



W19 Hybrid Audio Coding

High Frequency Reconstruction

HEIKO PURNHAGEN

DOLBY SWEDEN AB

142ND AES CONVENTION, 2017 MAY 20-23, BERLIN

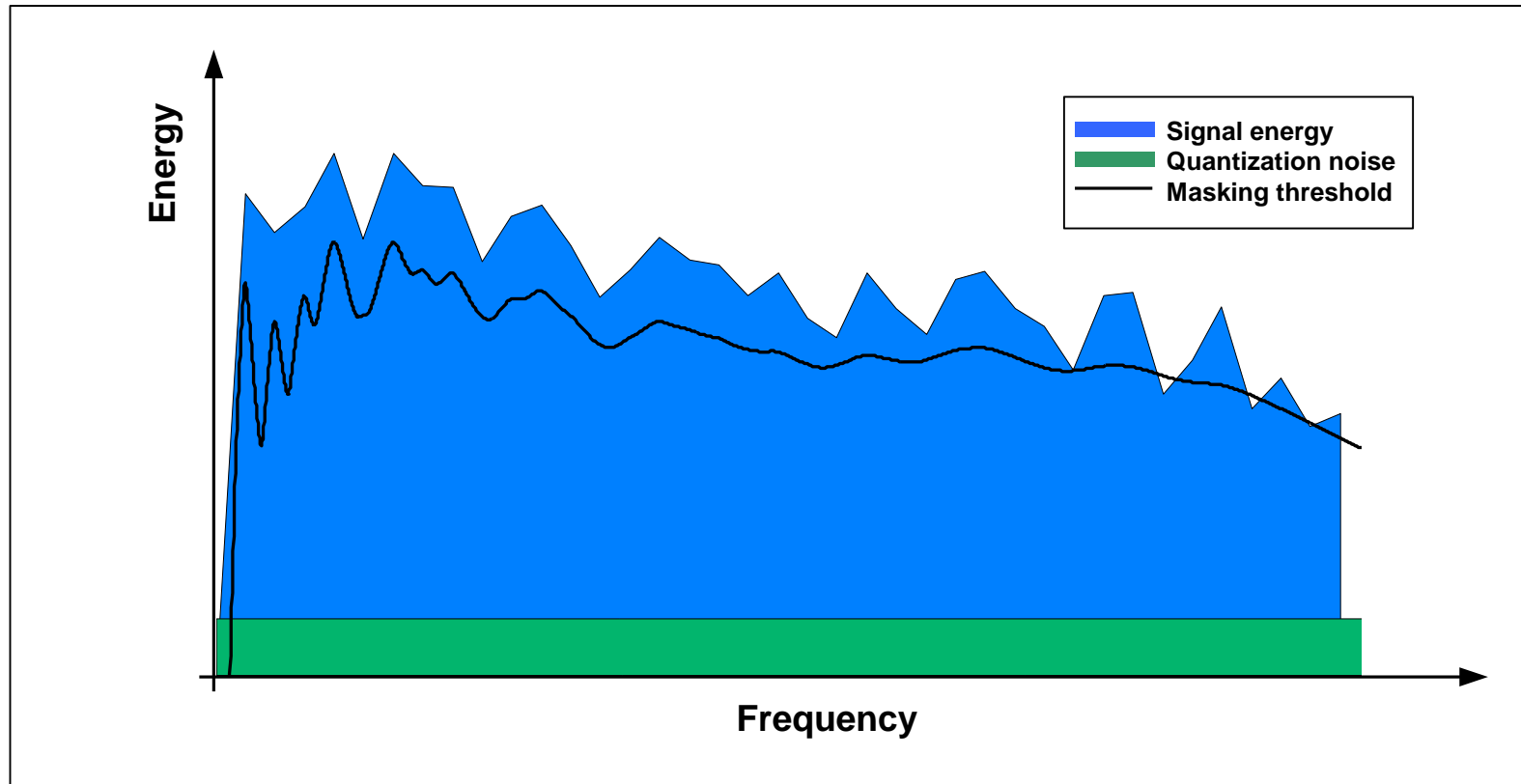
CLASSIC PERCEPTUAL AUDIO CODING: LIMITATIONS

Classic perceptual audio coding: filterbank/MDCT-based waveform coding

- Examples: mp3, AAC, Dolby Digital (AC-3), ...
- High bit rates: quantization noise below masked threshold, coding perceptually transparent
- Low bit rates: limited audio bandwidth (less annoying than audible quantization noise / artifacts)

CLASSIC PERCEPTUAL AUDIO CODING: LIMITATIONS

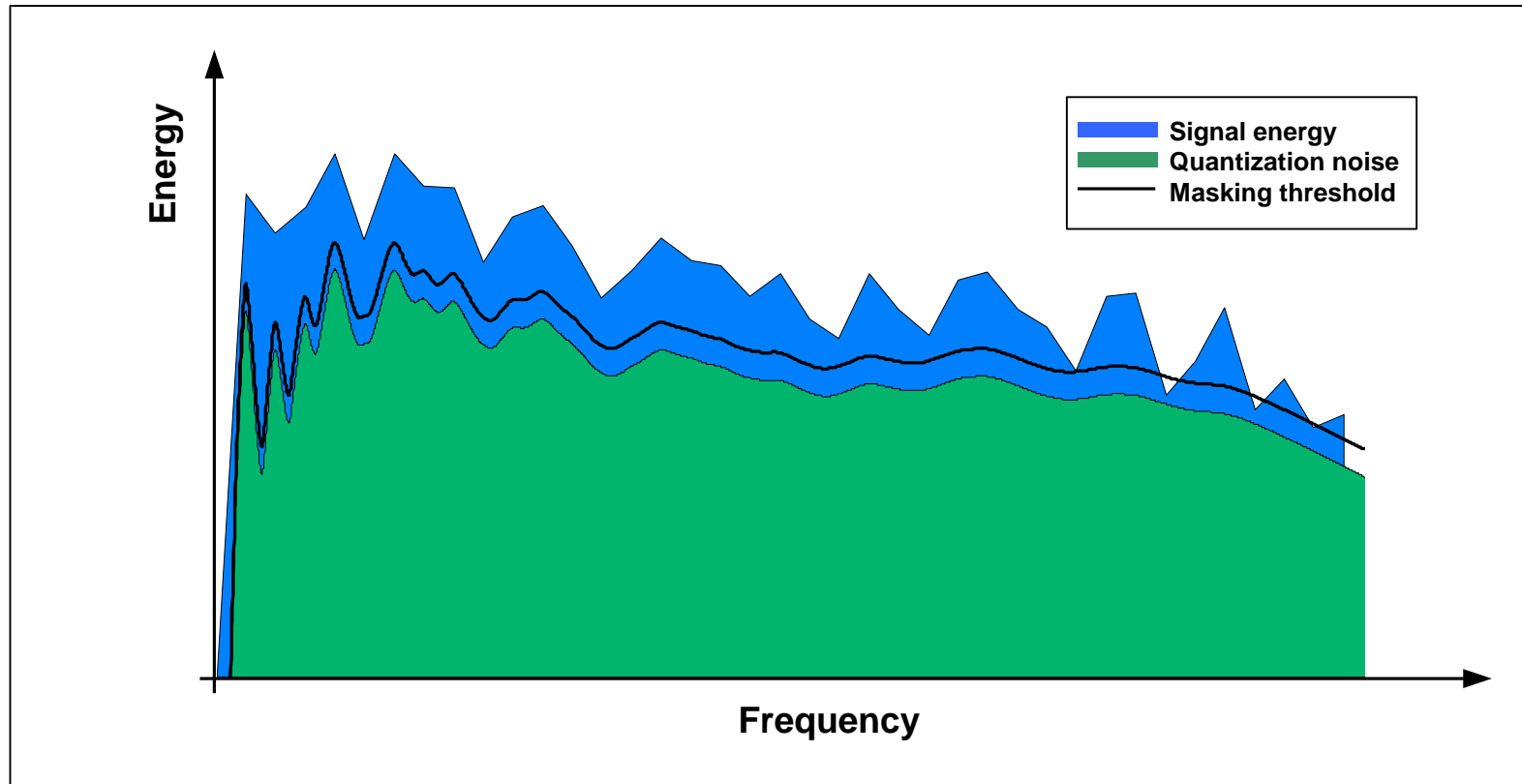
16 bit PCM



Visible blue area indicates bit rate

CLASSIC PERCEPTUAL AUDIO CODING: LIMITATIONS

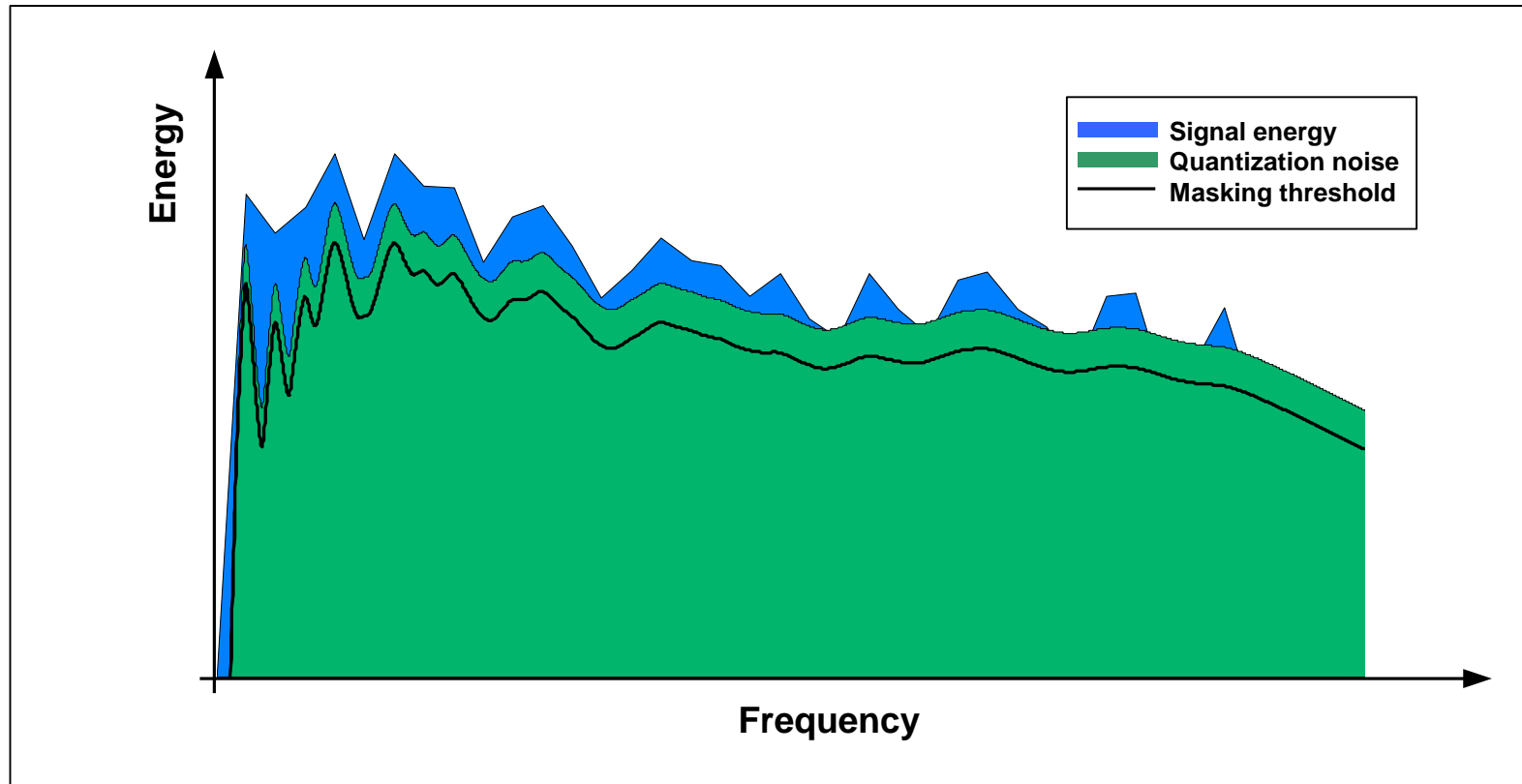
Bit rate sufficiently high



Visible blue area indicates bit rate

CLASSIC PERCEPTUAL AUDIO CODING: LIMITATIONS

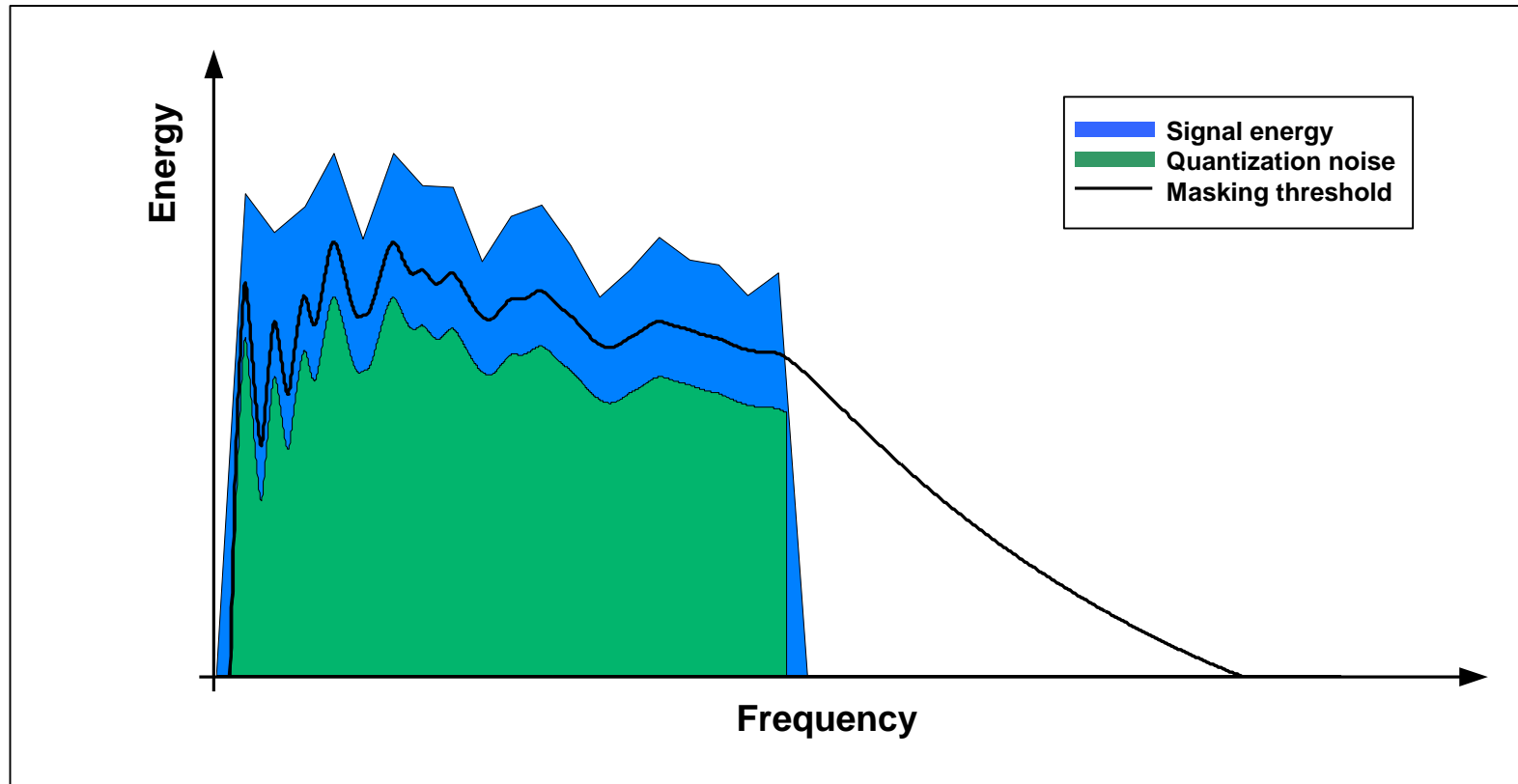
Bit rate too low: quantization noise audible



Visible blue area indicates bit rate

CLASSIC PERCEPTUAL AUDIO CODING: LIMITATIONS

Limit audio bandwidth to reduce coding artifacts



Visible blue area indicates bit rate

HIGH FREQUENCY RECONSTRUCTION / BANDWIDTH EXTENSION

Problem: Limited audio bandwidth of classic perceptual audio coding at low bit rates

Approach: Reconstruct missing high frequency band in decoder

- Example from analog age: aural exciter

Idea: Control high frequency reconstruction by side information conveyed from encoder

- Improve perceptual approximation of original signal

Result: Hybrid audio coding system

- Low frequency band: waveform coding (classic perceptual audio coding)
- High frequency band: parametric coding (high frequency reconstruction controlled by side information)

Two major elements of high frequency reconstruction

- High frequency component generation
- High frequency component shaping

HIGH FREQUENCY COMPONENT GENERATION

Synthesize high frequency components (controlled by side information)

- Noise
- Tonal components (sinusoids)
- Transients

Derive high frequency components from available low frequency band

- Upsampling: generate aliasing components
- Nonlinear processing: harmonic distortion, generate overtones etc.
- Replication of frequency bands: copy-up, modulation
- Harmonic transposition: generate overtones using phase vocoder technique (time-stretch / pitch-shift)

HIGH FREQUENCY COMPONENT SHAPING

Control spectral and temporal envelope of high frequency band

- Suitable time- and frequency resolution enables low side information bit rate
- Frequency-selective processing required (e.g. filterbank)

Control tonality/noisiness of high frequency band (if derived from low frequency band)

- Add noise components
- Add tonal components (e.g. sinusoids)
- Reduce tonality by flattening filter (whitening)

EXAMPLE: SPECTRAL BAND REPLICATION (SBR)

Overview

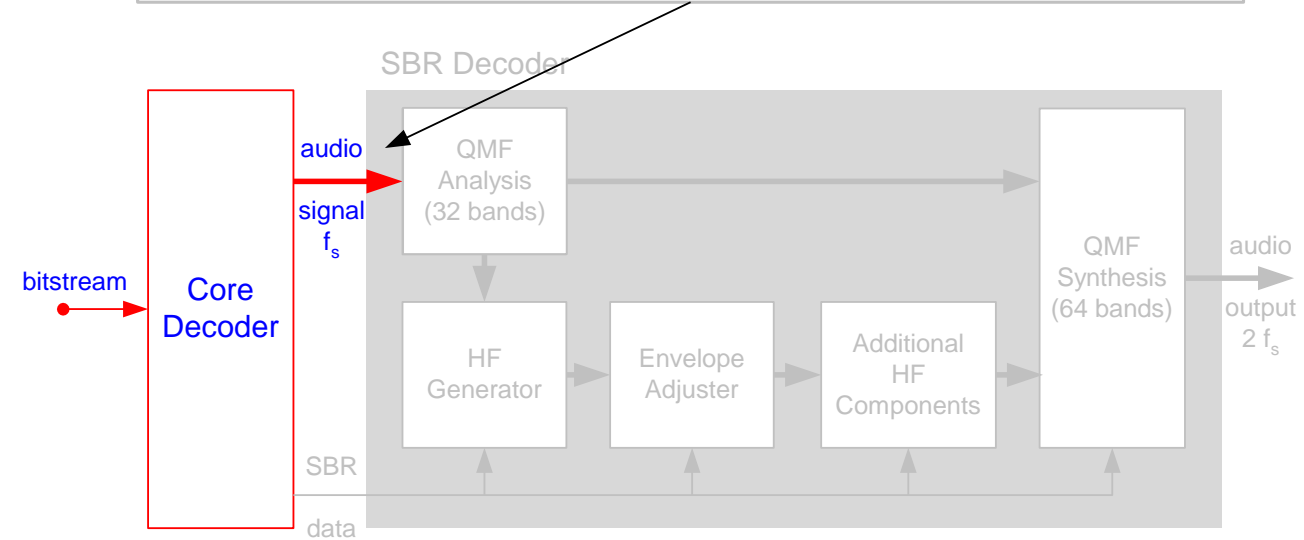
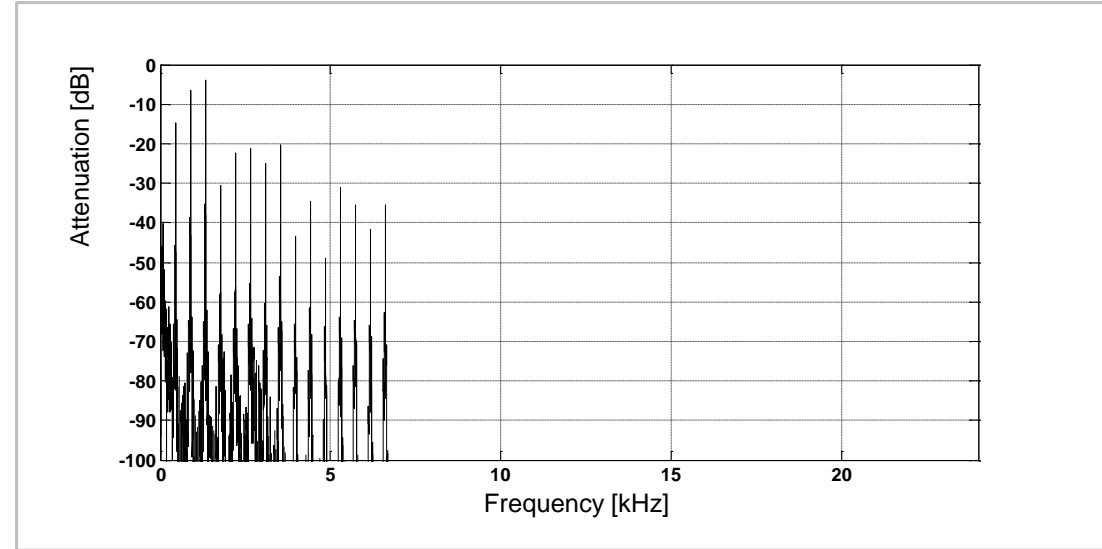
- Development started 1997 by Coding Technologies (since 2007 Dolby)
- AAC+SBR: aacPlus / High Efficiency AAC (HE-AAC)
- Standardized 2003 as part of MPEG-4 Audio

Details

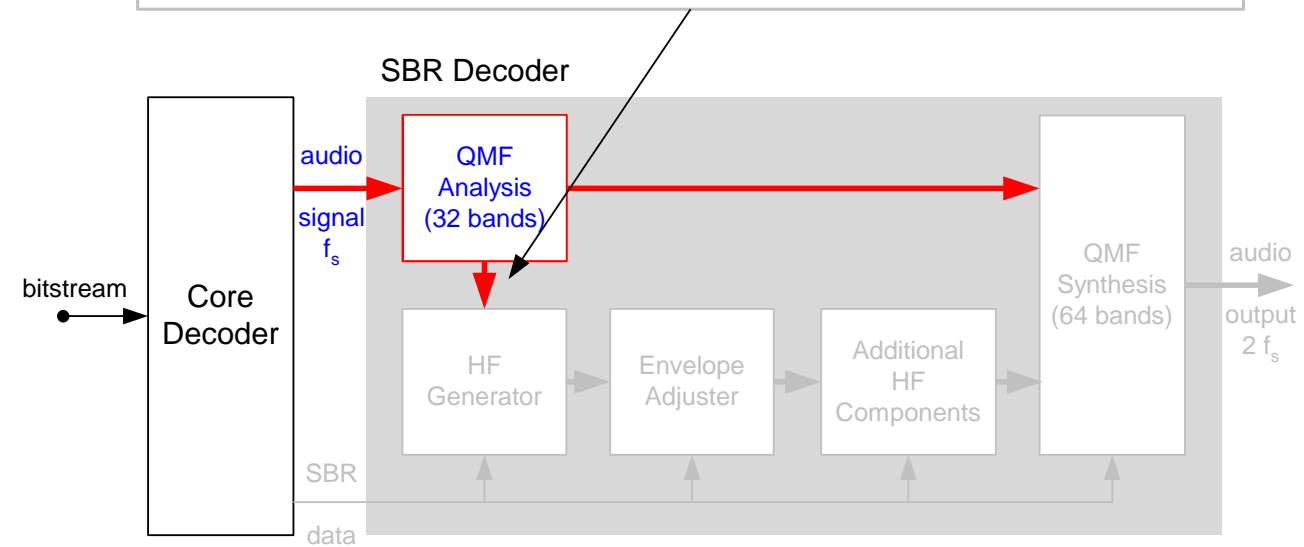
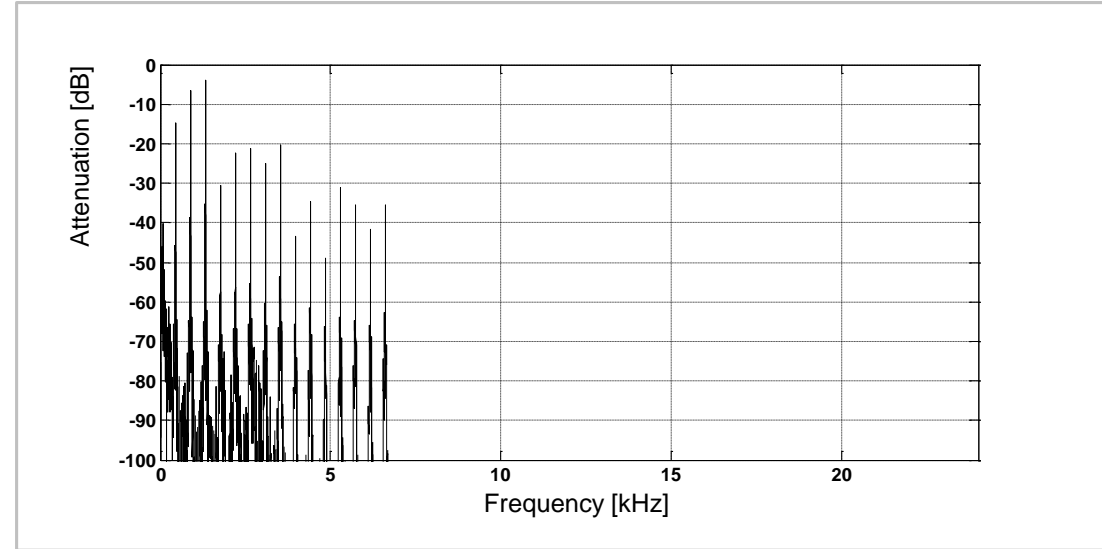
- 64-band QMF bank for high frequency generation (copy-up) and shaping
- Side information rate: typically 2 to 3 kb/s per channel
- Dual-rate system: cross-over frequency between waveform coding and HFR at max. half bandwidth



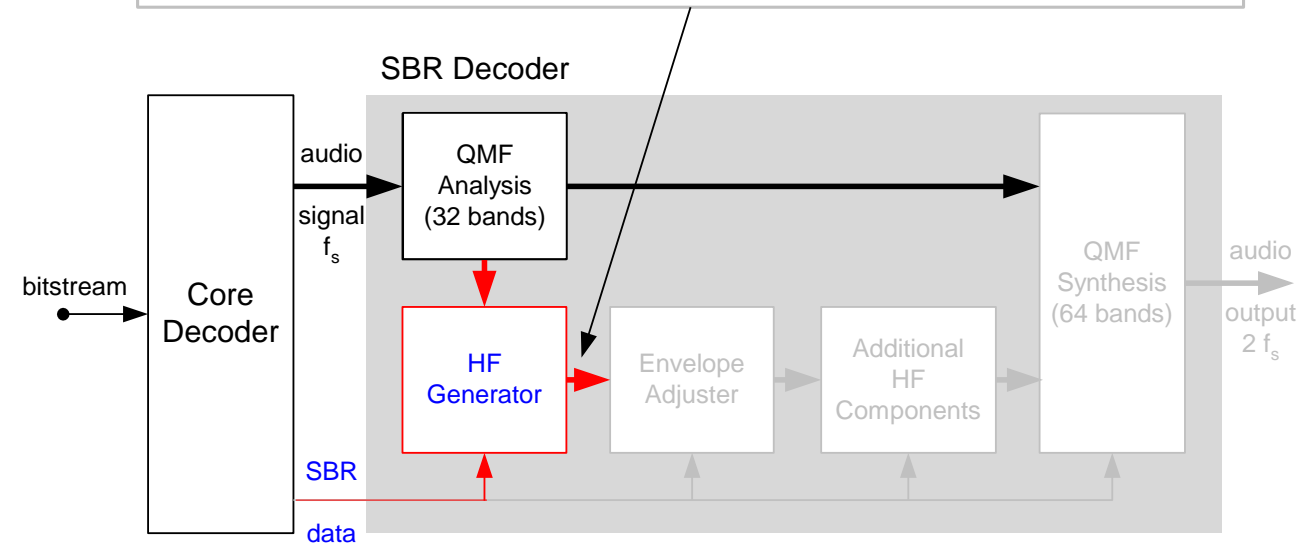
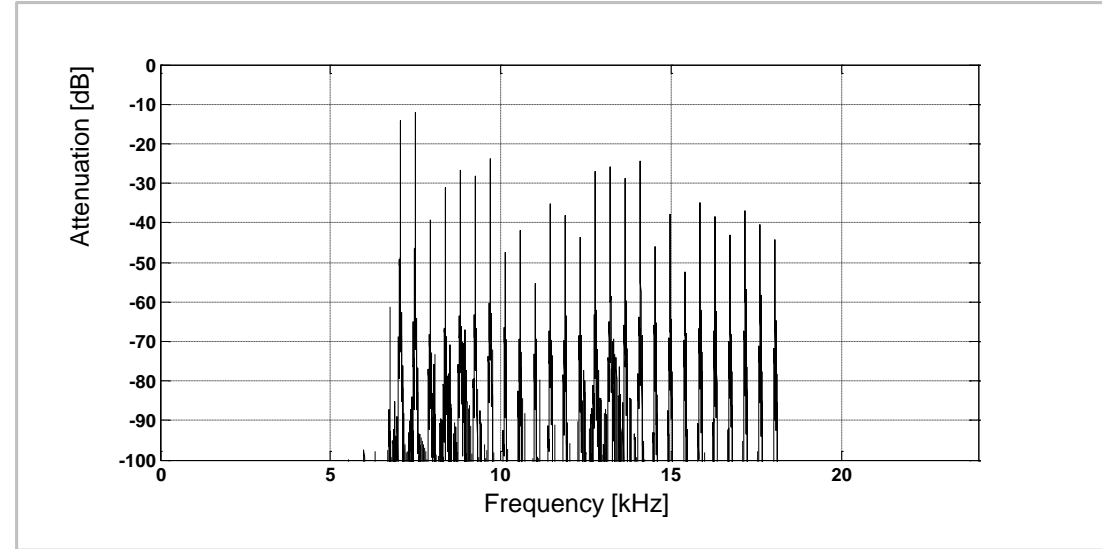
SBR DECODER



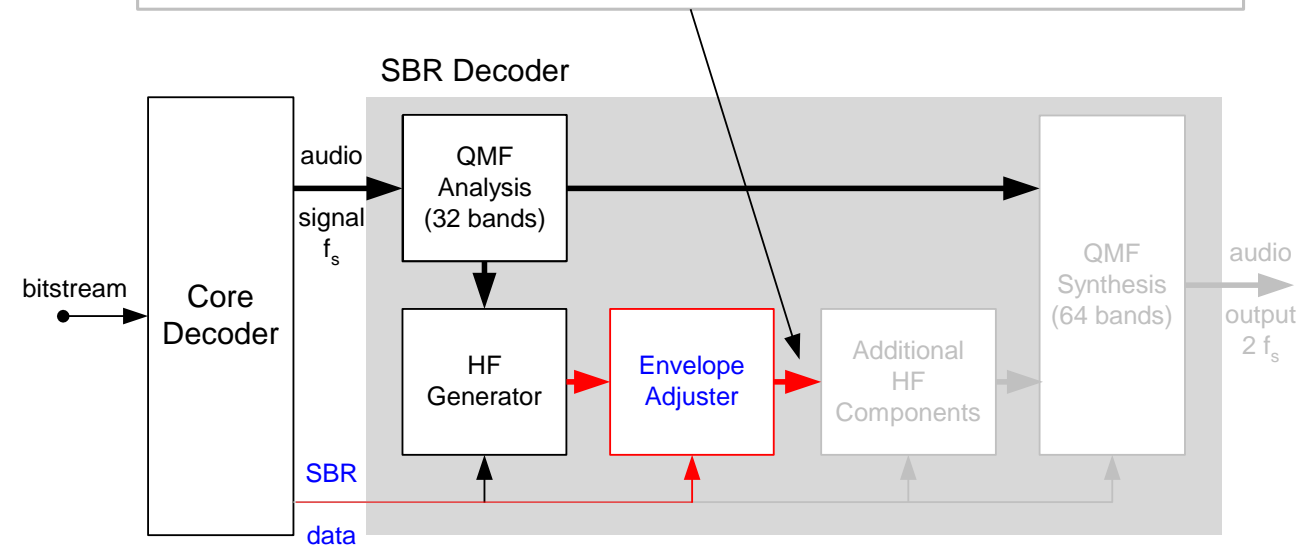
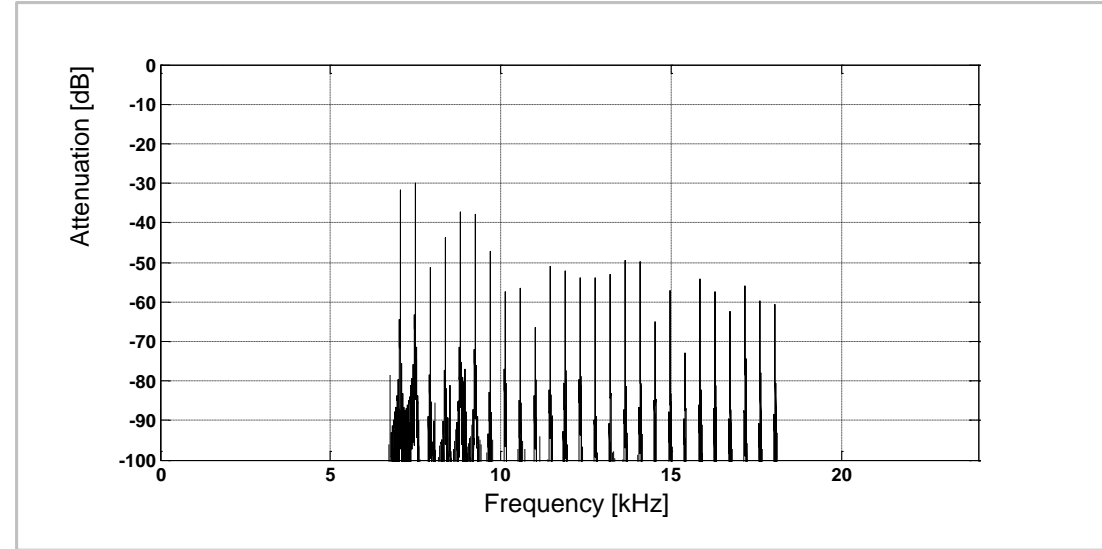
SBR DECODER



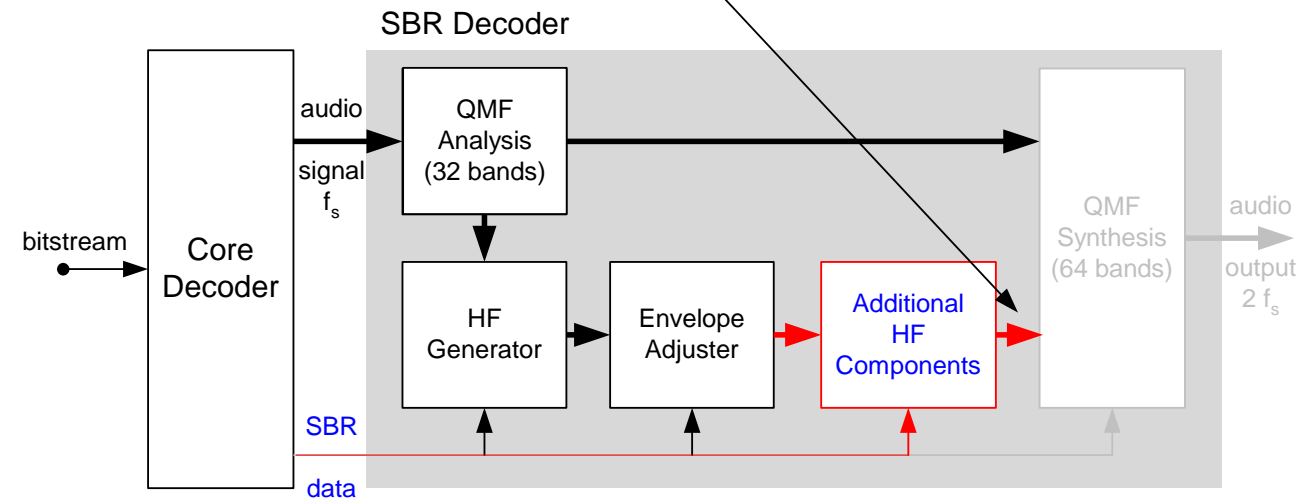
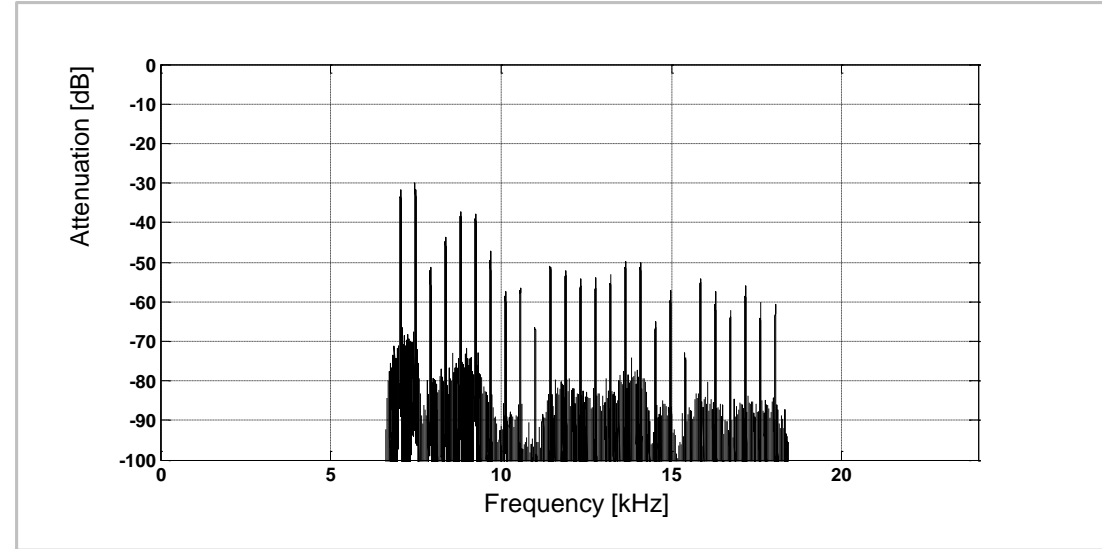
SBR DECODER



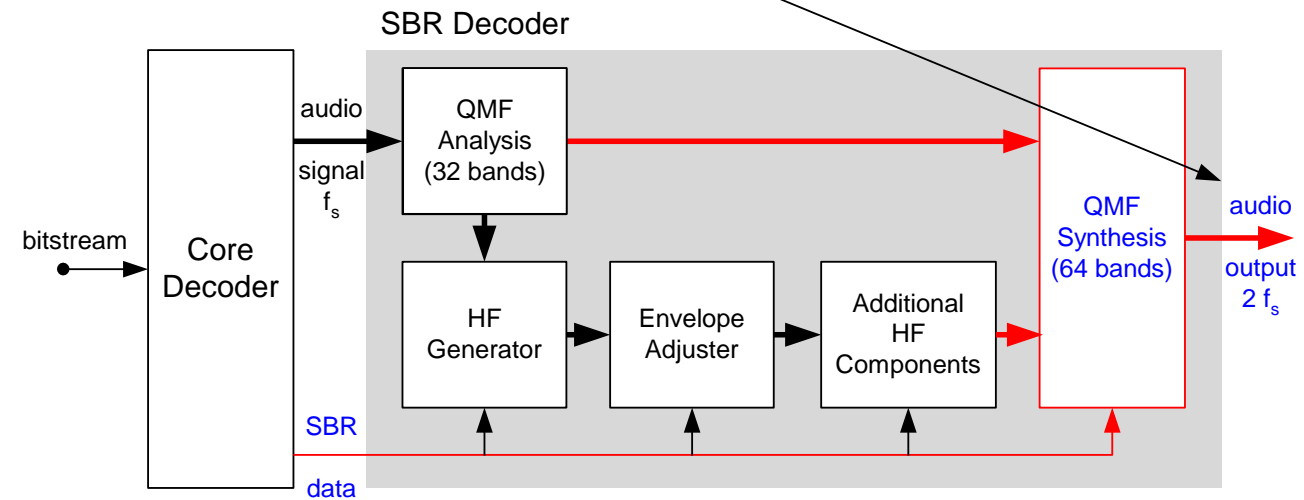
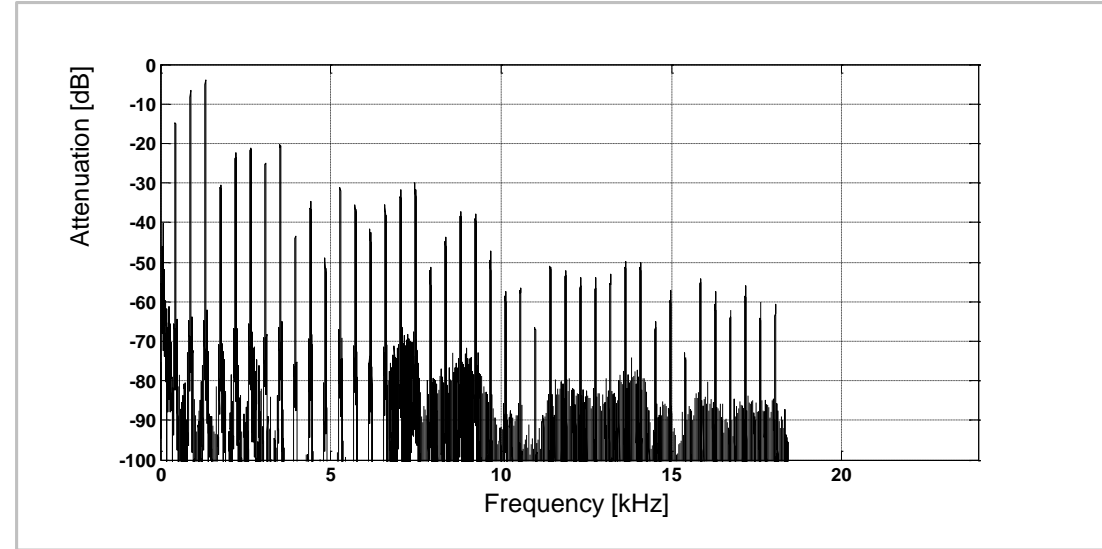
SBR DECODER



SBR DECODER



SBR DECODER



FURTHER EXAMPLES OF CODING SYSTEMS USING HFR

Dolby Digital Plus (Enhanced AC-3)

- Spectral Extension (SPX): copy-up and shaping in MDCT-domain

MPEG Unified Speech and Audio Coding (USAC)

- eSBR: includes harmonic transposition and other improvements

Dolby AC-4: Advanced Spectral Extension (A-SPX)

- Basic structure similar to SBR
- Single rate system: more than half bandwidth can be waveform-coded
- Interleaved coding: selected time or frequency sections of the high frequency band can be waveform-coded

