

MPEG-4 Scalable Audio Coding

Bernhard Grill

Fraunhofer Institute for Integrated Circuits (IIS)

Definitions

- Scalable Coding Capability to decode useful sub-sets of the coded bitstream
- Embedded / Layered Coding Usually the the subsets are ordered hierarchically
- Coding Layer One of these subsets
- Base Layer / Core Coder The first coding layer; stand alone decodable, containing no sub sets; only one base layer;
- Extension Layer Additional coding layer(s); not decodable without all coding layers lower in the hierarchy
- Core Coder The base layer is called core coder, if a coding scheme different to the extension layer is used

Types of Scalability

- **SNR / NMR (Noise to Mask Ratio) Scalability:**
 - “Extension layers improve the SNR/NMR of the coded signal”
- **Audio Bandwidth Scalability:**
 - “Extension layers increase the decodable audio band width”
- **Restriction of Generality:**
 - “Usage of a very low bit rate core coder optimized for special signals, e.g speech. This allows for good very low bit rate coding for speech signals. Additional layers provide good quality for all types of signals.”
- **Implementation Complexity:**
 - “Decoder complexity lower for base layer”

Application examples

- **Network based (packetized) transmission**
 - Requires routers which know about the importance of a packet
 - Less important (outer layer) packets may be dropped if the available bandwidth decreases
- **Broadcast**
 - The most important (inner layer) packets are transmitted with a better error protection scheme
- **Music data base**
 - High quality content is encoded and stored
 - Access to a lower quality version is possible without recoding to allow for pre-listening with a lower quality

Scalability in MPEG-4 Audio

- **Scalability within one of the MPEG-4 Coding Schemes**

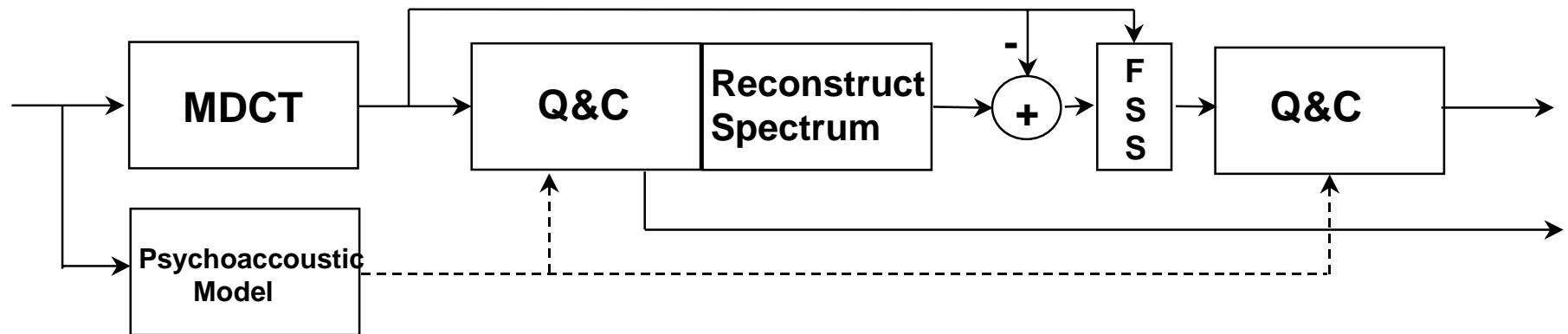
- Parametric Speech: 2 + 2 kbit/s
- CELP: Bitrate: 4..8 kbit/s + ~ 2 kbit/s steps
Bandwidth: 3.5 kHz -> 7 kHz
- General Audio Bitrate: 6..64 kbit/s + 8..64 kbit/s steps
Bandwidth: yes, multiple options
Coder Types: TVQ+TVQ, TVQ+AAC, AAC+AAC

- **Scalability based on the Combination of the MPEG-4 GA Coder and the MPEG-4 CELP Coder**

- Bitrate: 4..8 kbit/s CELP + ~2 kbit/s CELP
+ 8..64 kbit/s AAC Layer
- Bandwidth: 3.5 kHz CELP, extension by AAC layers

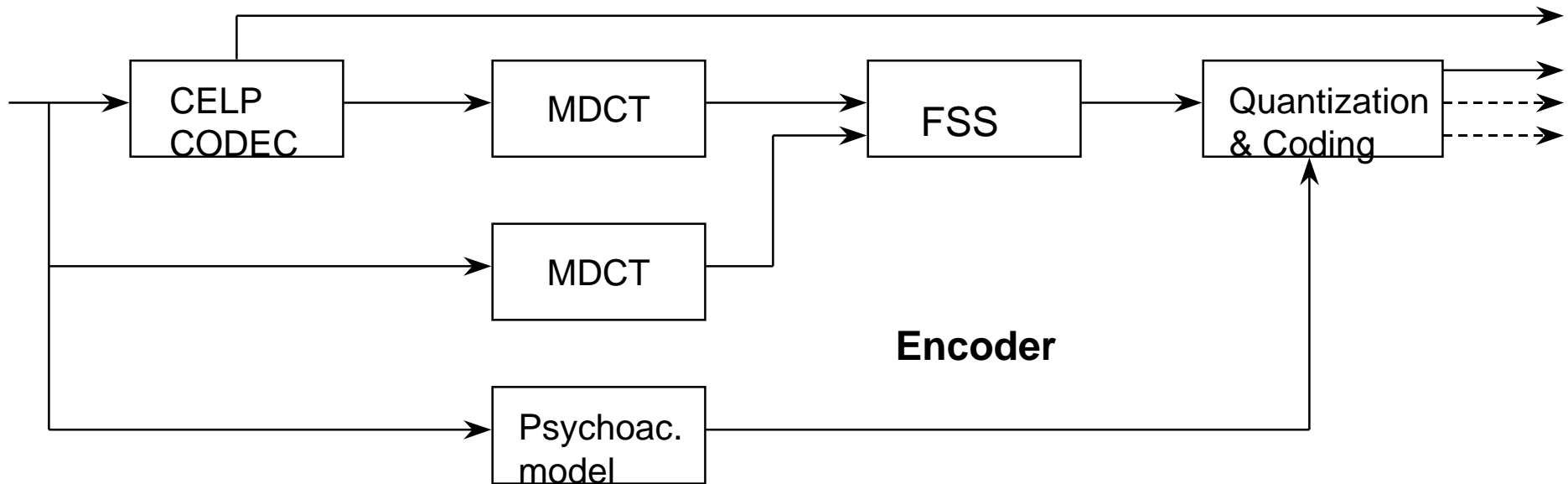
Scalable GA Coder (I)

Encoder Block Diagram



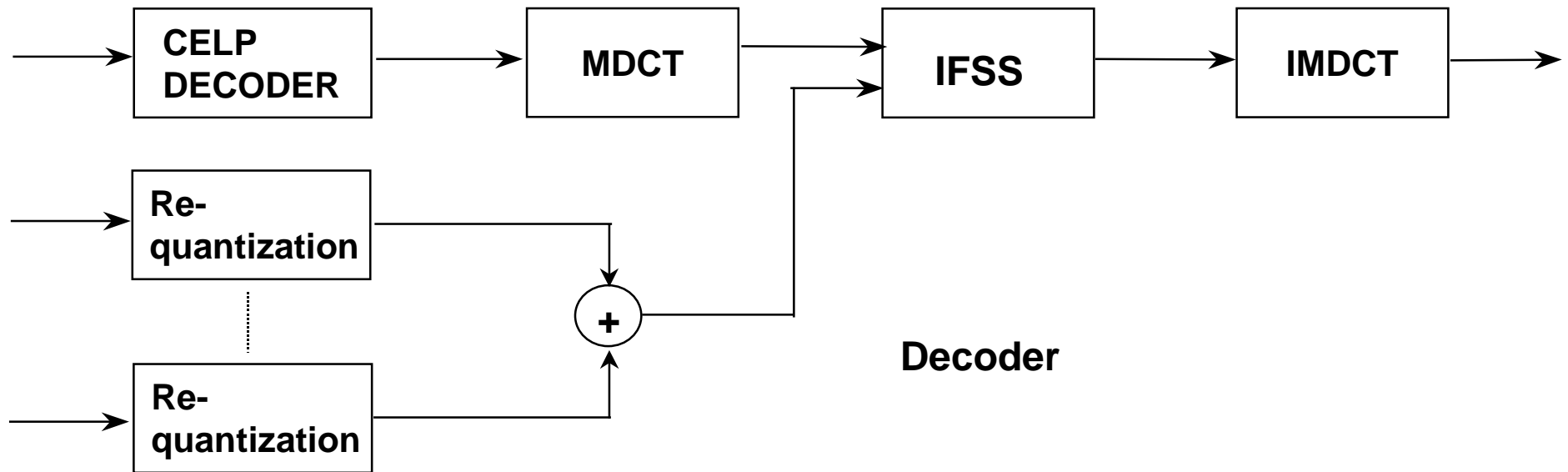
- Encoding of the error signal of an AAC Quantization and Coding (Q&C) module in a second, or third, or n-th similar quantization module in the frequency domain
- Useful for large enhancement steps of ≥ 8 kbit/s per step

Scalable GA Coder : Combination with Core Coder (I)



- Very low bitrate core coder (e.g. speech coder)
- Core coder typically operating at a lower sampling frequency
- MDCT used for efficient up-sampling

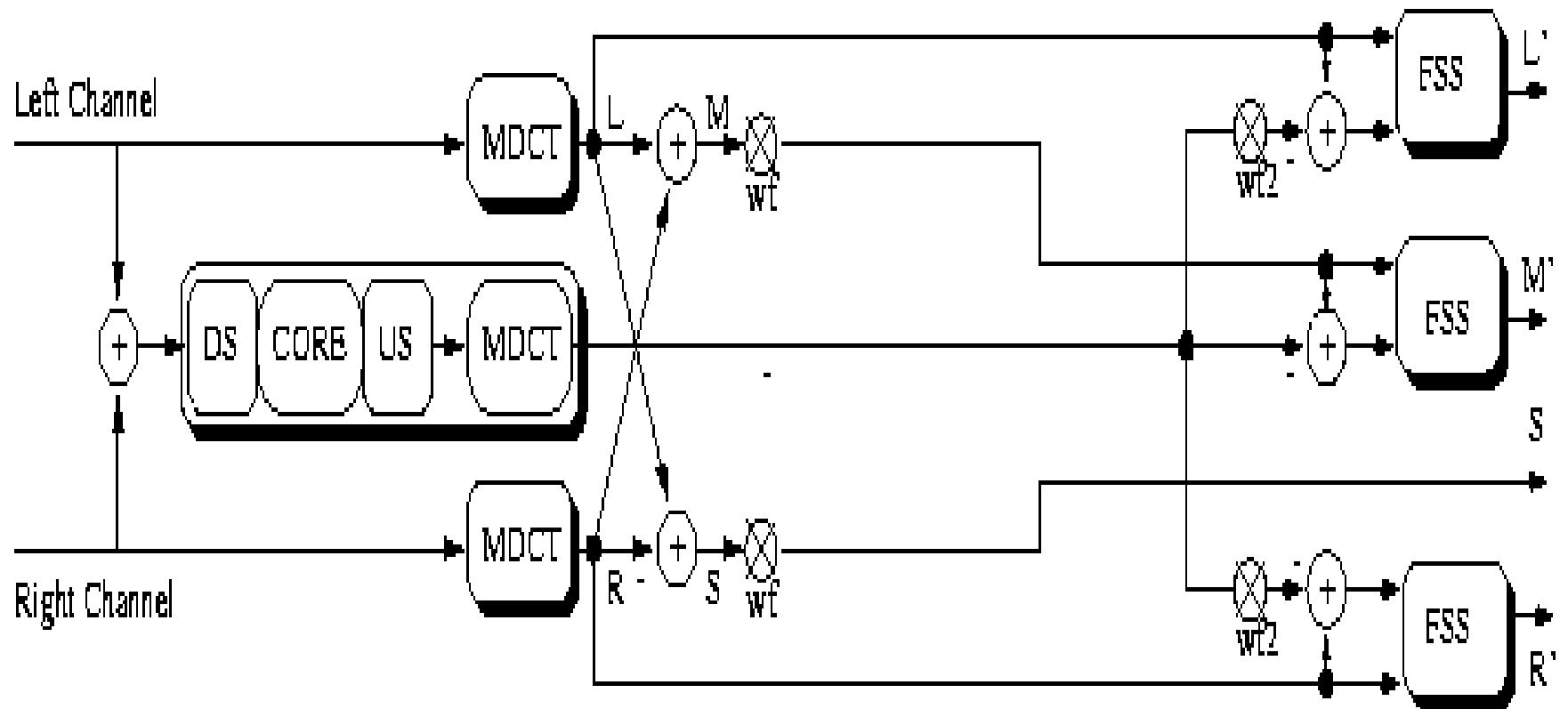
Scalable GA Coder : Combination with Core Coder (II)



Scalable Joint Stereo Coding

- **Four Basic Configurations**
 - All coding layers are GA-Layers and code a stereo signal
 - All coding layers are GA-Layers; Lower layers code a mono signal, higher layers a stereo signal
 - The base layer is a CELP coder; Extension layers are AAC stereo layers
 - The base layer is a CELP coder; Extension layers are AAC mono and stereo layers
- **MS- and Intensity-Stereo are possible**
 - No changes from IS to MS or MS to IS from layer to layer

Scalable Stereo Coding: Mono Core / Stereo



Scalable GA Coder : Typical Configurations

- Some successfully tested mono/mono combinations:

6 kbit/s CELP	+ 18 kbit/s AAC
6 kbit/s TwinVQ	+ 18 kbit/s AAC
8 kbit/s TwinVQ	+ 8 kbit/s TwinVQ

6 kbit/s CELP + 18 kbit/s + 24 kbit/s AAC

- Mono/stereo combinations

6 kbit/s mono CELP + 18 kbit/s mono + 24 kbit/s stereo AAC
24 kbit/s mono + 16 kbit/s stereo + 16 kbit/s stereo AAC
24 kbit/s mono + 72 kbit/s stereo AAC

- Stereo/stereo combinations

2 x 6 kbit/s mono CELP + 36 kbit/s stereo AAC