AC-3 and DTS

Prof. Dr.-Ing. Karlheinz Brandenburg

Fraunhofer IDMT & Ilmenau Technical University
Ilmenau, Germany

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Dolby Digital

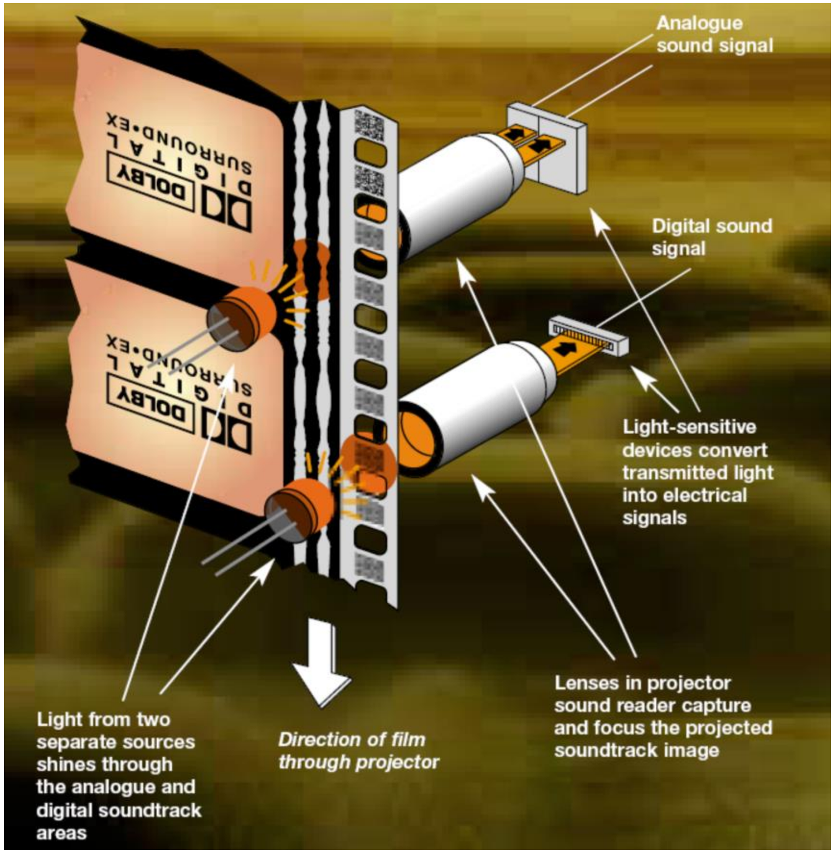
Source: Dolby Labs, Internet

- Dolby Digital (AC-3) was first commercially used in 1992
- Multi-channel digital audio for 35mm movie film material alongside the (optical) analog audio channel
- Perceptual coding with block length of 256 samples
- Additionally it is used in:
  - Laser Disc
  - ATSC High Definition Digital Television (HDTV)
  - DVB/ATSC Standard Definition Digital Television (SDTV)
  - DVD-Video/Audio
  - Internet-, Cable-, Satellite broadcasting
- For 5.1-channel audio, the bit stream is packed in the AES-EBU transmission format.
- Bit stream defined in ATSC "Digital Audio Compression Standard, Revision B": Doc. A/52B from June 2005; also E-AC3 included

TECHNISCHE UNIVERSITÄT ILMENAU

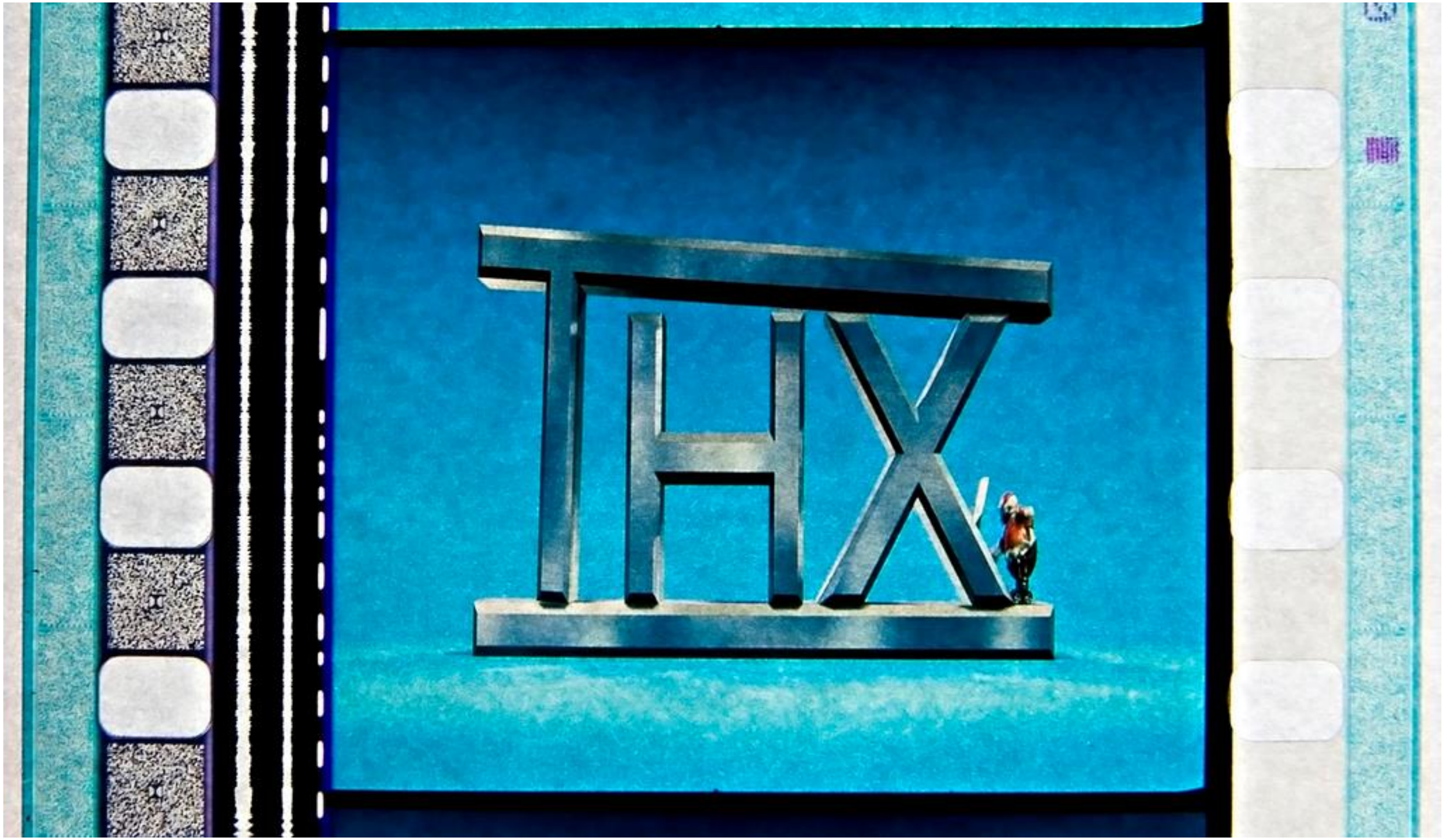
Fraunhofer IDMT

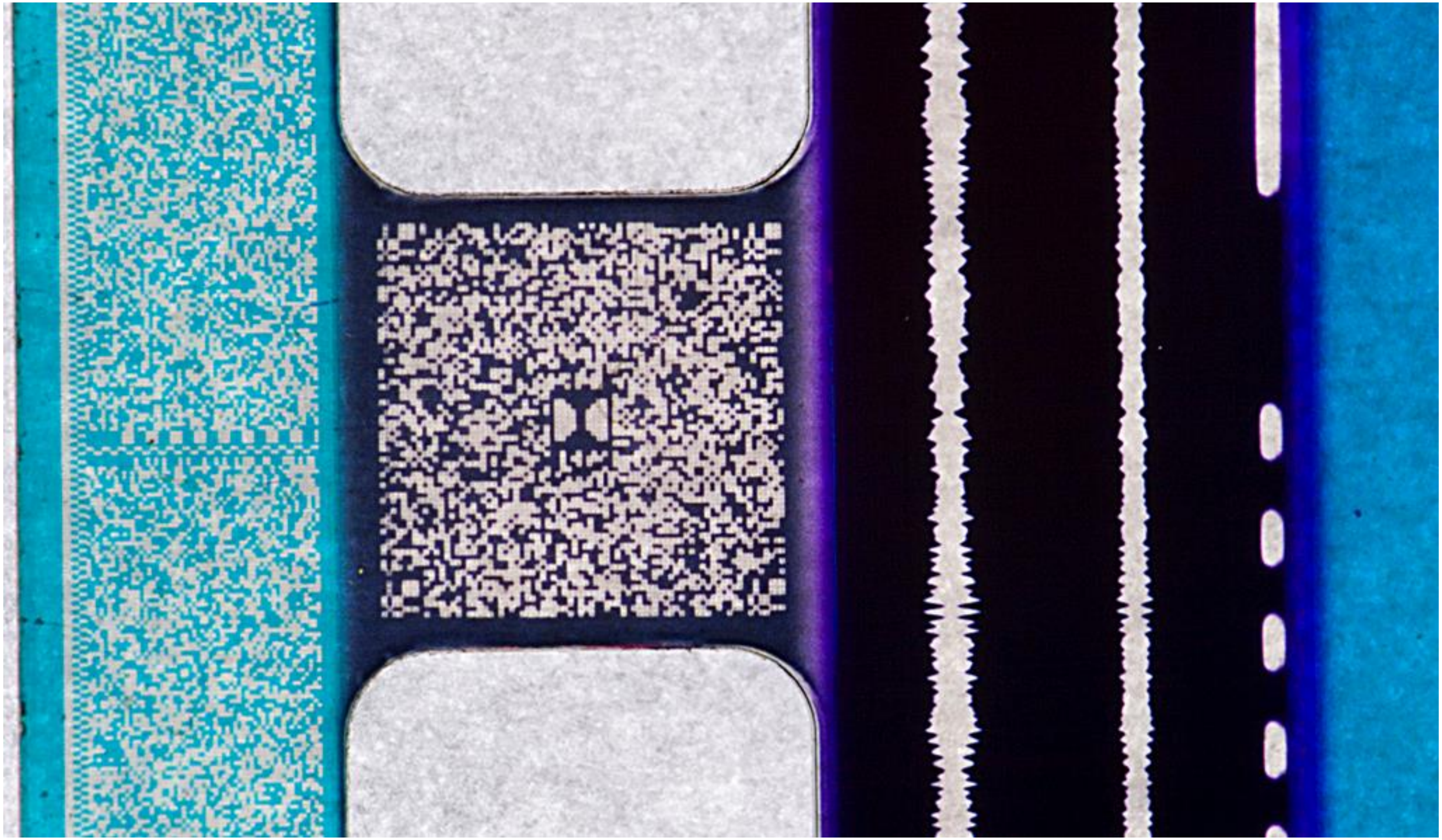
Dolby Digital embedded in a piece of film



TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT


THX



TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Dolby AC-3 (1)

Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage

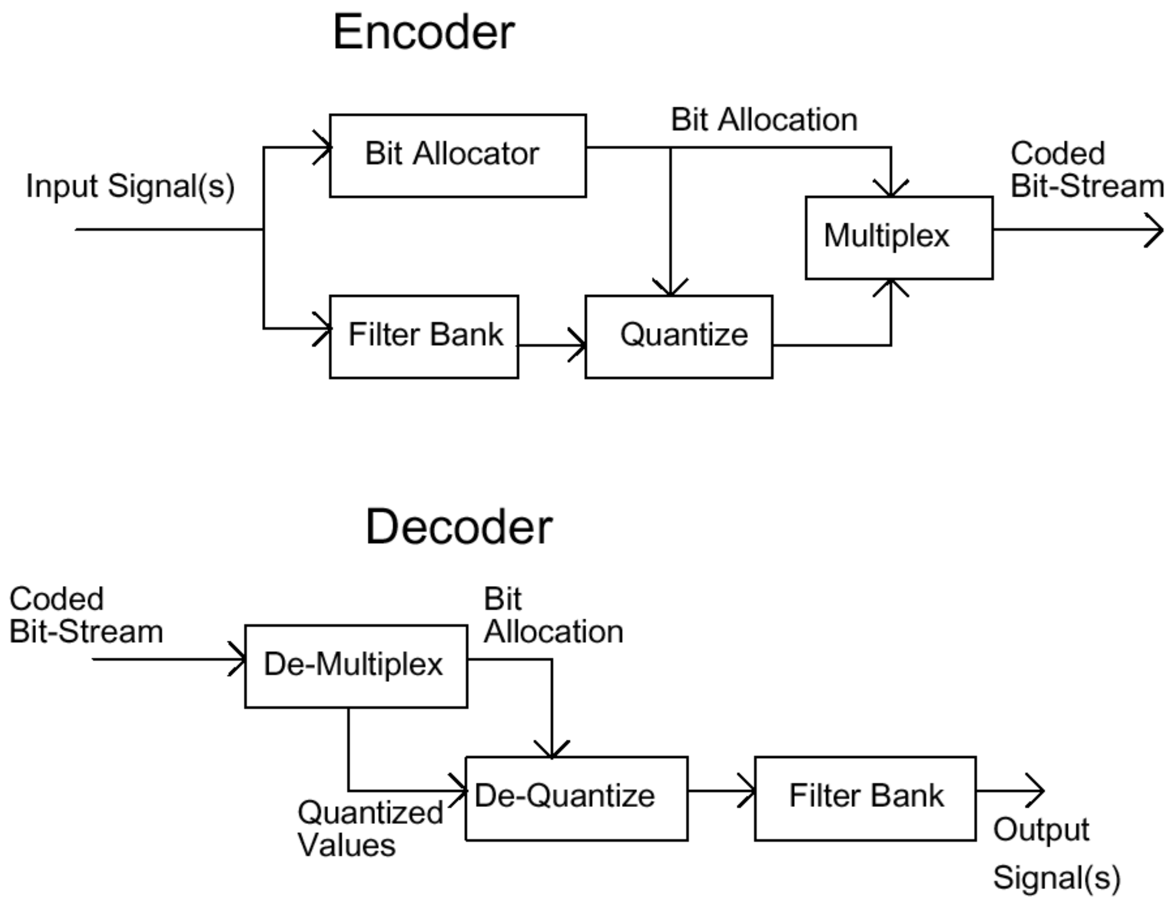
- Predecessors:
  - Dolby AC-1: low-cost, based on delta modulation
  - Dolby AC-2: transform based codec
- Lossy coder that uses psychoacoustics
- Special Features:
  - Use of a Variable Frequency Resolution Spectral Envelope
  - Hybrid Backward/Forward Adaptive Bit Allocation
- Primarily developed for multi-channel format for HDTV
- Based on ITU-R BS.775 that showed that 5 + 1 channels are enough for new digital audio system of movies (based on an analog Split-Surround-Format from 1979)

Dolby AC-3 (2)

Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage

- There is an useable data rate of 320kbps on 35mm movie film such that:
  - Audio compression must be used for 5.1 channel audio
  - The peak bit rate can not surpass 320kbps
- First film with AC-3: Star Trek VI (Dec. 1991)
- Transform:
  - Fielder windowing (aka KBD-Window)
  - Window length 512 Samples (10.66ms@48kHz) with 50% overlap: 256 Spectral values
  - Oddly Stacked Time-Domain Alias-Cancellation Filter Bank from Princen and Bradley
  - With signal transients (attacks) block switching is used to half the block length.
  - Frequency resolution: 93,75 Hz
  - No „Critical Bands“ like in MP3

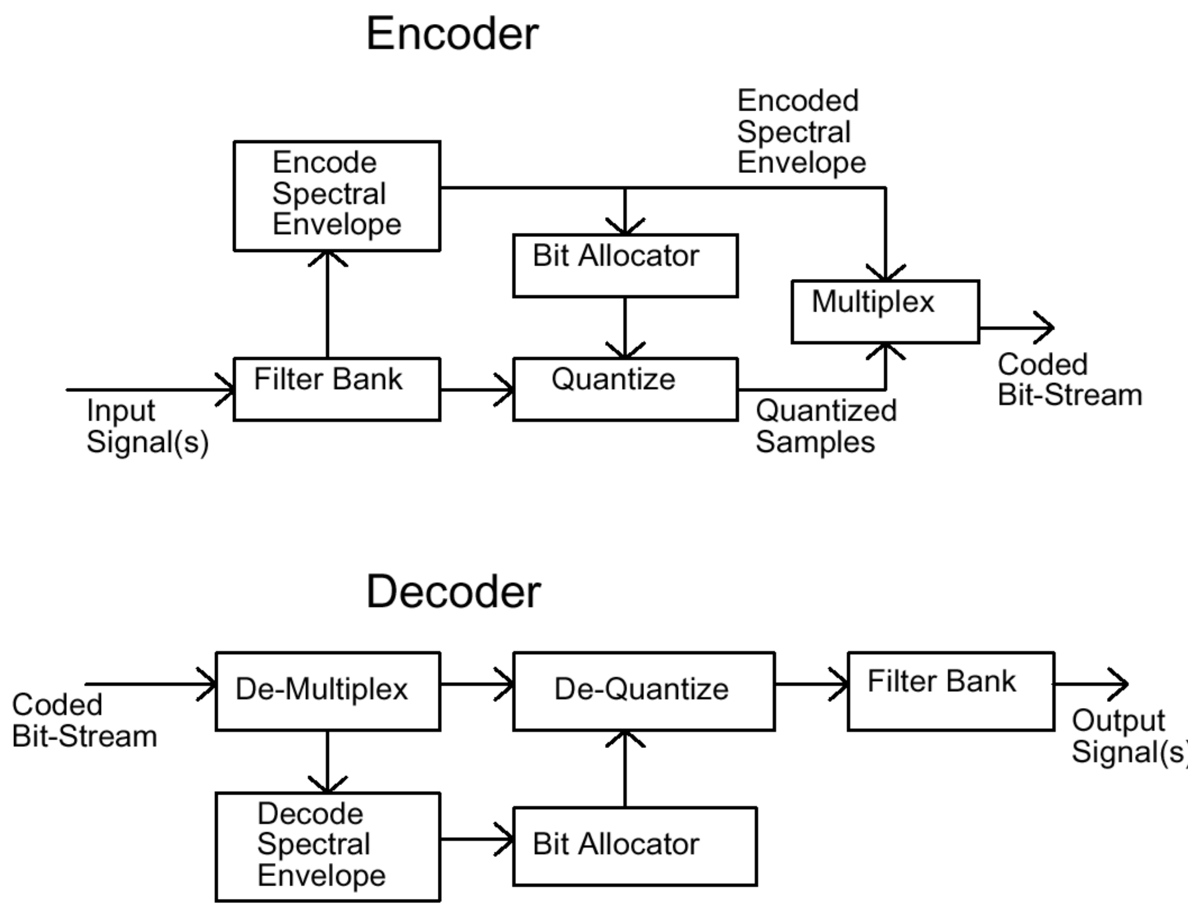
Dolby AC-3: Forward Adaptive Bit Allocation



Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

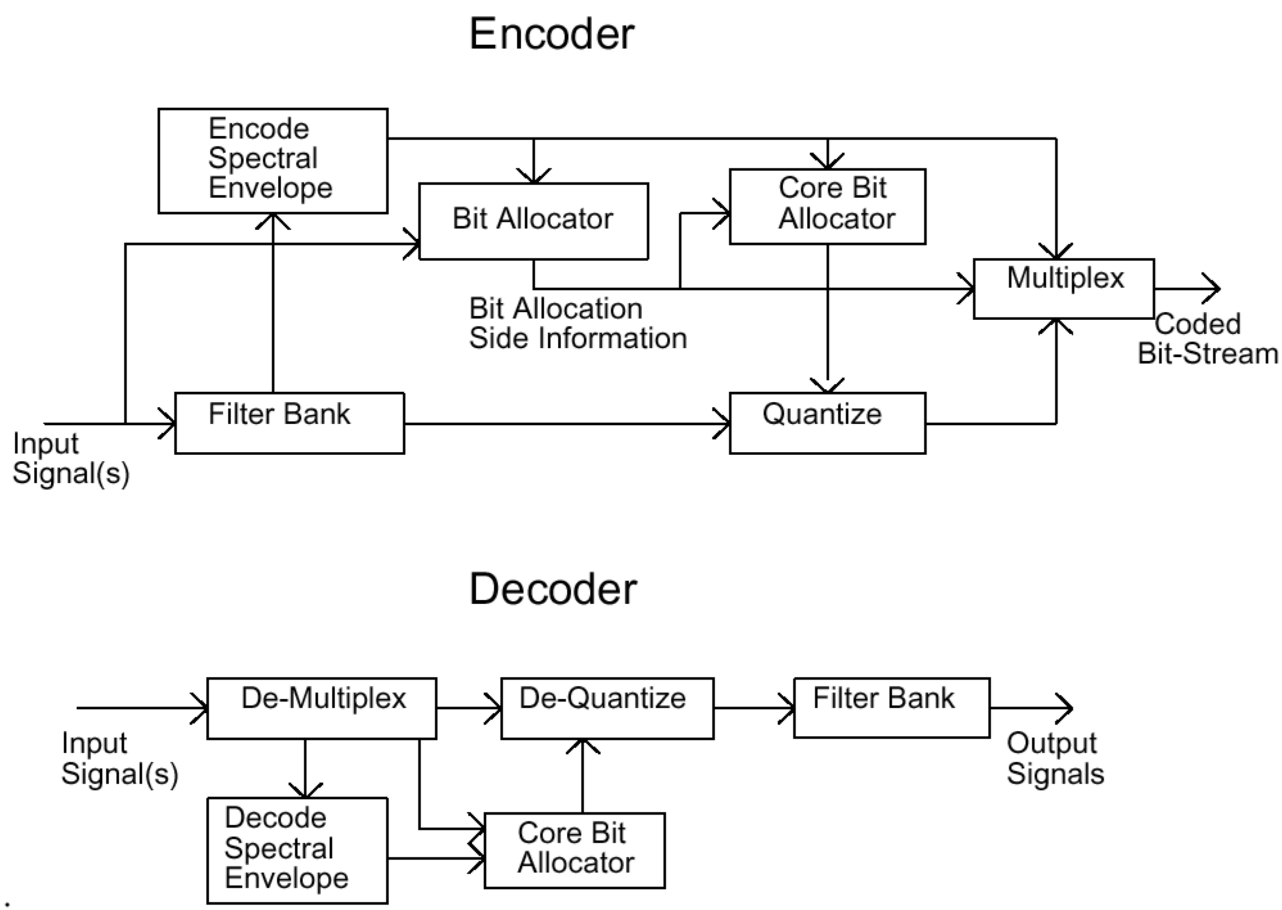
Dolby AC-3: Backward Adaptive Bit Allocation



Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage

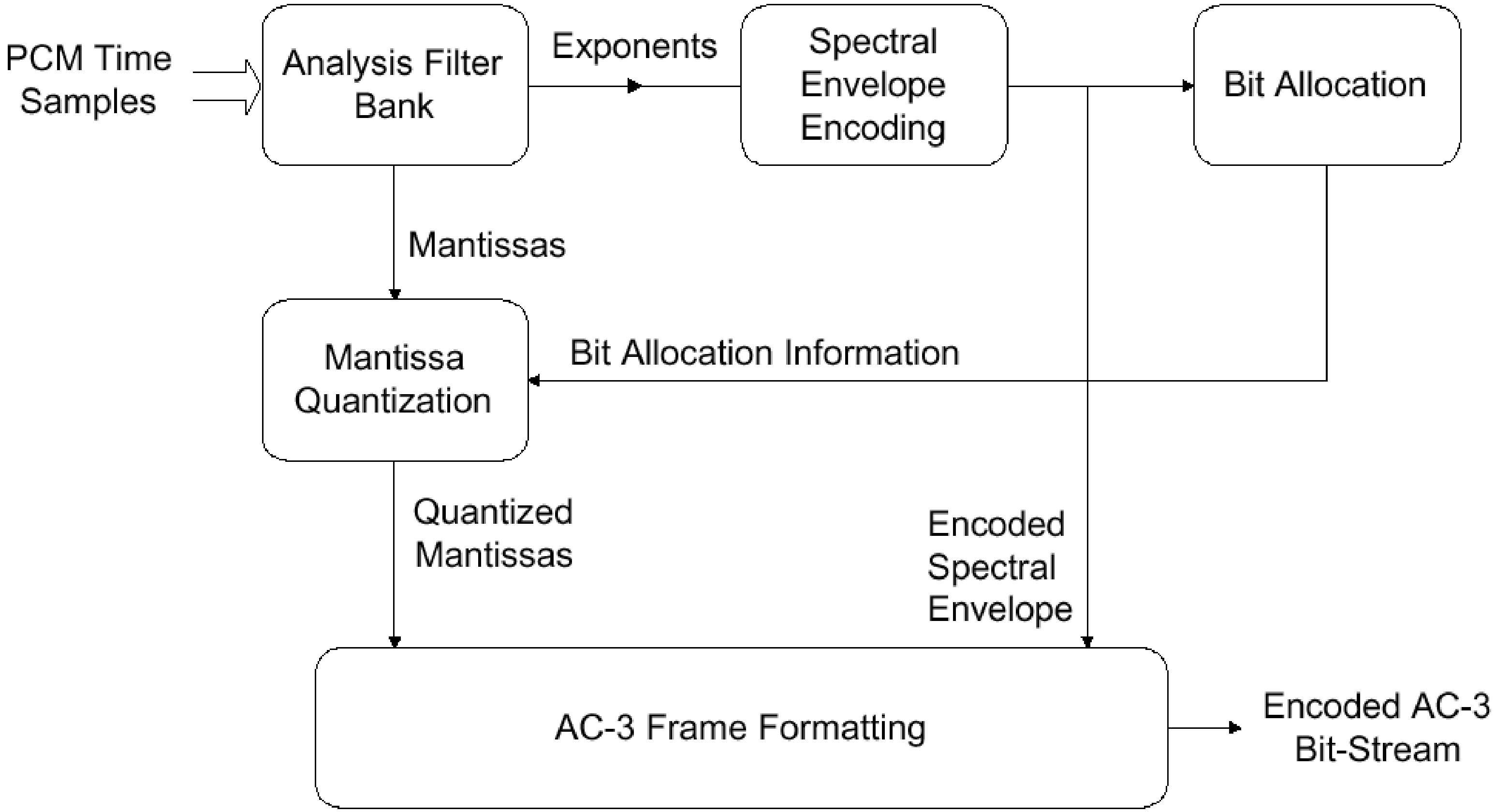
Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

Dolby AC-3: Hybrid Backward/Forward ABA



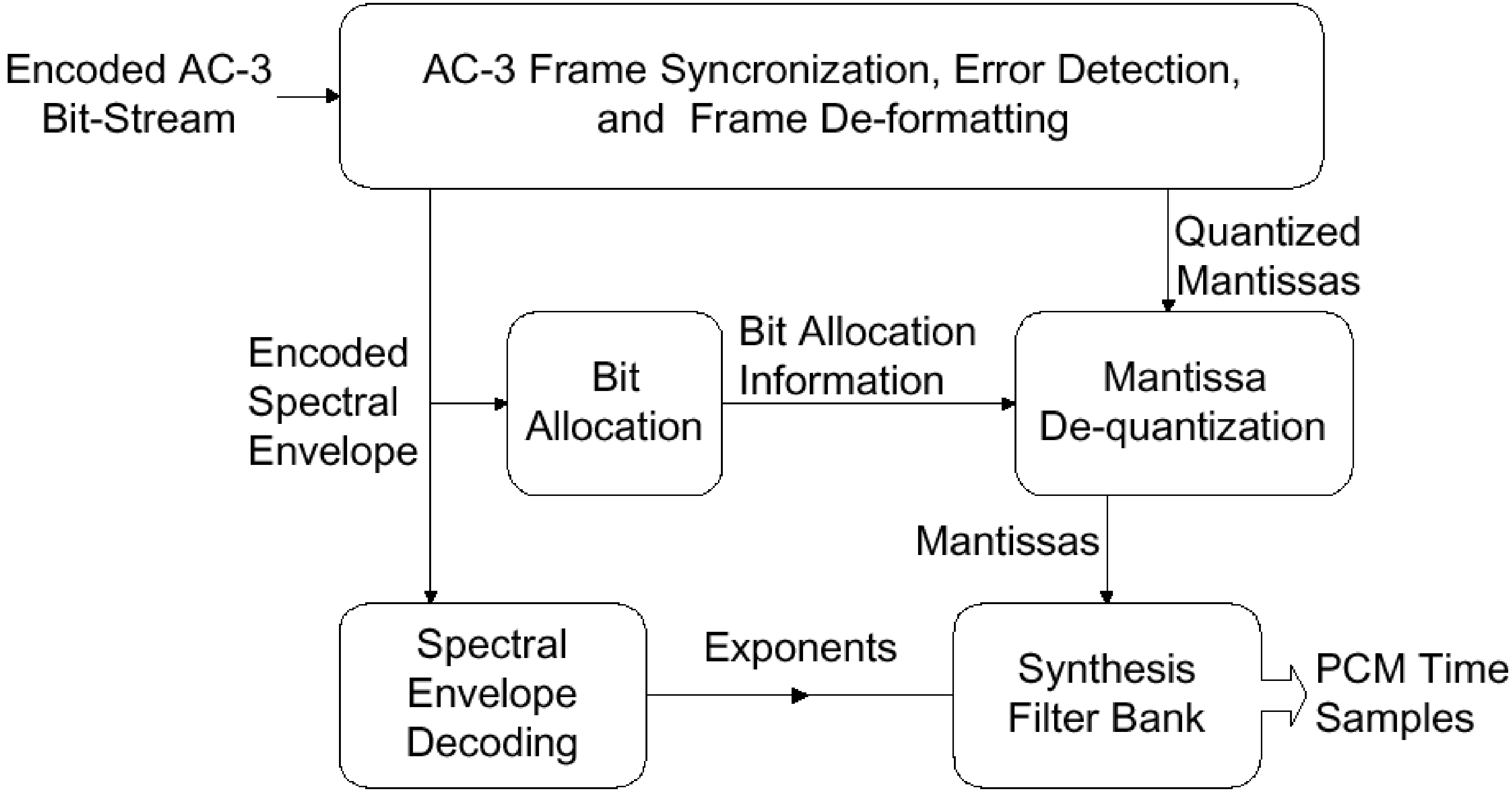
Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage

Dolby AC-3: Encoder



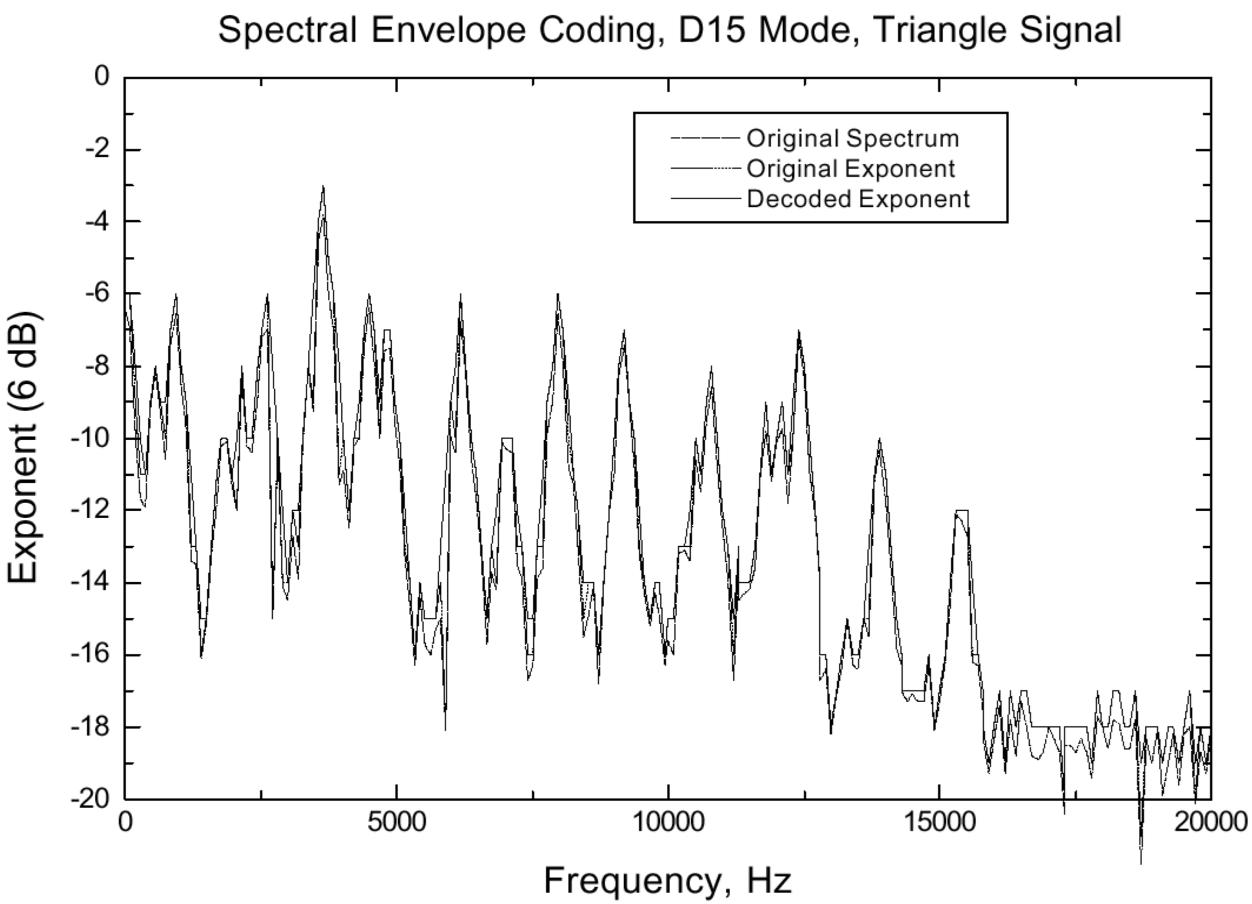
Source: Advanced Television Systems Committee: „Digital Audio Compression Standard (AC-3)“, Nov. 94

Dolby AC-3: Decoder



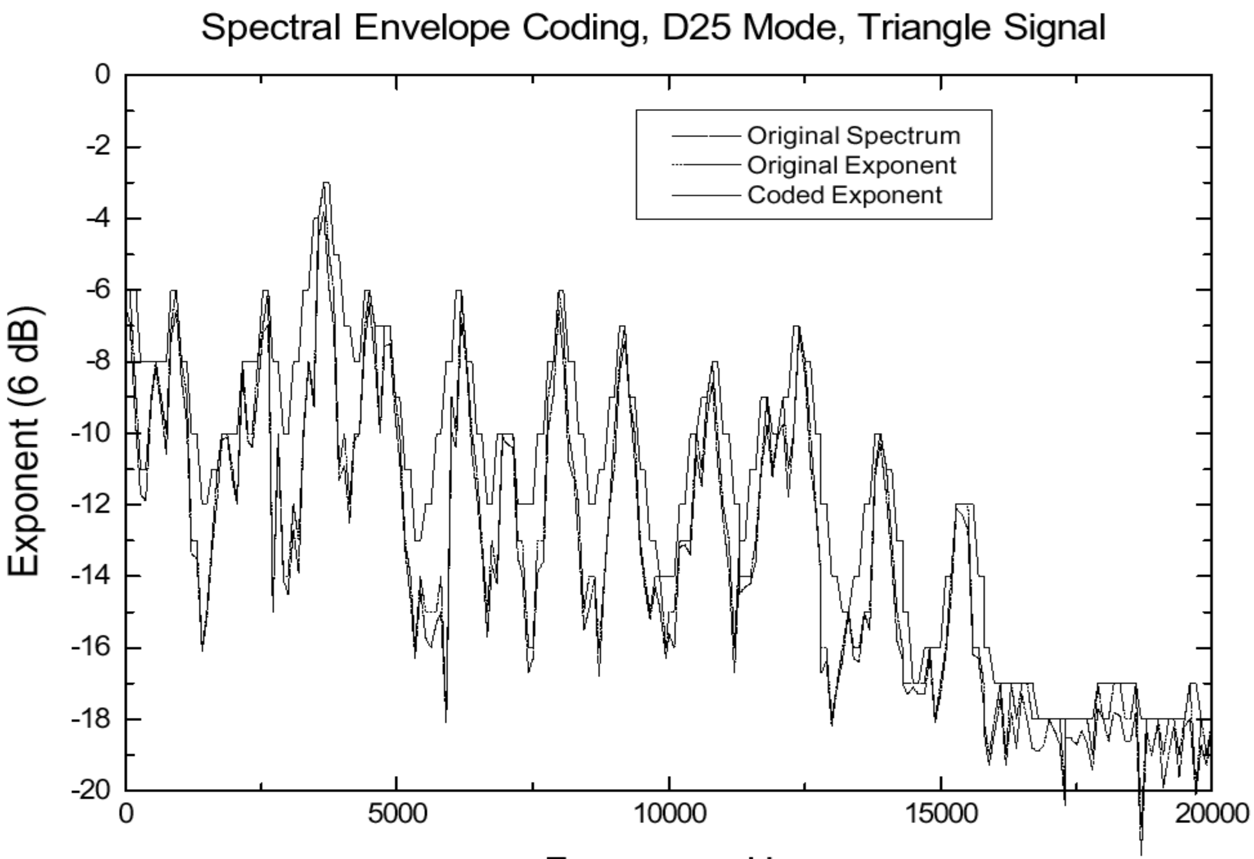
Source: Advanced Television Systems Committee: „Digital Audio Compression Standard (AC-3)", Nov. 94

Dolby AC-3: Spectral Envelope (1)



Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage"

Dolby AC-3: Spectral Envelope (2)



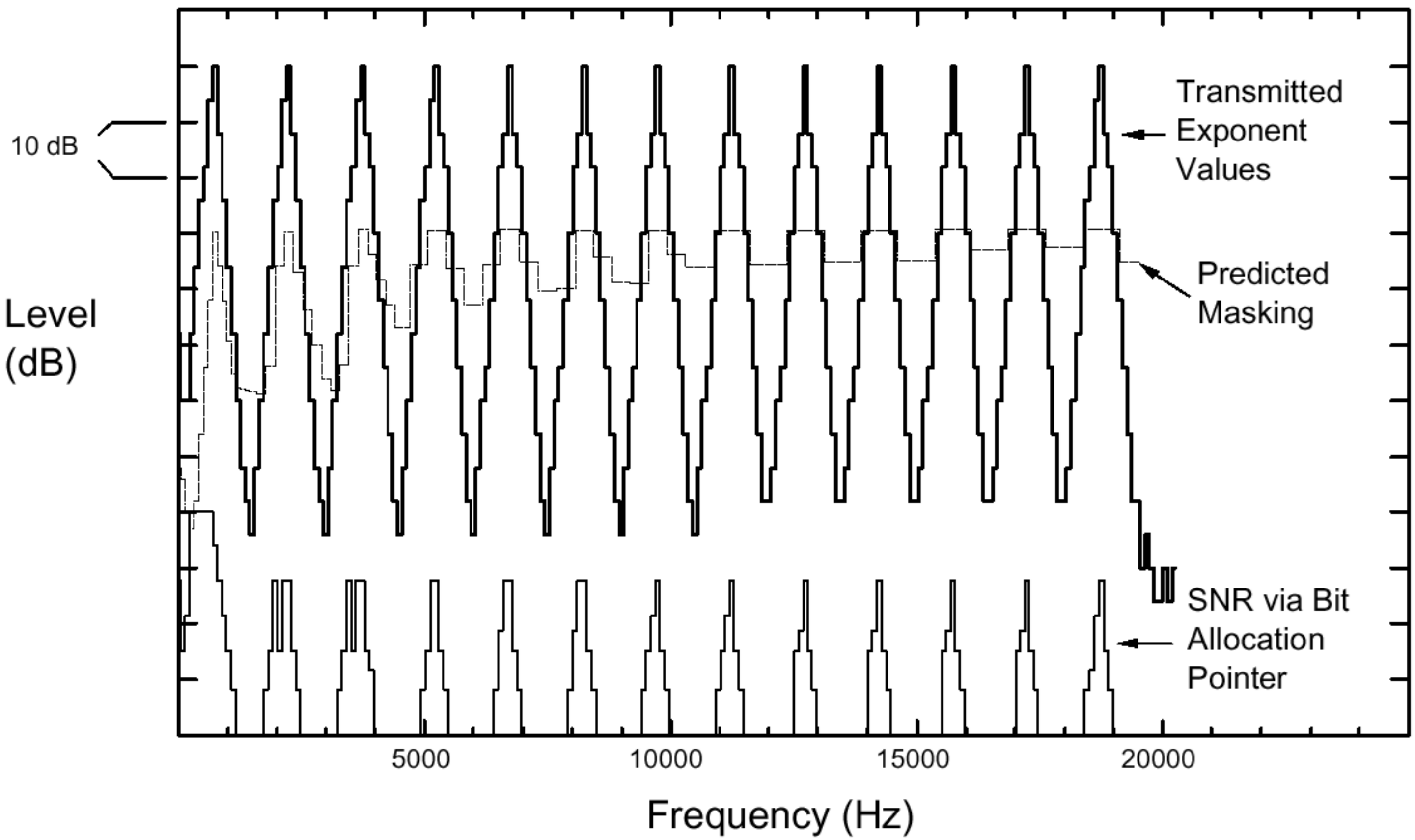
Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage"

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

TECHNISCHE UNIVERSITÄT ILMENAU

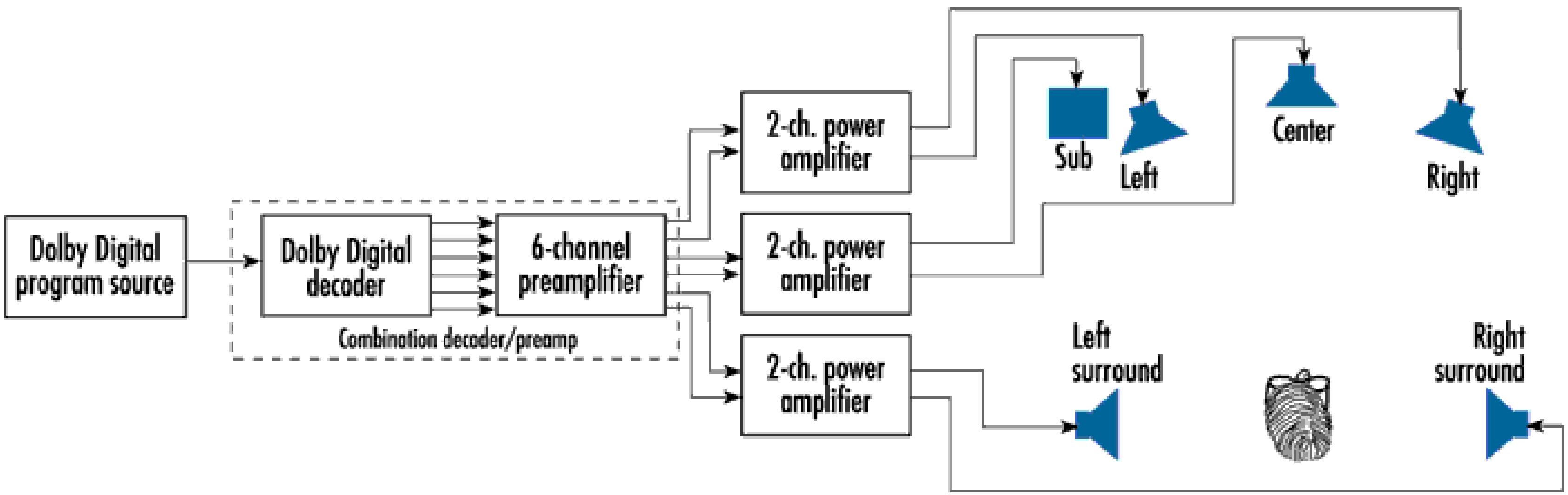
Fraunhofer IDMT

Dolby AC-3: Bit Allocation



Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage"

Dolby Digital Setup



Source: Dolby Labs, Internet

Dolby Digital Enhancement for 6.1-channel audio

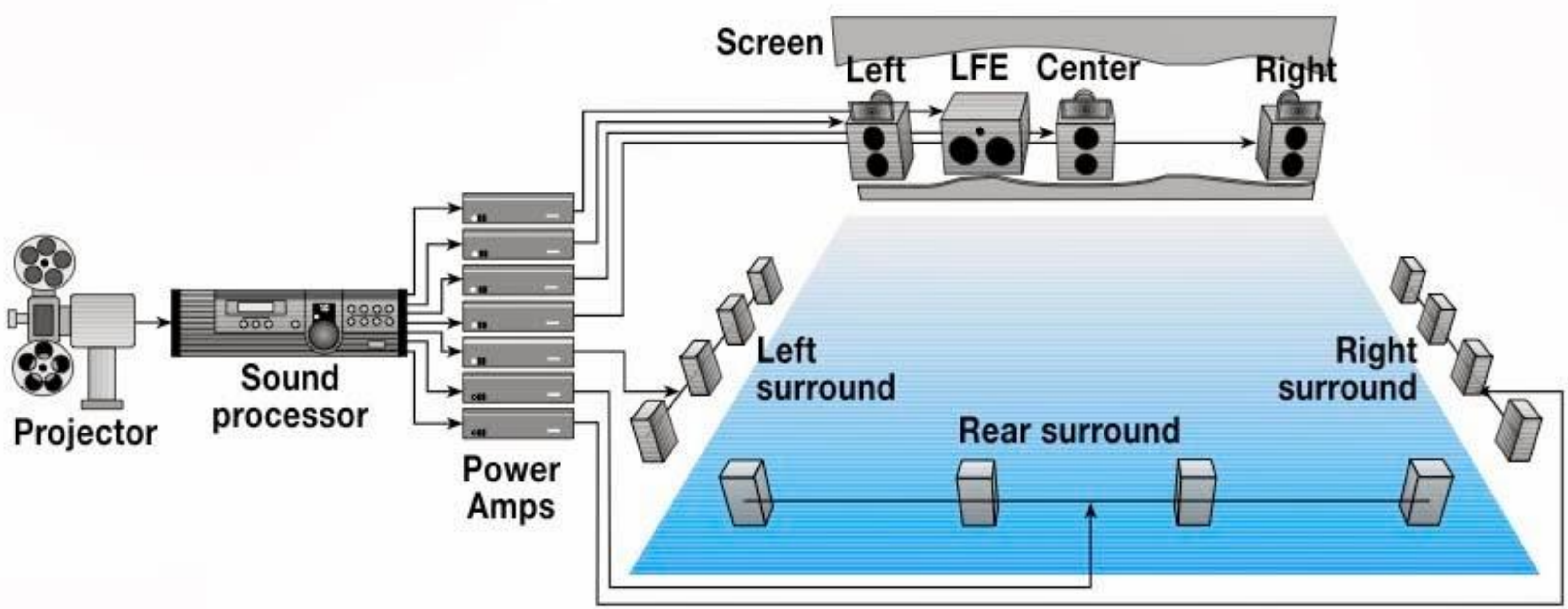


Figure 6: Dolby Digital Surround EX playback

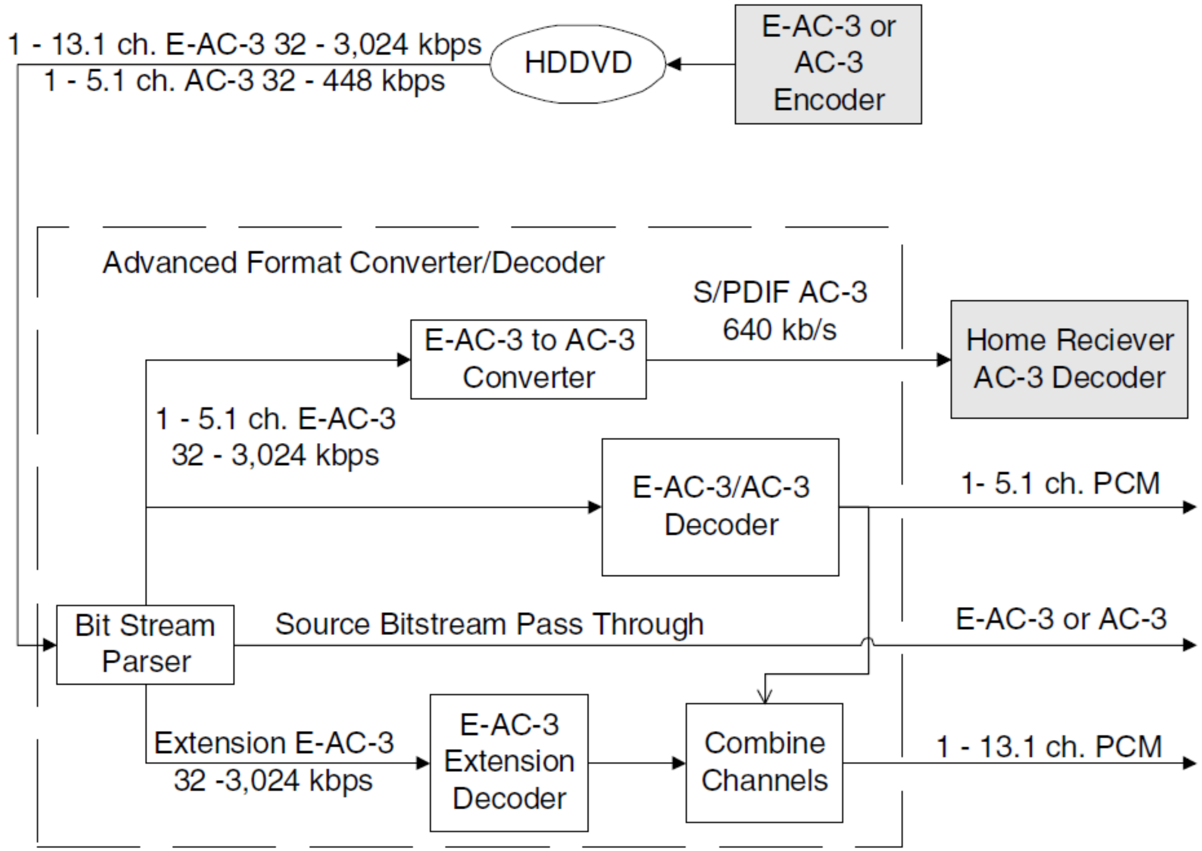
Source: Dolby Labs, „AC-3: Flexible Perceptual Coding for Audio Transmission and Storage"

Dolby Digital Plus (E-AC-3, Enhanced AC-3): Main features

- Greater range of data rates: 32kbps – 6.144 Mbps, fine-grain data rate resolution
- 13.1 channel support
- High resolution hybrid filter bank (AC-3 filter combined with 2nd stage DCT -> 1536 coeffs or subbands)
- New quantization tools
- Improved channel coupling (similar to BCC)
- Spectral extension tool (similar to SBR)
- Transient pre-noise processing
- Based on AC-3: low-loss and low-complexity conversion from E-AC-3 to AC-3

Source: L.D. Fielder et al., "Introduction to Dolby Digital Plus, an Enhancemnet to the Dolby Digital Coding System", 117th AES Convention

E-AC-3: Decoder setup example



Source: L.D. Fielder et al., "Introduction to Dolby Digital Plus, an Enhancemnet to the Dolby Digital Coding System", 117th AES Convention

E-AC-3: Adaptive Hybrid Transform (AHT)

- Based on AC-3 MDCT (256 coeffs or subbands, KBD window with alpha factor 5.0) for easy interoperability, filter length N=512
- 2nd stage DCT Type 2 with M=6 subbands, resulting in 1536 coeffs or subbands $\rightarrow$ higher frequency resolution for stationary signals

MDCT:
$$X(k,m) = \sum_{n=0}^{N-1} x\left(n + m\frac{N}{2}\right) \cdot w(n) \cos\left(\frac{2\pi\left(k+\frac{1}{2}\right)\cdot\left(n+\frac{N}{4}+\frac{1}{2}\right)}{N}\right)$$

k ... MDCT frequency index $\rightarrow k = 0, \dots, N/2 - 1$
n ... MDCT sample index
m ... MDCT blockindex $\rightarrow m = 0, \dots, L-1$; L: total number of Blocks

DCT-2:
$$X^{AHT}(kM + l) = \frac{\sqrt{2}}{M} C_l \sum_{m=0}^{M-1} X(k,m) \cdot \cos\left(\frac{(2m+1)l\pi}{2M}\right);$$

l ... DCT-II frequency index $\rightarrow l = 0, \dots, M-1$
m ... DCT-II sample index

$$C_l = \begin{cases} \frac{1}{\sqrt{2}}, l = 0 \\ 1, l \neq 0 \end{cases}$$

Source: L.D. Fielder et al., "Introduction to Dolby Digital Plus, an Enhancemnet to the Dolby Digital Coding System", 117th AES Convention

E-AC-3: Spectral Extension Tool

- Parametric description of high frequency region of the high frequency subband coeffs, which then are transmitted as this parametric description
- Spectral extension bands approx. match Critical Bands
- For each band an energy ratio and a noise blending parameter is calculated

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

E-AC-3: Spectral Extension (1)

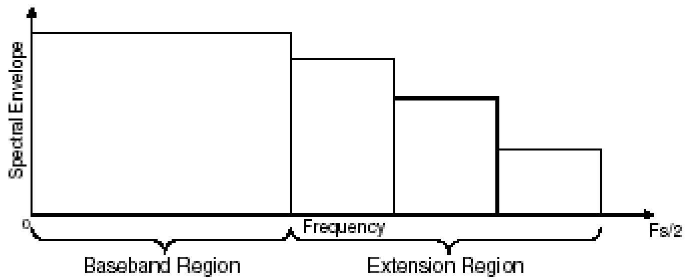


Fig 1: Original Spectrum

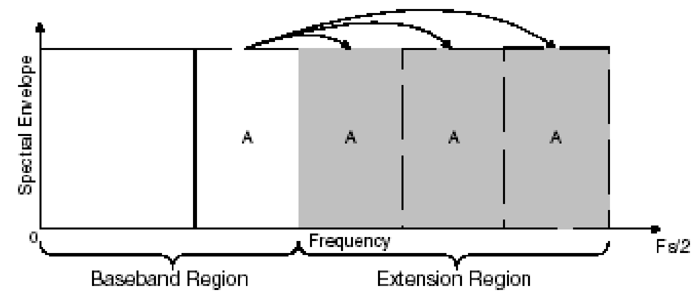


Fig 2: Decoder: Translation

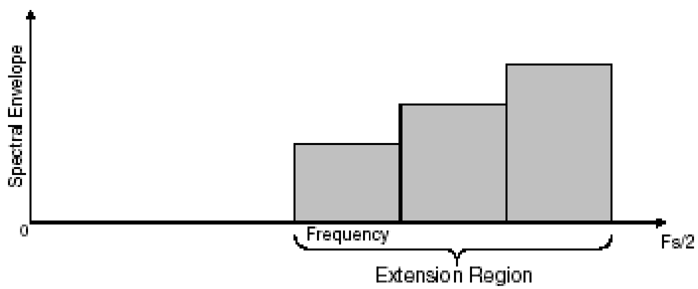


Fig 3: Decoder: Noise spectrum multiplied by blending function

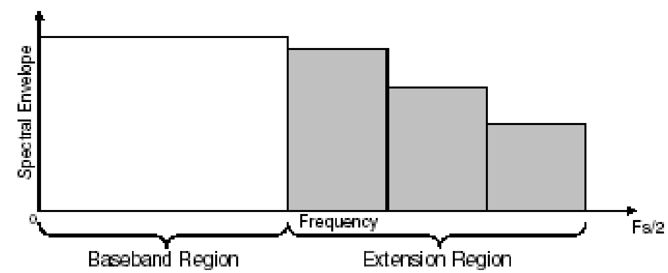


Fig 4: Decoder: Translated spectrum, multiplied by inverse blending function

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

E-AC-3: Spectral Extension (2)

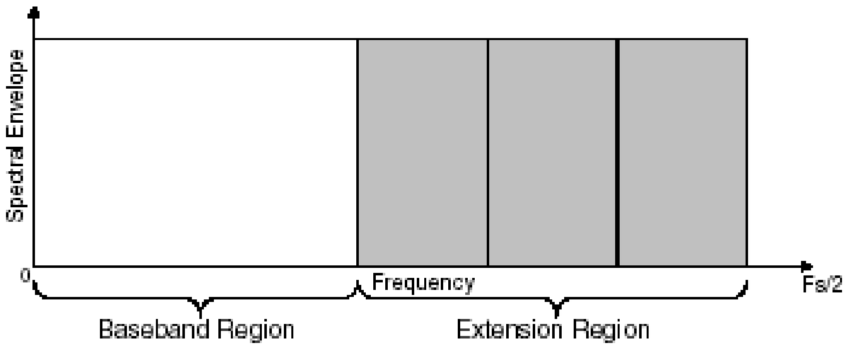


Fig 5: Decoder: Blended spectrum (blending of noise and translated spectrum)

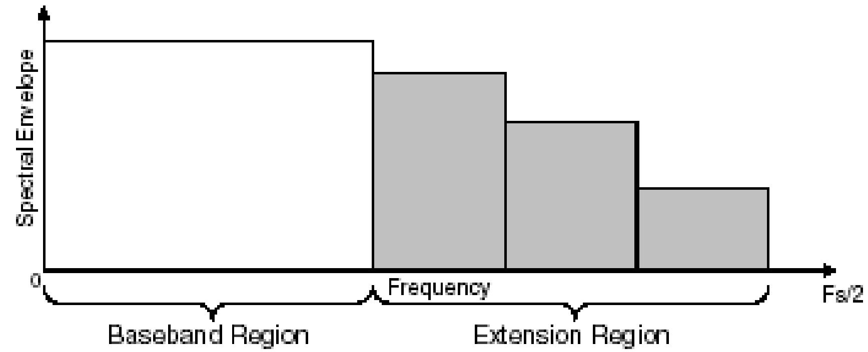
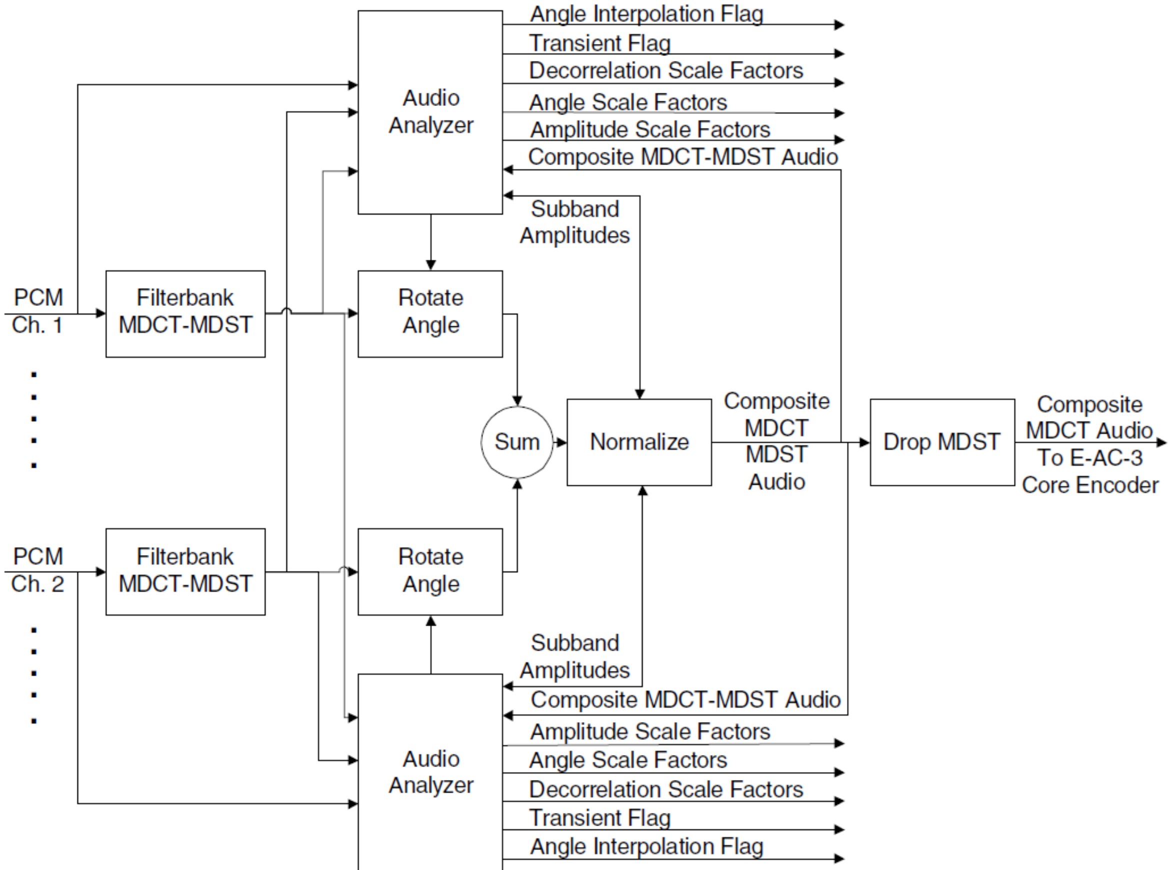


Fig 6: Decoder: Final spectrum, multiplication by transmitted energy ratios

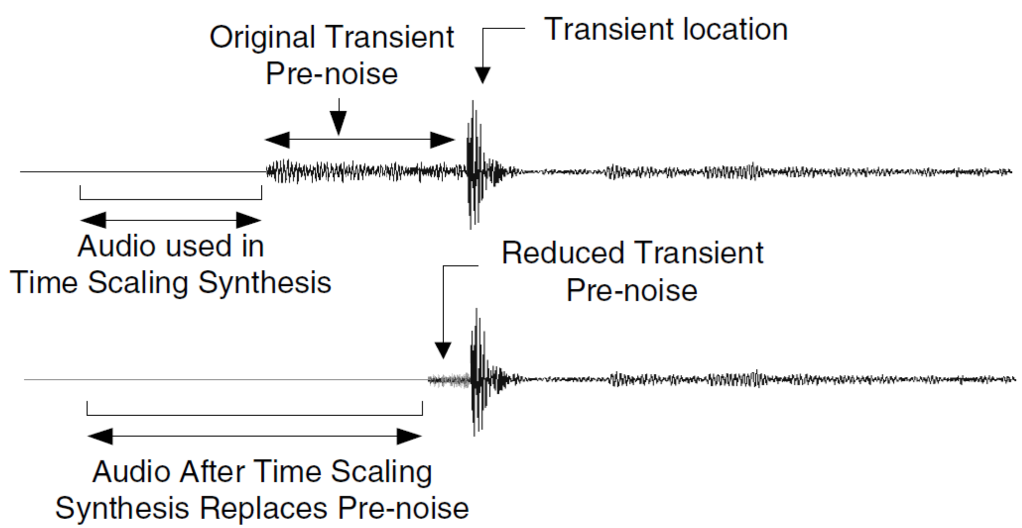
E-AC-3: Enhanced Coupling



Source: L.D. Fielder et al., "Introduction to Dolby Digital Plus, an Enhancemnet to the Dolby Digital Coding System", 117th AES Convention

E-AC-3: Transient Pre-Noise Processing

- reduces pre-echo artifacts with a time-domain strategy
- a time-scaled part of the signal substitutes quantization noise just before a transient



Source: L.D. Fielder et al., "Introduction to Dolby Digital Plus, an Enhancemnet to the Dolby Digital Coding System", 117th AES Convention

Digital Sourround

DTS

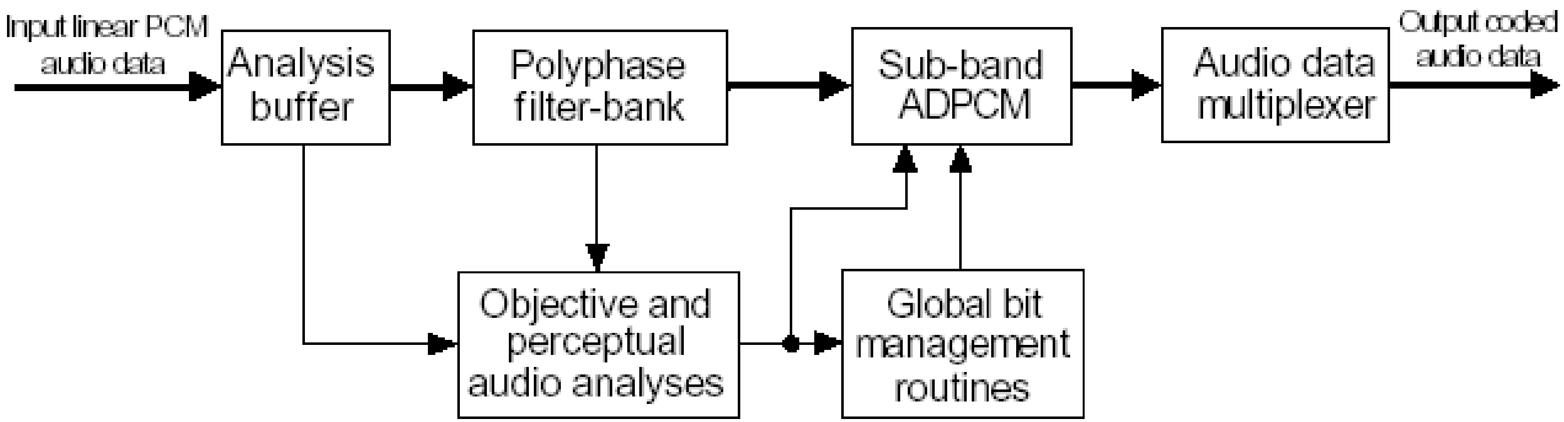
DTS: Coherent Acoustics coding or Digital Surround®

- Intended for entertainment and professional use
- Optional coding scheme for DVD
- Part of the Blue Ray audio standard
- Audio data rates from 8 to 512 kbit/s/channel
- Sampling rates up to 192 kHz / 24 bit
- 5.1 core coder with up to 1536 kbit/s

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

DTS: Encoder overview

Two main stages: polyphase filtering and subband-ADPCM



Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

DTS: Polyphase Filter Bank

- 32 subbands
- Frames of 256, 512, 1024, 2048 or 4096 samples
- Long frames mainly used for low bit rates (coding efficiency)
- Two filter banks: perfect reconstruction (high bit rates) and near-perfect reconstruction (lower bit rates)

Type	Taps	Transitional bandwidth [Hz]	Stop-band rejection [dB]	Ultimate Rejection [dB]	Reconstruction resolution [dB]
NPR	512	300	110	120	90
PR	512	350	85	90	145

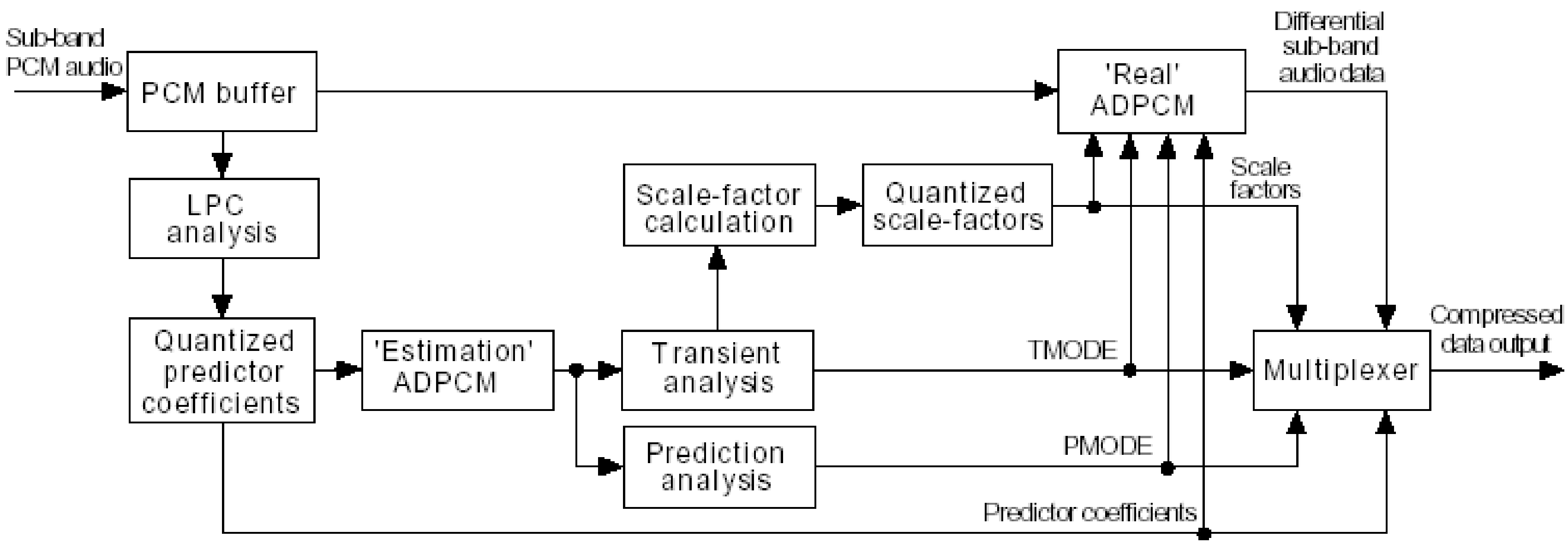
Example: Polyphase filter banks at 48 kHz

DTS: Subband Adaptive DPCM

- Reduce sample-to-sample correlation within each subband
- Disengageable within each subband, if simple PCM renders better results
- Forward prediction based on LPC analysis

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

DTS: Subband Adaptive PCM (block diagram)

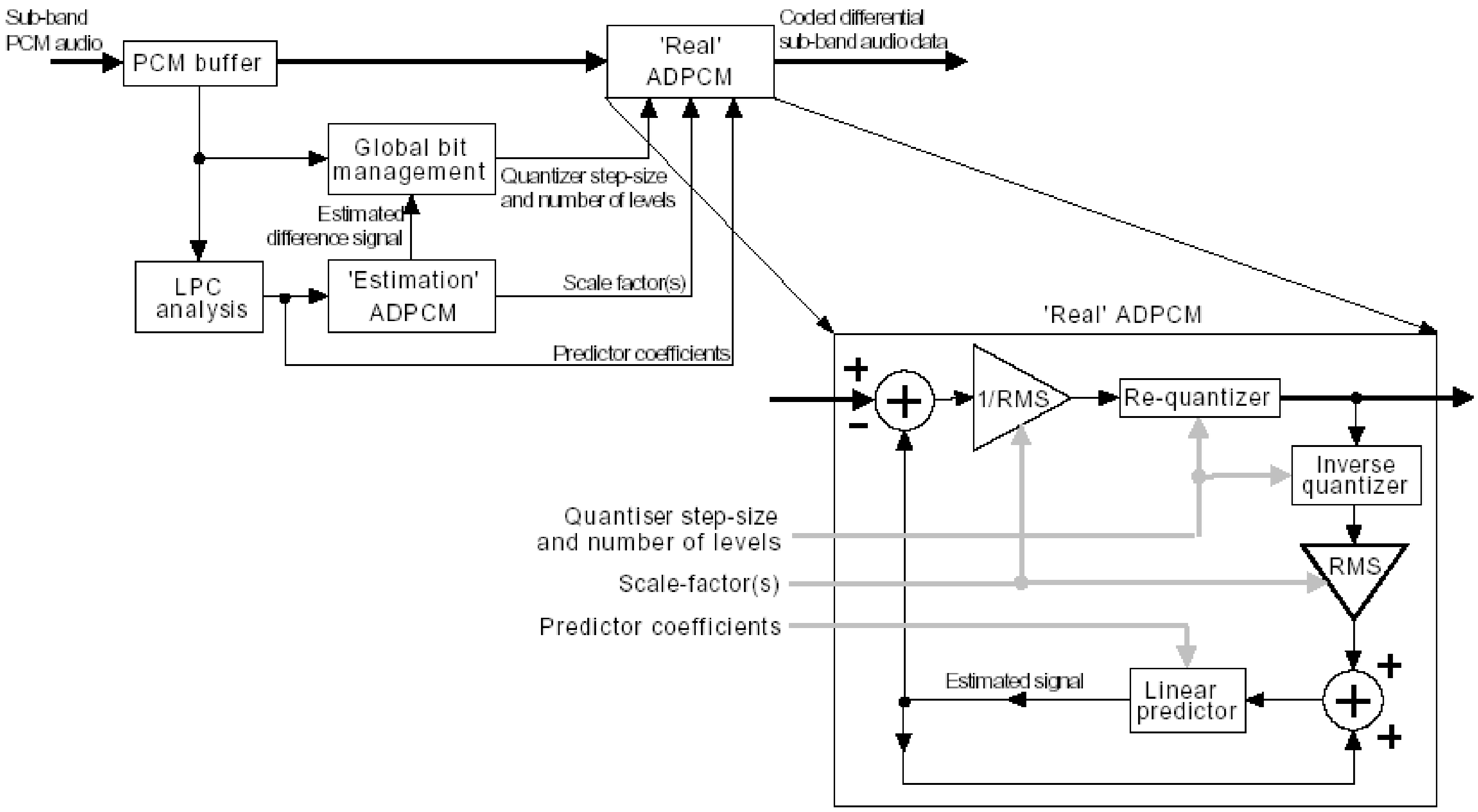


DTS: Quantization and Bit Allocation

- 28 different mid-tread quantizers up to 16,777,216 levels
- Psychoacoustically controlled
- Optional table-based entropy coding at low bit rates

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

DTS: Quantization and Bit Allocation (block diagram)

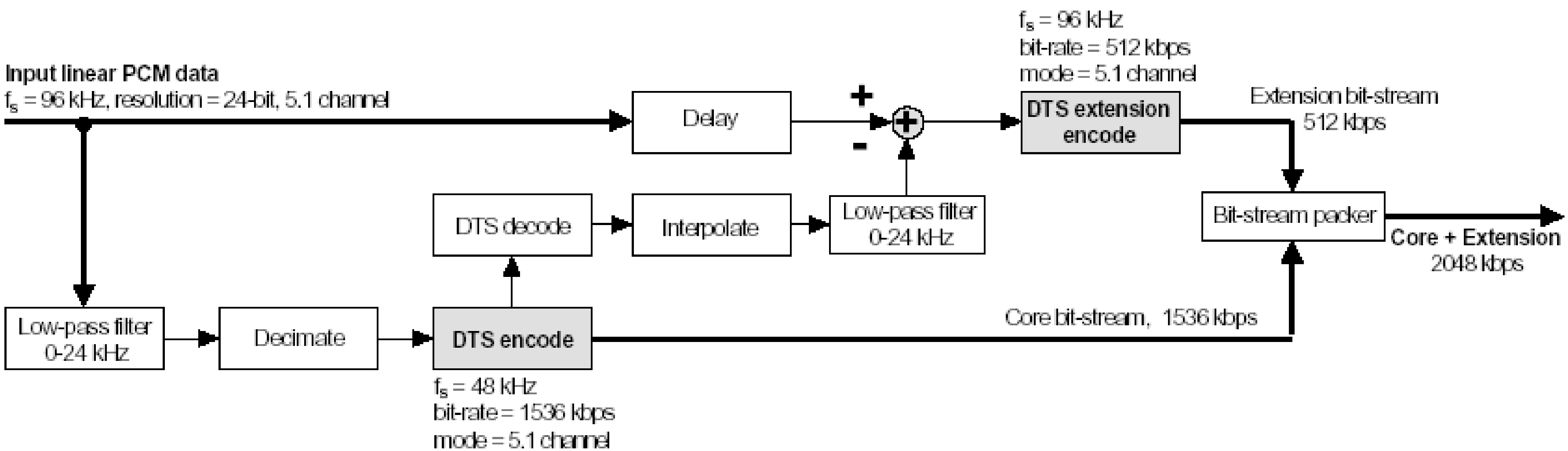


Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

DTS: Example for the use of the extension audio data

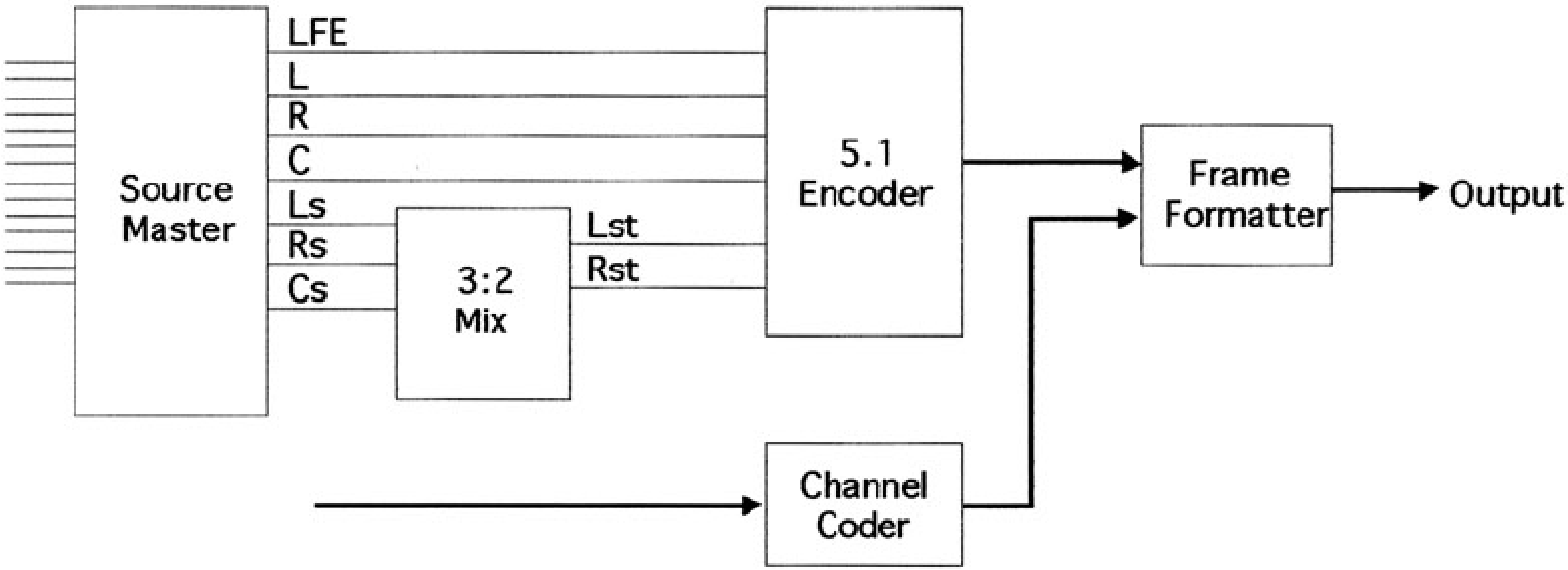


DTS-ES: discrete 6.1 multi channel coding

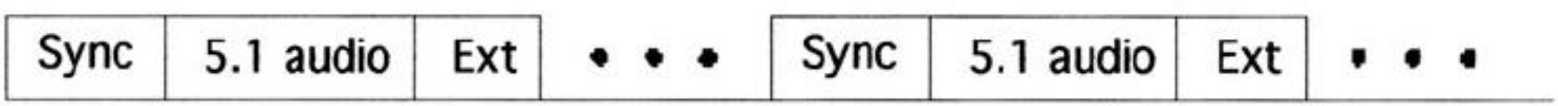
- 5.1 channel DTS core + additional Center Surround channel
- Additional channel is transmitted using Extension Audio Data
- Backwards compatible to 5.1 DTS core coder
- Three possible decoder setups possible:
  - 5.1 decoding with phantom source
  - Matrix decoding of Center Surround Channel
  - Discrete 6.1 decoding by evaluating Extension Audio Data

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

DTS-ES: Encoder block diagram and Bit Stream



DTS-ES Encoder



DTS-ES Bit Stream

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

DTS-HD

- DTS Digital Surround (DTS 5.1 core) mandatory for HD-DVD and Blu-Ray
- DTS-HD is optional for (HD-DVD outdated) Blu-Ray
- DTS-HD is a set of extensions to DTS core, encompassing DTS core, DTS-ES, Neo:6 and DTS 96/24
- Lossless audio coding possible

Ogg Vorbis

- Ogg project started 1993 to provide a license-fee free audio coder/decoder
- Ogg: file transport protocol
- Vorbis: audio coder
  - Psycho-acoustically controlled forward adaptive monolithic codec based on MDCT
  - Inherently variable bit rate coder
  - Provides no framing, synchronization or error protection by itself (therefore use Ogg for file transport, RTP for multicast)
  - Low-complexity decoder, but high memory usage due to non-static probability models
  - Huffman and VQ codebooks are transmitted within bit stream header

Windows Media Audio (WMA)

- Proprietary Audio Coder developed by Microsoft
- Collection of profiles for different applications:
  - WMA 9: most scenarios, backwards compatible to WMA 8, about 20% lower data rate, VBR possible
  - WMA 9 professional: 24 bit/96 kHz audio, 7.1 channels, 128-768 kbps, stereo downmix available
  - WMA 9 voice: speech content at low bit rates (<20 kbps)
  - WMA 9 lossless: compression depending on input audio, used for high-quality archiving purposes

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

WMA: main features

- MDCT (or MLT) based
- Multiple numbers of frequency lines (128, 256, 512, 1024, 2048)
- Sinusoidal shaped windows, transition windows and “bridge” windows (“soft” transition between long and short blocks)
- Uniform quantization within scale factor bands
- M/S coding frame-by-frame instead of scale-factor-band-wise
- Bit reservoir available (1-pass and 2-pass coding)

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

Next Lexture:

Low Delay Coding
14.01.2019

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de