Coding of Stereophonic Signals

Prof. Dr.-Ing. Karlheinz Brandenburg

Fraunhofer IDMT & Ilmenau University of Technology
Ilmenau, Germany

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Introduction

Goal: produce spatial acoustic impression

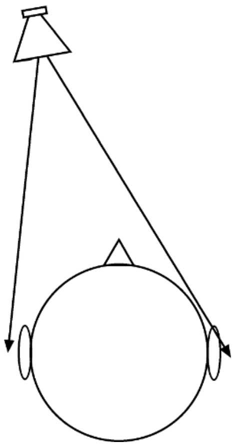
Approach:

- Look at information the ear is using
- Psycho-acoustic effects

Psycho-Acoustics - Spatial Hearing (1)

Interaural time difference (ITD)

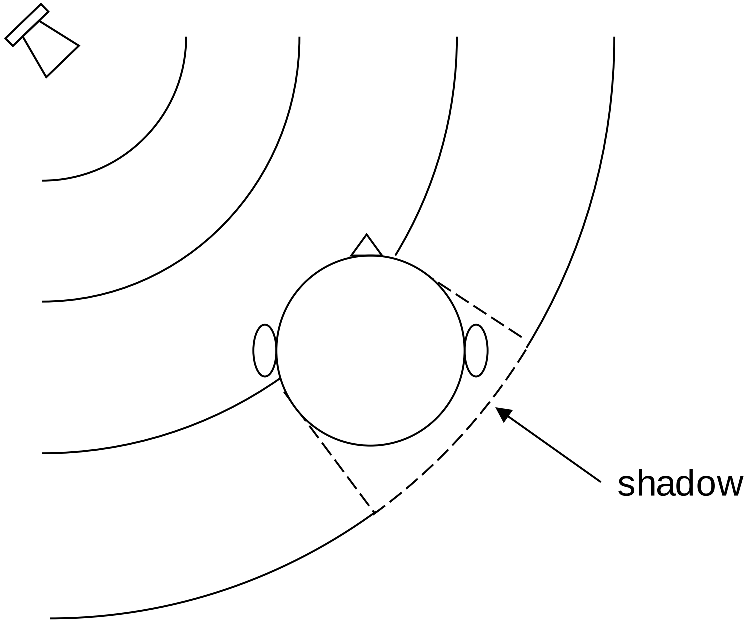
- different ear arrival times due to different propagation paths



Psycho-Acoustics – Spatial Hearing (2)

Interaural level difference (ILD)

- Different sound pressure levels due to shadowing of the head (for high frequencies)



Psycho-Acoustics – Spatial Hearing (3)

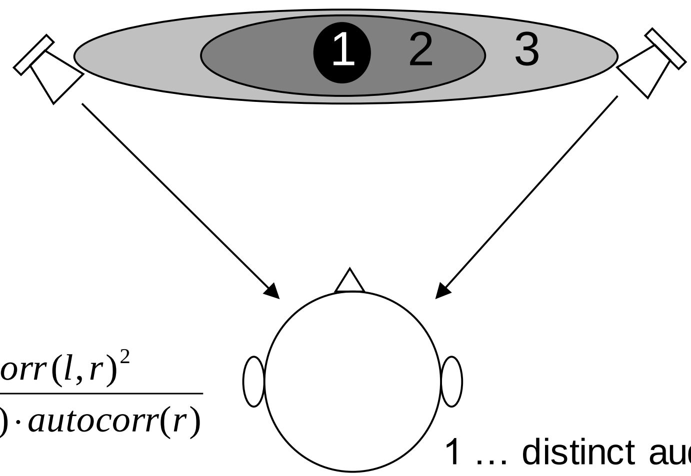
Interaural Coherence (IC)

→ degree of "similarity" between left and right ear entrance signals

Inter-Channel Coherence (ICC)

→ degree of "similarity" between left and right source signal

width of auditory object increases (1-3) as IC/ICC between left and right ear/channel decreases



$$IC = \frac{\left|\sum_k x_l(k) x_r^*(k)\right|}{\sqrt{\sum_k x_l(k) x_l^*(k) \cdot \sum_k x_r(k) x_r^*(k)}} = \frac{crosscorr(l,r)^2}{autocorr(l) \cdot autocorr(r)}$$

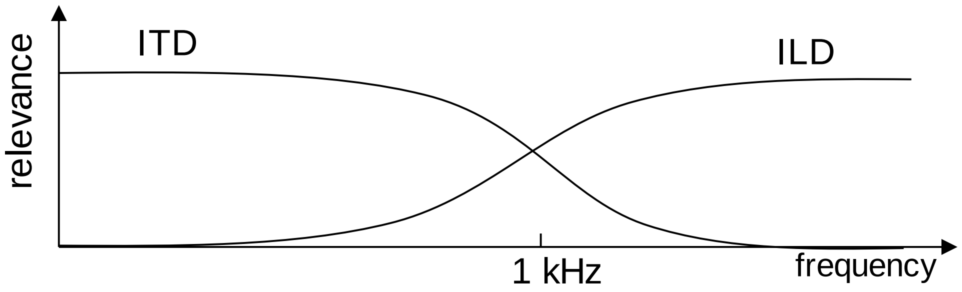
X_l and X_r denote complex valued signals resulting from filter bank

1 … distinct auditory object; IC = 1
2 … 0 < IC < 1
3 … diffuse auditory object; IC = 0

Psycho-Acoustics Irrelevance (4)

Duplex theory

- ITD dominate localization at low frequencies
- ILD at high frequencies
- Transition between ITD and ILD dominance is about 1-2kHz



ILD and ITD can be intercharged within limits

BUT: precedence effect (Haas-Effect)!

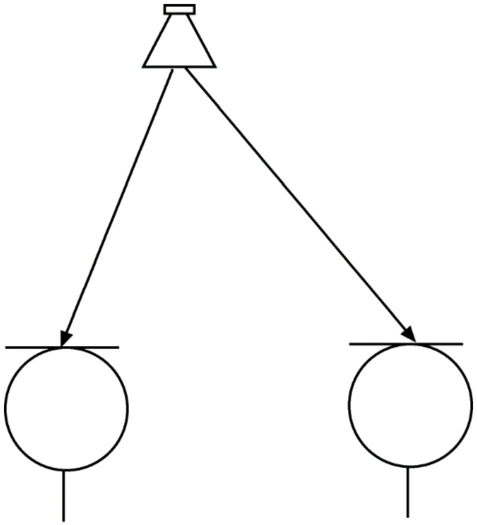
Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

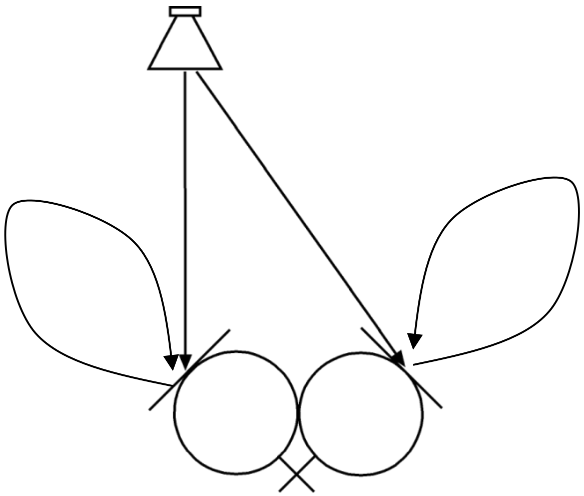
Recording

Delay stereo



Intensity stereo

(pan pot)



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

Coding

Coding: Distribution of spatial quantization noise?

- signal and noise from same apparent spatial position → optimum solution
- Signal and quantization noise from different spatial positions: quantization noise becomes more audible

Difference in masking levels:

Binaural Masking Level Difference (BMLD) between 0dB to 18dB

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

BMLD experiment

- Masker (audio signal) identical to both ears

1. Test tone (quantization noise) identical for both ears, at masking threshold

   → inaudible

2. Same test tone, but different phase relationship between ears

   → audible

- Difference in masking threshold: BMLD

Separate Stereo Coding

- Simplest method for stereo coding
- Problem: e.g. audio signal in center
  - signal appears in center
  - quantization noise is uncorrelated from sides
  - → unmasking can result (BMLD)
- Redundancy between channels not used for compression

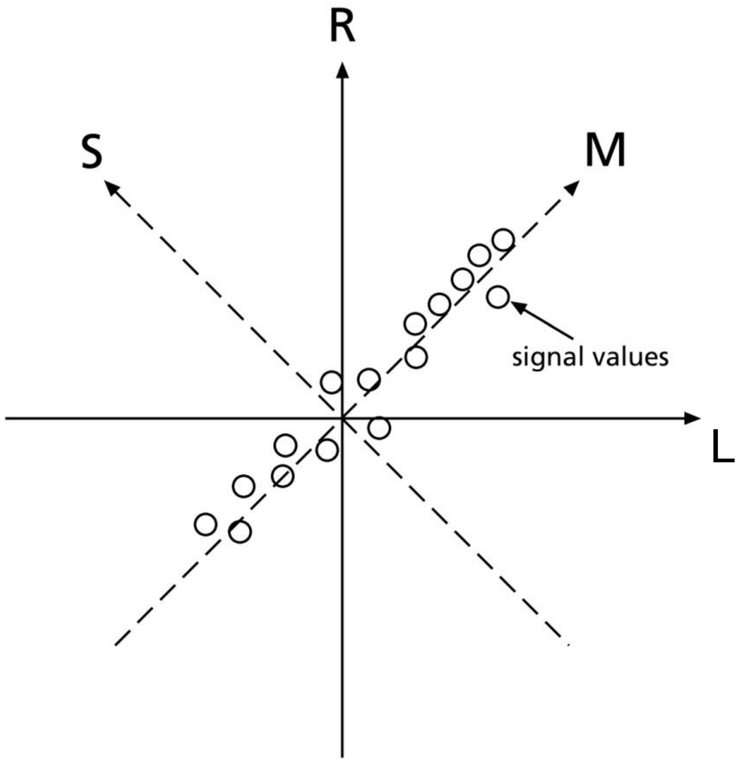
Mid/Side (M/S) Coding (1)

- Matrixing, sum/difference
- For encoding: $M = \frac{L+R}{2}$ $S = \frac{L-R}{2}$
- For decoding: $L = \mathrm{M} + \mathrm{S}$ $R = M - S$
- Invertible in the absence of quantization
- Quantization noise: with a factor of $\sqrt{2}$
  - energy conservation

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

Mid/Side (M/S) Coding (2)



TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Mid/Side (M/S) Coding (3)

- For a centered signal S=0:
  - Few bits
  - Little quantization distortion
- Reduced bit-rate (redundancy reduction)
- Quantization noise centered, like signal
- Works for predominantly centered signals

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

Mid/Side (M/S) Coding (4)

- Problem: assume signal only in left channel
  - L=A, R=0
- Matrixing, quantization noise n_1, n_2:

$$M = A\frac{1}{2} + n_1, \quad S = A\frac{1}{2} + n_2$$

→ unnecessary bits needed

- Decoder:

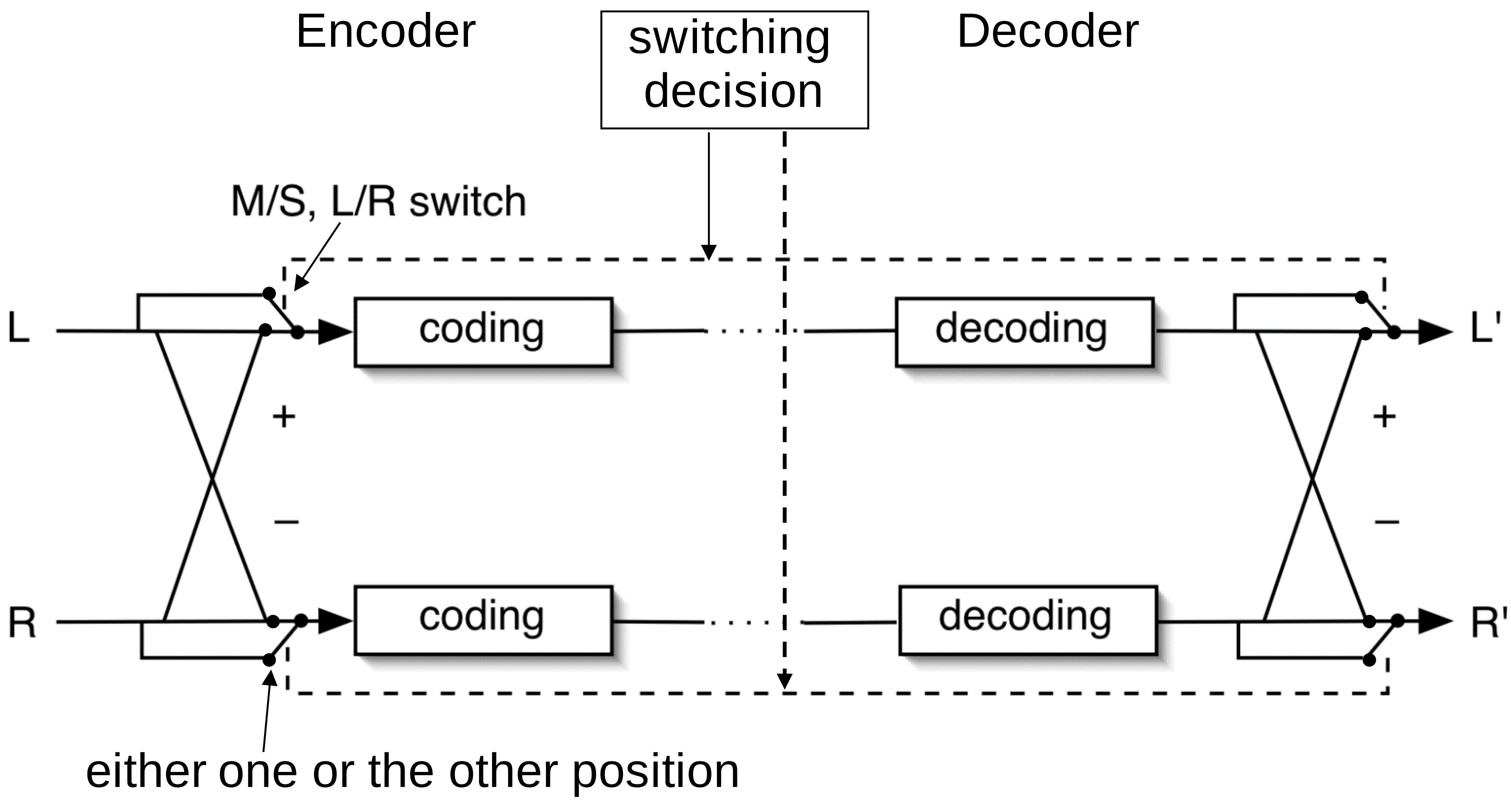
$$L = A + n_1 + n_2, \quad R = n_1 - n_2$$

→ noise spreads to other channel

- For signals with little correlation between channels: switch to separate left/right coding necessary

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

Mid/Side (M/S) Coding (5)

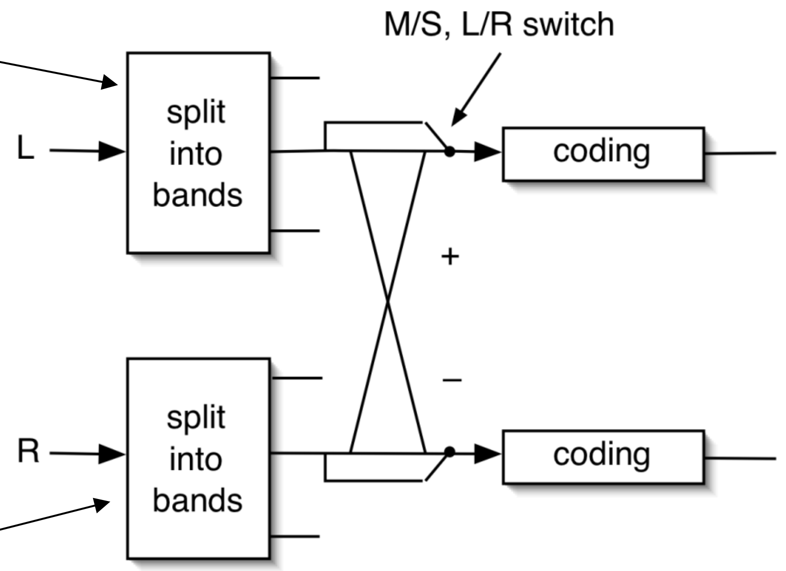


Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

Mid/Side (M/S) Coding (6)

More efficient if M/S, L/R switch is independent in subbands

Somewhat separating different sources or instruments, as much as the ear needs



e.g. groups of subbands of the MDCT of the audio coder

Intensity Stereo (1)

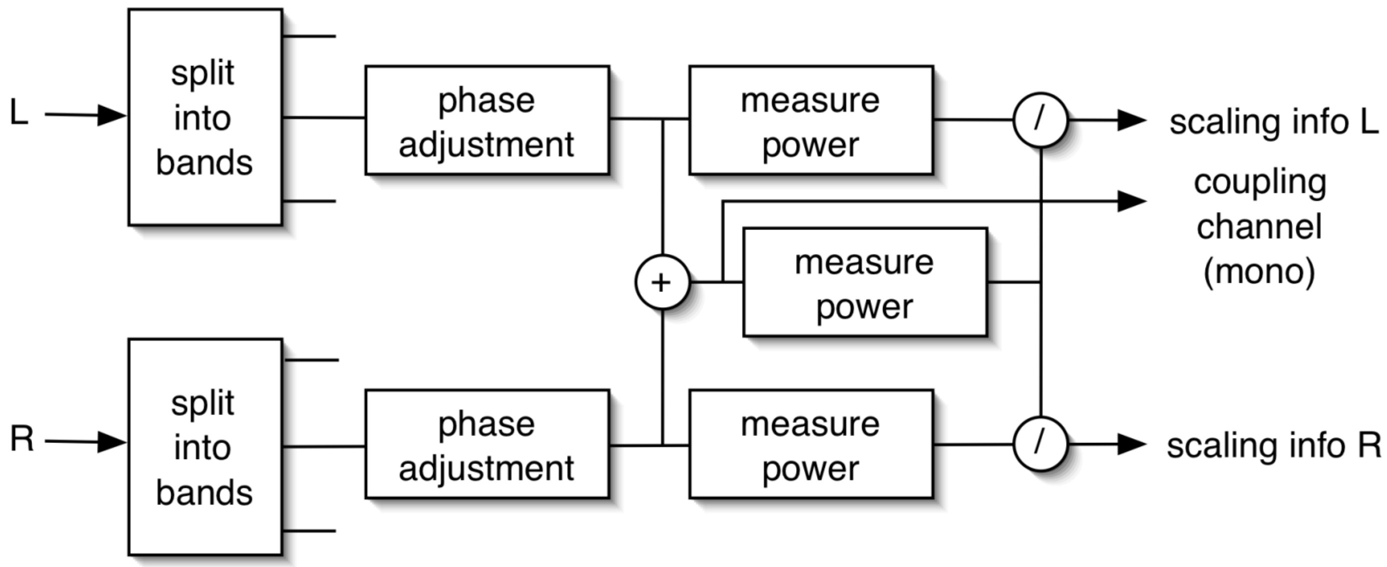
- M/S for higher quality
- Goal for intensity: reduce overhead for stereo → bit reduction
- Lossy coding → ILD
- Usually used above a certain frequency (4 kHz)

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

TECHNISCHE UNIVERSITÄT ILMENAU

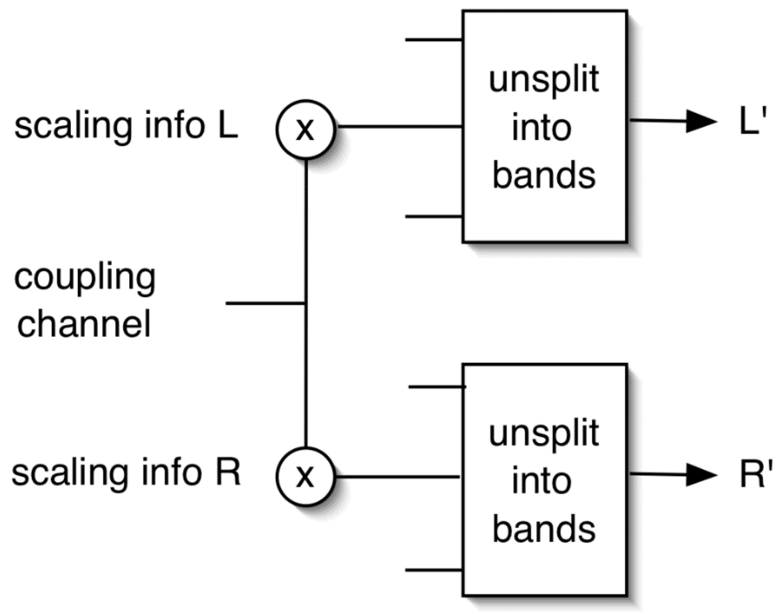
Fraunhofer IDMT

Intensity Stereo (2) - Encoder

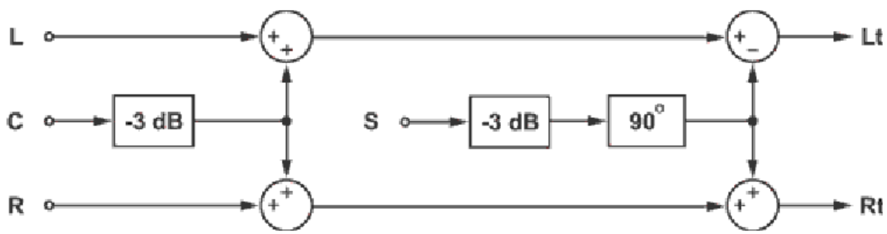


Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

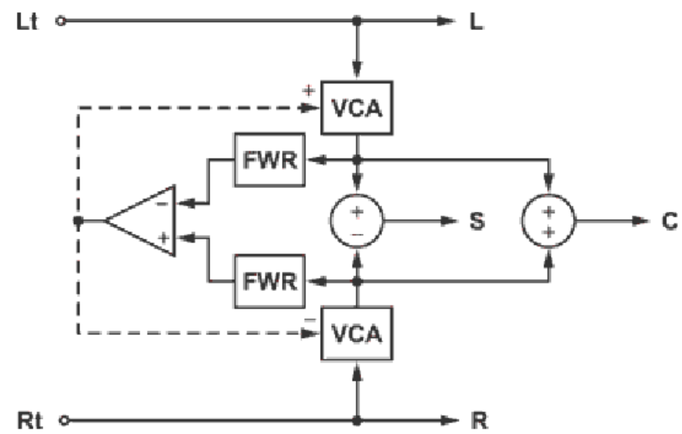
Intensity Stereo (3) - Decoder



Interaction Intensity Stereo ←→ Dolby Pro-Logic



Simple four-input Pro-Logic encoder



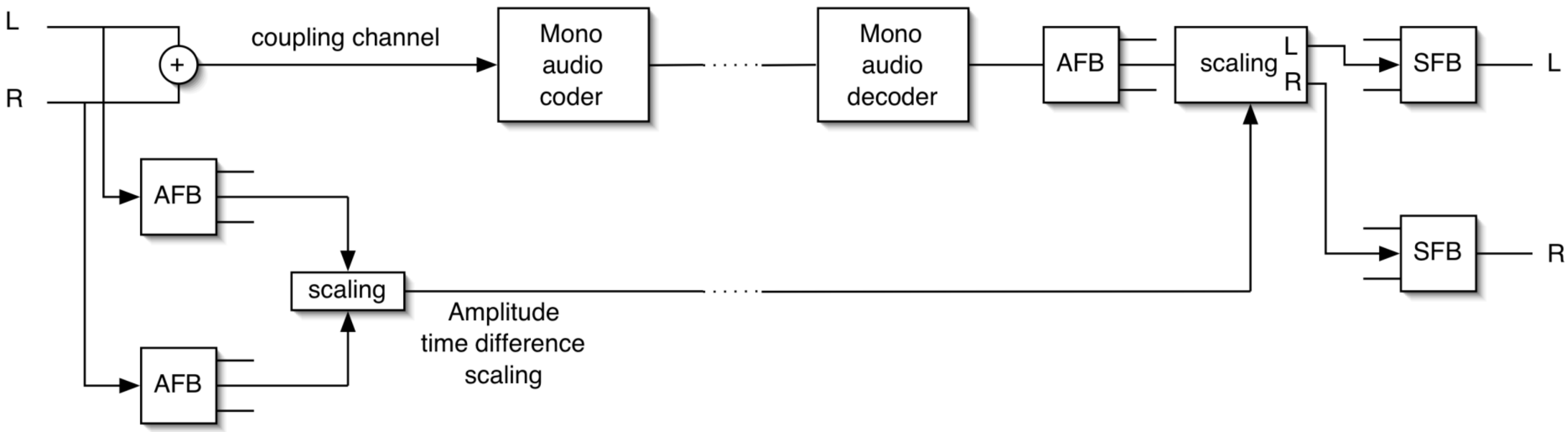
Decoder with Voltage Controlled Amplifier (VCA) and Full Wave Rectification (FWR)

- Pro-Logic: (analog) matrix surround en-/decoding system
- Center and Surround channels are mixed into stereo signal
- Problem: loss of phase information by coding Lt and Rt with intensity stereo coders (crosstalk!)

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

Variants of intensity stereo coding – BCC (1)

- Binaural Cue Coding BCC
  - Lucent/Agere



Variants of intensity stereo coding – BCC (2)

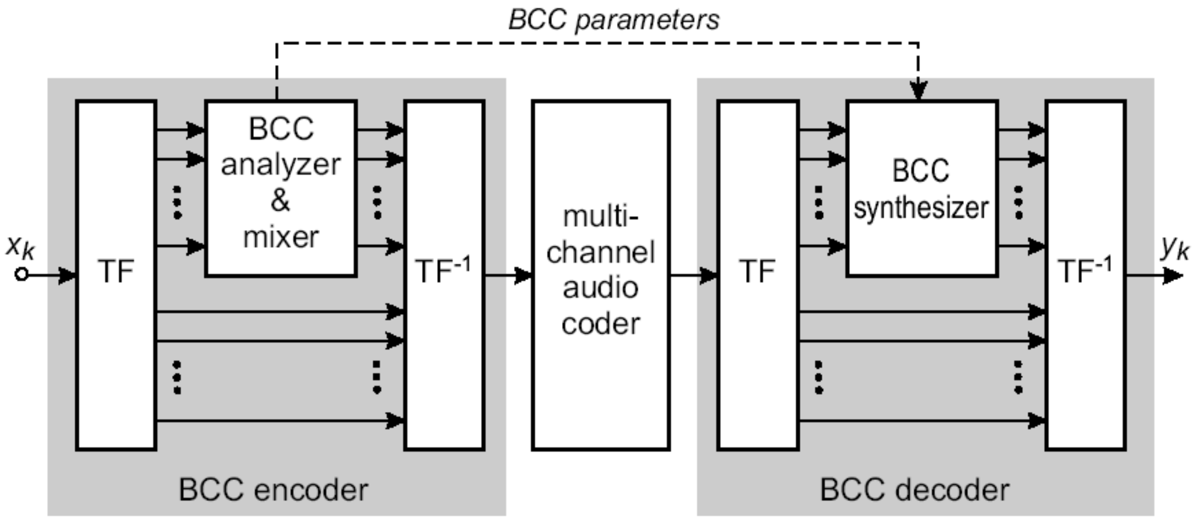
- Differences/Advantages

  intensity stereo ←→ Binaural Cue Coding

  - Sub-band decomposition/time-frequency tradeoff is separate from mono audio codec
  - no lower cutoff frequency, because ITD and ILD and IC are transmitted and resynthesized (remember: duplex theory)
  - better reconstruction of stereo images

Hybrid Coder using Scalable BCC

TF – time-frequency transform

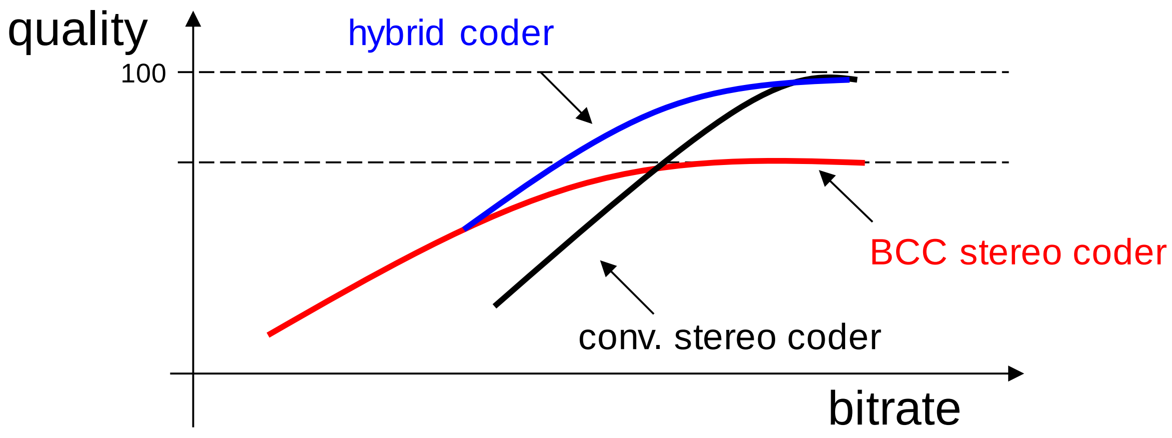


Source: Baumgarte et alt., 116th AES Convention, Berlin

- BCC parameters
  - ITD (interaural time difference)
  - ILD (interaural level difference)
  - IC (interaural coherence)

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

Hybrid Coder using Scalable BCC



Source: Baumgarte et alt., 116th AES Convention, Berlin

- Maximum achievable quality of parametric stereo coders limited due to limited model accuracy
- Hybrid approach closes gap between BCC and conventional coders

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

Parametric Stereo (1)

- Parametric Stereo
  - Is supported within MPEG-4, within the HE-AAC (high-efficiency) profile
  - Delivers ‘good’ quality at bitrates as low as 24-32 kbps, and ‘excellent’ quality around 48kbps on a MUSHRA score (Multi Stimulus with Hidden Reference and Anchors – ITU-R BS.1534-2)
  - Is part of the current “state of the art” MPEG Audio codecs (see later)

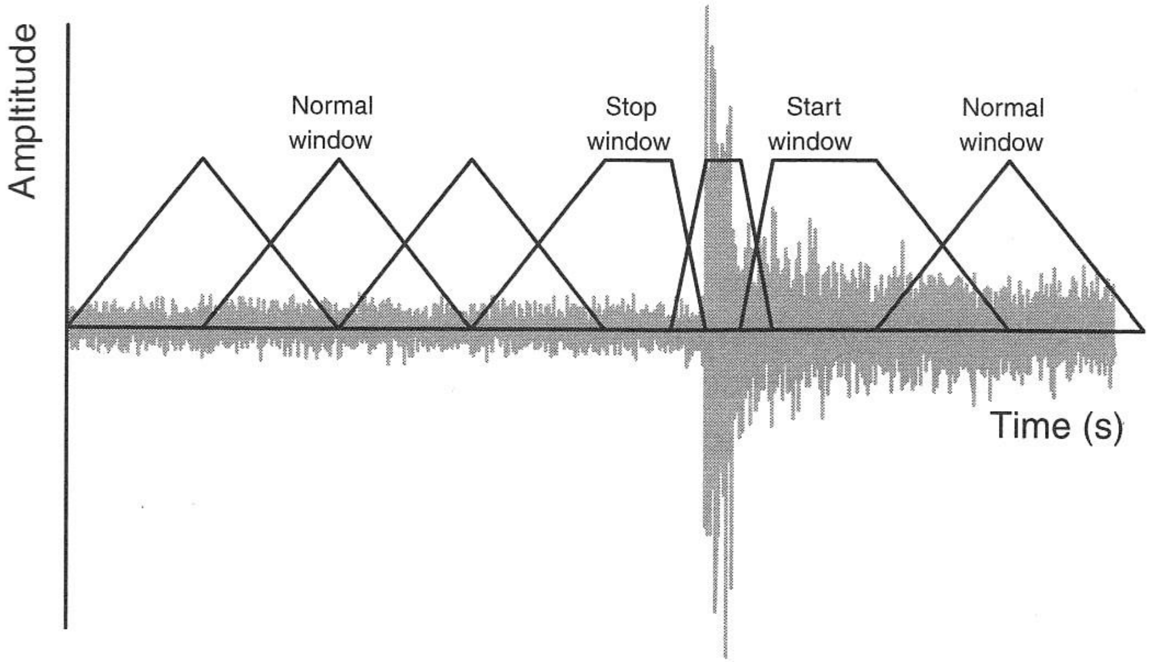
Parametric Stereo (2)

- Parametric Stereo vs. BCC
  - based on identical principles, but different implementation aspects and engineering choices
  - PS supports dynamic segmentation of incoming audio (variable parameter update rate, depending from spatial stability → Haas effect modeling possible)
  - ICTD (inter-channel time difference) is replaced by IPD (inter-channel phase difference) to allow parameterization of out-of-phase signals
  - DFT replaced by QMF filter bank
  - Alternative synthesis for reconstructing out-of-phase signals

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

Parametric Stereo (3)

- Time/Frequency resolution

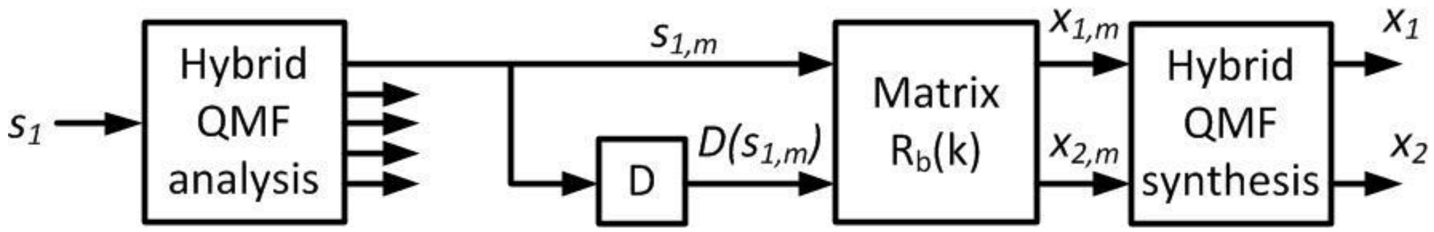


Dynamic window switching in case of transients

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

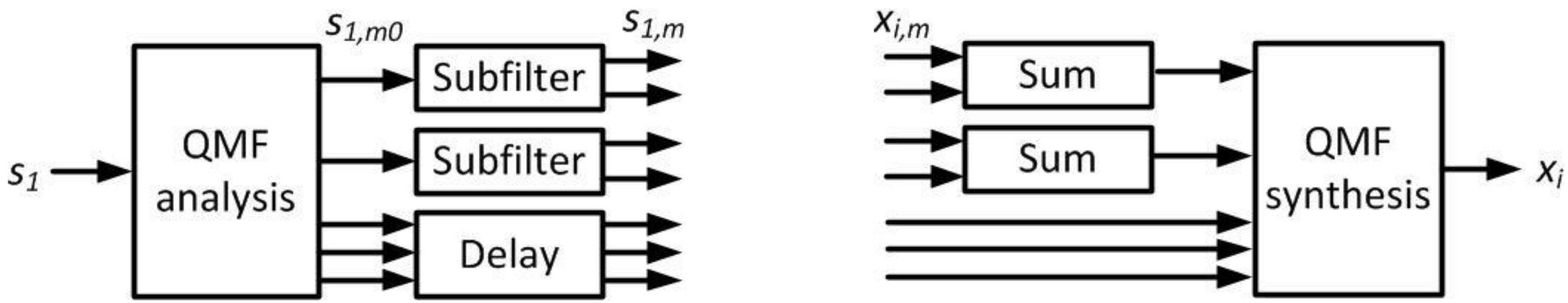
Parametric Stereo (4)

- Parametric Stereo Decoder
  - Two-stage hybrid quadrature mirror analysis filter bank (QMF; extension to the filter bank used in SBR)
  - Incoherent signal is generated by the decorrelator D (i.e. Laudrisen decorrelator) by convolution of the mono input with an all-pass filter
  - Matrixing: mixing and phase-adjustment process
  - Two-stage hybrid QMF synthesis



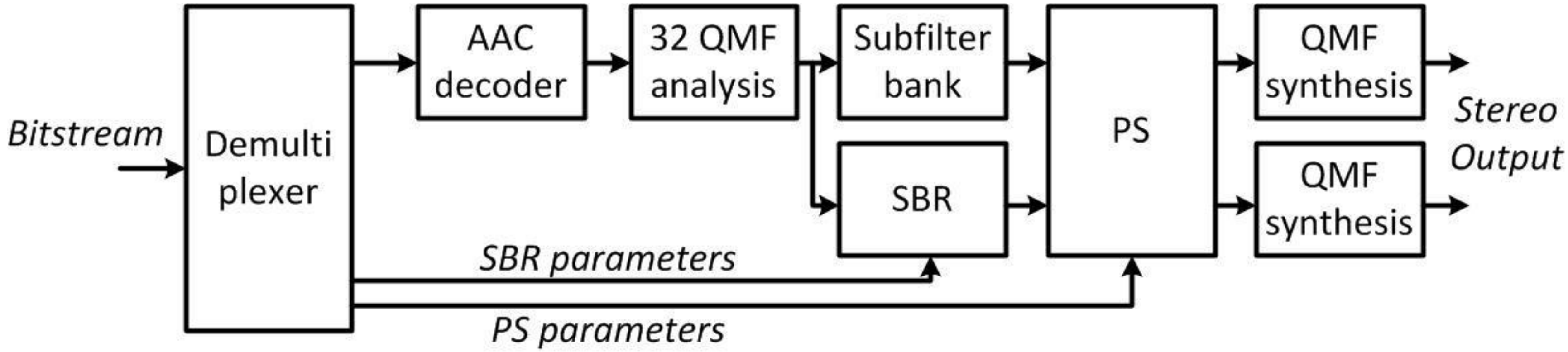
Parametric Stereo (5)

- Structure of the QMF filter banks
  - lowest QMF sub-bands are filtered through a sub-filterbank of Order N in order to enhance frequency resolution
  - Remaining sub-bands signals are delayed by N/2 samples



Parametric Stereo (6)

- PS in HeAAC
  - SBR and PS operate in virtually the same QMF domain → very effective combination resulting in significant complexity reduction
  - Delay of SBR process is identical to the delay caused by sub-filterbank

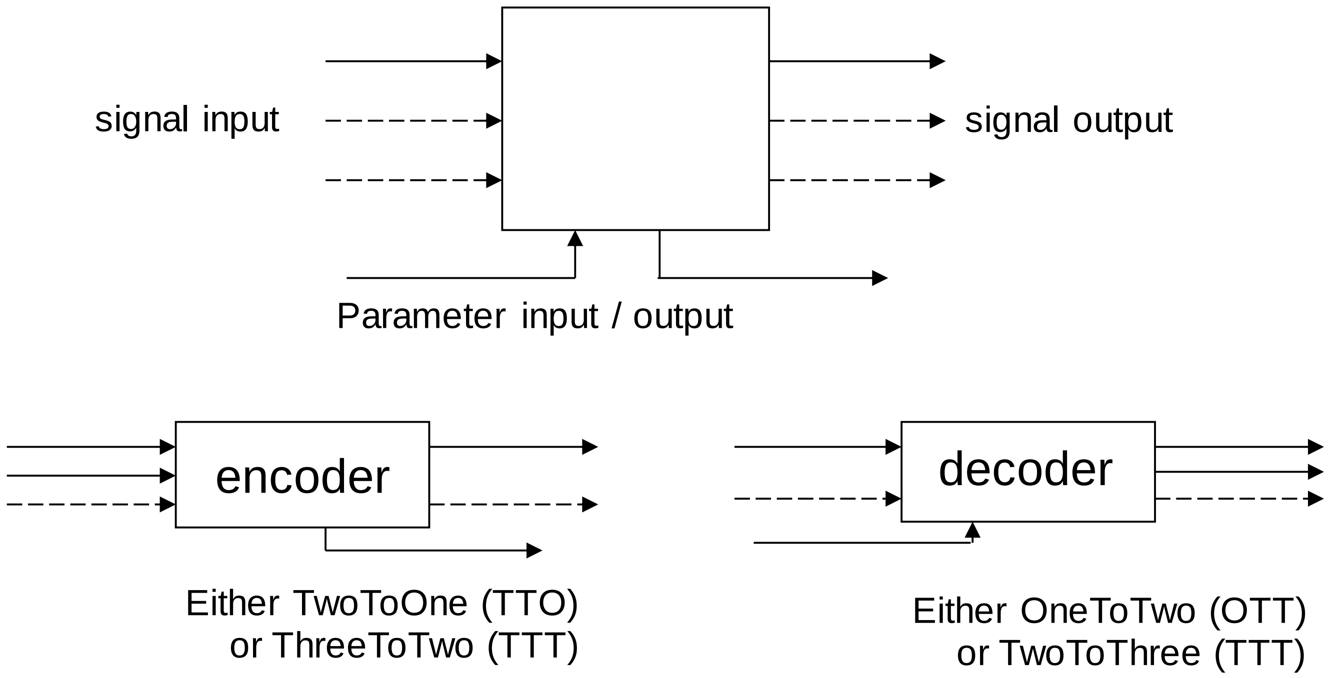


MPEG Surround (1)

- MPEG Surround
  - based on Parametric Stereo
  - spatial parameters enable modification of certain aspects of the downmix
    - matrixed surround compatible
    - artistic downmixes
    - binaural rendering
  - channel configuration of encoder can be different from the channel configuration of the spatial decoder (e.g. rendering of 4.0 from a 5.1 signal configuration without having to decode all 5.1 channels first)
  - Residual coding possible, to enable MPS to support higher quality

MPEG Surround (2)

- Elementary building block

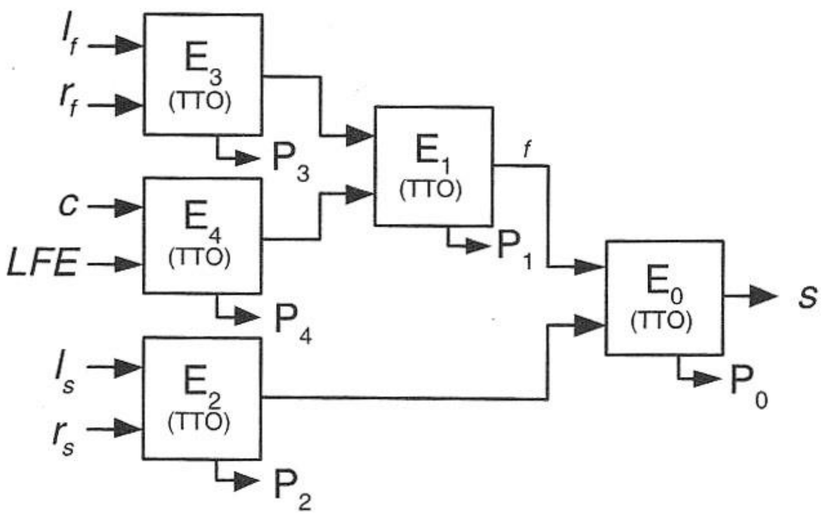


Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

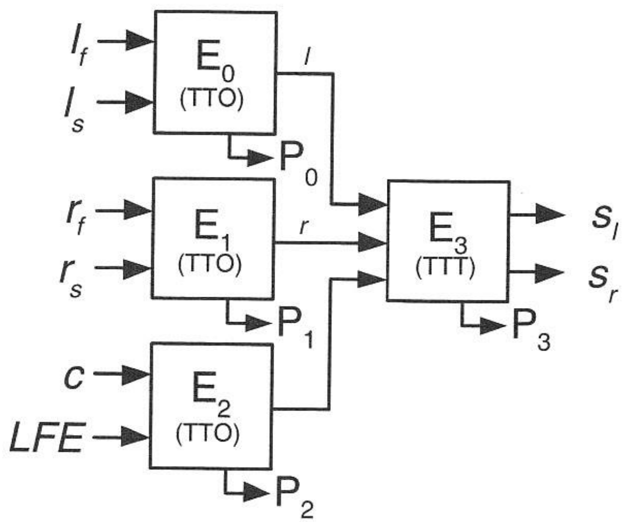
MPEG Surround (3)

- Channel configuration for 5.1 downmix

  $P_x \quad x = 0..4$ Parameter sets with spatial information



downmix to mono

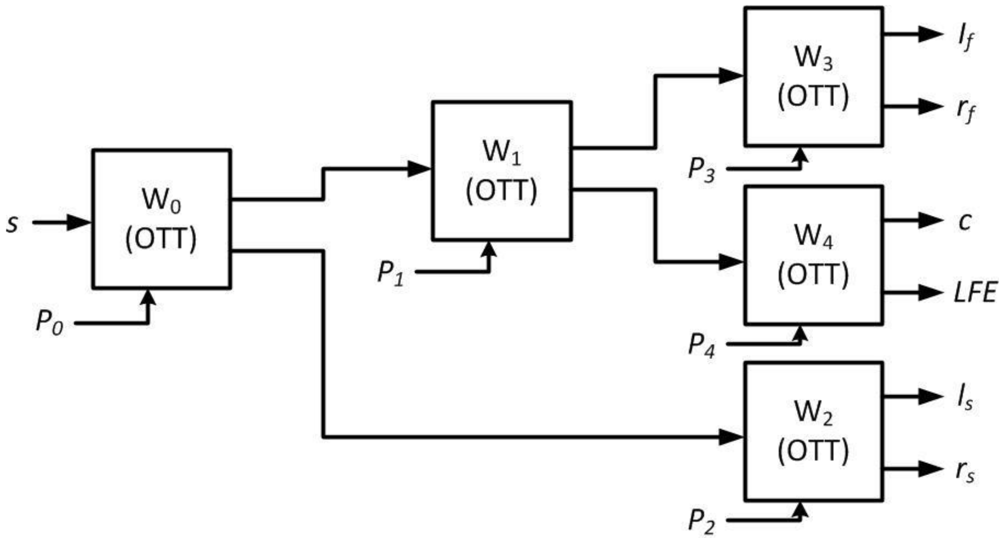


downmix to stereo

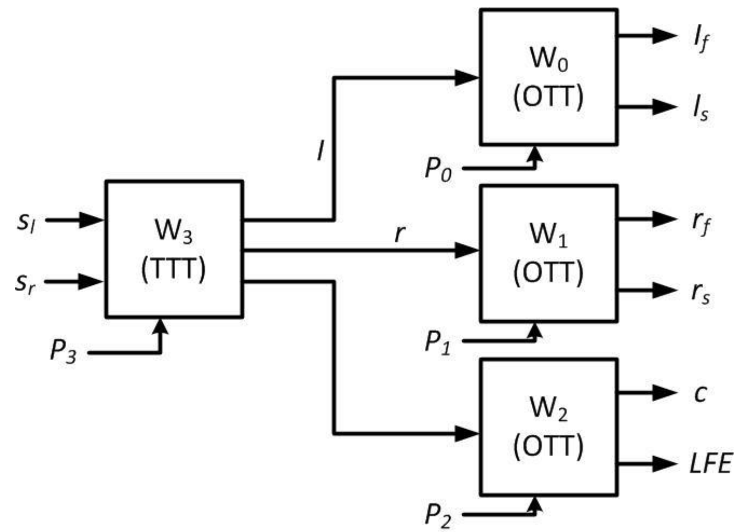
MPEG Surround (4)

- Channel configuration for 5.1 upmix

  $P_x \quad x = 0 \ldots 4$ Parameter sets with spatial information



upmix from mono



upmix from stereo

MPEG Surround (5)

- Matrixed surround conversion block
  - Method to create a *pseudo* surround experience based on a stereo downmix with specific downmix properties
  - Conversion from conventional downmix to matrixed downmix: post-processing stage of the encoding tree
  - Conventionally: downmix such that surround signals are in anti-phase → impossible to retrieve original input channels
  - MTX conversion: dynamically varying, invertible matrix; dependent on the spatial parameters →decoder can 'undo' the processing

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

MPEG Surround (6)

- Coding of residual signals
  - TTO and TTT encoding can create residual signals
  - They can be bit-efficiently encoded and transmitted along with downmix and spatial parameters
  - This allows full waveform reconstruction

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

Applications – What is used Where

	PAC	Lossless Coders	MP3	AC3	AAC
M/S full band		X	X		
M/S sub band	X			X	X
Intensity	X		X	X	(X)
PS					X
MPEG Surround			X	X	X

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

Conclusions

- M/S stereo is the most widely used stereo tool
  - used at medium bit-rates and above
  - Produces lossless stereo image
- Intensity stereo is used at lower bit-rates
  - Lossy
  - Stereo image can lose detail
- MPEG Surround is the most recent development
  - Low bit rate, multichannel audio
  - Backwards compatibility with stereo codecs possible

→ Useful literature: J. Breebaart, C. Faller: "Spatial Audio Processing - MPEG Surround and other Apllications", Wiley, Chichester, 2007

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

next lecture:

18.01. - Parametric Coding of High-Quality Audio

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT