CODING OF FINE GRANULAR AUDIO SIGNALS USING HIGH RESOLUTION ENVELOPE PROCESSING

Florin Ghido[1,2], Sascha Disch[1,2], Jürgen Herre[1,2], Franz Reutelhuber[1], Alexander Adami[2]

[1] Fraunhofer Institut für Integrierte Schaltungen (IIS), Am Wolfsmantel 33, 91058 Erlangen, Germany

[2] International Audio Laboratories Erlangen, Am Wolfsmantel 33, 91058 Erlangen, Germany

Contact: sascha.disch@iis.fraunhofer.de

Problem Statement

High Resolution Envelope Processing (HREP):

- Improved perceptual coding for traditionally difficult signals containing dense transient events (applause, raindrops, etc.)
- Signal pre-/post-processing based on gain control principle
- **Preserves the temporal fine structure and subjective quality of applause-like signals**

Solution Description

HREP pre-processing:

- Split of input signal into low pass (LP) and high pass (HP) parts via DFT; subtraction in time domain to obtain HP part
- Application of a **time-dependent scalar gain** to HP part
- Addition of processed HP part and LP part

HREP side information:

- HREP analysis block (not depicted) estimates parameters
- ***Beta_factor***, 3 bits, once per frame
- ***Scalar gains g(k)***, max. 3 bits per gain, adaptive entropy coding
- **Example: assuming 1024 sample frame with 16 blocks of 128 samples, 50% overlap, the maximum side information size is 3+16×3 = 51 bits/frame**

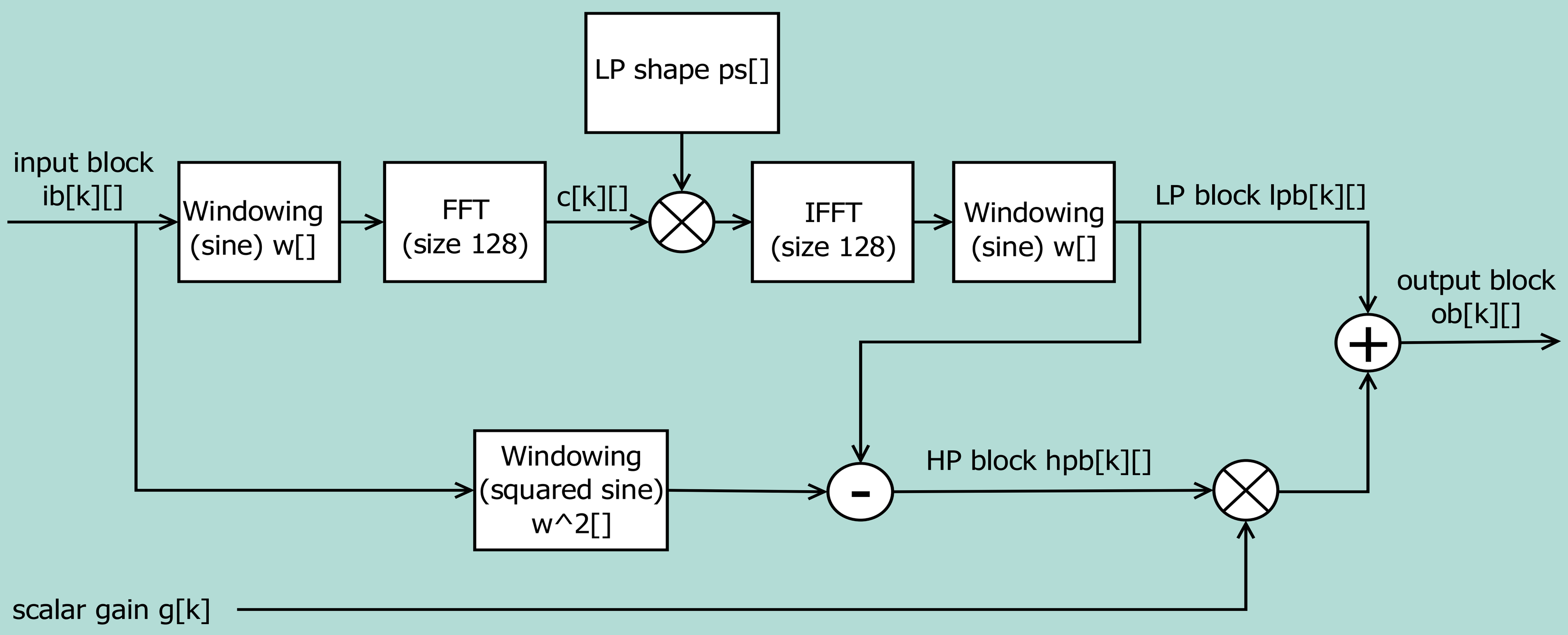

Figure 1: HREP pre-processing

Relation to Temporal Noise Shaping (TNS)

- TNS is applied *after* down-mixing into transport channels; unlike HREP, it cannot preserve each channel's envelope
- Combination of HREP with TNS is recommended
- HREP for sound textures with densely spaced transients
- TNS for prominent claps, single transients and note onsets
- Reduced TNS activation rate without losing its benefits

Computational Complexity

- Computational complexity at the decoder side dominated by the calculation of a DFT/IDFT pair for each block
- Cost of a DFT/IDFT pair at 48 kHz sampling frequency (real valued, size 128, half-overlap): **2.72 Mops per signal**

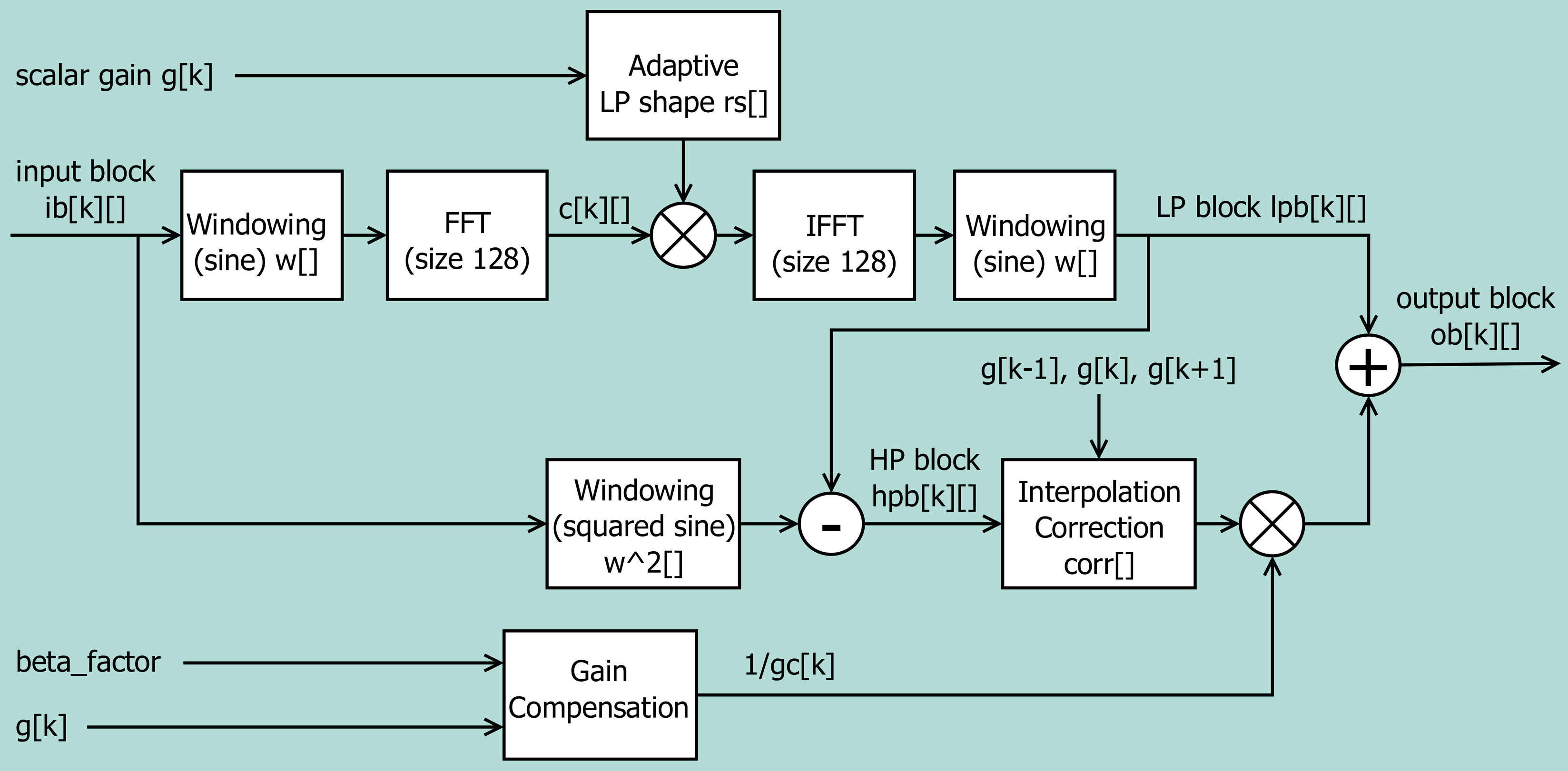

Figure 2: HREP post-processing

Evaluation

Methodology	MUSHRA ITU-R BS.1534	
Test Conditions	"nohrep": Reference Quality Encoder "hrep": HREP enabled configuration	
Test Signals	1: ARLapplause 2: Exercise 3: HeavyApplause 4: Intro	5: Klatschen 6: MediumApplause 7: SallyBrown 8: applse
Bit Rate/Ch.	128 kbps stereo (high quality)	
No. of listeners	14	
Bit Rate/Ch.	48 kbps stereo (medium quality)	
No. of listeners	15	

Table 1: Subjective quality evaluation parameters

Results

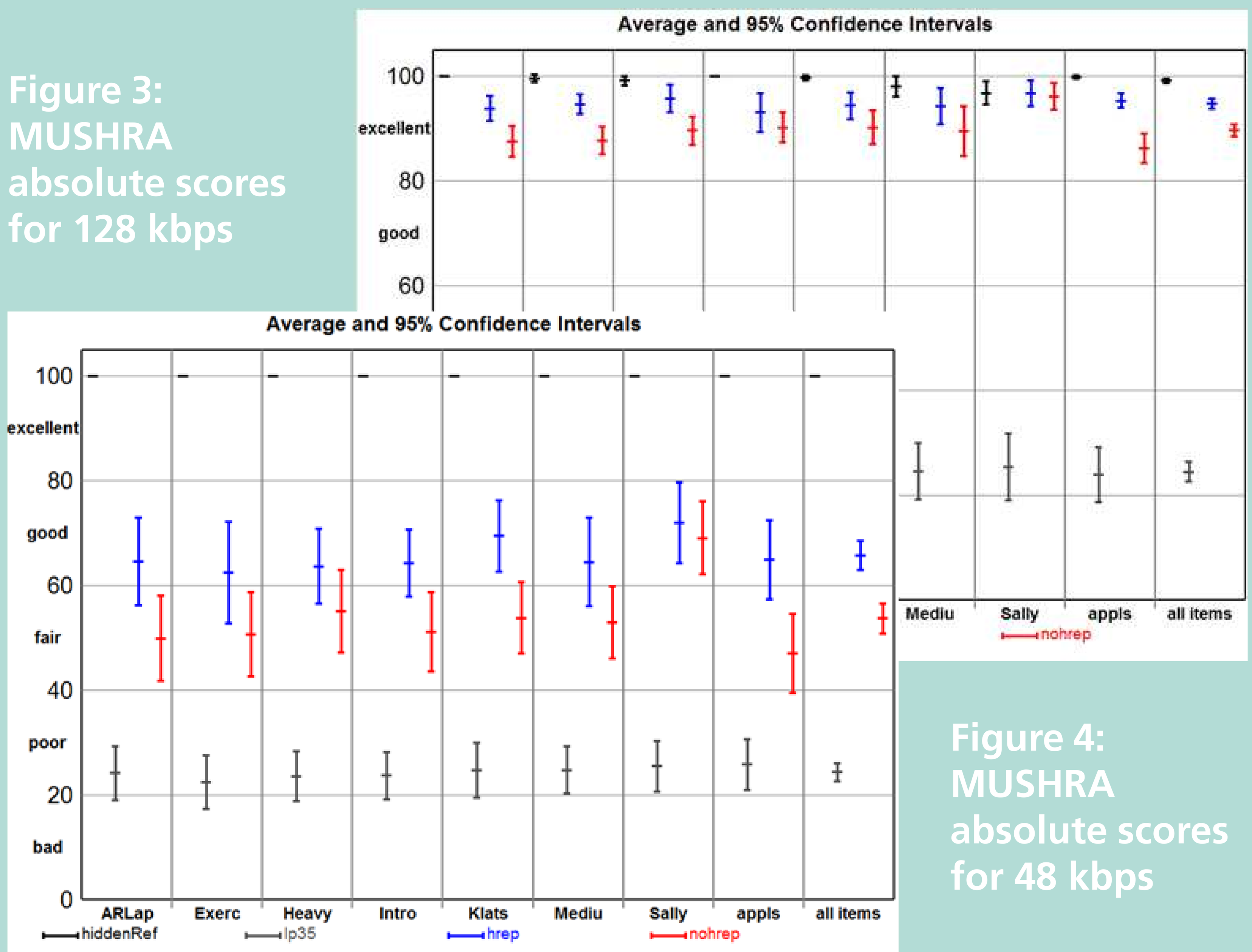

Figure 3: MUSHRA absolute scores for 128 kbps

Figure 4: MUSHRA absolute scores for 48 kbps

Conclusion

- HREP is a tool for improved perceptual coding of signals with many dense transient events
- **Moderate computational complexity** at the **decoder** side (approx. **2.72 Mops** per signal)
- **Improvement of** approx. **12 MUSHRA points at 48 kbps stereo** (MPEG-H 3D Audio codec with HREP activated)
- HREP is part of MPEG-H 3D Audio Second Edition standard

IN ZUSAMMENARBEIT MIT

Fraunhofer IIS

AUDIO LABS