

Bernhard Grill

# MPEG-4 Scalable Audio Coding



- Overview**
- Introduction to “Scalable Audio Coding”
    - What is scalable audio coding
    - What is it good for
  - Overview of the Scalability Options in MPEG-4 Audio
  - The MPEG-4 Scalable CELP Coder
  - The MPEG-4 Scalable General Audio Coder
  - Results of Listening Tests
  - Demonstrations of a Real-Time Player



## 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



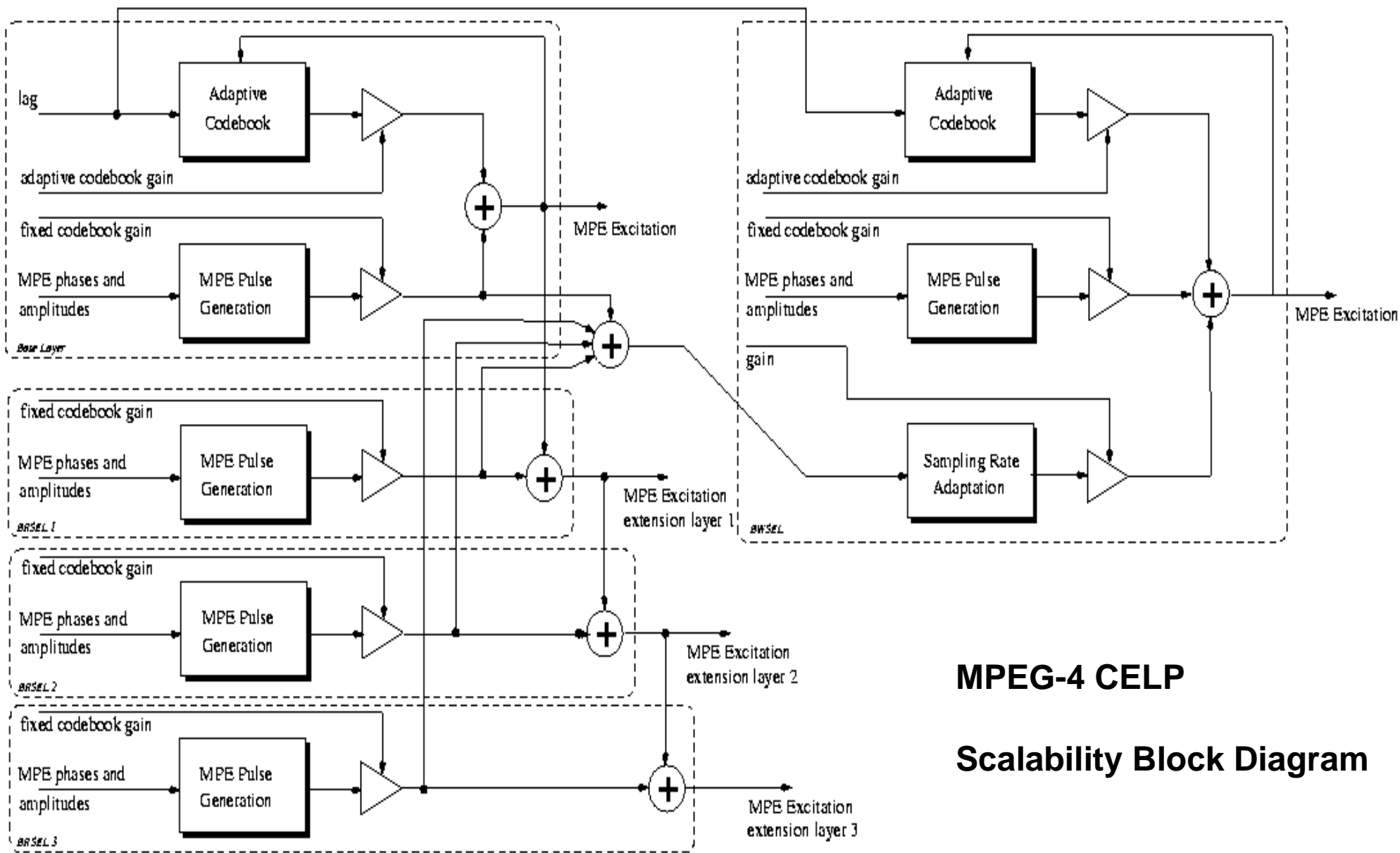
## Scalability in MPEG-4 Audio

- **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

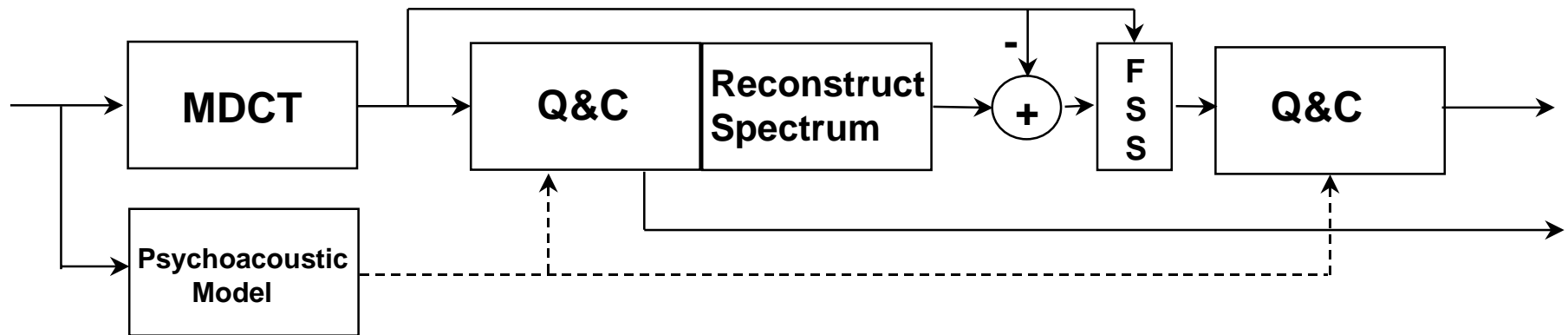


## MPEG-4 CELP

### Scalability Block Diagram

## Scalable GA Coder (I)

Encoder Block Diagram

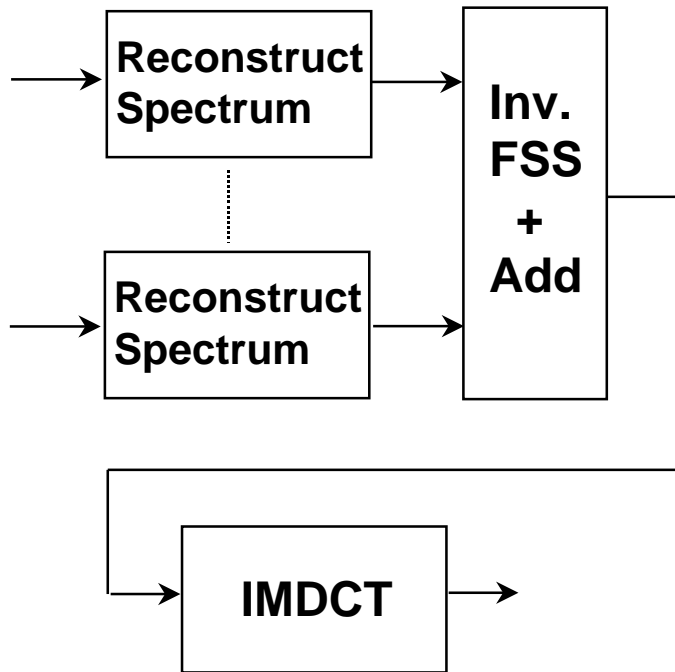


- Encoding of the error signal of an AAC or Twin-VQ Quantization and Coding (Q&C) module in a second, or third, or n-th similar quantization module in the frequency domain
- Solutions using only AAC, or only Twin-VQ modules possible
- Additionally, Twin-VQ / AAC combinations defined
- Useful for large enhancement steps of  $\geq 8$  kbit/s per step



## Scalable GA Coder (II)

### Decoder Block Diagram



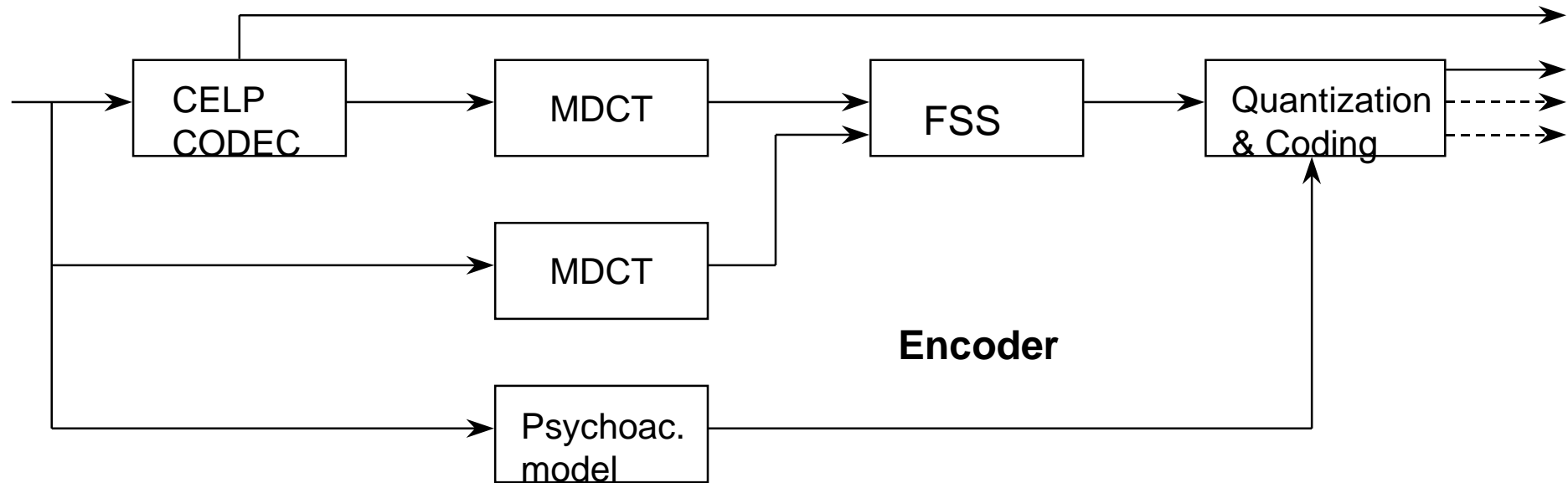
- **Twin-VQ Q&C Modules**

- 8 kbit/s fixed step size Vector Quantizer (VQ) modules
- optional 6 kbit/s in first layer
- first choice for a 6 or 8 kbit/s base layer for the coding of general audio signals

- **AAC Q&C modules**

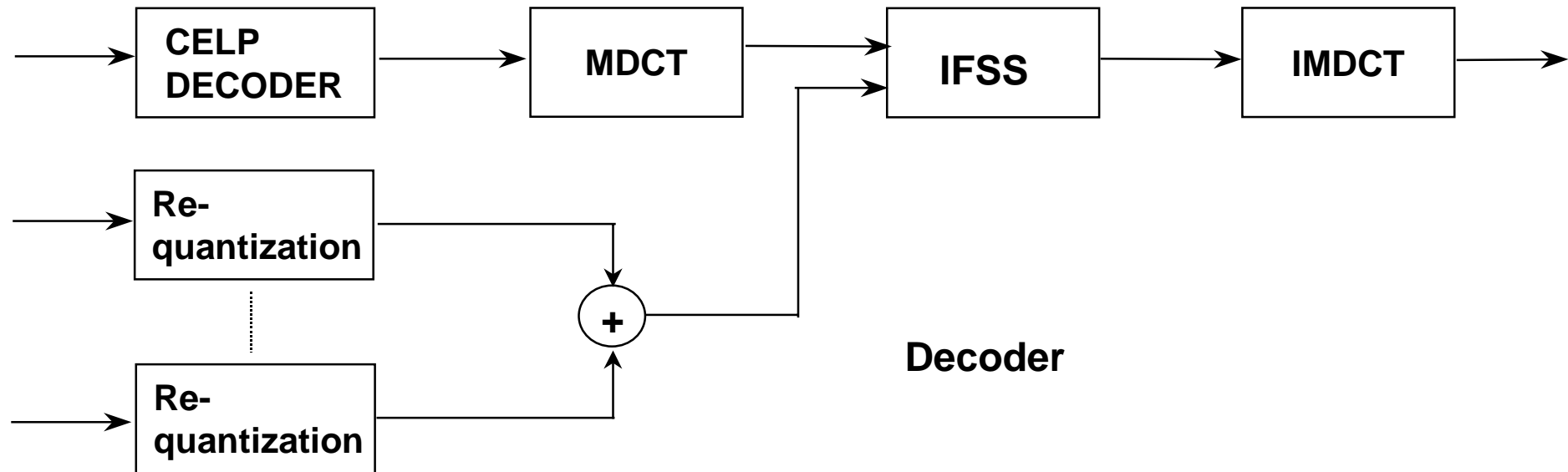
- Any step size possible
- Reasonable step sizes from 8 to >64 kbit/s
- The same end quality can be achieved as from a single step AAC coder
- However, a higher bit rate may be required for the same audio quality

## 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)



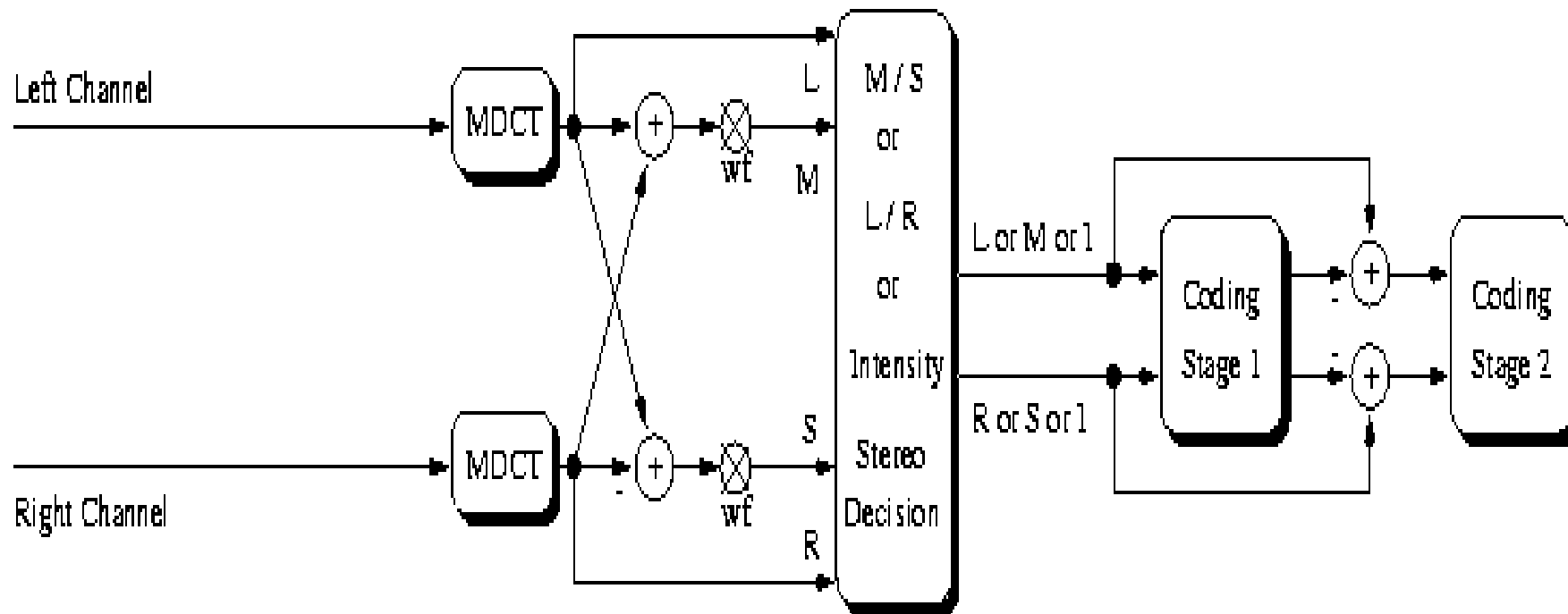
## Scalable GA Coder: Combination with Core Coder(III)

- Different core coders are possible
  - MPEG-4 CELP 4.9 or 6.0 or 8.0 kbit/s
  - MPEG-4 Parametric Music Coder HILN 6 e.g. kbit/s
  - ITU G.723.1 or G.729
  - Proposed: GSM Coders
- GA coder operates at an alternative frame length of 960 (instead of 1024) samples to match the frame length of the speech coders
  - Gives nice numbers for the T/F (AAC) coder frame length. E.g. 20 ms instead of 21.333... ms at 48 kHz sampling rate
  - Alternative frame length of 960 samples also available for a stand alone T/F (AAC) coder in MPEG-4

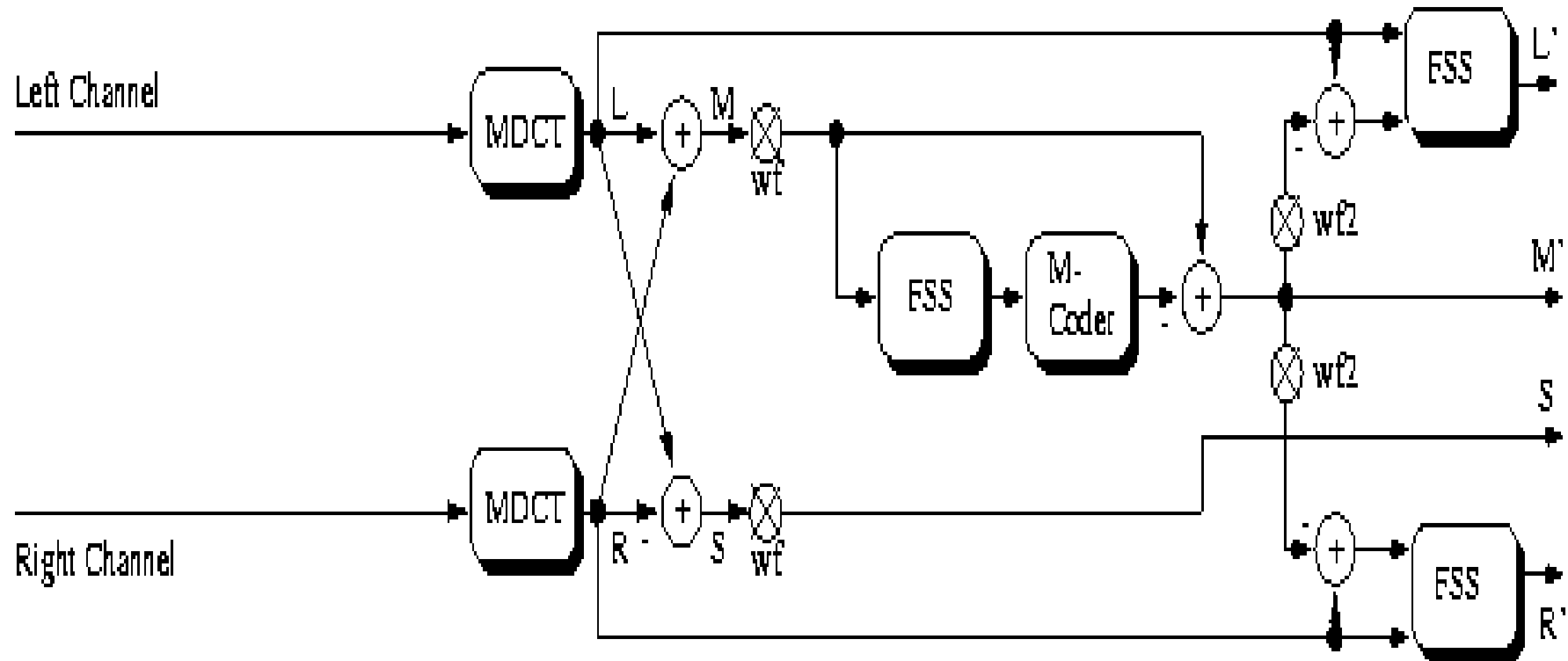
# Scalable Joint Stereo Coding

- **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

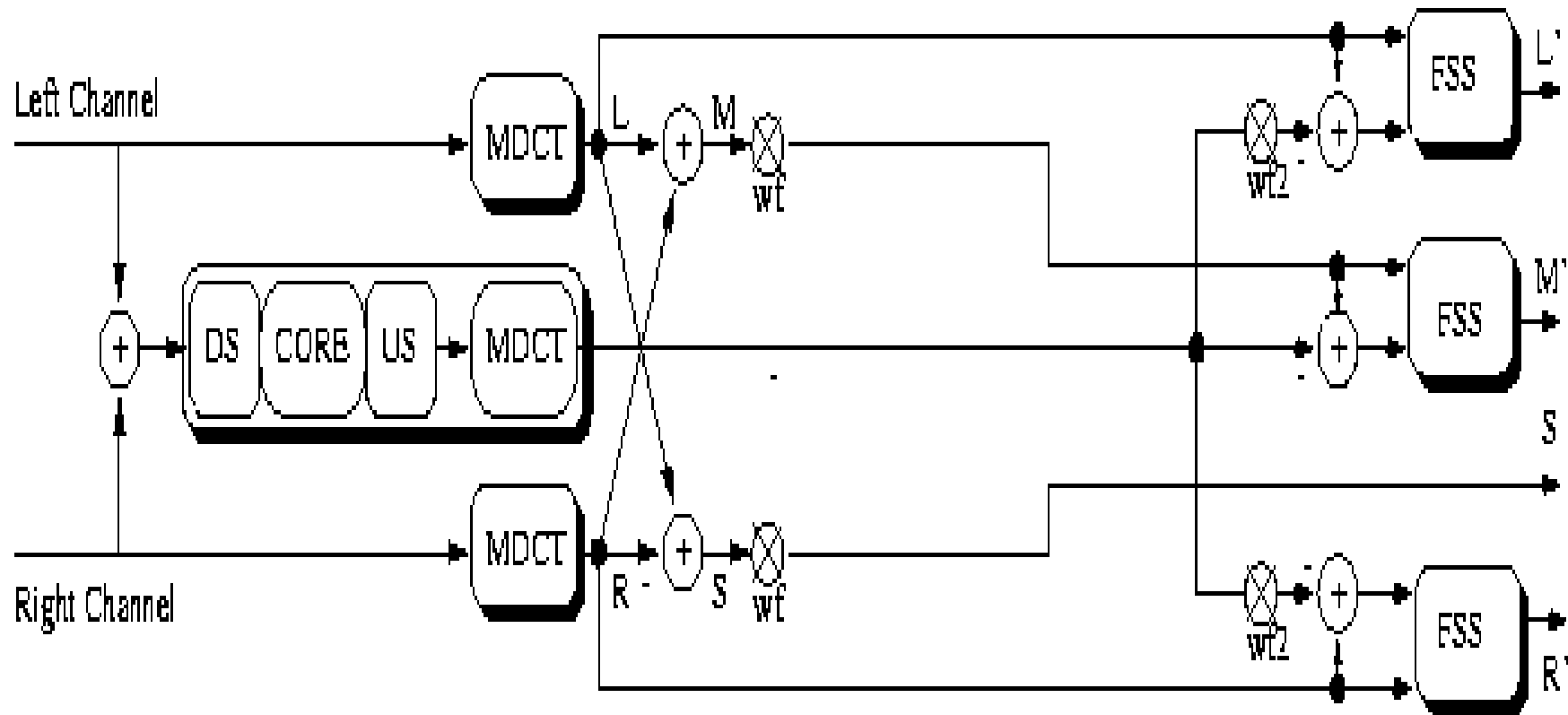
# Scalable Stereo Coding: Stereo / Stereo



# Scalable Stereo Coding: Mono / Stereo

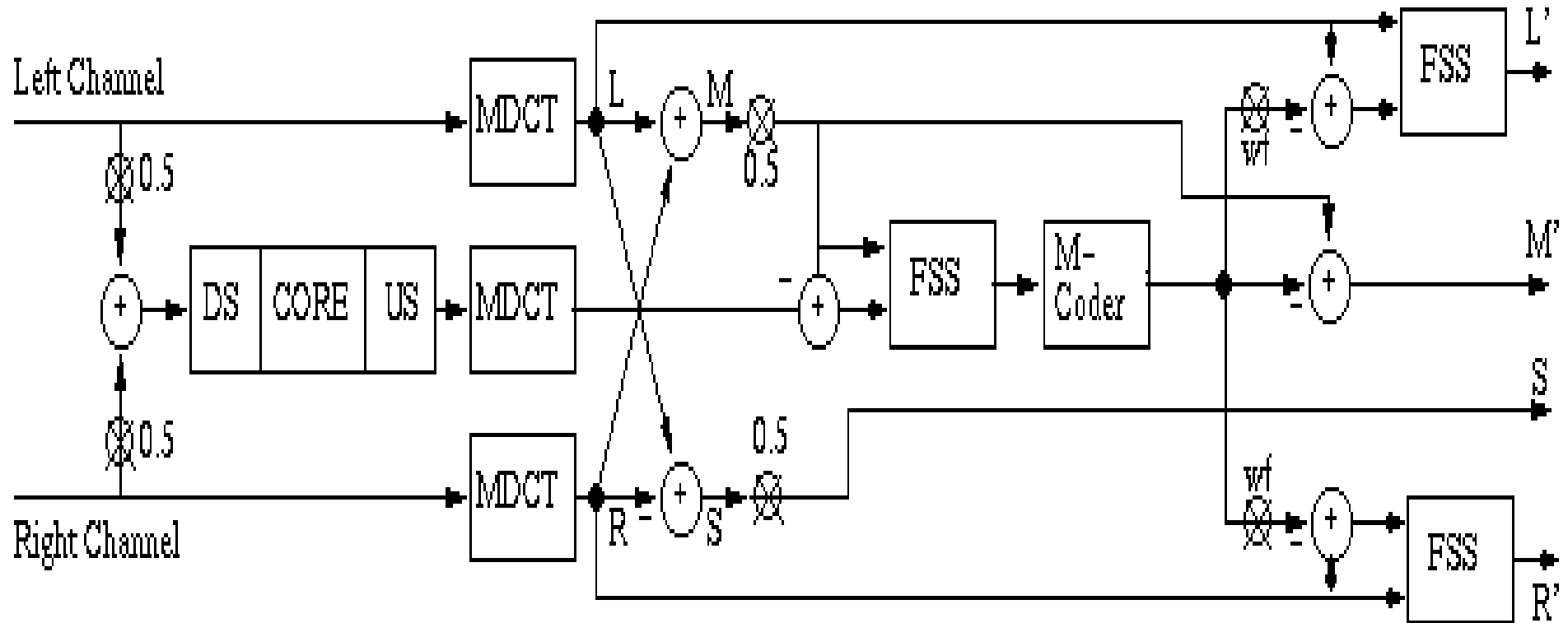


# Scalable Stereo Coding: Mono Core / Stereo





# Scalable Stereo Coding: Mono Core/Mono GA/Stereo GA



## 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

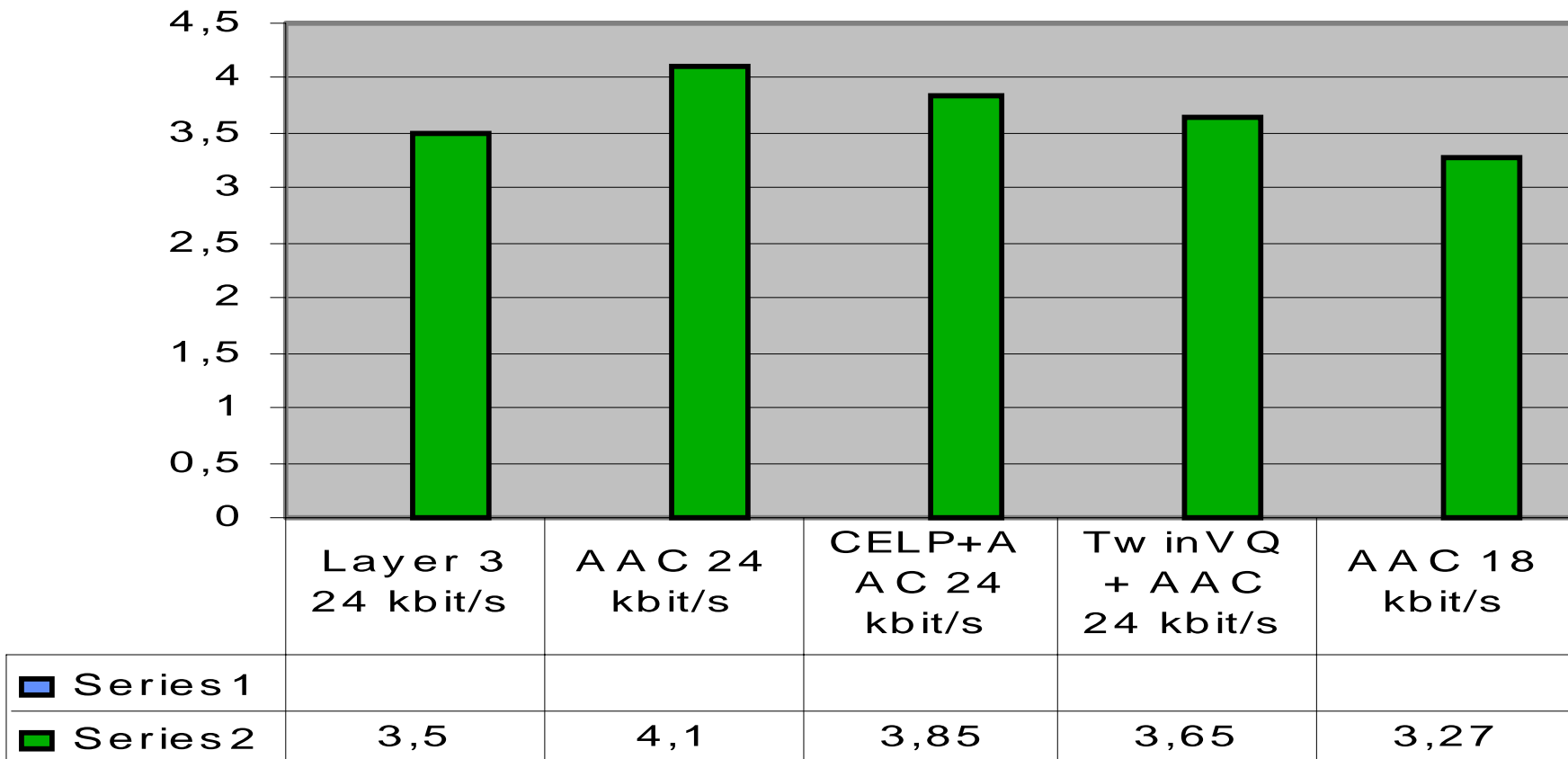
- Mono/stereo combinations

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

# Results (I) Mono Configurations

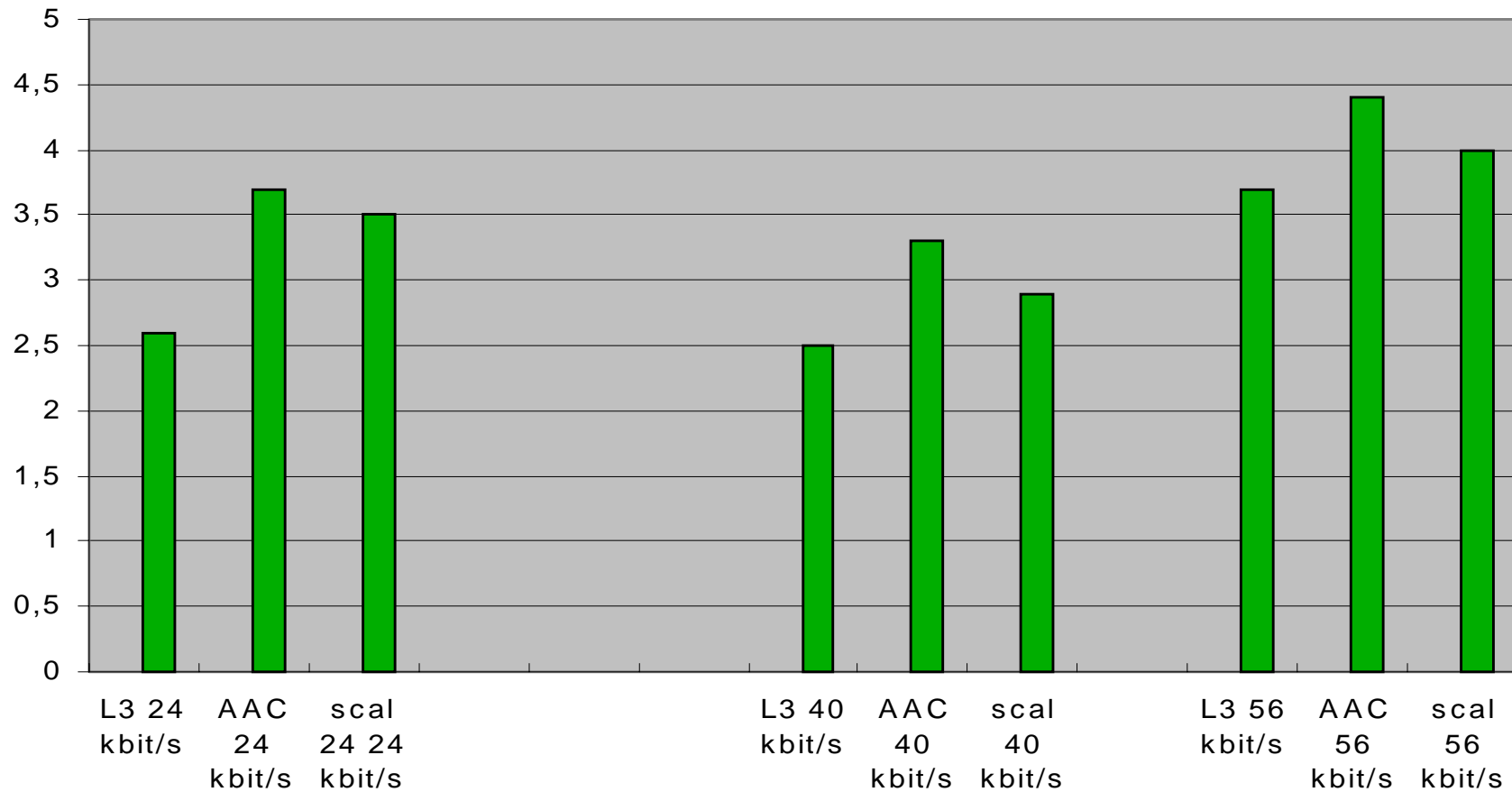


gr1 6/98



## Results (II)

## Mono / Stereo Configuration



gr1 6/98



# Conclusions

- Concept of scalable audio coding verified in formal listening tests
- The scalable coding modes make the MPEG-4 GA coder the most versatile coding system available today
  - Highest quality coding at bitrates of 64 kbit/s
  - Integrated coder for optimum performance at very low bit rates
  - Additional flexibility with the availability of the scalable coding modes
  - Unique capabilities through the availability of the mono-stereo coding modes
- Overall complexity within the limits of today's hardware