Audio Coding – Overview and Latest Developments in MPEG-4

Heiko Purnhagen, Senior Research Engineer

Coding Technologies

Medieteknikdagarna Norrköping

2004-04-01

coding technologies

Outline

- About Coding Technologies (CT)
- Perceptual Audio Coding Overview
  - MPEG Standards: Layer II, mp3, AAC
- CT's Audio Coding Enhancement Tools
  - Spectral Band Replication (SBR)
  - Parametric Stereo (PS)
- Codecs and Applications
  - mp3PRO, aacPlus
  - Broadcasting, Internet, …

coding
technologies

About Coding Technologies

- 1997: Founded as Coding Technologies Sweden by Lars Liljeryd
- 2000: Merged with spin-off from Fraunhofer IIS („Home of mp3"), Germany
- 2004: 32 employees, 3 office locations
  - Nuremberg, Germany
  - Stockholm, Sweden
  - Mountain View, California, USA
- http://www.codingtechnologies.com/

coding
technologies

Perceptual Audio Coding – Motivation

- Digital Audio (e.g. Compact Disc)
  PCM: 44.1 kHz, 16 bit, stereo -> 1.4 Mbit/s

- Perceptual Audio Coding
  -> compact audio representation … exploiting:
  - Signal redundancy
    -> e.g. transform coding ("MDCT")
  - Perceptual irrelevance
    -> masking effect ("psychoacoustic model")

- "transparent" quality at 128 kbit/s
  -> compression ratio ~1:12

coding
technologies

Perceptual Audio Coding – Standardization

ISO/IEC "Moving Picture Experts Group" (MPEG)

- MPEG-1 Audio   (1992)
  -> Layer II (DAB), Layer III (mp3)
- MPEG-2 Audio   (1994)
  -> Layer II Multichannel (DVD)
- MPEG-2 AAC   (1997)
  -> Advanced Audio Coding (iTunes, …)
- MPEG-4 Audio   (2000)
  -> natural/synthetic speech and audio objects
     (3GPP, QuickTime, …)

coding
technologies

Perceptual Audio Coding – Limitations

- Audio Demonstration
  - Original PCM stereo 1.5 Mbit/s
  - AAC stereo   128 kbit/s   1:12
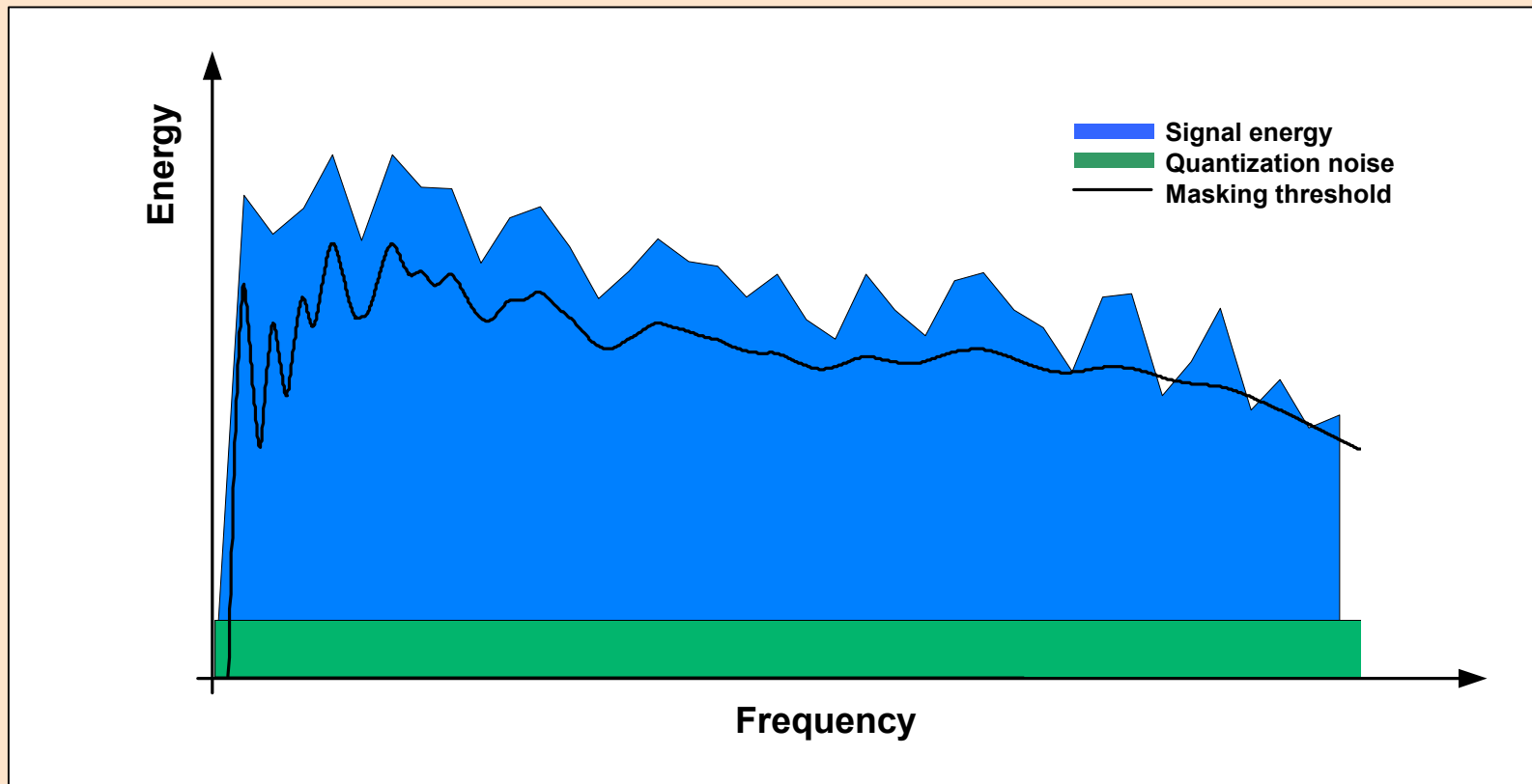 - Difference signal (23 dB SNR)
- But …
  - AAC stereo   48 kbit/s   1:32
  - AAC mono   24 kbit/s   1:64

-> limited audio quality for dial-up Internet or mobile receivers (radio, cellular phones)
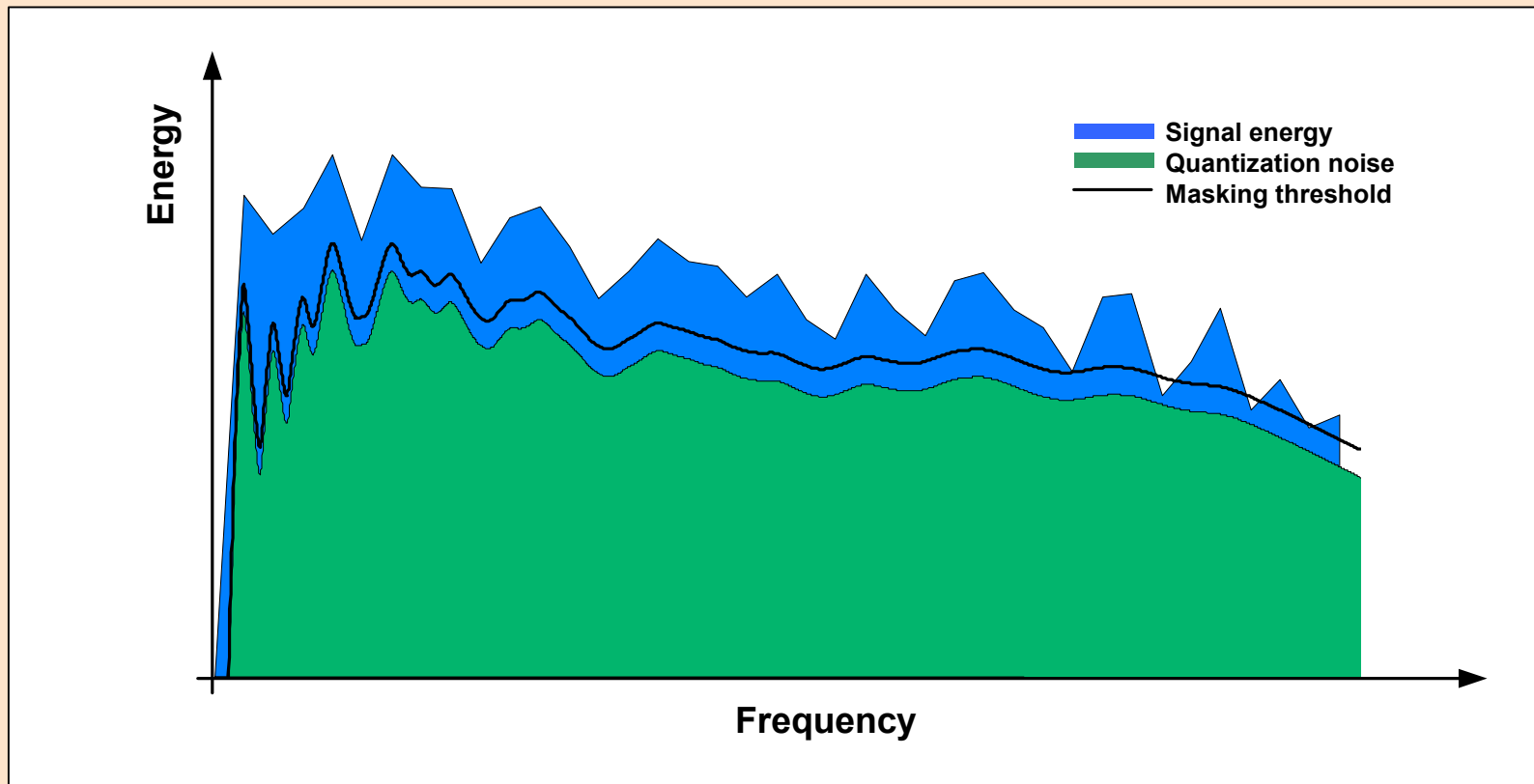
Perceptual Audio Coding – Background

16 bit PCM

Legend:
- Signal energy
- Quantization noise
- Masking threshold

Energy (vertical axis) / Frequency (horizontal axis)

Visible blue area indicates bit rate

coding technologies

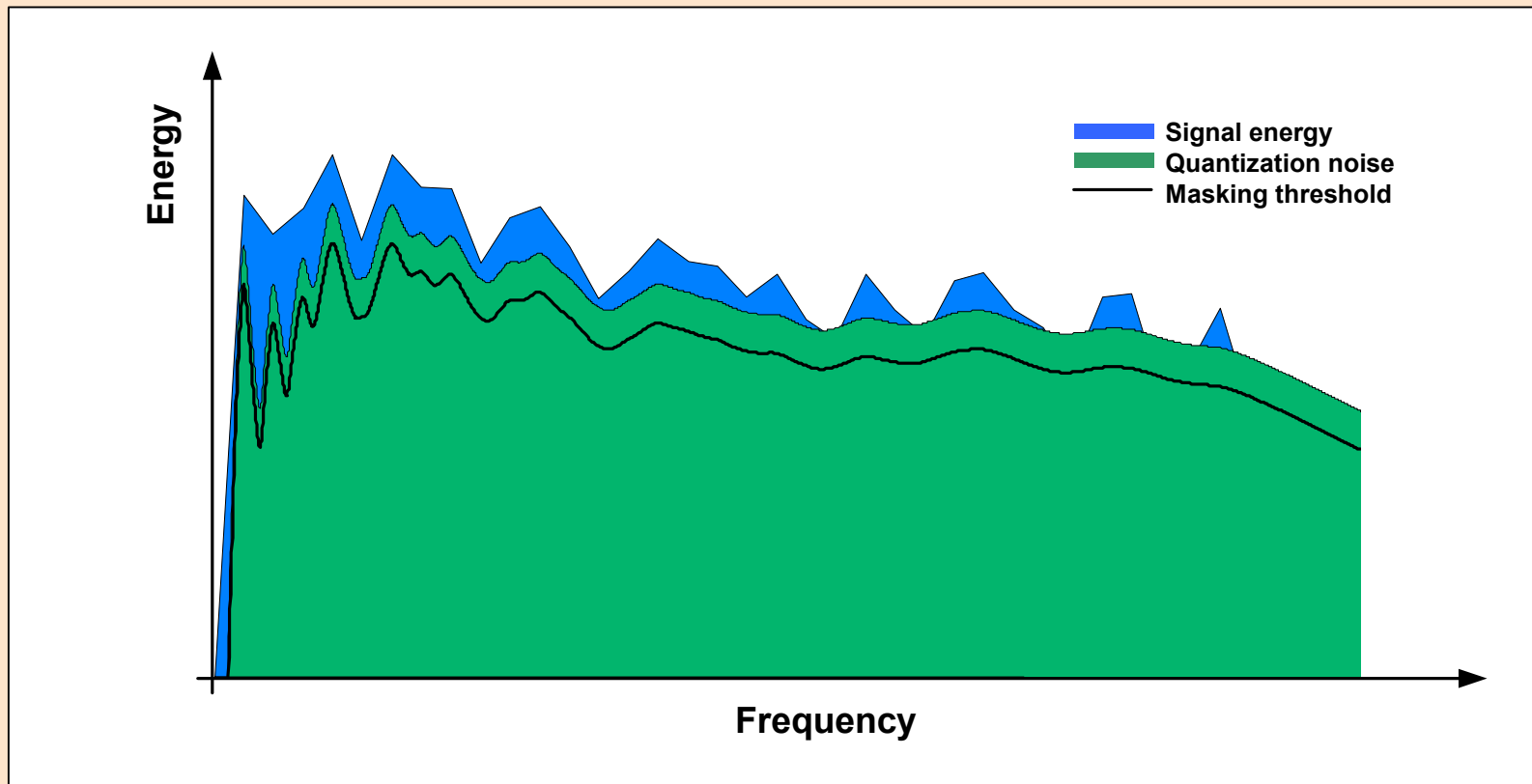# Perceptual Audio Coding – Background

bit rate sufficiently high

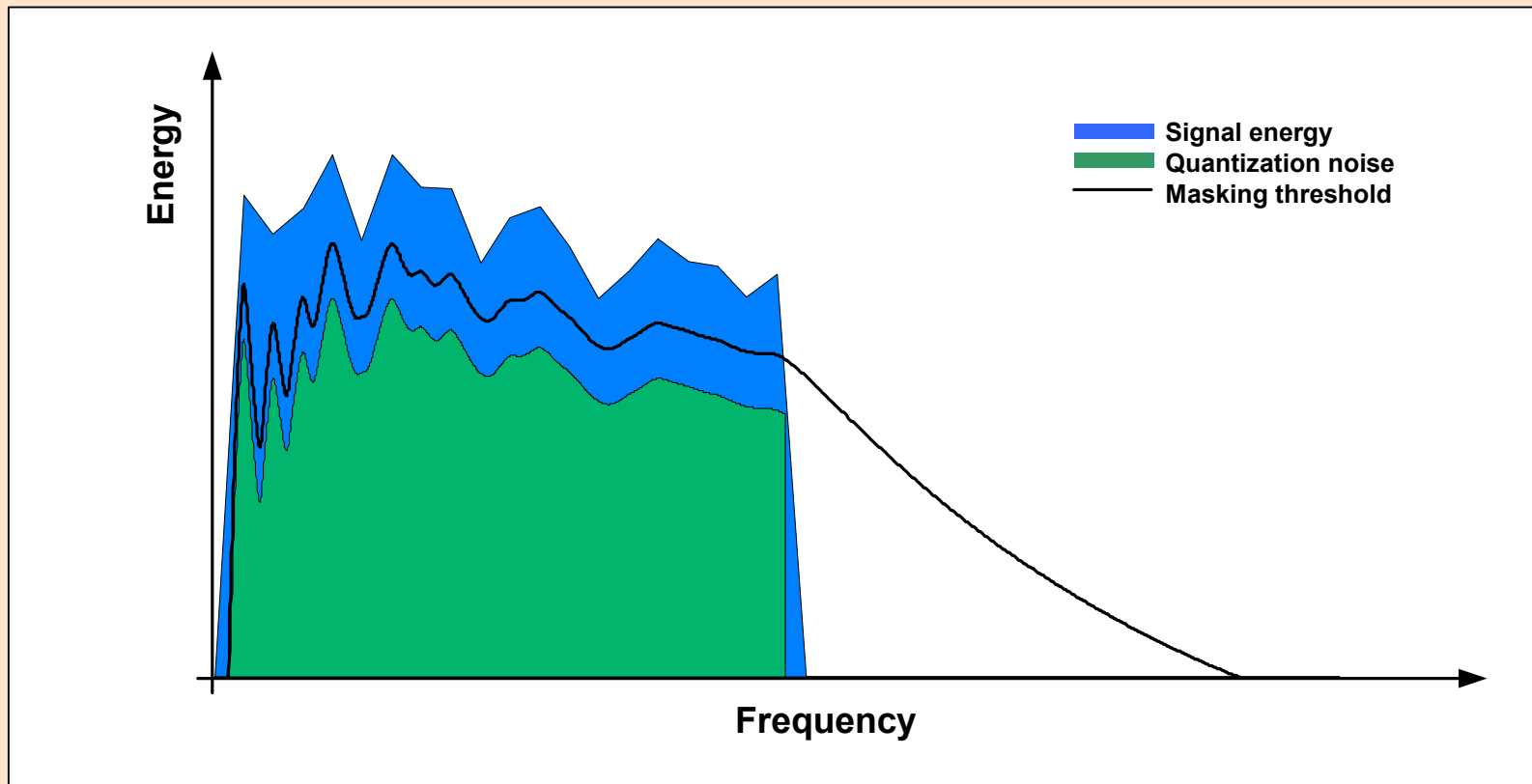Visible blue area indicates bit rate

Perceptual Audio Coding – Background

bit rate too low -> quantization noise audible

Visible blue area indicates bit rate

coding technologies

Perceptual Audio Coding – Background

-> limit audio bandwidth to reduce coding artifacts

Visible blue area indicates bit rate

Audio Coding Enhancement – Approach

- Problem
  - Conventional coding of high frequency band "expensive" (needs many bits)
  - -> limited audio bandwidth at low bit rates
- Approach
  - Reconstruct high frequency band from low frequency band
  - Side information (2 to 3 kbit/s per channel) to control reconstruction
  - -> Spectral Band Replication (SBR)

coding technologies

Spectral Band Replication (SBR)

reconstruct high frequency band by transposition
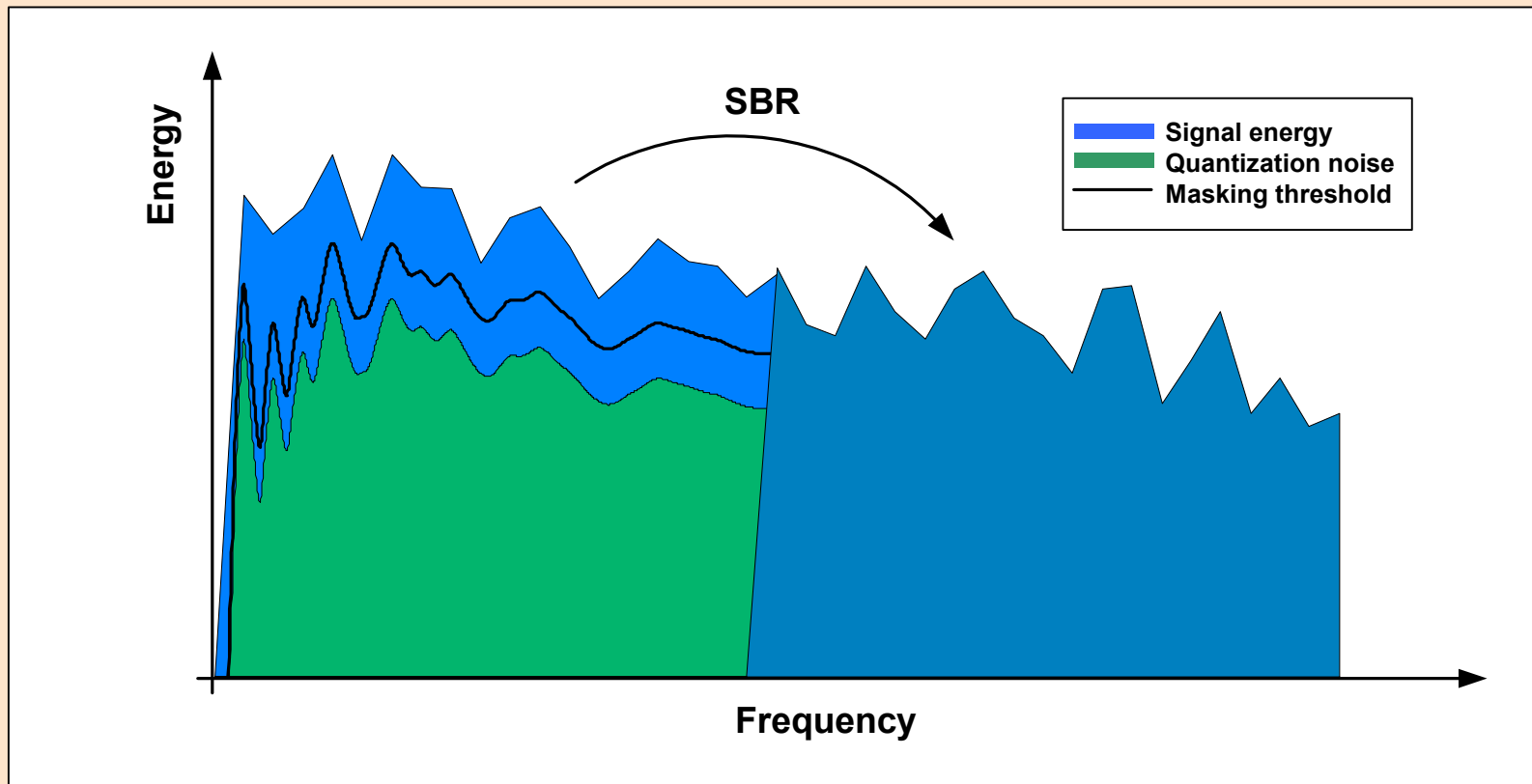
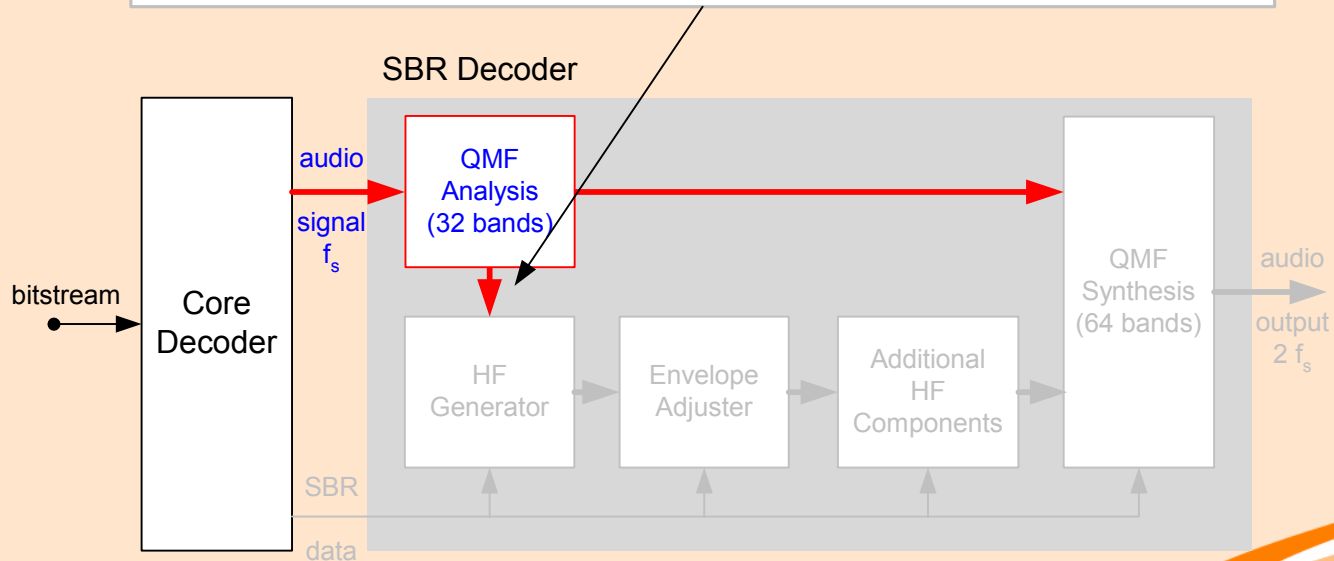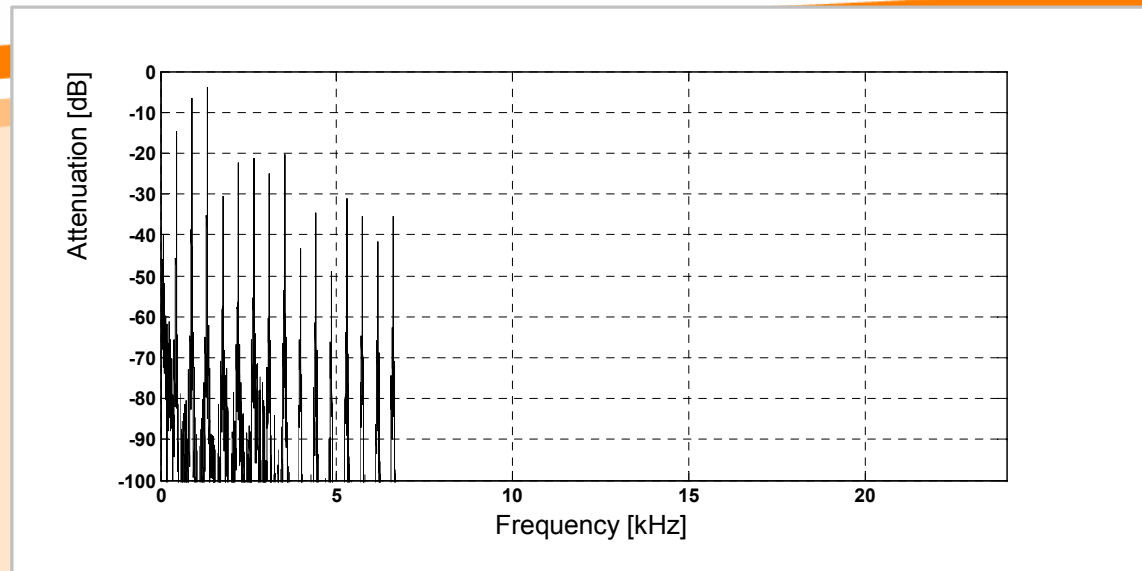# Audio Decoder with SBR
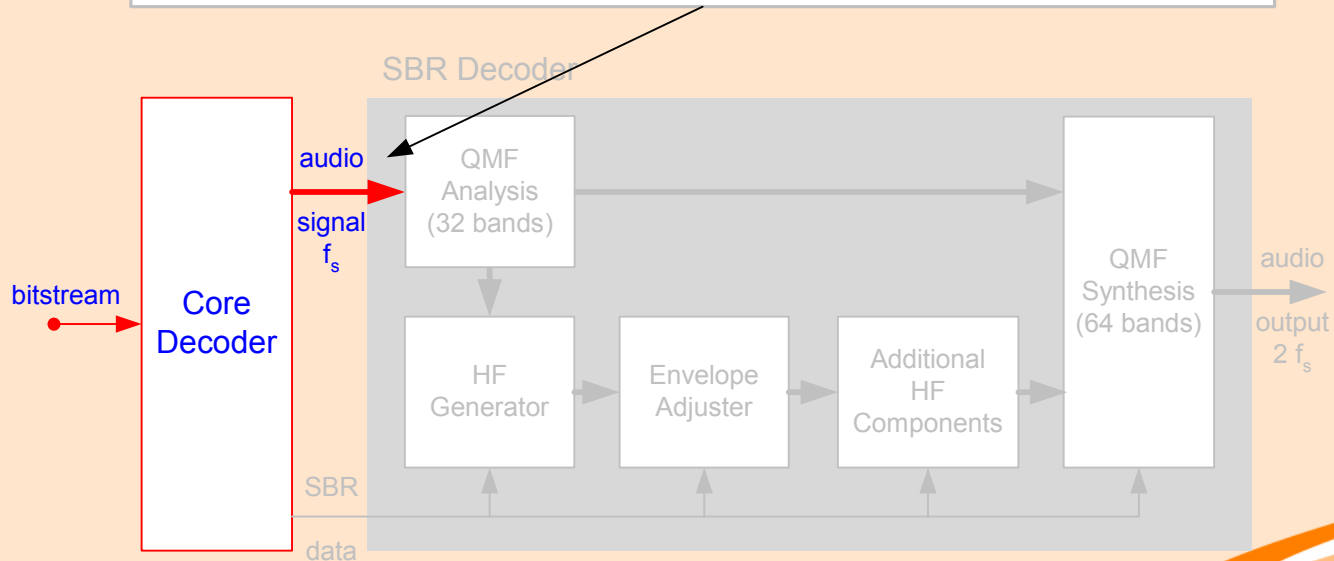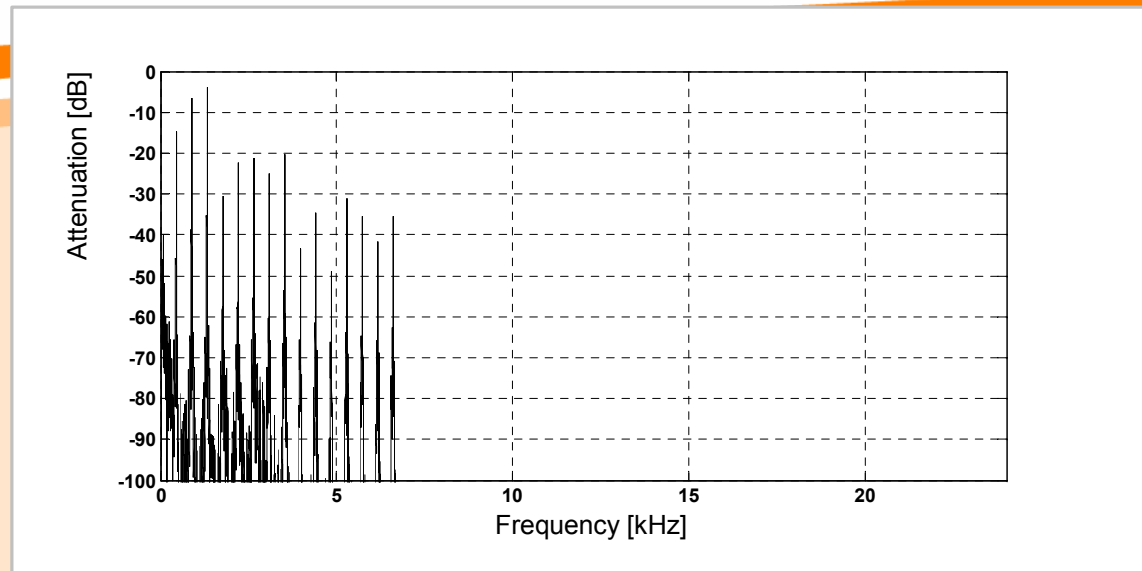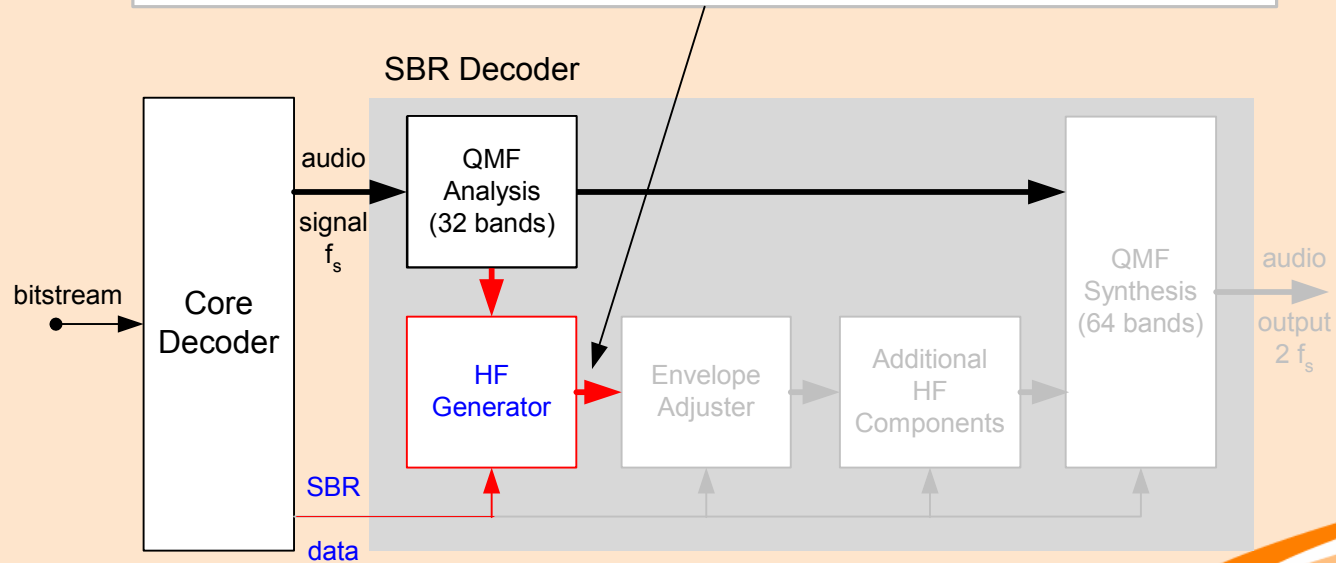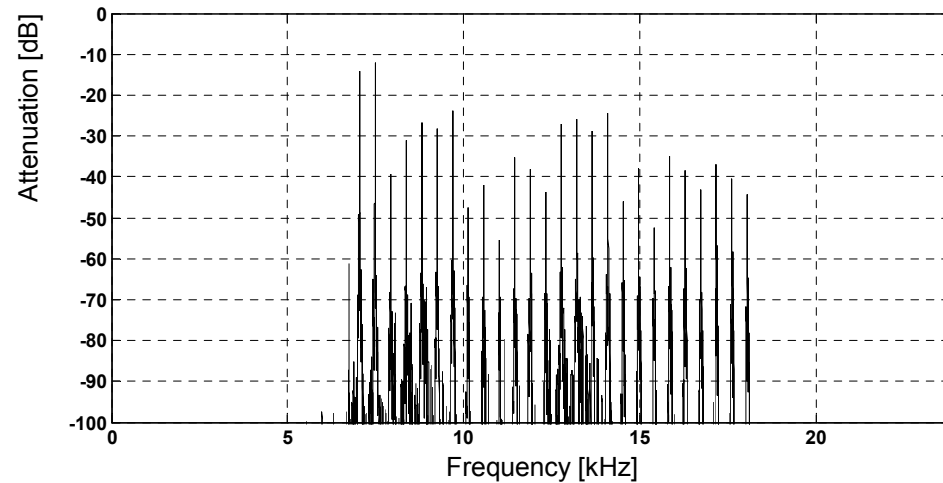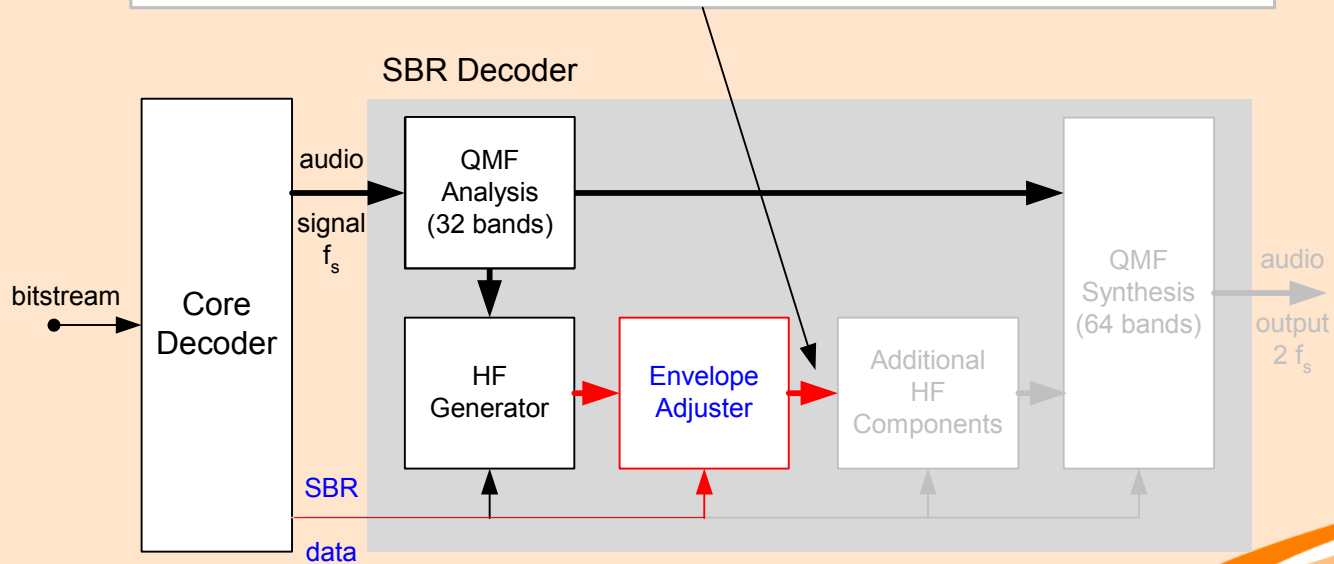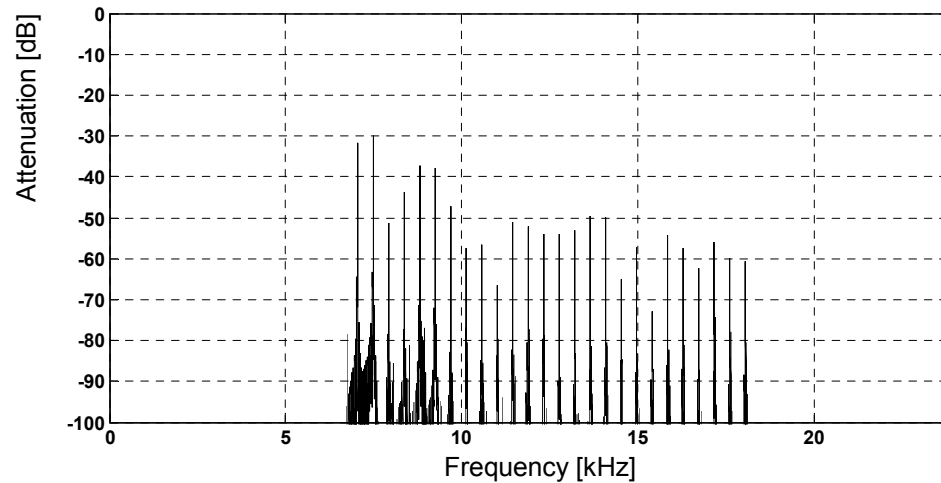

Core coder uses ½ sampling rate

Audio Decoder with SBR

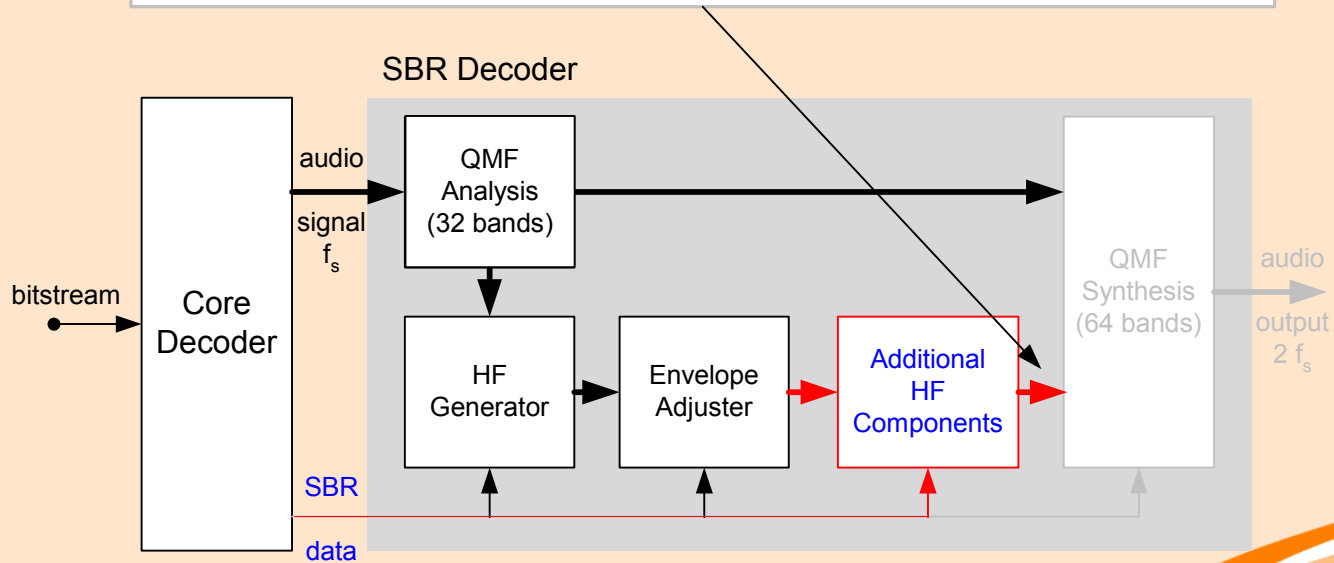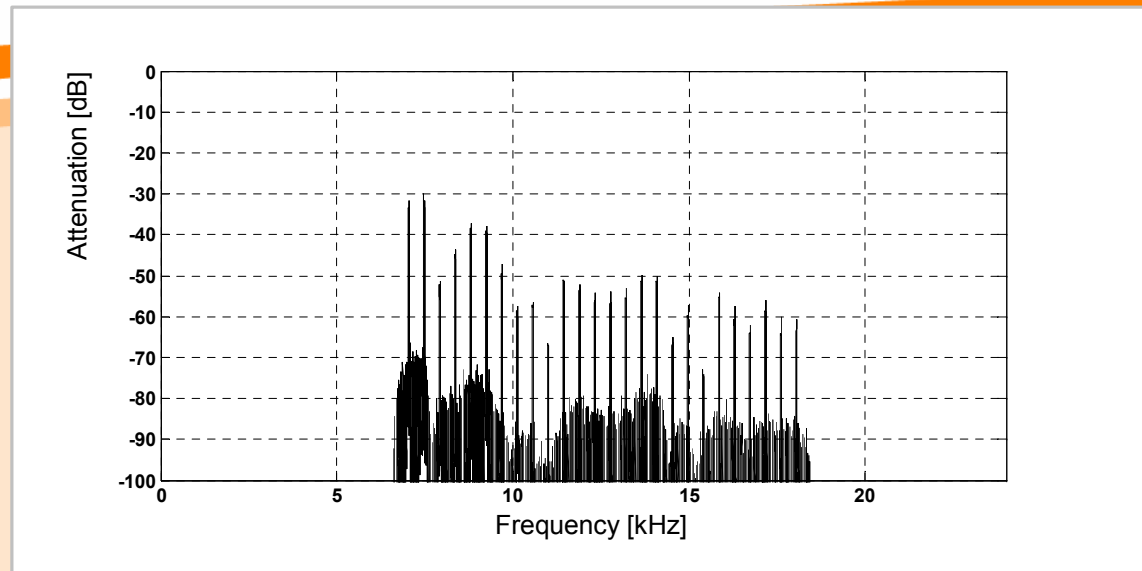Complex QMF enables alias-free processing
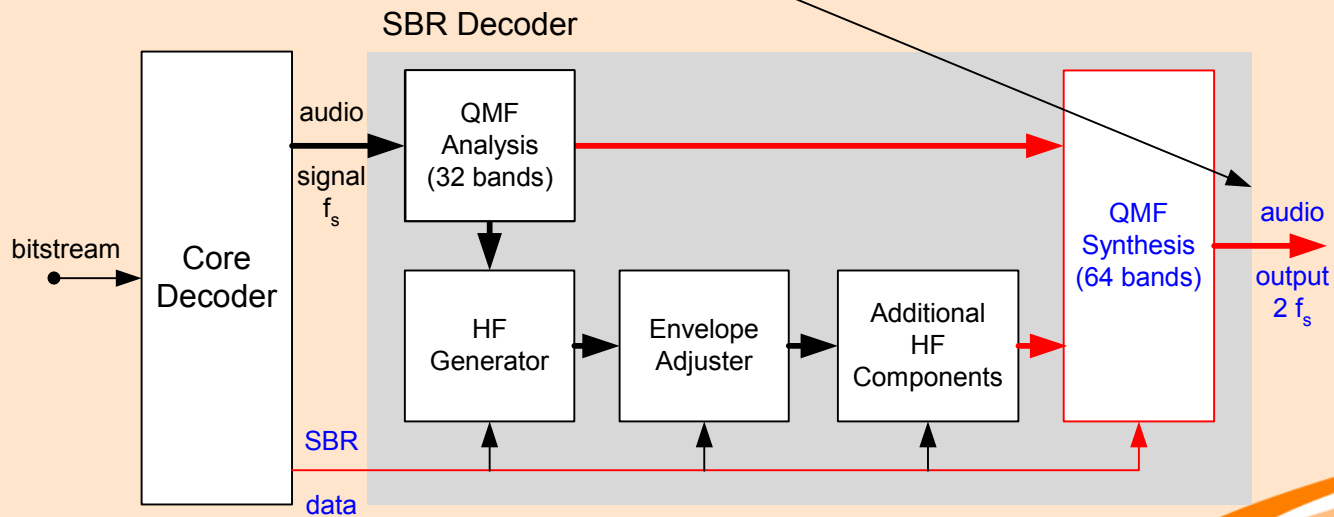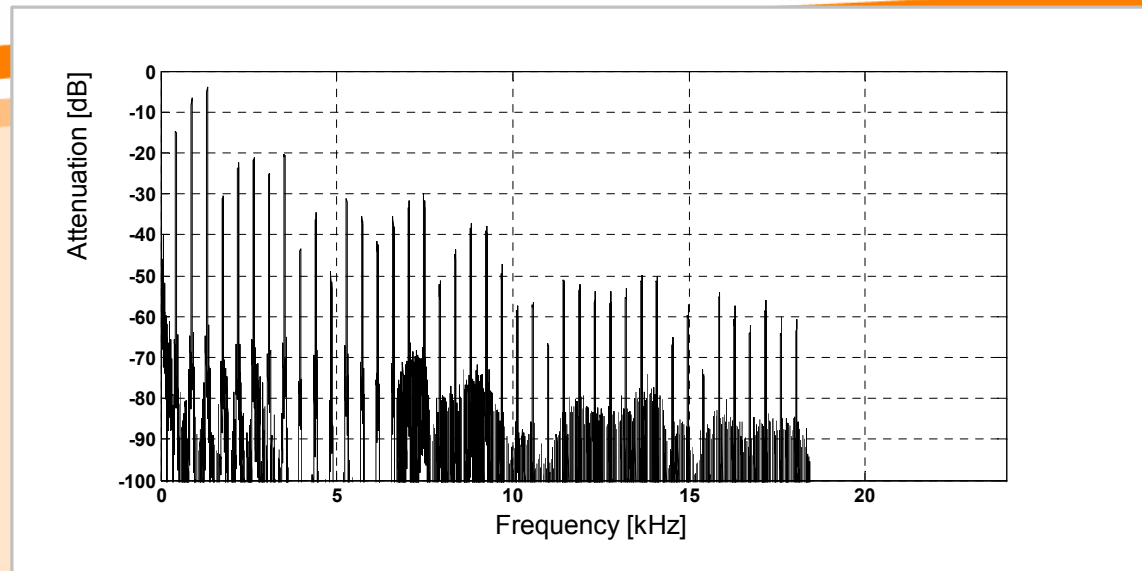
Audio Decoder with SBR

Audio Decoder with SBR

Audio Decoder with SBR

Audio Decoder with SBR

Coding Technologies' Codecs with SBR
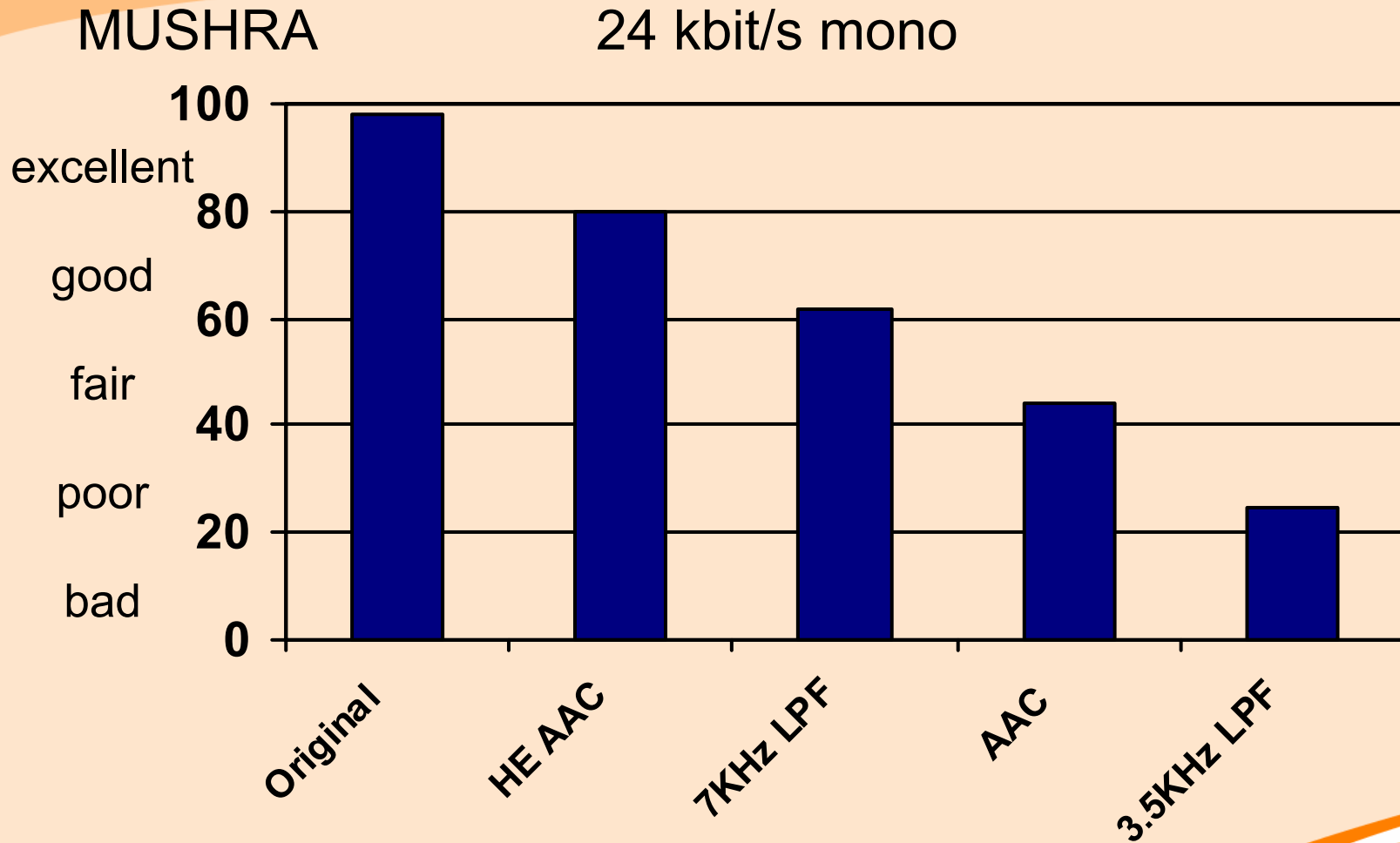
mp3PRO = mp3 + SBR

aacPlus = AAC + SBR

In 2003, aacPlus was standardized as
MPEG-4 High Efficiency AAC (HE-AAC)

coding
technologies

MPEG Verification Test

MUSHRA   24 kbit/s mono

LPF: low-pass filtered "anchor"

coding technologies

EBU Internet Audio Test (2002)

MUSHRA   48 kbit/s stereo

LPF: low-pass filtered "anchor"

coding
technologies
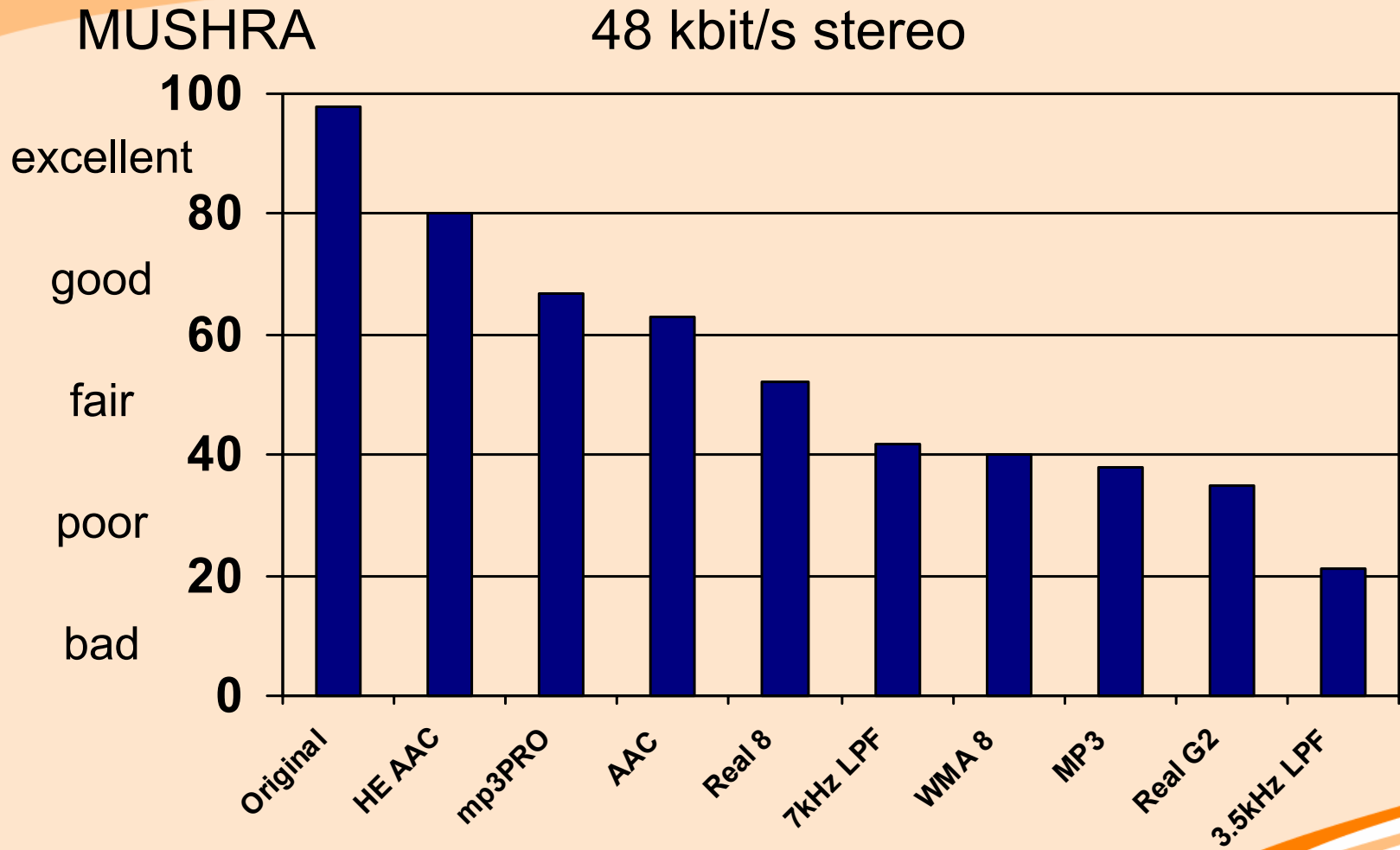
SBR – Results

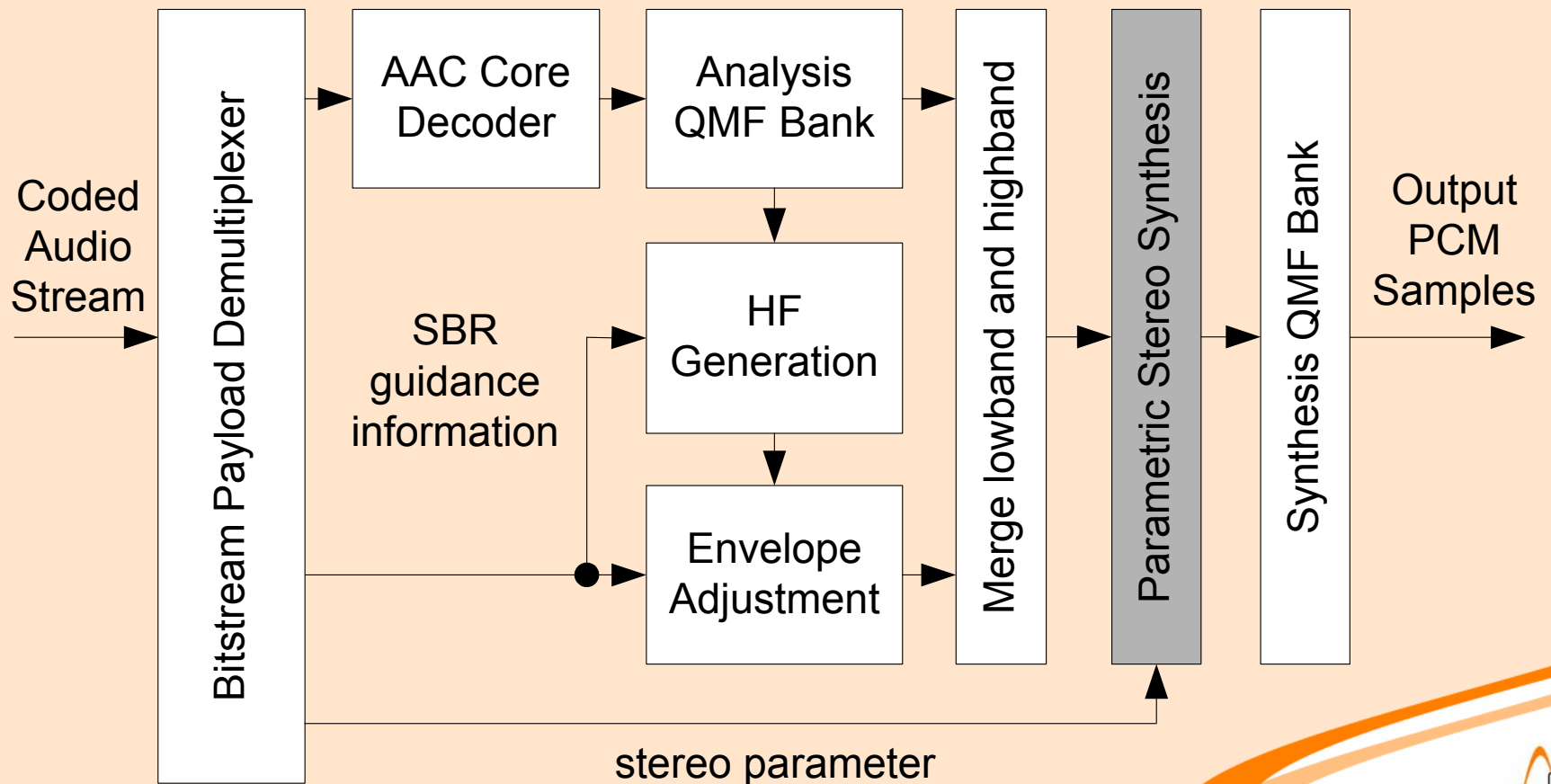- SBR enables bit rate reduction by 30% to 50% compared to full bandwidth core coder
- mp3PRO
  - Stereo broadcast quality at 64 kbit/s
- aacPlus (MPEG-4 High Efficiency AAC)
  - Stereo broadcast quality at 48 kbit/s
  - 5.1 surround at 128 kbit/s
- Forward & backward compatible with core coder
- Decoder complexity (MIPS) same as full bandwith core coder due to dual-rate approach

coding □
technologies

Parametric Stereo (PS)

- Problem
  - Mono coding better than stereo at low bit rates
- Approach
  - Transmit mono signal + stereo side information
  - Reconstruct stereo signal in decoder
- Stereo Parameter (per subband and time slot)
  - Interchannel Intensity Difference (IID) -> "pan"
  - InterChannel Correlation (ICC) -> "ambience"
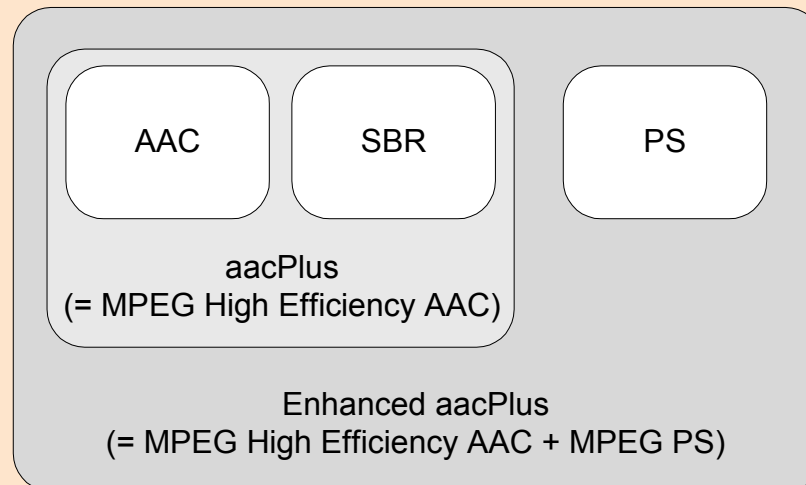- Approx. 2 to 3 kbit/s stereo side information

coding technologies

Combining Parametric Stereo with SBR

Parametric Stereo re-uses QMF bank from SBR

Parametric Stereo – Results

- Parametric Stereo (PS) enables bit rate reduction by 25% to 35% compared to stereo aacPlus
- AAC + SBR + PS = Enhanced aacPlus
- PS Tool now standardized in MPEG-4
- Decoder complexity (MIPS) same as stereo aacPlus

AAC   SBR   PS

aacPlus
(= MPEG High Efficiency AAC)

Enhanced aacPlus
(= MPEG High Efficiency AAC + MPEG PS)

coding
technologies

Applications – mp3PRO

- Demo: http://www.mp3PROzone.com/
- Philips Streamium
- Thomson/RCA Lyra
- MusicMatch, CoolEdit, Nero, Winamp, …

Applications – aacPlus

- Digital Radio Mondial (DRM) – AM radio
- XM Satellite Radio – USA, 121 channels
- Standardized: MPEG-4 High Efficiency AAC
- GPRS Mobile Music Download (O2, Vodafone)
- RealPlayer 10, Nero, Orban, …

Applications – Enhanced aacPlus

- Enhanced aacPlus (AAC + SBR + PS)
  -> the most efficient audio codec available today!
- Recommended Codec in 3GPP "Release 6"

- Audio Demonstration
  - Original PCM stereo 1.5 Mbit/s
  - AAC stereo   48 kbit/s   1:32
  - aacPlus stereo   48 kbit/s   1:32
  - Enhanced aacPlus   24 kbit/s   1:64

Conclusions

- CT's Audio Coding Enhancement Tools
  - Spectral Band Replication (SBR)
  - Parametric Stereo (PS)
- Enhanced aacPlus (AAC + SBR + PS)
  - High quality stereo at 24 kbit/s (1:64)
  - Standardized in MPEG-4
- Applications
  - Broadcast, Internet, Mobile Communication

coding
technologies