Hannover uni

MPEG audio

An Overview of MPEG-4 Audio Version 2

Heiko Purnhagen

Laboratorium für Informationstechnologie

University of Hannover, Germany

Hannover uni

MPEG audio

Outline

- Introduction
- Review of MPEG-4 Audio Version 1
- New Tools of MPEG-4 Audio Version 2
- Outlook
- Conclusions

Hannover uni

MPEG audio

Introduction: MPEG-4

- New multimedia applications demand ...
  - efficient & flexible coding of natural & synthetic audiovisual content

  => Development of MPEG-4 Standard: "Coding of audiovisual objects"

- MPEG-4 Audio Version 1 (Oct. 1998)
  - audio and speech coding @ 2 .. 64+ kbit/s/ch

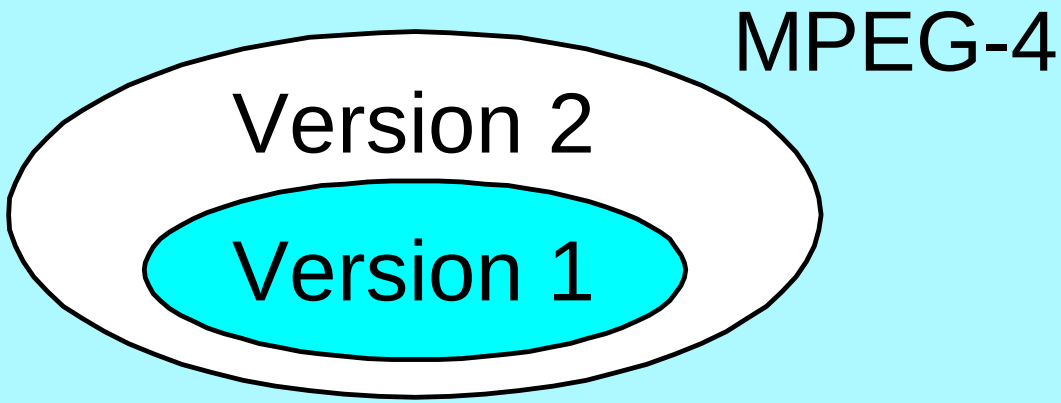
- MPEG-4 Audio Version 2 (Dec. 1999)
  - additional functionalities

Introduction: MPEG-4 Versions

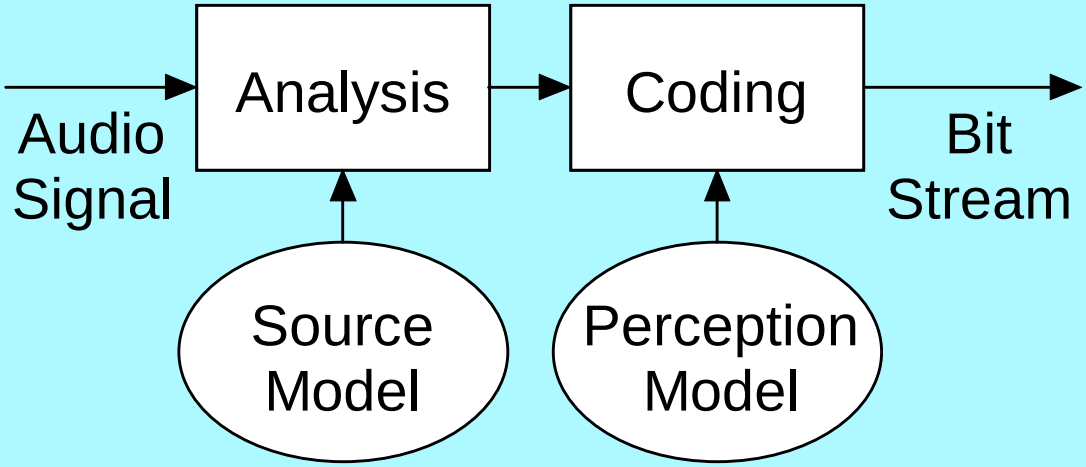
- Problem: Tight schedule for MPEG-4

  => several promising tools not mature in time

- Solution: MPEG-4 Version 2 (Amendment 1)
  - backward compatible extension of Version 1
  - new tools added => additional functionalities



Version 1: Overview



- MPEG-4 Audio Requirements:
  - efficient coding (various content types / bit rates)
  - other functionalities (e.g. scalability)

=> Combination of coding techniques required

  - utilise different source and perception models

Hannover uni

MPEG audio

Version 1: Tools

- MPEG-4 Version 1
  - Audio Tools: coding of audio objects

	Speech	General Audio
Natural	• HVXC (param.) (2 .. 4 kbit/s) • CELP (NB+WB) (4 .. 24 kbit/s)	• TwinVQ (6 .. 16 kbit/s/ch) • AAC (+scalable) (16 .. 64+ kbit/s/ch)
Synthetic	• TTS-Interface	• SA (incl. MIDI)

  - Systems Tools: composition of audio objects
    => audio scene (mixing, effects: Structured Audio)

Hannover uni

MPEG audio

Version 2: Overview

- New Audio Tools
  - Error Robustness
  - Low-Delay Audio Coding
  - Small Step Scalability
  - Parametric Audio Coding
  - CELP/HVXC Silence Compression
- New Systems Tools
  - Environmental Spatialisation
  - MP4 File Format

Hannover uni

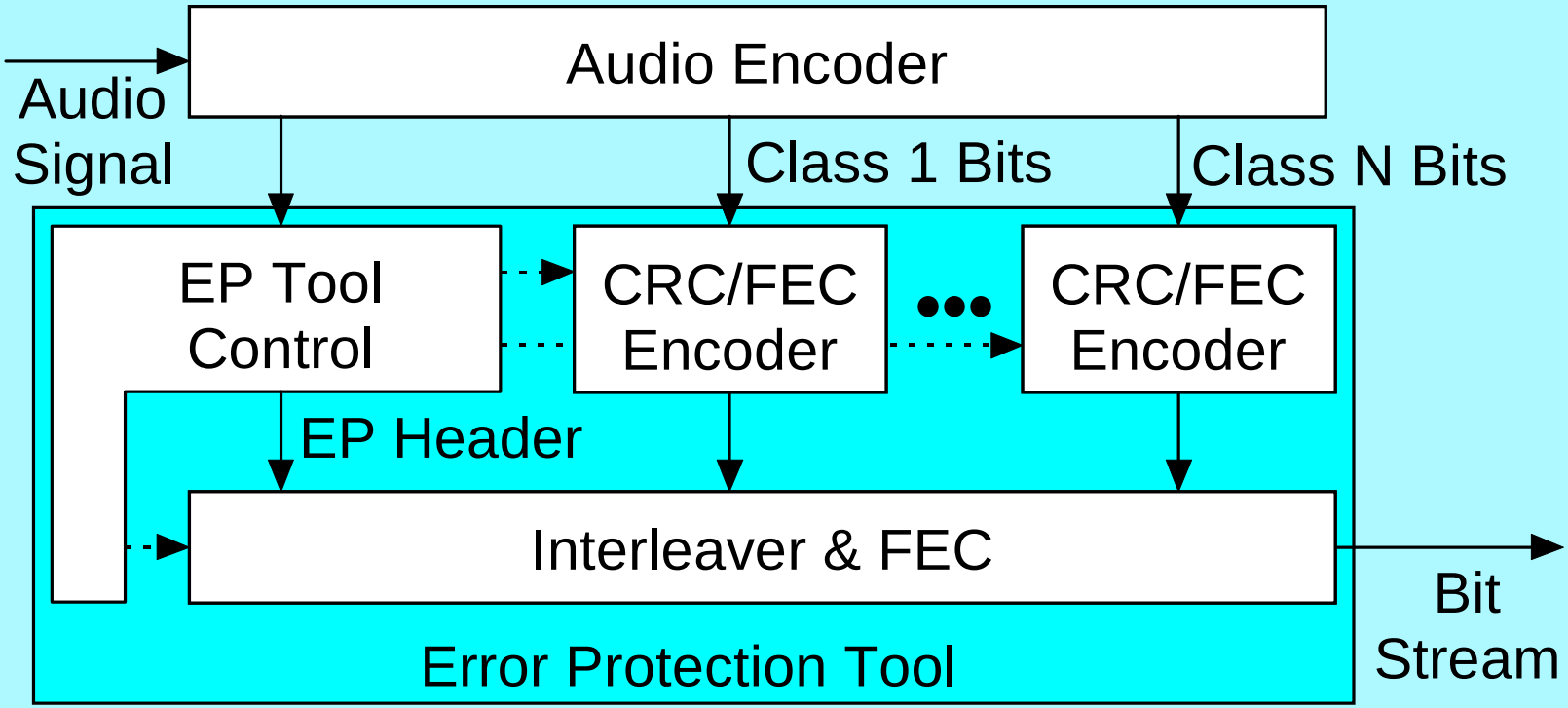
MPEG audio

Version 2: Error Robustness

- Goal: Transmission on error-prone channels
- Approach: Error Robustness Tools
  - Error Protection tool
  - Error Resilience for source coding tools

Version 2: Error Robustness

- Error Protection Tool
  - Unequal Error Protection: error sensitivity classes
  - Cyclic Redundancy Check / Forward Error Correction



Hannover uni

MPEG audio

Version 2: Error Robustness

- Error Resilience for source coding tools
  - Error Resilience for AAC
    e.g. Huffman Codeword Reordering
    => less error propagation
  - Error Resilience for other coding tools

  => reduced detoriation in case of bit errors

- Error Concealment in decoder (not normative)

Hannover uni

MPEG audio

Version 2: Error Robustness

- **Demo:** AAC 48 kHz mono @ 64 kbit/s
  random bit error rate: 10^-3
  - no Error Robustness
    (frame sync only)
  - Error Protection, Resilience, and Concealment
    (19% mean bit rate overhead)

Hannover uni

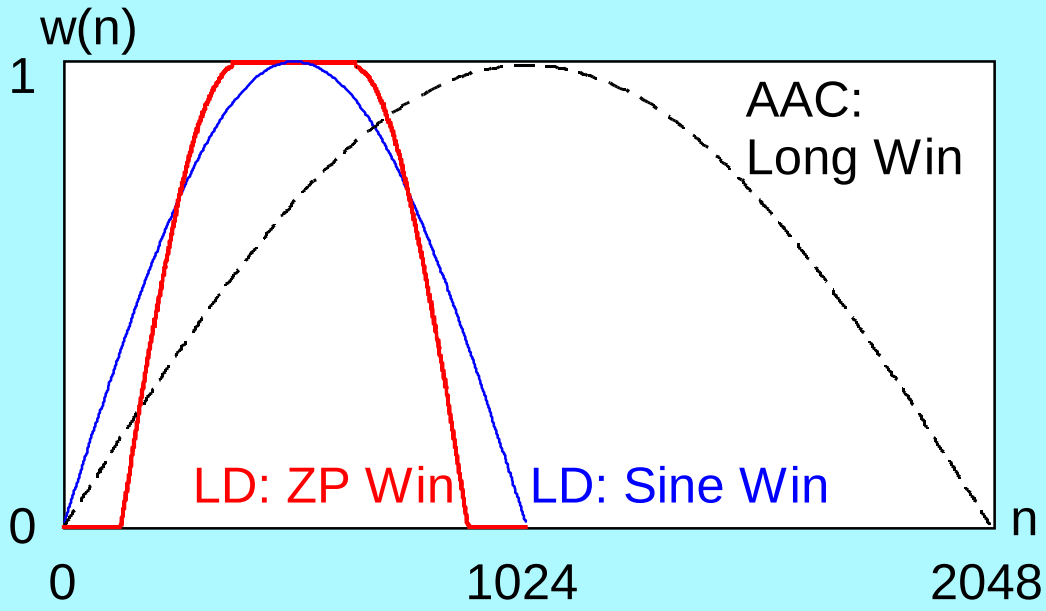
MPEG audio

Version 2: Low-Delay Audio Coding

- Goal: Low-delay general audio coding
  - realtime bi-directional communication
  - AAC algorithmic delay:
    e.g. 24 kHz @ 24 kbit/s
    => 110 ms + max. 210 ms (bit reservoir)
- Approach: Low-Delay Audio Coder
  - derived from AAC

Version 2: Low-Delay Audio Coding

- Low-Delay Audio Coder (modified AAC)
  - frame length & filterbank delay
    => 1/2 window size
  - "look-ahead" for window-switching
    => Zero-Padded window for transients
  - bit reservoir
    => not used



=> 20 ms algorithmic delay (48 kHz)

Hannover uni

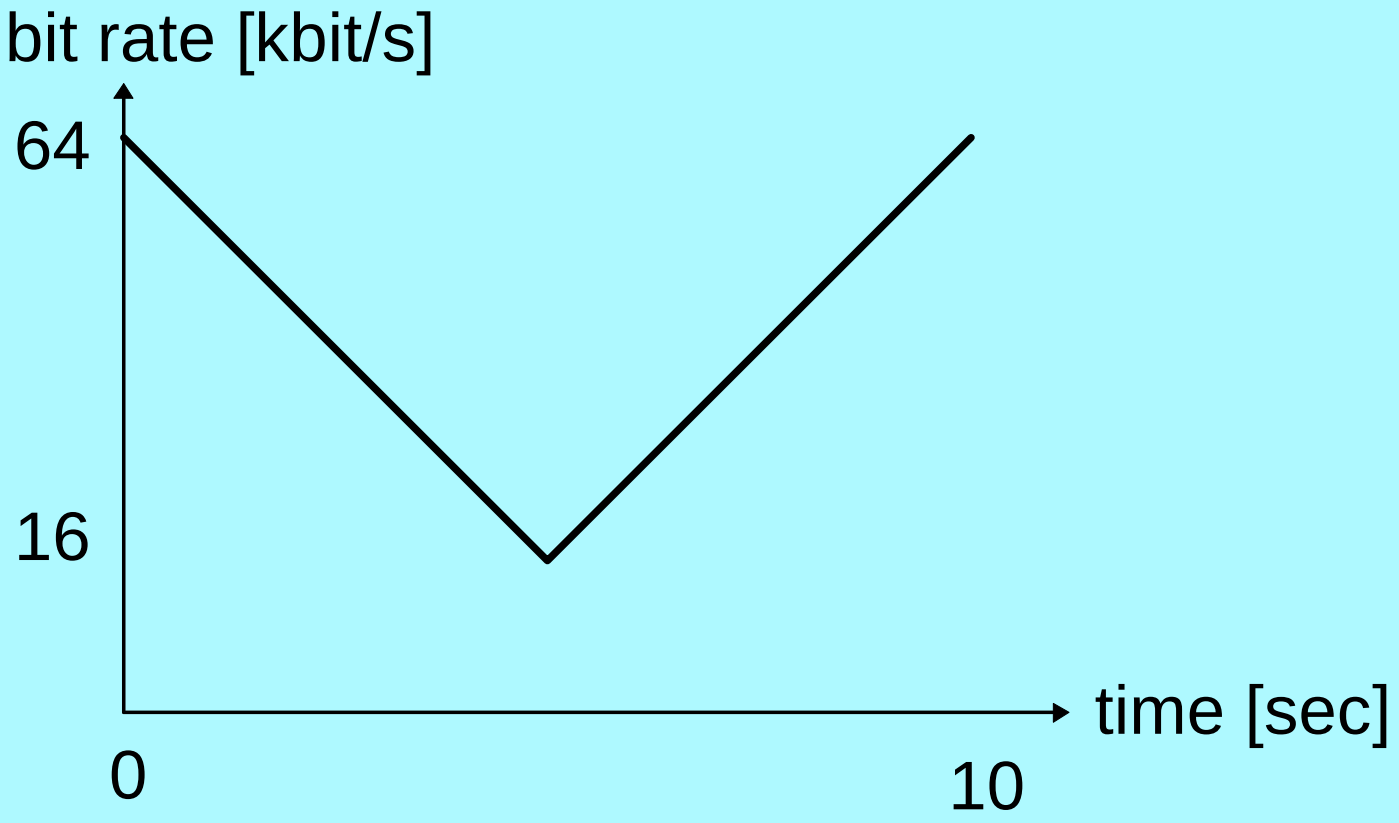
MPEG audio

Version 2: Small Step Scalability

- Goal: Small step bit rate scalability
  - AAC scalability: typ. 16 kbit/s enhancement layers
- Approach: Bit-Sliced Arithmetic Coding
  - combined with AAC:
    BSAC replaces AAC Huffman coding
    => 1 kbit/s/ch enhancement layers
  - Principle:
    transmit bit-slices with most significant bits first
    enhancement: less significant bits (finer quant.)
    higher frequency bands

Version 2: Small Step Scalability

- **Demo:** BSAC 48 kHz mono @ 64 kbit/s



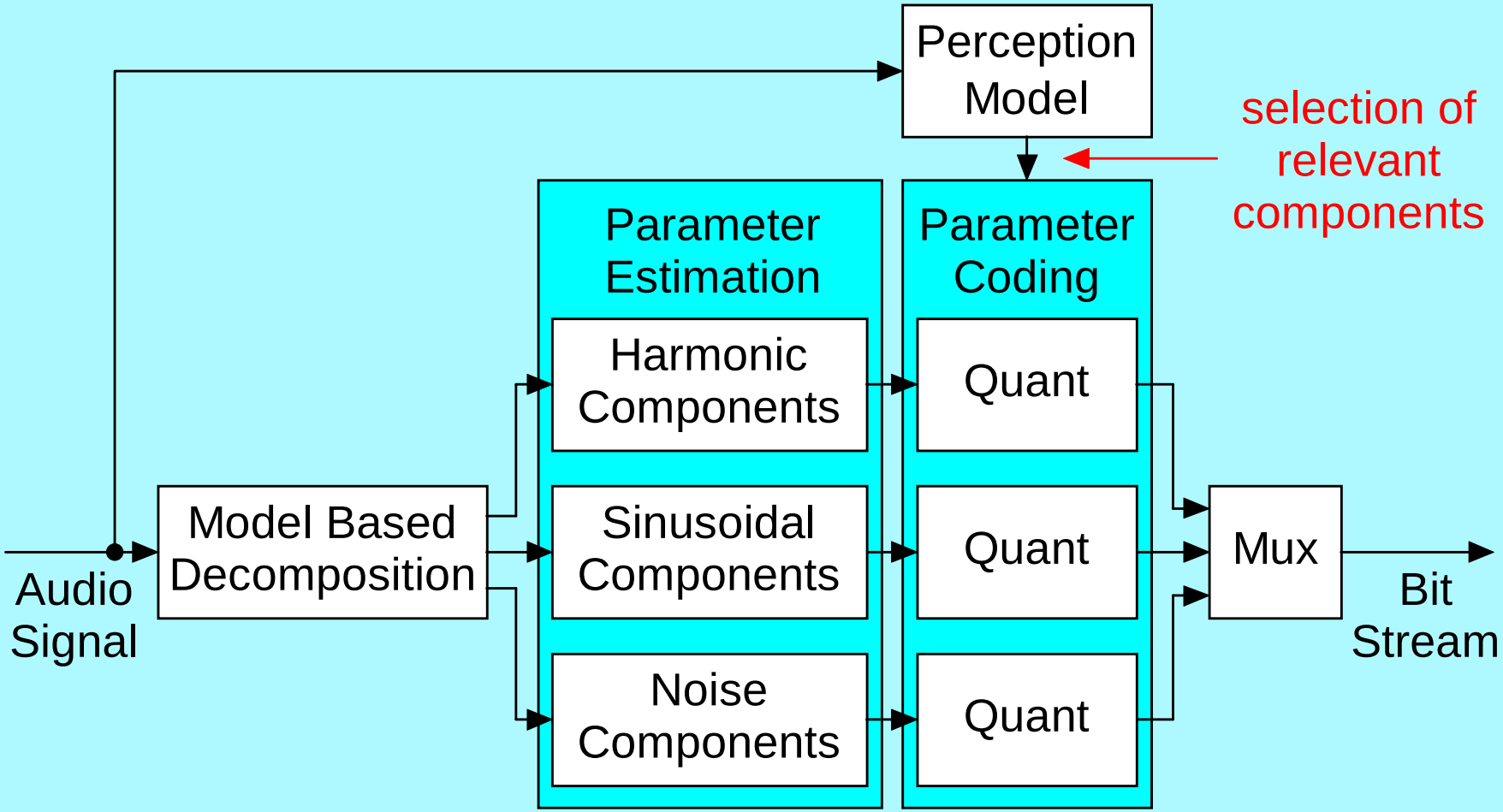
Hannover uni

MPEG audio

Version 2: Parametric Audio Coding

- Goal: Very low bit rate audio coding
  - source model for music ?
- Approach: Parametric Signal Representation
  - Signal decomposition into components:
    "Harmonic and Individual Lines plus Noise" (HILN)
  - Functionalities:
    very low bit rate (4 .. 16 kbit/s)
    speed and pitch change (decoder)
    bit rate scalability

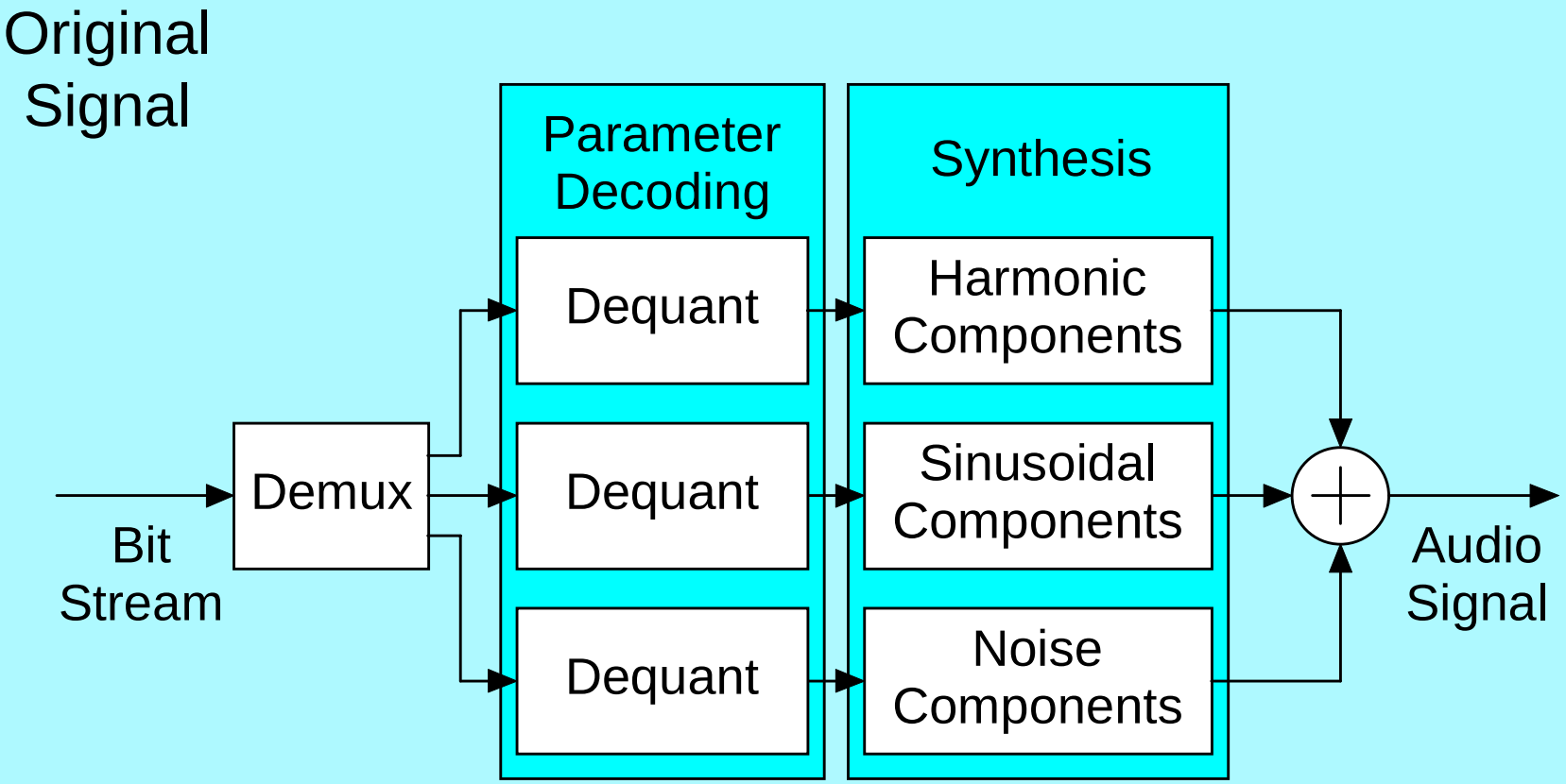
Version 2: Parametric Audio Coding



Parametric Audio Encoder (HILN)

Version 2: Parametric Audio Coding

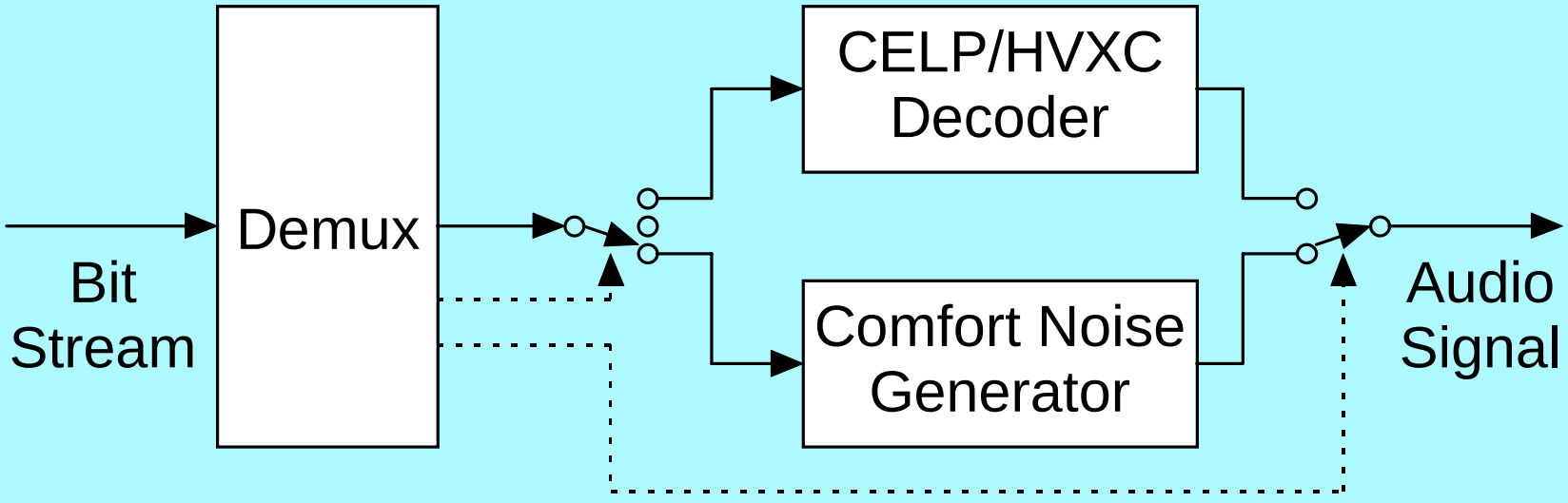
- **Demo:** HILN 16 kHz mono @ 6 kbit/s



Parametric Audio Decoder (HILN)

Version 2: Silence Compression

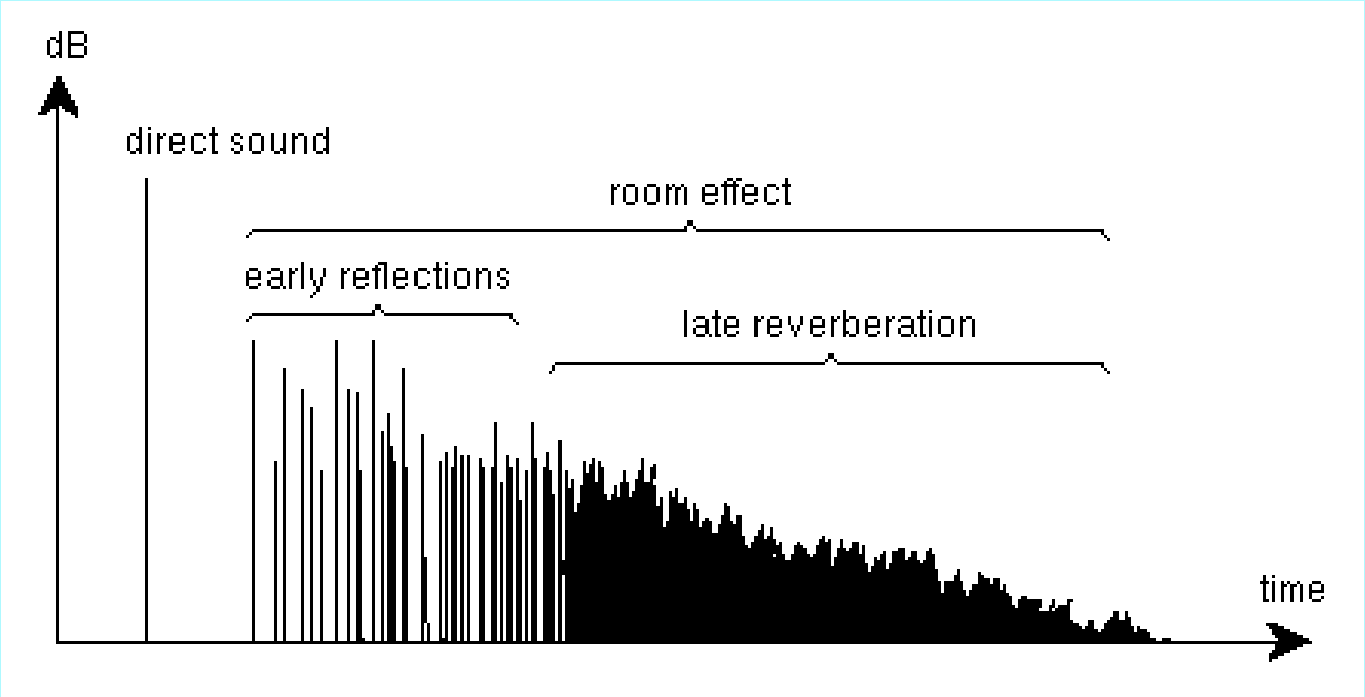
- Goal: Bit rate reduction if no speech activity
  - bi-directional communication: ~50% silence
- Approach: HVXC/CELP Silence Compression
  - generate comfort noise if no speech activity



Version 2: Environmental Spatialisation

- Goal: Efficient and flexible "3-D" audio scenes
  - SA-based composition not easy for "3-D" scenes
- Approach: Physical or perceptual description

typical room response



Hannover uni

MPEG audio

Version 2: Environmental Spatialisation

- Physical approach
  - description of acoustical properties of environment (room geometry, sound source position, ...)
  - corresponding audio and visual scene e.g. 3-D virtual reality
- Perceptual approach
  - high-level perceptual description of "audio scene" (room reverberance, source presence, ...)
  - audio and visual scene independent e.g. movie-like applications

Hannover uni

MPEG audio

Version 2: Other Tools

- MPEG-4 File Format (MP4)
  - flexible format for:
    interchange, editing, presentation
  - based on QuickTime
- Backchannel
  - e.g. adaptive streaming

Hannover uni

MPEG audio

Outlook

- Version 2 Profiles and Levels
  - currently under discussion
- Version 2 Verification Test
  - scheduled for autumn 1999
- Encoder Optimisation (not normative)
  - e.g. automatic segmentation of speech / music (audio objects are transparent ...)

Hannover uni

MPEG audio

Conclusions

- MPEG-4 Audio Version 2
  - backward compatible extension of Version 1
  - additional functionalities by new tools
- MPEG-4 Version 2 finalised: Dec. 1999
- New work item: MPEG-7
  "Multimedia Content Description Interface"

Hannover uni

MPEG audio

Conclusions

- MPEG Audio Web Page

  `http://www.tnt.uni-hannover.de/project/mpeg/audio/`

- Further Demonstrations
  - Error Robustness (for AAC)
  - Low-Delay Audio Coding (realtime en-/decoder)
  - Small Step Scalability (BSAC)
  - Parametric Audio Coding (HILN)