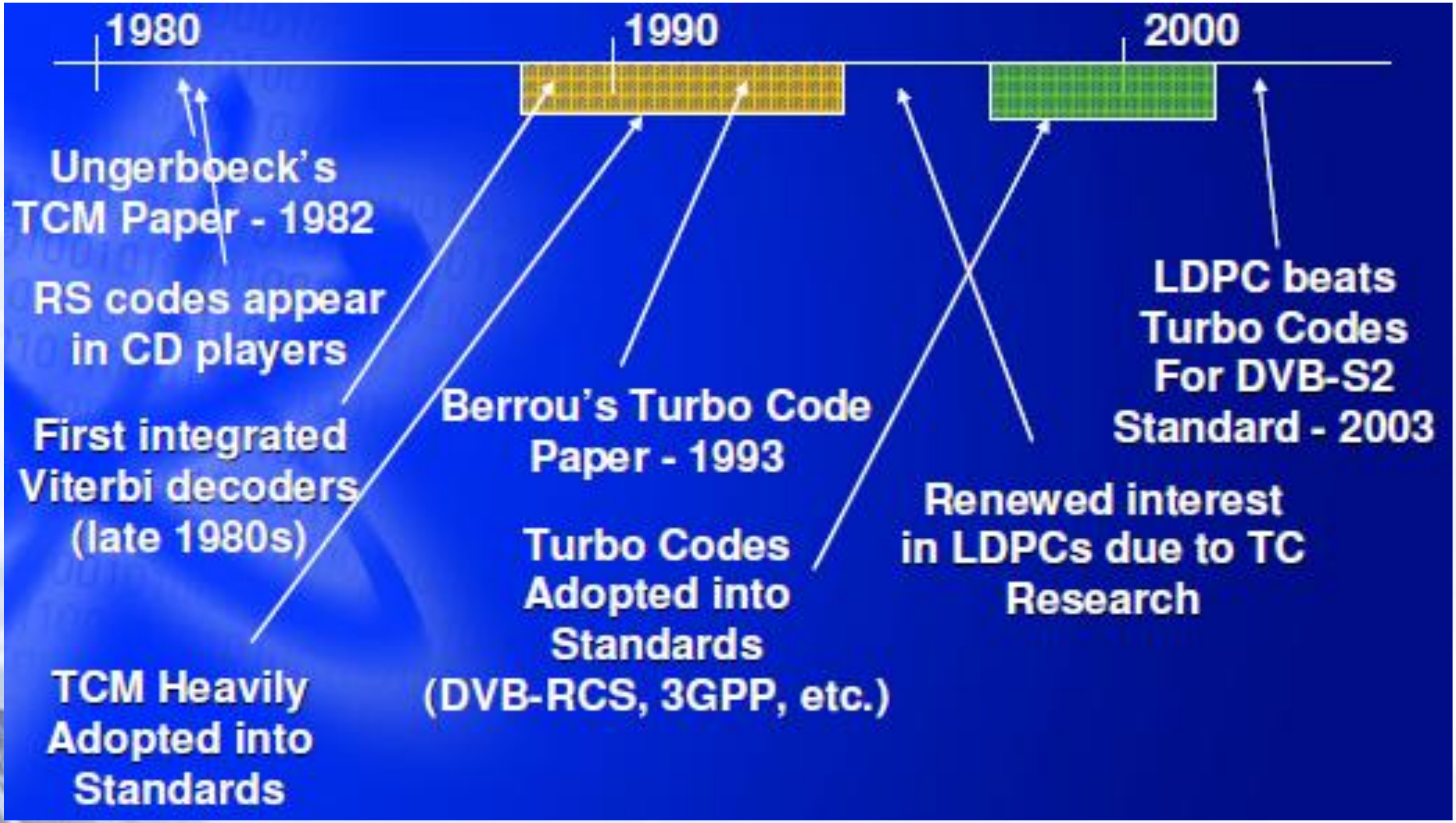
2. Block codes: BCH, Hamming, Reed-Solomon, Turbo, Turbo Product, LDPC, fountain codes and BICM codes (BICM combines LDPC code and modulation). Essentially all iteratively-decoded codes are block codes.

Error Correction Codes History



Source: Intel tutorial

Error Correction Codes History



Source: Intel tutorial

Block Codes

A block code is any code defined with a finite codeword length:

– The code rate **R = k/n**, where

  - k is the lengths of useful information bits

  - n is the total length codeword.

– If the codeword is constructed by appending parity symbols to the payload Data Field, it is called a **"systematic" code**.

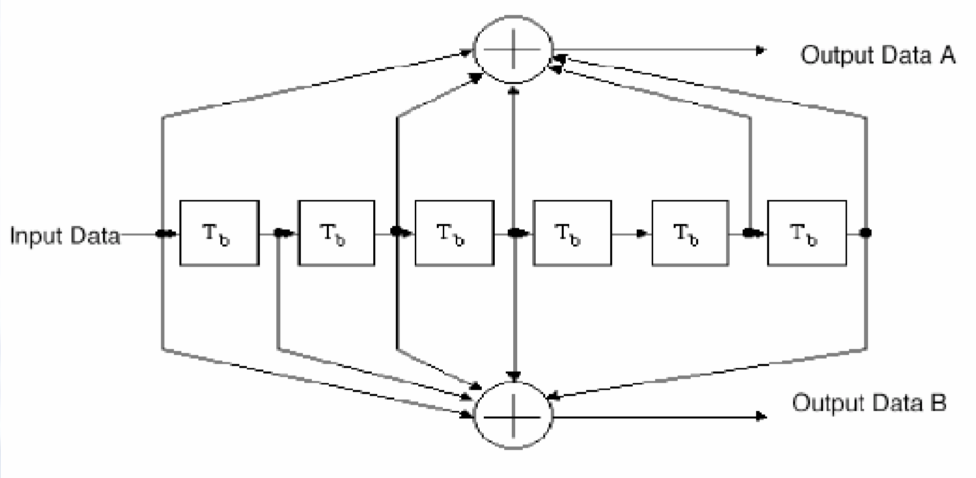
– Examples of block codes: BCH, Hamming, Reed-Solomon, Turbo, Turbo Product, LDPC.

– Essentially all iteratively-decoded codes are block codes.

Convolutional Codes

Convolutional codes are generated using a shift register to apply a polynomial to a stream of data:

- The convolutional code can be systematic, if the data is transmitted in addition to the redundancy bits, but it often isn't.

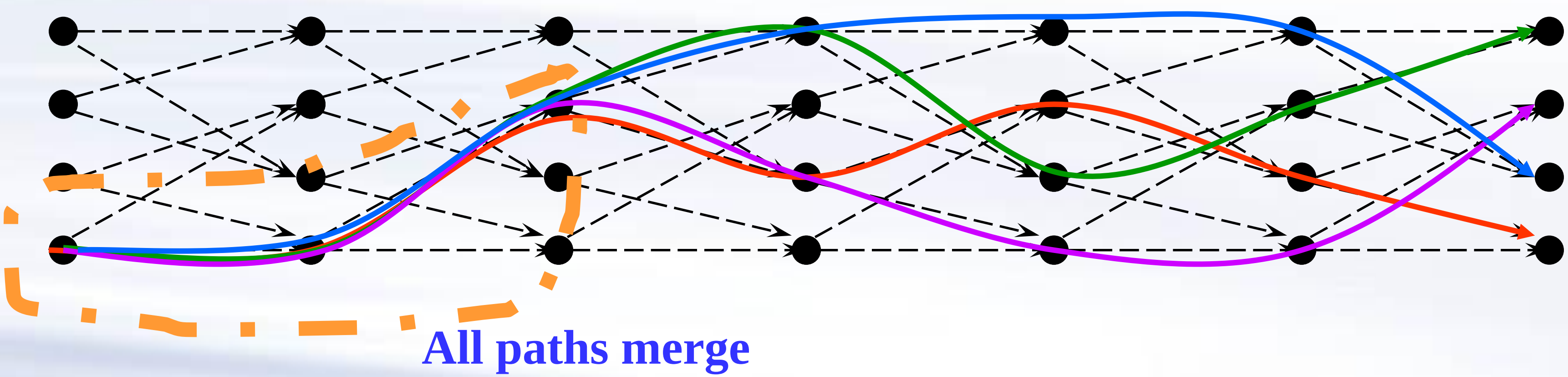


- The code above is naturally R = ½. By deleting selected output bits, or "**puncturing**" the code, the code rate can be increased to 2/3, 3/4, 7/8…...

Convolutional Codes

Convolutional codes are typically decoded using the Viterbi algorithm

- The Viterbi decoder is a Maximum Likelihood Sequence Estimator, that estimates the encoder state using the sequence of transmitted codewords.

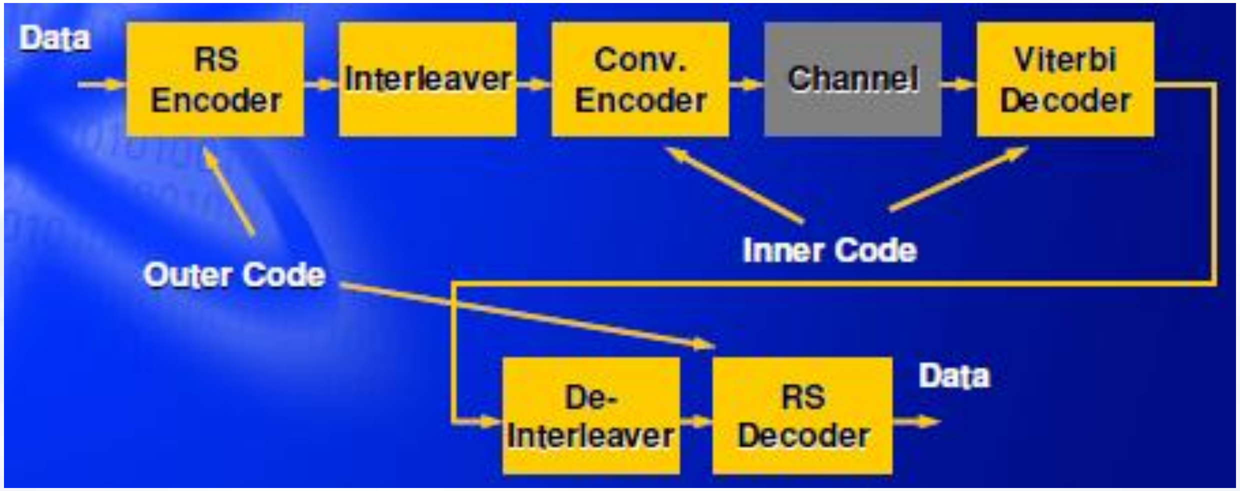


- Pros: Viterbi algorithm works well under noisy and fading channel;
- Cons: when making a mistake, it can lose track of the sequence and generate burst errors until it reestablishes code lock.

Concatenated Codes

Concatenation of an inner convolutional code with an outer block code

– A very common and effective coding structure as shown below:



– The convolutional code works better in low SNR range, but its BER curve roll-off slowly and can have irreducible error floor in fading channel.

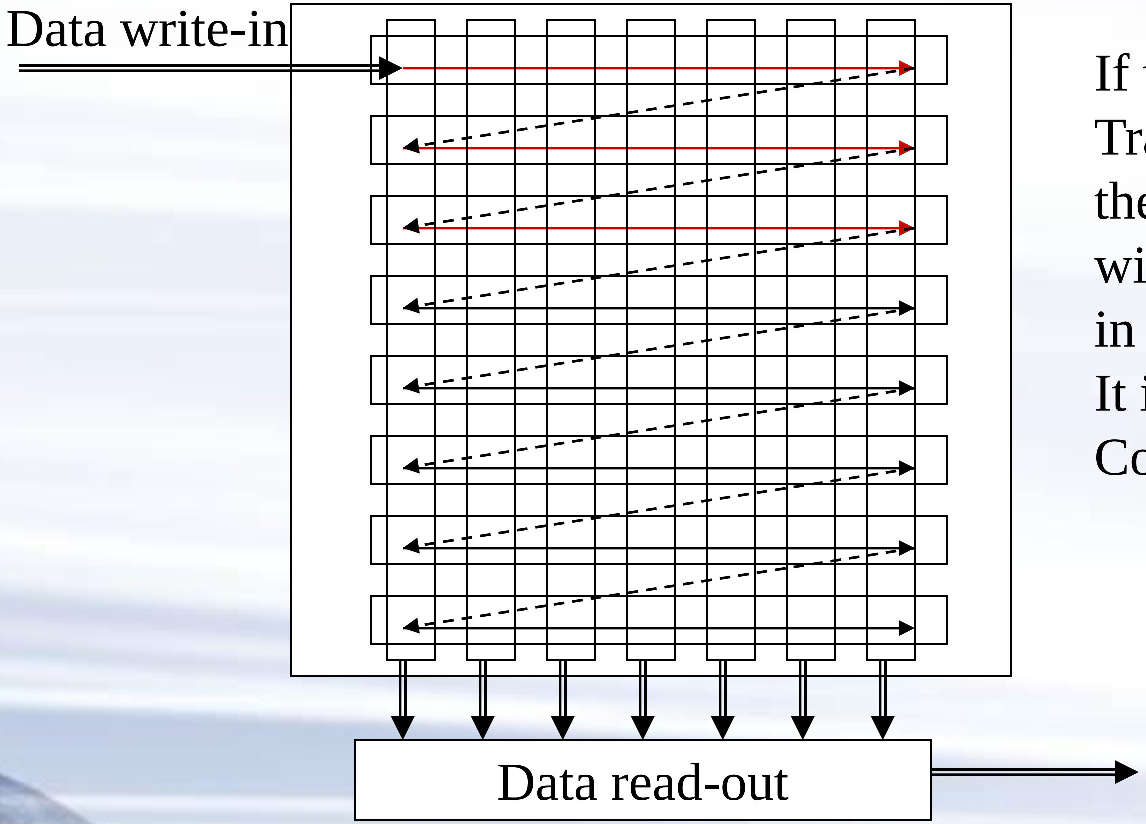
– The Reed-Solomon code has fast BER roll-off and well suited to correct the bursty output errors common with a Viterbi decoder.

– An interleaver can be used to spread the Viterbi output error bursts across multiple R-S codewords for burst error correction.

What is an Interleaver

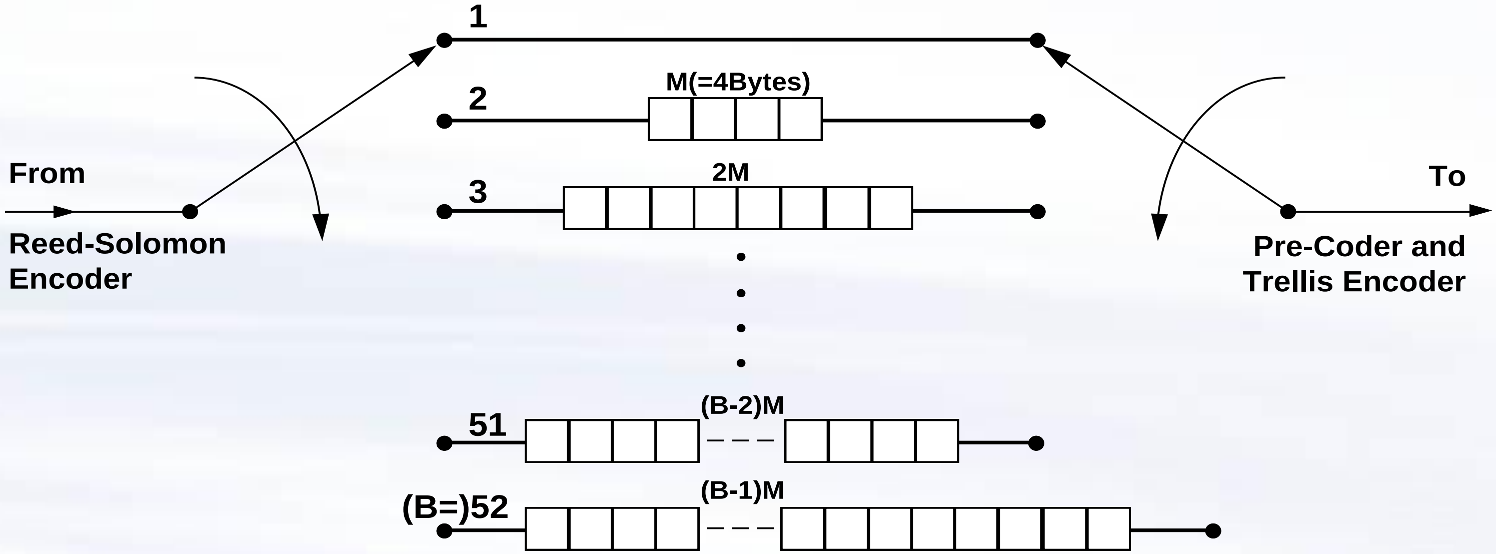
- Interleaver and de-interleaver are a pair of devices that can break up (or de-correlate) the long burst of error symbols, occurred in transmission or generated by inner decoder, so that they can be easily corrected by error correction code;
- The simplest interleaver is an “orthogonal memory”, which read-in data symbol line-by-line and read-out column by column;
- There are different type of interleavers, such as, covolutional interleave, block interleaver, random interleaver, etc.

A Simple Block Interleaver



If there are three lines of
Transmission errors,
the interleaver out put
will only have 3 symbols
in consecutive errors.
It is easier for decoder to
Correct them.

Convolutional Interleaver in ATSC A/53



M=4, B=52, N=208, R-S Block =207, BXM=N

Code Performance Measurement

- Bit Error Rate (BER), Symbol Error Rate (SER), Packet Error Rate (PER), or Block Error Rate;
- Signal to Noise Ratio (SNR) or E_b/N_o, where
  - E_b is energy per bit,
  - N_o is power spectrum density (power per Hertz);
- BER vs S/N or E_b/N_o curve is the most important performance measurement;
- Other measurement: complexity, delay, etc.

S/N vs. E_b/N_o

E_b is energy per bit
R_b is bit rate [bits/s]
BW is channel bandwidth
N_o is the noise spectrum density

$$\frac{S}{N} = \frac{E_b \cdot R_b}{N_0 \cdot BW} = \frac{E_b}{N_0} \cdot \frac{R_b}{BW}$$

$$\frac{S}{N}(dB) = \frac{E_b}{N_0}(dB) + 10 \cdot \log\left(\frac{R_b}{BW}\right)$$

$$\frac{E_b}{N_0}(dB) = \frac{S}{N}(dB) - 10 \cdot \log\left(\frac{R_b}{BW}\right)$$

E_b/N_o is independent of data rate and bandwidth!

S/N vs. E_b/N_o

$$\frac{E_b}{N_0}(dB) = \frac{S}{N}(dB) - 10 \cdot \log\left(\frac{R_b}{BW}\right)$$

Example: for ATSC system,
S/N = 15 dB, BW = 6 MHz, R_b = 19.4 Mbps,
E_b/N_o = 15 – 10 log(19.4/6) = 9.9 dB

When say S/N, system bandwidth MUST be specified.
For North American TV industry, BW is always 6 MHz.
For telecom industry, BW is always 3dB bandwidth.
ATSC system 3dB BW = 5.38 MHz, system S/N = 15.5 dB
ATSC channel BW = 6.0 MHz, system S/N = 15.0 dB.

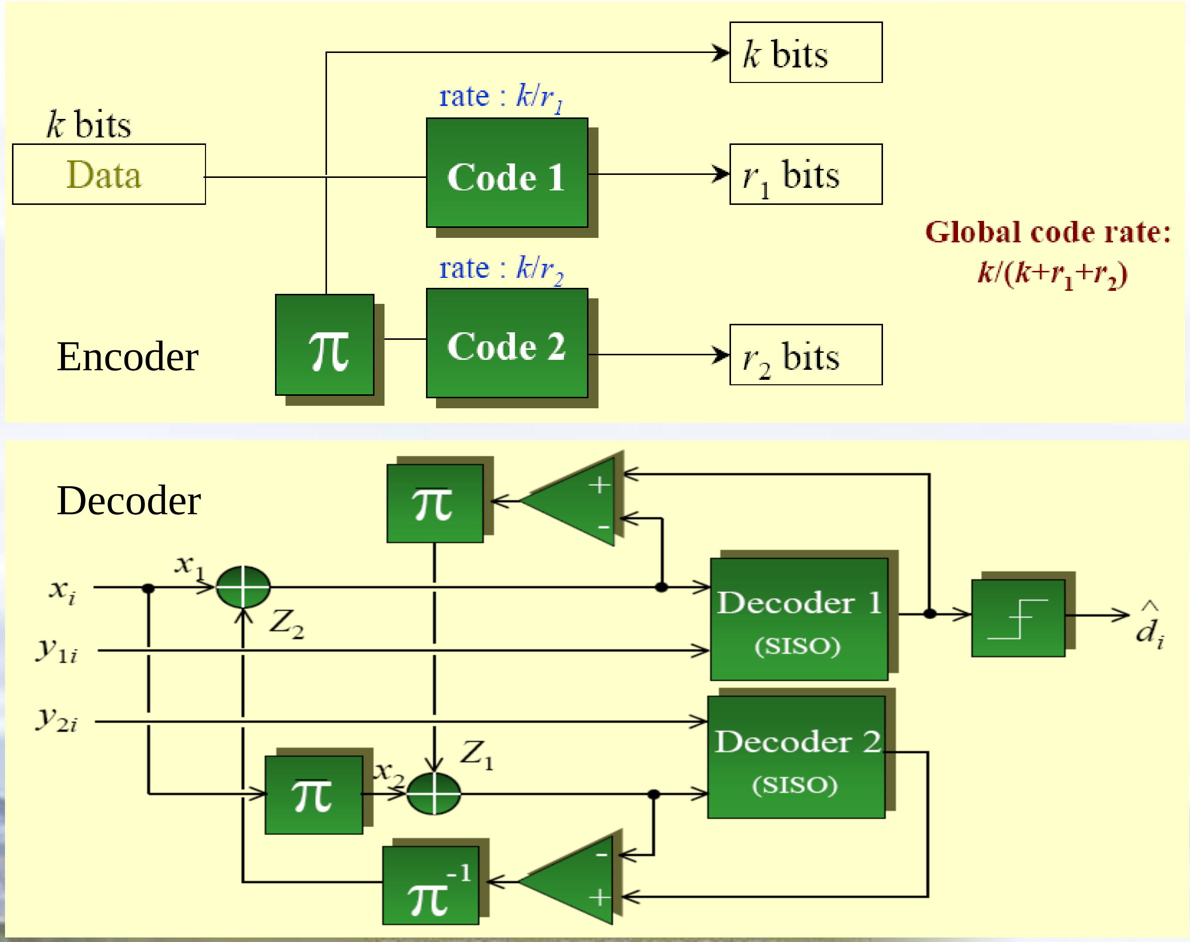
E_b/N_o is bandwidth and bit rate independent!!!

Turbo Codes

- History
  - Introduced by Berrou, Glavieux, and Thitimajshima (from Telecom-Bretagne, former ENST Bretagne, France) in their paper: "*Near Shannon Limit Error-correcting Coding and Decoding: Turbo-codes*" published in the Proceedings of IEEE International Communications Conference (ICC'1993).
  - Turbo-codes are also called Parallel (or Serial) Concatenated Convolutional Code (PCCC or SCCC).
  - Since the introduction of the original parallel turbo codes (PCCC) and iterative decoding algorithm in 1993, many other classes of turbo code have been discovered, including serial versions (SCCC) and repeat-accumulate (RA) codes.
  - Iterative Turbo decoding methods have also been applied to more conventional FEC systems, including Reed-Solomon/convolutional code structure.

Turbo Codes

- Encoder-Decoder Structure

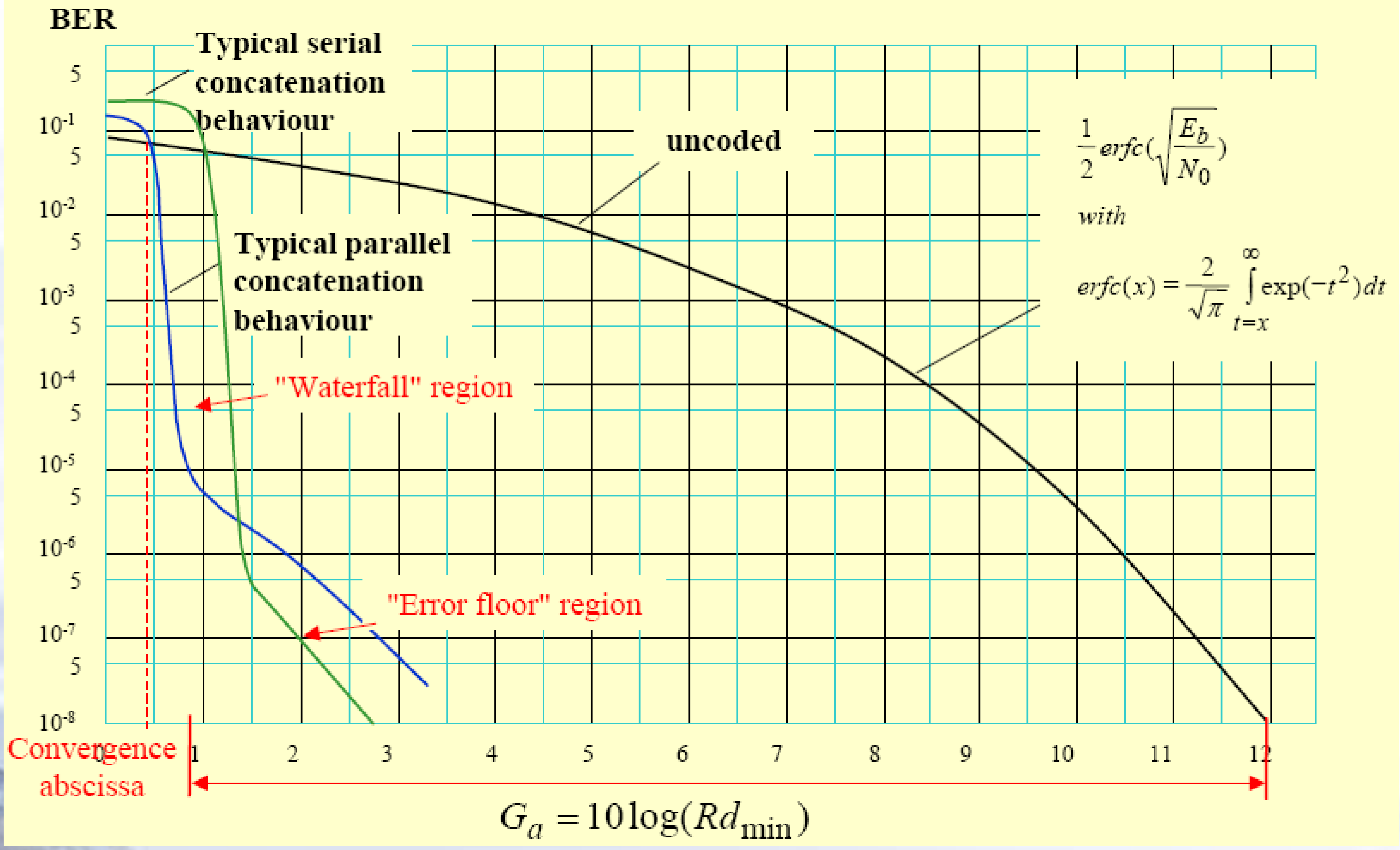


PCCC:
Parallel
Concatenated
Convolutional
Code.

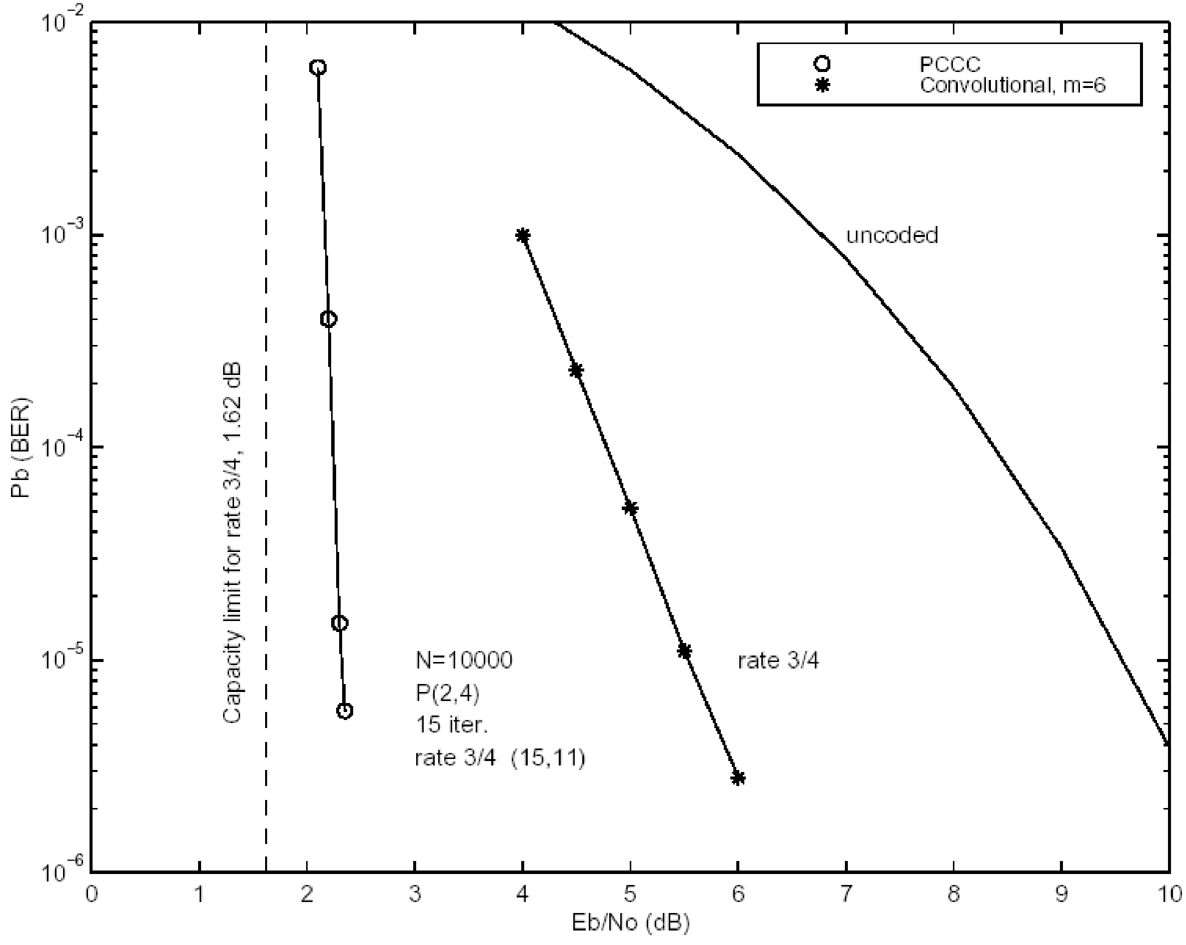
π: is an interleaver

Turbo Codes

Performance (BER vs E_b/N_o, PCCC vs SCCC)



Turbo Codes



Rate 3/4 PCCC bit error rate performance.

Turbo Codes

- Turbo codes have good performance in low coding rate, $R < 0.5$, cases. It is often used in mobile communication systems for low SNR fading channel (3GPP, LTE, etc.);

- The existence of "error-floor" requires an outer code (R-S or BCH code) to be used for error-free transmission.

LDPC Codes

- History

The Low-Density Parity-Check (LDPC) code concept was invented by Robert G. Gallager in his doctoral dissertation at MIT in 1963. Deemed impractical to implement when first developed (that was pre-silicon and microprocessor days), LDPC codes were forgotten, but they were re-discovered in 1996.

In the last few years, the advances in LDPC codes have surpassed turbo codes in terms of error floor and performance in the higher code rate range, leaving turbo codes better suited for the lower code rates (mobile communications).

LDPC Codes

- Parity Check Matrix

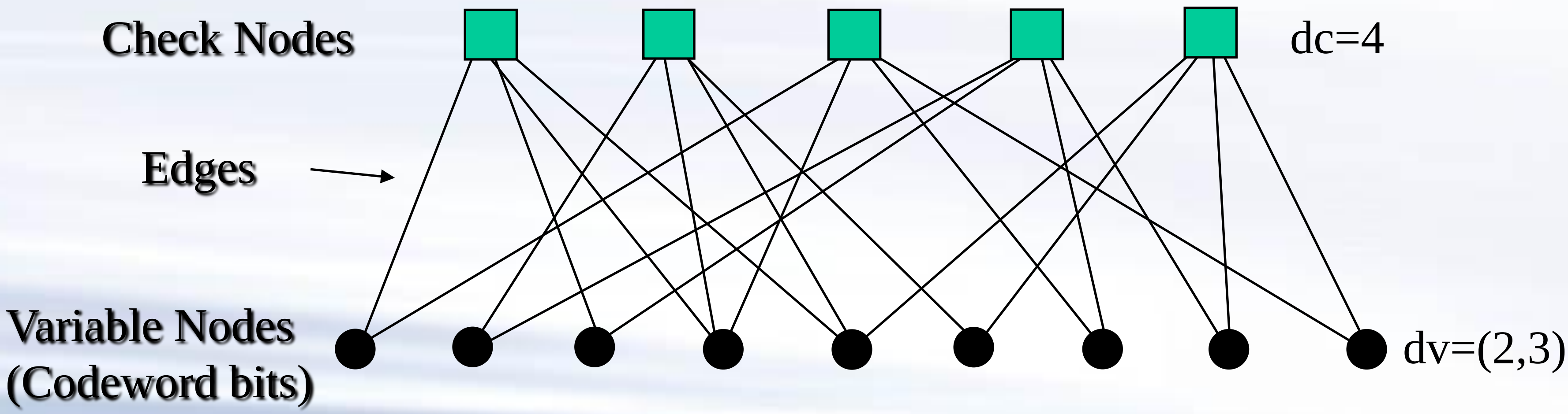
$$
\boldsymbol{H} = \begin{matrix}
0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0
\end{matrix}
$$

(dv=3,dc=4)

  - An LDPC code is called regular code, if its column and row weights (dv,dc) are constant for each column and row.
  - Otherwise, it is called irregular LDPC code. A special class of irregular LDPC codes, namely QC-LDPC, prevails in most current communications and broadcasting standards.

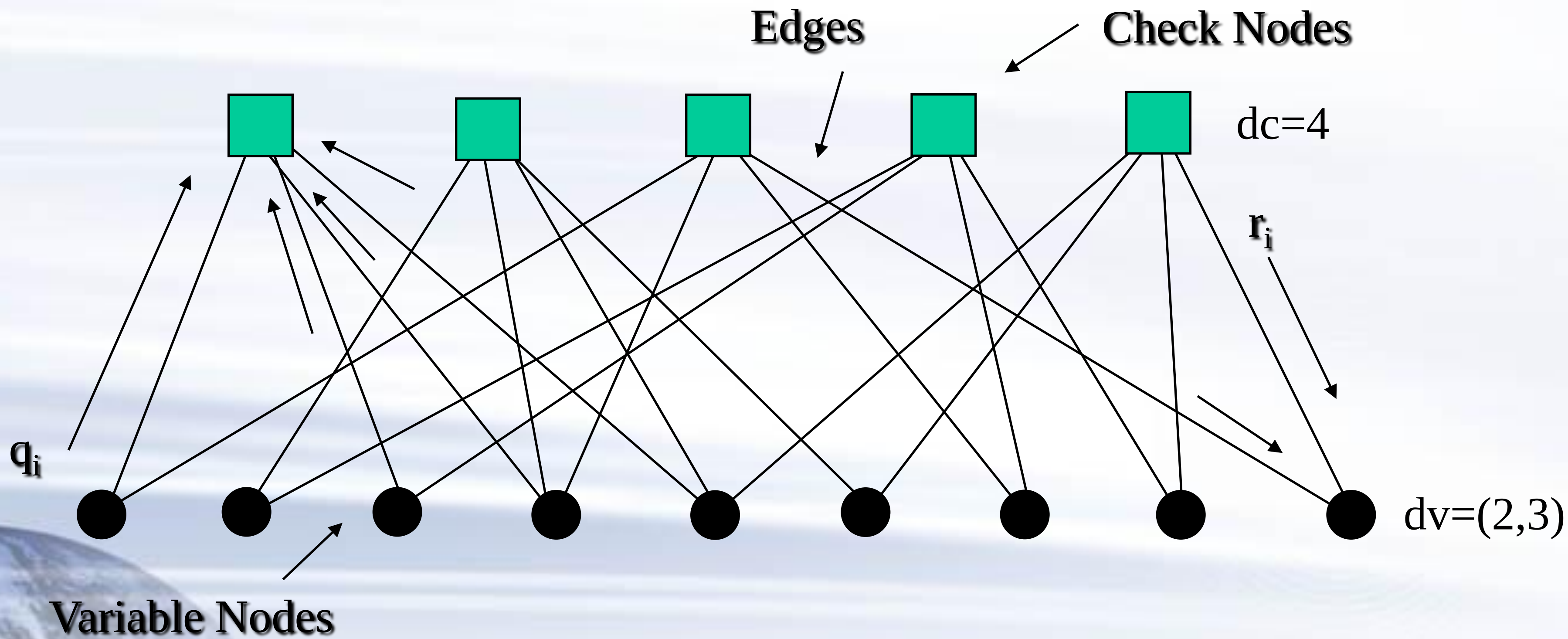
LDPC Codes

- Bipartite Graph (*equivalent to **H***)



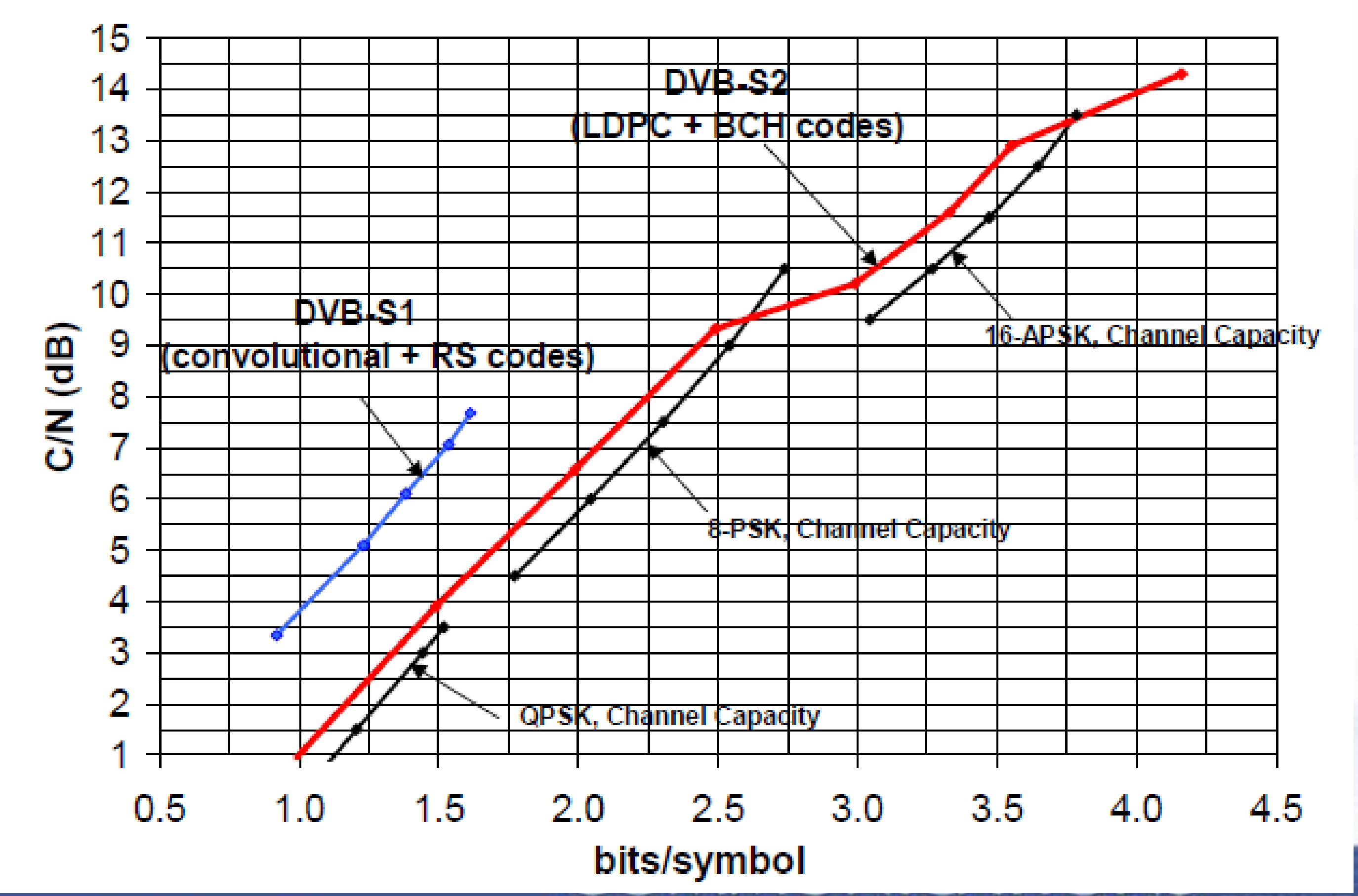
LDPC Codes

- Iterative Decoding (Belief Propagation)



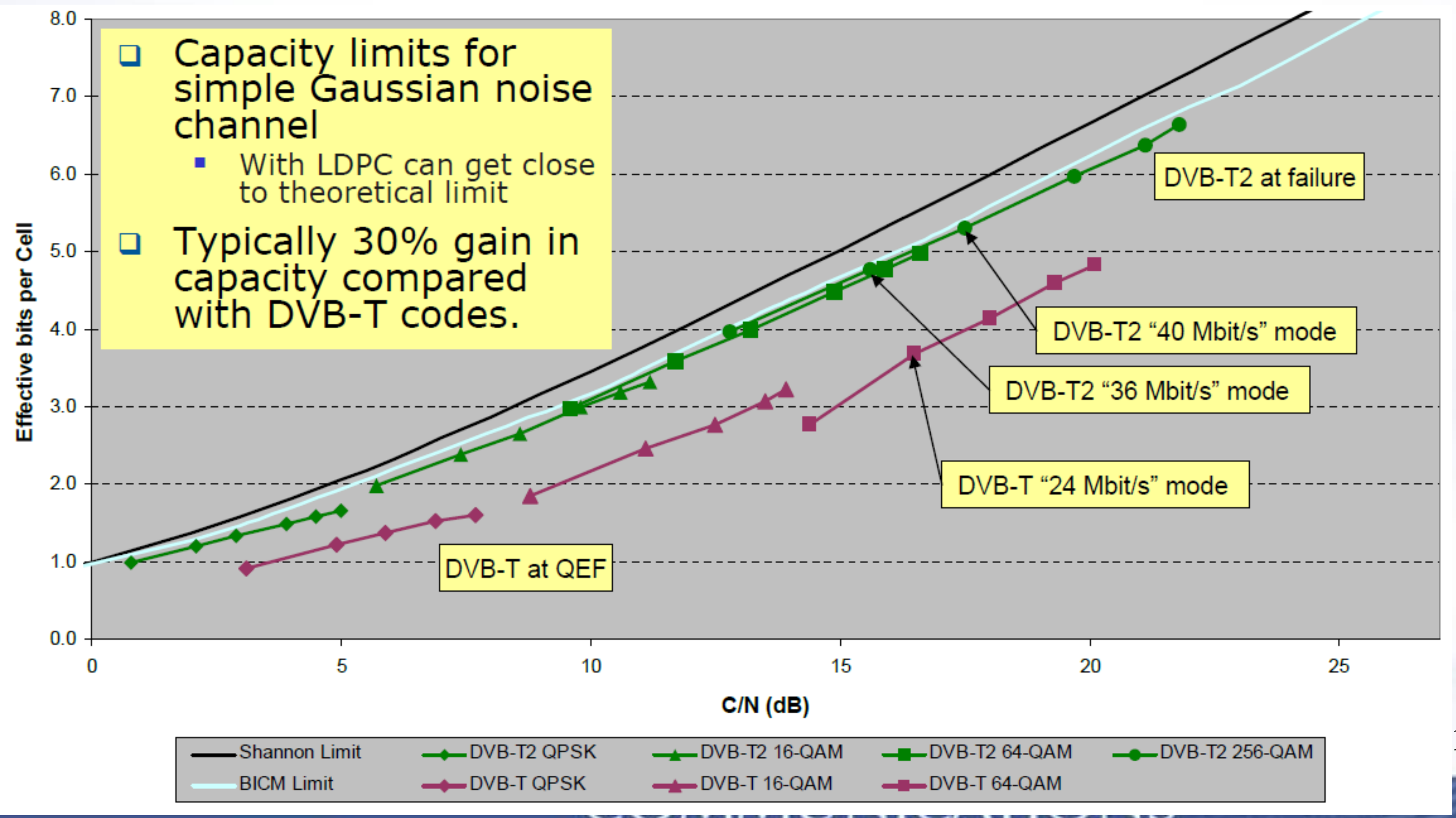
LDPC Codes

Performance of DVB-S vs. DVB-S2



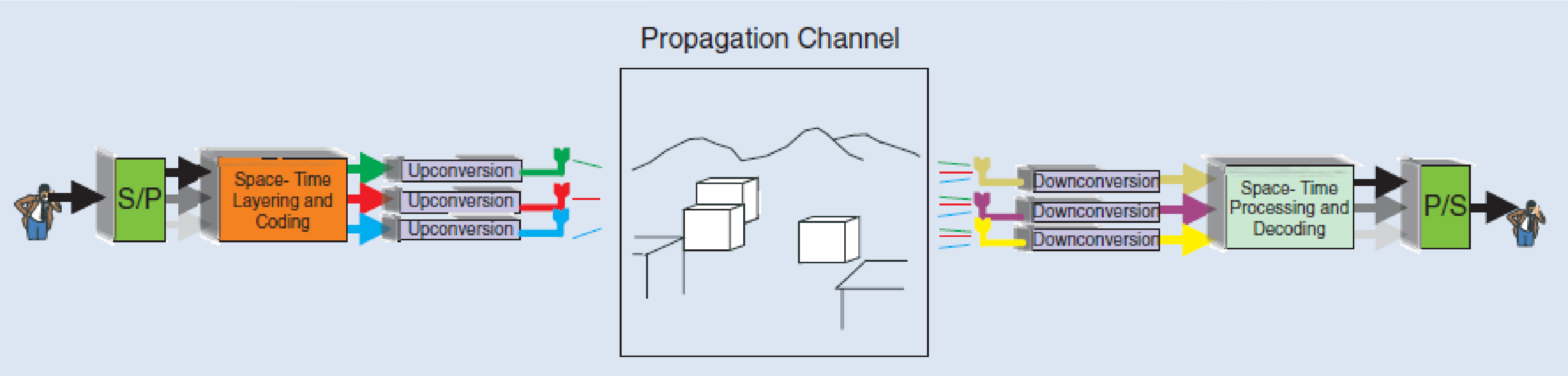
LDPC Codes

Performance of DVB-T2 vs. DVB-T



Space-Time Coding

- MIMO Technology
  - BLAST - Bell Labs Layered Space-Time Architecture
  - STTC - Space-Time Trellis Codes
  - STBC - Space-Time Block Codes (good for broadcasting)



Space-Time Coding

- STBC

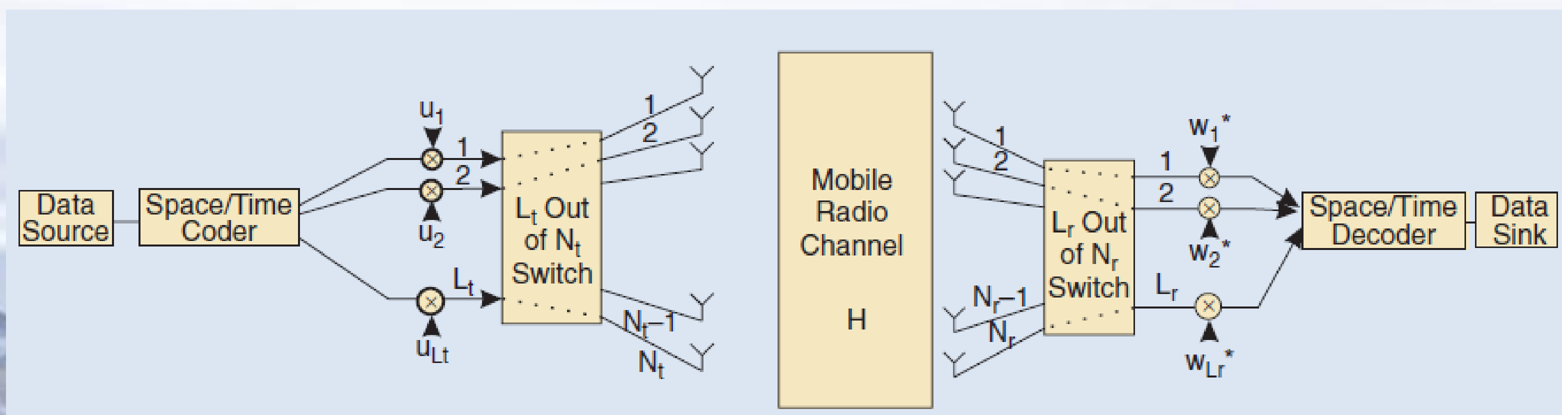
General Case

$$\begin{array}{cc} & \text{transmit antennas} \\ \text{time-slots} & \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n_T} \\ s_{21} & s_{22} & \cdots & s_{2n_T} \\ \vdots & \vdots & & \vdots \\ s_{T1} & s_{T2} & \cdots & s_{Tn_T} \end{bmatrix} \end{array}$$

Alamouti Codes

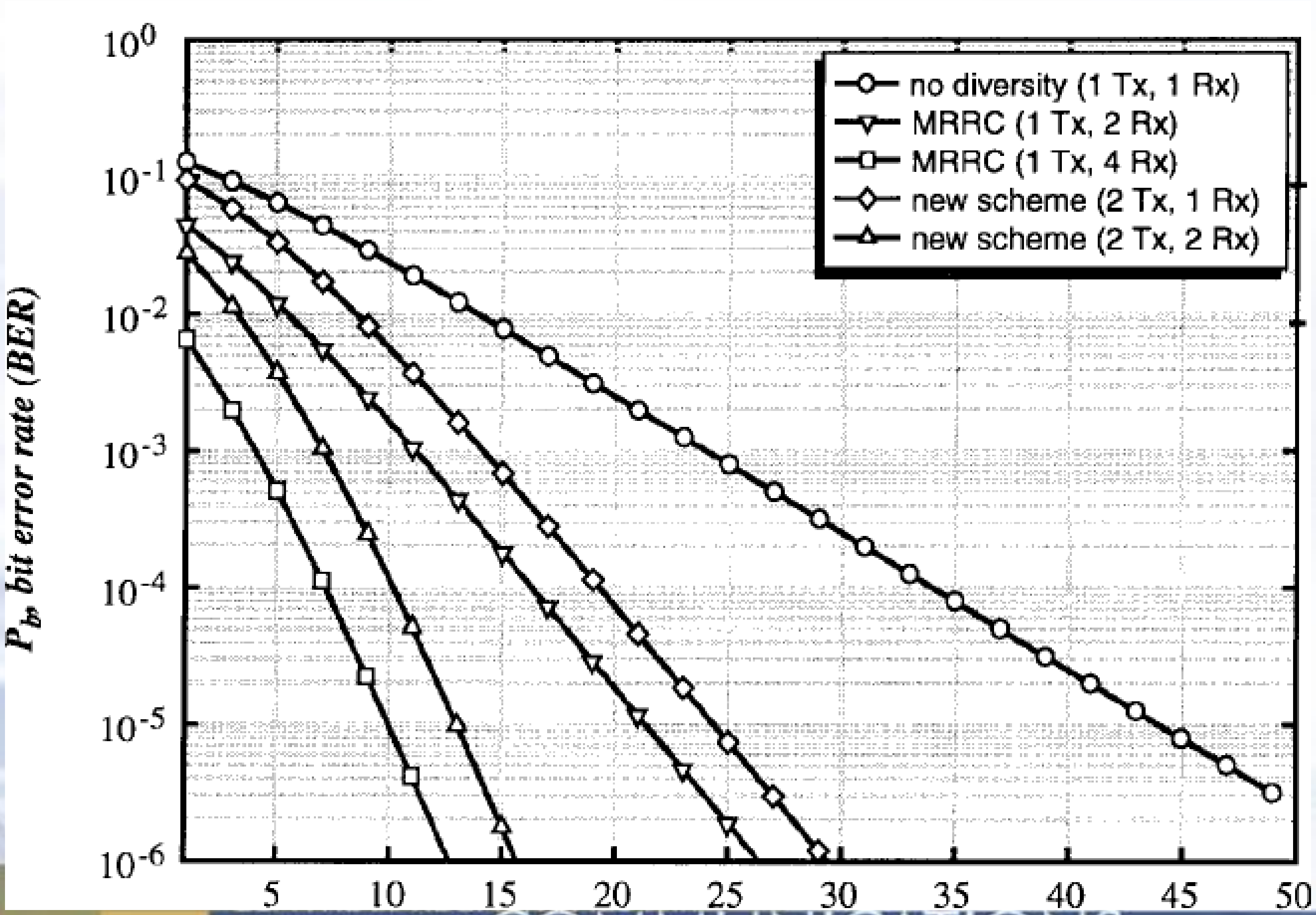
(2x2)

$$C_2 = \begin{bmatrix} c_1 & c_2 \\ -c_2^* & c_1^* \end{bmatrix}$$



Space-Time Coding

- Performance of Alamouti Codes



Error Correction Codes in Wireless Communications and Broadcasting

- ATSC: TCM + R-S;
- DVB-T and ISDB-T: Convolutional + R-S;
- DTMB: LDPC(8k) + BCH;
- DVB-T2: LDPC (16/64k) + BCH;
- Wi-Max and WiFi: LDPC(0.5/1/2k) + BCH;
- ATSC M/H: Turbo + R-S/CRC product codes;
- 3GPP and LTE: Turbo + R-S;

Joint Source and Channel Coding

- Based on information theory's **separation principle**, there is nothing to be gained from joint data compression and channel coding. This principle holds for stationary channel and source provided that the delay is unbounded.
- The separate design of source and channel codes allows for diverse sources to share the same digital media (Satellite, Cable, terrestrial, IPTV, Internet).
- However, in the real world, source and channel are not stationary. Meanwhile, the communication systems always have finite block-length data package. The separation principle does not apply!!!
- Therefore, based on the information theory, there will be some **residual redundancy left in the source encoded data** sequence, due to constraints on delay or complexity.
- a joint source and channel decoder can **exploit these residual redundancies to reduce the overall system error rate**.

Joint Source and Channel Coding

So far Joint Source and Channel Coding (JSCC) has never been implemented in practical system due to the following reasons:

- JSCC is closely related to the content of multimedia source. Therefore, it is difficult to develop a universal JSCC solution to handle a variety of content in today's multi-service communication system/network.
- JSCC requires a special combination of certain source and channel coding schemes, which leads to high computational complexity.

JSCC remains a academic research topic.

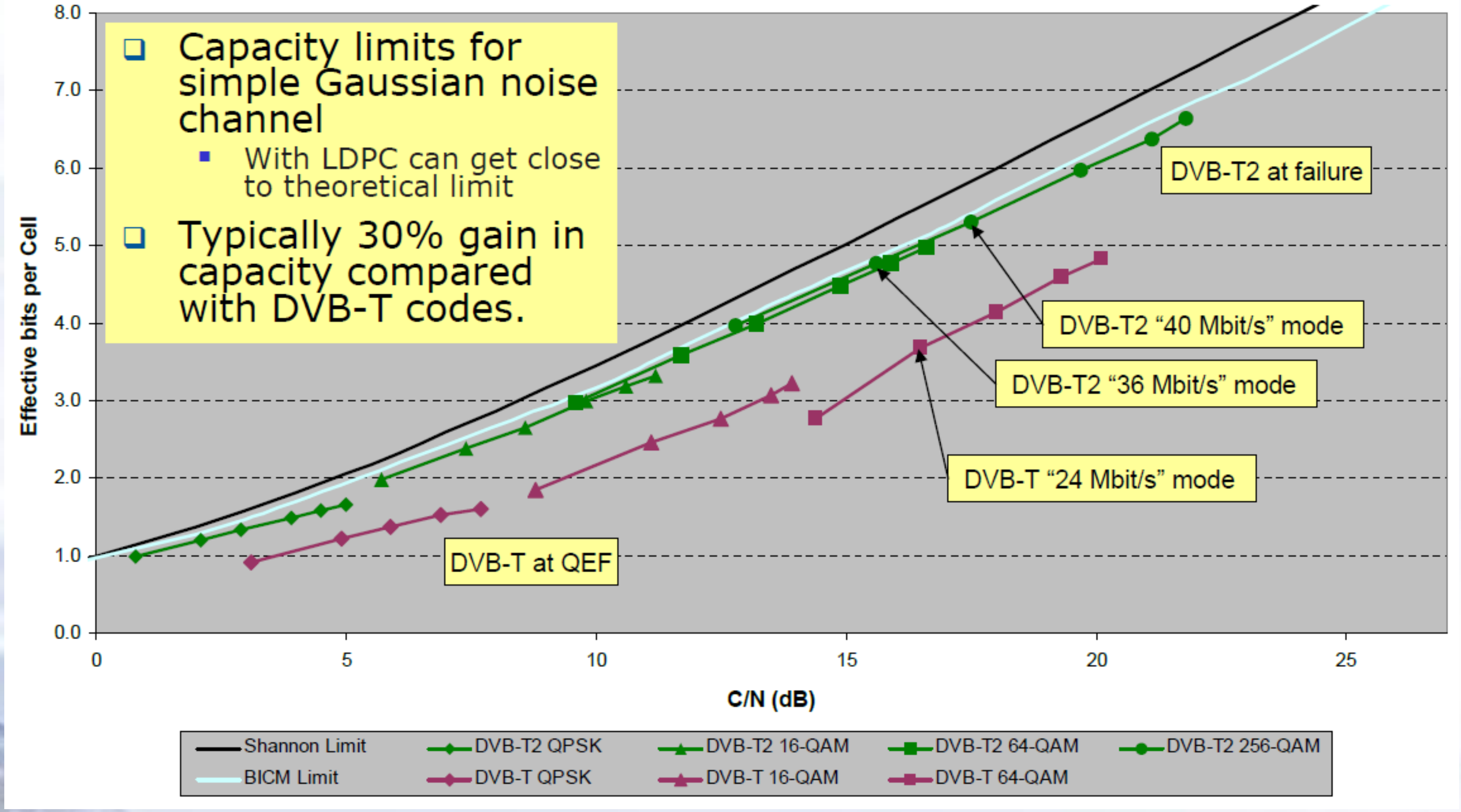
It has the potential to provide better performance system, particularly in mobile wireless environments.

Cross-Layer Recursive Forward Error Correction Coding (X-Rec FEC)

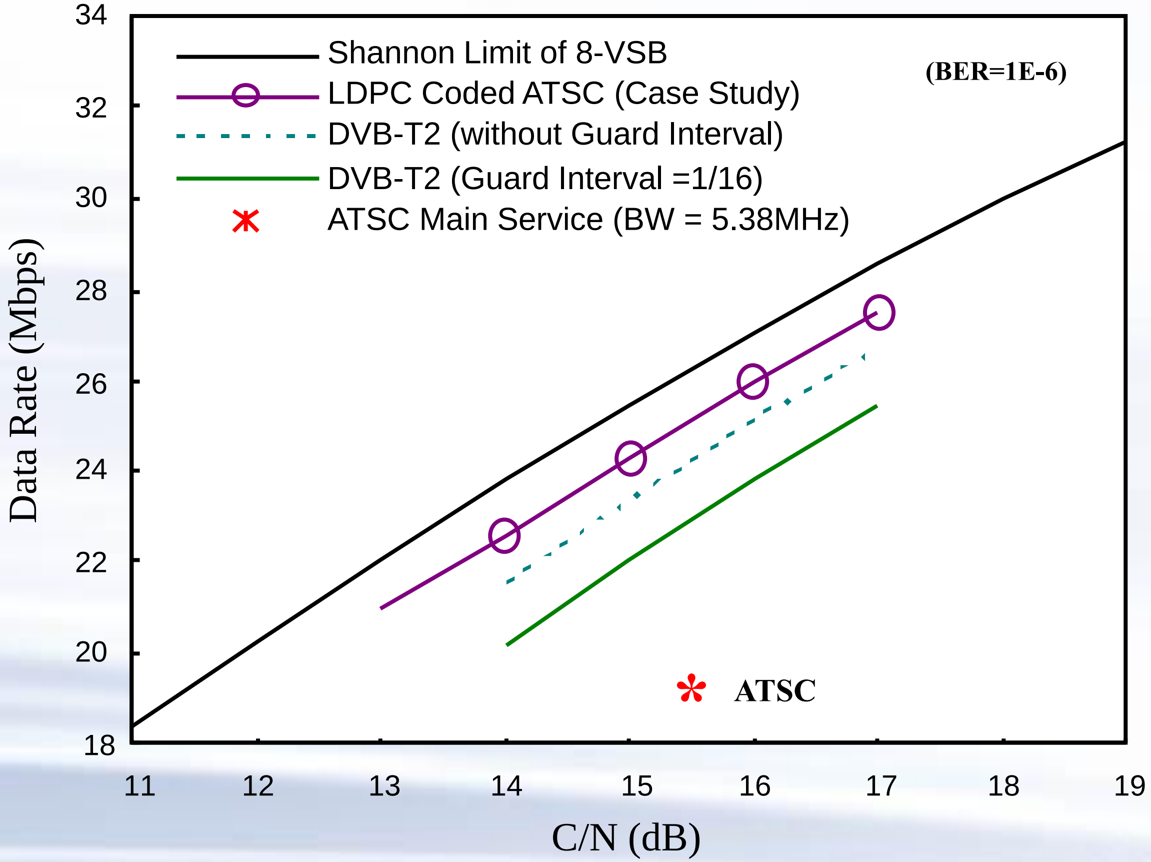
A CRC on going research project: it is a standard independent decoding scheme, that exploits the cross-layer redundancy information to improve the receiver performance.

LDPC code

Performance of DVB-T2 vs. DVB-T



Performance Comparison



Note: C/N calculation using BW = 5.38 MHz, or BW = 6 MHz with 0.5 dB implementation margin.

Cross-Layer Recursive Forward Error Correction Coding (X-Rec FEC)

- Based on information theory and the research in Joined Source and Channel Coding (JSCC): the residual redundancy left in the source encoded data can be exploited to reduce the decoding error rate.
- However, the JSCC only investigated the combination of source coding (presentation layer) and channel coding (physical layer). A modern digital communication system contains many additional layers between physical and presentation layers, such as transport layer, data link layer, network layer. Each of those layer might create redundancy information, which can be used to improve the error correction performance.

ATSC Mobile DTV Standard

Presentation Layer

AVC H.264 video coding and AAC-HE audio coding
Closed captioning

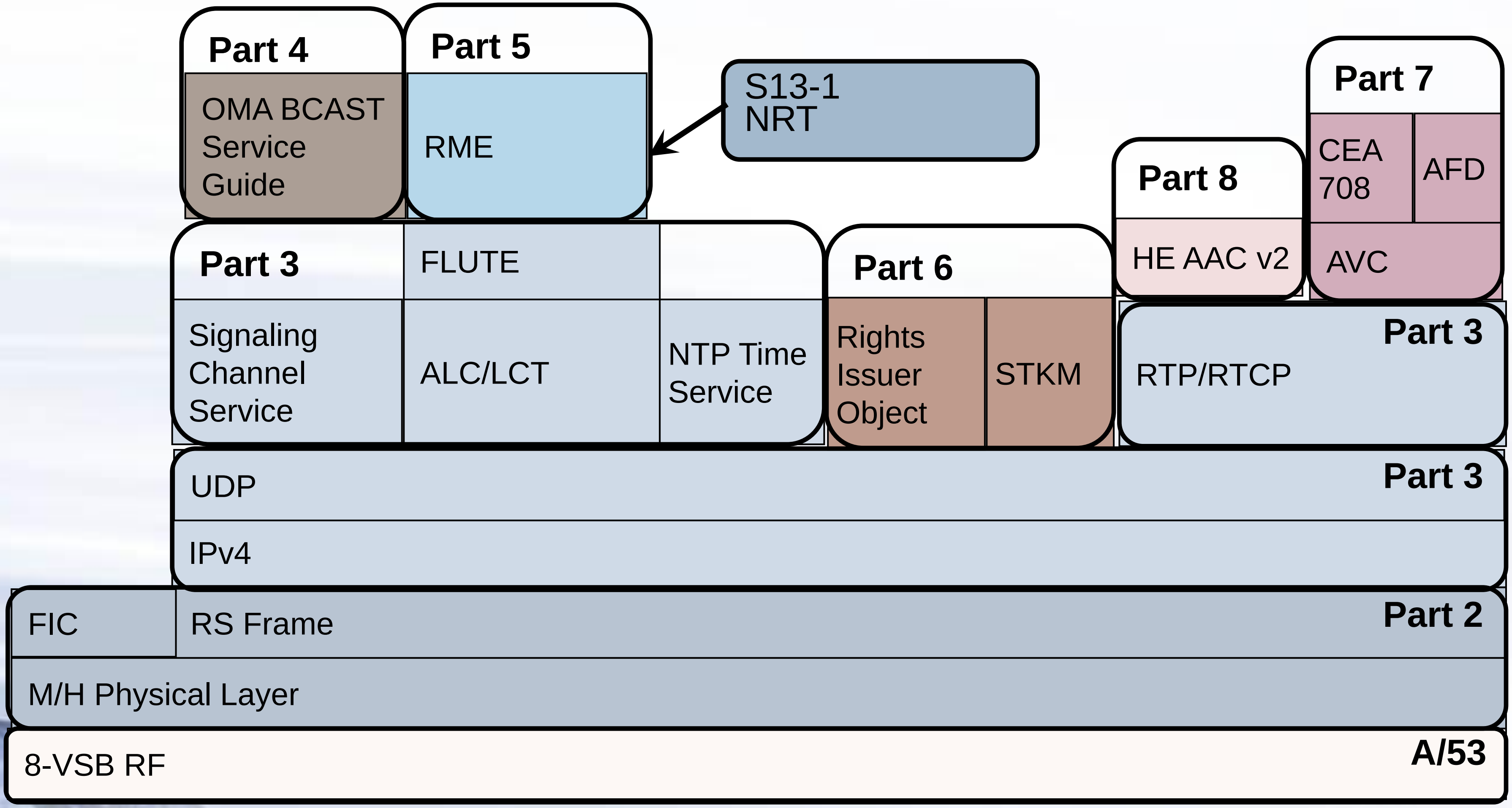
Management Layer

Transport – IP carried over ATSC MPEG-2 transport
Streaming and non-real time file transfer – NRT under development in ATSC
Electronic Service Guide - based on OMA BCAST

Physical Layer

RF transmission and forward error correction
Compatibility with legacy 8-VSB receivers/decoders

A/153 Protocol Stack



What are the redundancy from different layers?

(Using ATSC M/H as an example)

- There are information that needs to be sent out periodically for (switching on) receiver to quickly obtain synchronization and start decoding the data.
- Physical layer: Sync symbols, NULL packets;
- Management layer: IP headers, Packet ID, Meta data, SFN sync data, CA data, tables (SMT, SLT, GAT, CIT, RRT, etc.);
- Presentation layer: MPEG overheads (picture size, GoP size, etc.), audio overheads, program guide, service information, station logo, etc.
- There are estimated about 5-6% data in a DTV stream are sent out periodically, which are “known” or “deterministic” bits that can be used to improve the error correction. They are spread in all layers.

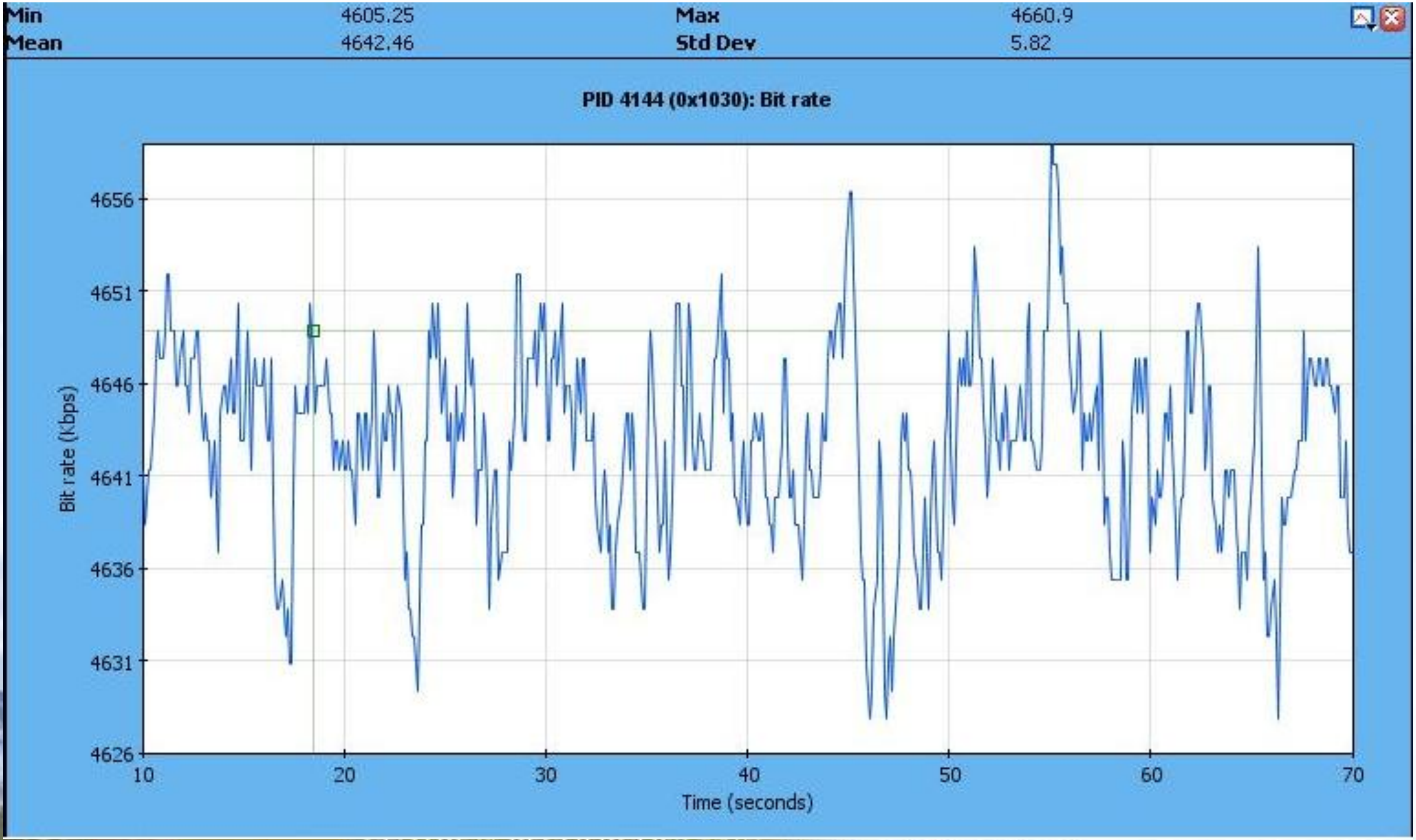
NULL packets

- The MPEG video coding output data rate always fluctuates (Variable Bit Rate), due to video content, GoP size, I-, P-, and B-frames;
- Since the (cable, satellite, and terrestrial) broadcasting channels are constant data rate channel, an output buffer is used to smooth out the data rate;
- Buffer over-flow is not allow, which causes service interrupt. When there is under flow, NULL packets are used to fill the gap;
- NULL packets are deterministic data and can be used for error correction. There are about 5-10% NULL packets in DTV stream.

Percentage of Null Packet in Five ATSC Stations

Channel	Station	% of Null packets	Content of Null	Comments
20	SUN-TV	40%	0xFF	SDTV
22	CBC	6.7%	0xFF	
25	R-C	6.3%	0xFF	French
27	OMNI-1	5.3%	0xAA	
66	OMNI-2	4.7%	0xAA	

A H.264 Video Encoder TS Output @ 5Mbps



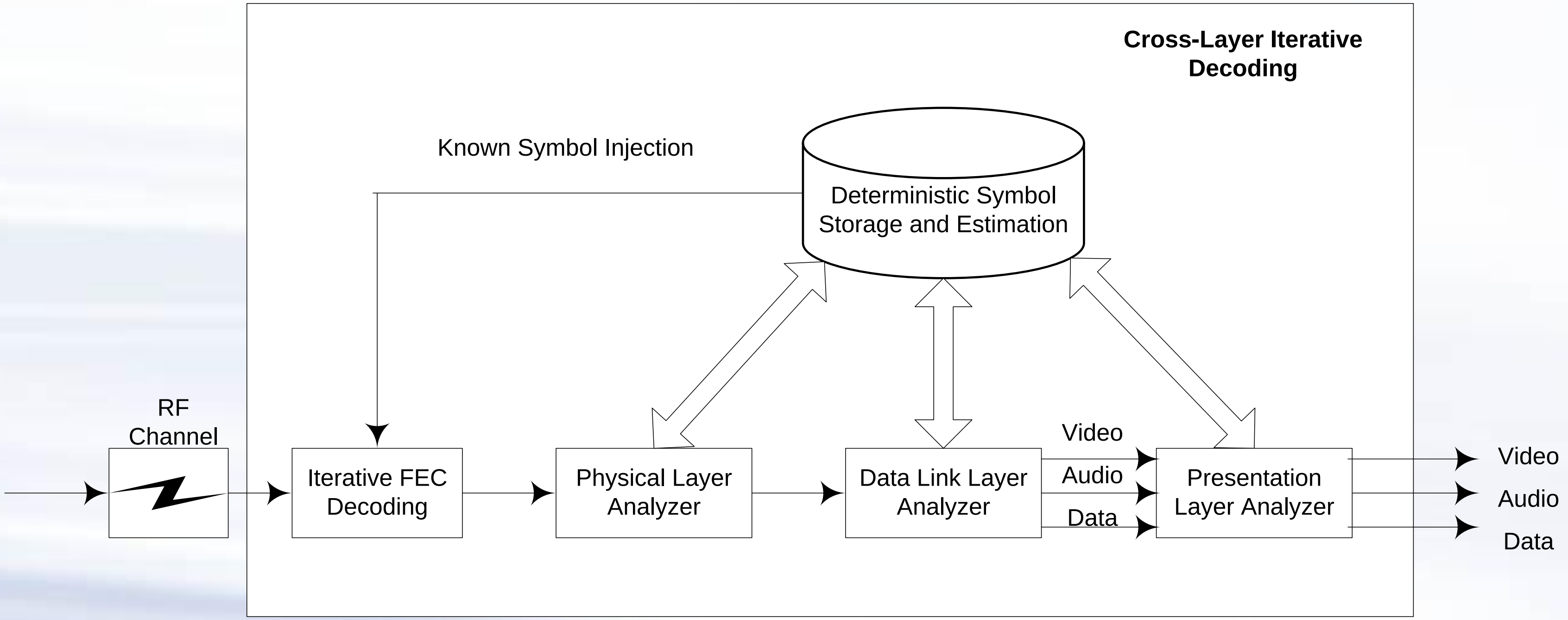
ATSC M/H R-S frame payload organization

<table>
<tr><td colspan="2">SMT</td><td>SLT</td><td>GAT</td><td>CIT</td><td>RRT</td></tr>
<tr><td colspan="6">RRT</td></tr>
<tr><td></td><td colspan="5">Video pkt(s)</td></tr>
<tr><td colspan="6"></td></tr>
<tr><td></td><td colspan="5">Audio pkt</td></tr>
<tr><td colspan="6">Video pkt(s)</td></tr>
<tr><td colspan="6">……….</td></tr>
<tr><td>Stuffing bytes</td><td colspan="2">RTCP (video) pkt</td><td colspan="2">RTCP (audio) pkt</td><td>NTP pkt</td></tr>
<tr><td colspan="6">Stuffing bytes</td></tr>
<tr><td colspan="6">……….</td></tr>
</table>

Cross-Layer Recursive Forward Error Correction Coding (X-Rec FEC)

- In our proposed cross-layer forward error-correction (X-Rec FEC) decoding, the known (or deterministic) information bits or sequences may come from the layers higher than the physical layer, e.g., data link layer, network layer, transportation layer, and presentation layer.
- The deterministic information bits can include physical layer's sync bits, training sequences, MPEG overheads, NULL TS packets; data layer's IP address, packet overheads; presentation layer's program guide, service information, audio/video coding overheads. They can be gradually identified and used in the iterative decoding process to improve the decoding performance.
- The proposed cross-layer decoding is different from the conventional Joint Source-Channel Coding (JSCC). Our scheme addresses the packetizing and transmission redundancy while JSCC makes use of the information redundancy in source coding process. As a result, the implementation of JSCC requires a combination of special source coding and channel coding algorithms; our scheme, however, can have good backward compatibility to most existing communication system with a variety of source and channel coding algorithms.

Cross-Layer Recursive Forward Error Correction Coding (X-Rec FEC)



Physical Layer:

Sync Symbol
Training Squence
SFN Sync Information

Data Link Layer:

Packet ID
Network Overhead
Mata Data

Presentation Layer:

Program Guide
Service Information
MPEG Header
Audio Header

X-Rec FEC Decoder Complicity and Application

- X-Rec FEC receiver might not be much more complex than a regular receiver, since all of those know bits (sync bits, packet overheads; program guide, service information, audio/video coding overheads and NULL packets) are present in a regular receiver. They just need to be identified, stored and used properly, when needed;
- When there is no transmission error, X-Rec FEC algorithm does not need to be applied (a green algorithm);
- X-Rec FEC is expected to work well in mobile environment and ideal for Software Defined Receiver;
- X-Rec FEC can apply to all existing systems (ATSC, DVB, ISDB, DTMB, WiMax, WiFi, 3GPP, LTE, etc. Some can achieve more gains some with less).

How is the S/N improvement achieved

- For a coding system with rate R = 0.75 code, if there are 10% of known bits that can be identified, those bits do not need to be error corrected. Therefore the real coding rate is reduced to

  $R_1 = 0.75 \times (1 - 10\%) = 0.675$;

- A code with R_1=0.675 will have lower E_b/N_o than a code with R = 0.75. Meanwhile, the system data rate R_b is also reduced by 10%;
- Based on formula: $$\frac{S}{N}(dB) = \frac{E_b}{N_0}(dB) + 10 \cdot \log\left(\frac{R_b}{BW}\right)$$ both E_b/N_o and R_b are reduced – the S/N is reduced.

X-Rec FEC can seamlessly change the coding rate

- Since finding the known bits is equivalent to change of the coding rate, if we deliberately insert some known bits in the transmitter side (e.g. Null packets or some know sequences), we could seamlessly vary the coding rate.
- For LDPC code, each code rate (e.g., R=1/2, 2/3, 3/4, 5/6 in DVB-T2 and WiMax) needs a matching decoder. Although there are methods to share the silicon circuits, a receiver with four LDPC code rates will be much more complicated than the one with one code rate. X-Rec FEC can use one code and achieve different code rate by carefully inserting know bits. This will greatly reduce the receiver complexity.
- **Research on X-Rec FEC is still on going at CRC.**

Data Rate vs. Robustness for Terrestrial DTV

Data Rate vs. Robustness

- The Information theory is about strike a balance between data rate and reliability;
- Recently, DVB-T2 and NHK Super Hi-vision both aimed at higher data rate for higher spectrum efficiency, which might lead to less robust reception?
- **Question**: Can terrestrial broadcasting system win the data rate race over satellite/Cable/IPTV?

Data Rate vs. Robustness

- The Information theory is about strike a balance between data rate and reliability;
- Recently, DVB-T2 and NHK Super Hi-vision both aimed at higher data rate for higher spectrum efficiency, which might lead to less robust reception?
- **Question**: Can terrestrial broadcasting system win the data rate race over satellite/Cable/IPTV?
- Unlikely!!!

Data Rate vs. Robustness

- The real advantage of terrestrial broadcasting – Mobile and Portable reception – “anywhere, anytime, at home, on the go”.
- How to strike a balance between spectrum efficiency and reception robustness? Are there any alternatives?

Data Rate vs. Robustness

- Recent advance in video coding (H.264 and HEVC) and multiplexing technologies could achieve double or quadruple the effective data rate – increasing the number of services, in comparison to MPEG-2 based system;
- Recent advance in error correction code techniques could increase the data rate by 30-50%;
- There are other promising techniques: diversity, cross-layer error correction;
- More efficient use of spectrum can mean higher data rate, or more reception robustness. Which way terrestrial broadcasting should go?

Recommendation

- Using advanced audio/video coding to achieve better spectrum efficiency and more services available within 6 MHz channel;
- Using advanced error correction system to achieve more robust reception;
- Additionally, using diversity technologies and distributed transmission can lead to a greener terrestrial broadcast DTV eco-system.

Thank You

Yiyan Wu

Communications Research Centre Canada

yiyan.wu@crc.ca

613-998-2870