Multiresolution STFT Phase Estimation with Frame-Wise Posterior Window Length Decision

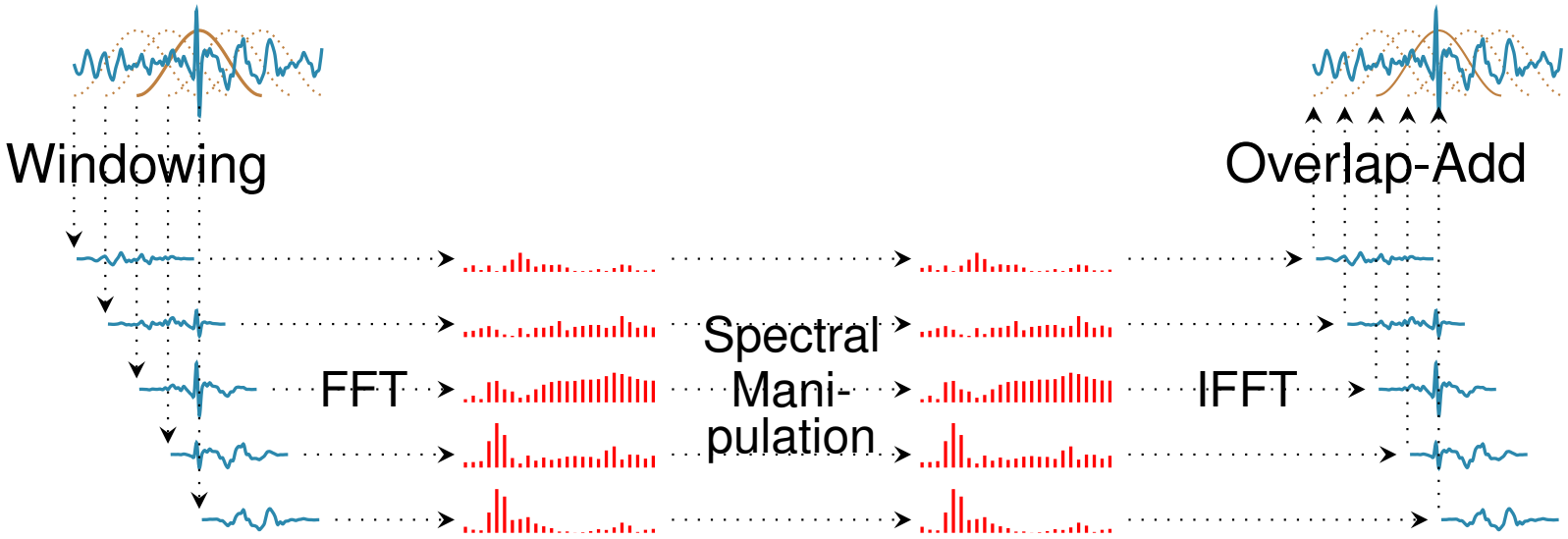
Volker Gnann

Institut für Nachrichtentechnik
RWTH Aachen University

Overview

- Motivation
- RTISI Phase Estimation
- Parallel Multi-Resolution RTISI Phase Estimation
- Posterior Minimax Decision
- Experiments and Results

Motivation: What Is Phase Estimation Good For?

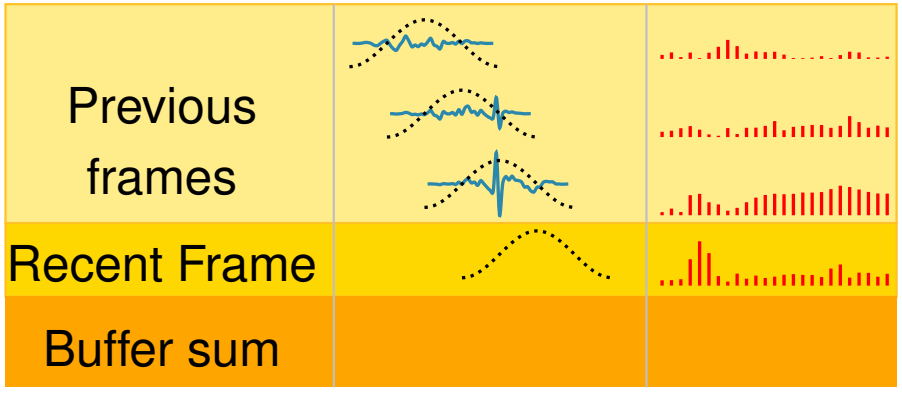


FFT Coefficients are complex

- $C = |C| \cdot e^{j\varphi}$ mit $|C|$ = Magnitude, φ = Phase
- Magnitude **and** phase are required for reconstruction: ♫, ♫
- In many applications no (or just an approximate) phase is available.

RTISI Phase Estimation

RTISI = Real-Time Iterative Spectrogram Inversion (Zhu, Beauregard, Wyse, IEEE ASLP, 2007)



- Consecutive frames are stored in consecutive rows.
- Every row stores:
  - recent time-domain signal estimation
  - magnitude spectrum

Initialization

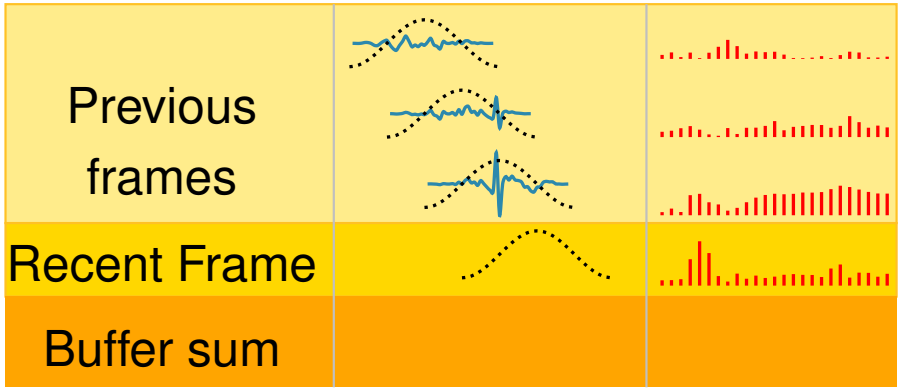
- neutral: zeros
- better: phase unwrapping

Basis algorithm

- Combination of
  - recent magnitude spectrum
  - phase spectrum of the buffer sum
- Result → time domain
- Window, store back

RTISI Phase Estimation

RTISI = Real-Time Iterative Spectrogram Inversion (Zhu, Beauregard, Wyse, IEEE ASLP, 2007)



- Consecutive frames are stored in consecutive rows.
- Every row stores:
  - recent time-domain signal estimation
  - magnitude spectrum

Initialization

- neutral: zeros
- better: phase unwrapping

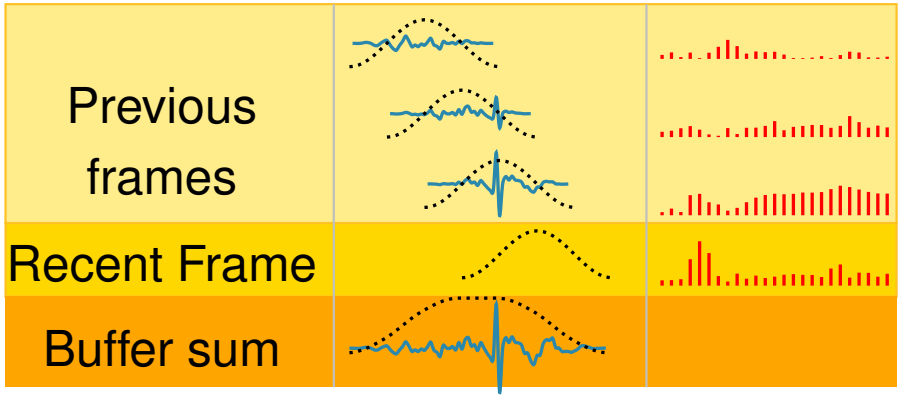
Basis algorithm

- Combination of
  - recent magnitude spectrum
  - phase spectrum of the buffer sum
- Result → time domain
- Window, store back

Volker Gnann | Institut für Nachrichtentechnik IENT | RWTH Aachen University

RTISI Phase Estimation

RTISI = Real-Time Iterative Spectrogram Inversion (Zhu, Beauregard, Wyse, IEEE ASLP, 2007)



- Consecutive frames are stored in consecutive rows.
- Every row stores:
  - recent time-domain signal estimation
  - magnitude spectrum

Initialization

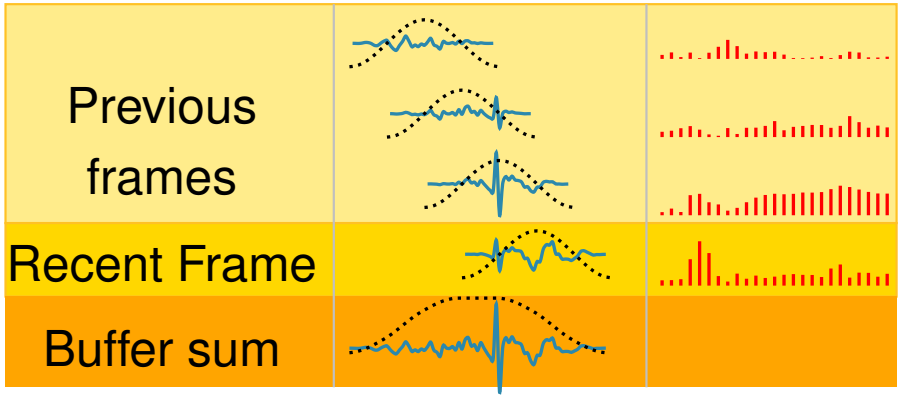
- neutral: zeros
- better: phase unwrapping

Basis algorithm

- Combination of
  - recent magnitude spectrum
  - phase spectrum of the buffer sum
- Result $\rightarrow$ time domain
- Window, store back

RTISI Phase Estimation

RTISI = Real-Time Iterative Spectrogram Inversion (Zhu, Beauregard, Wyse, IEEE ASLP, 2007)



- Consecutive frames are stored in consecutive rows.
- Every row stores:
  - recent time-domain signal estimation
  - magnitude spectrum

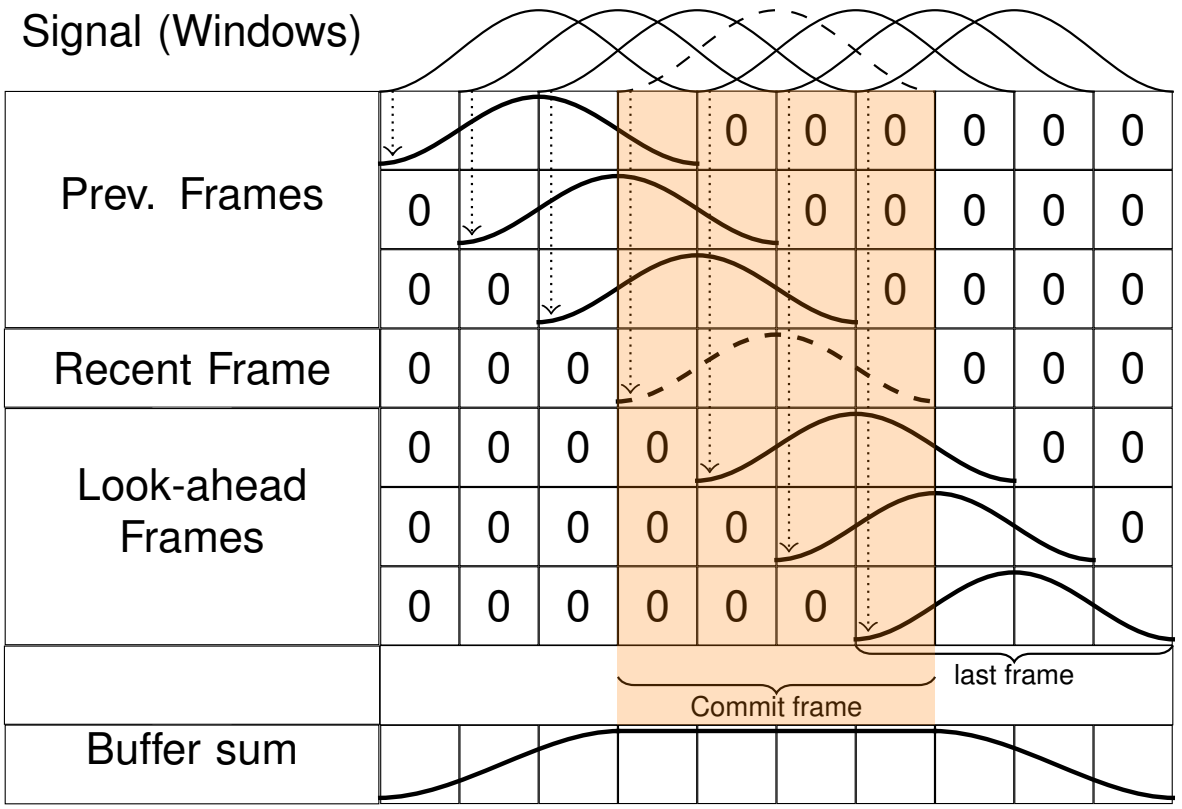
Initialization

- neutral: zeros
- better: phase unwrapping

Basis algorithm

- Combination of
  - recent magnitude spectrum
  - phase spectrum of the buffer sum
- Result $\rightarrow$ time domain
- Window, store back

RTISI with Look-Ahead

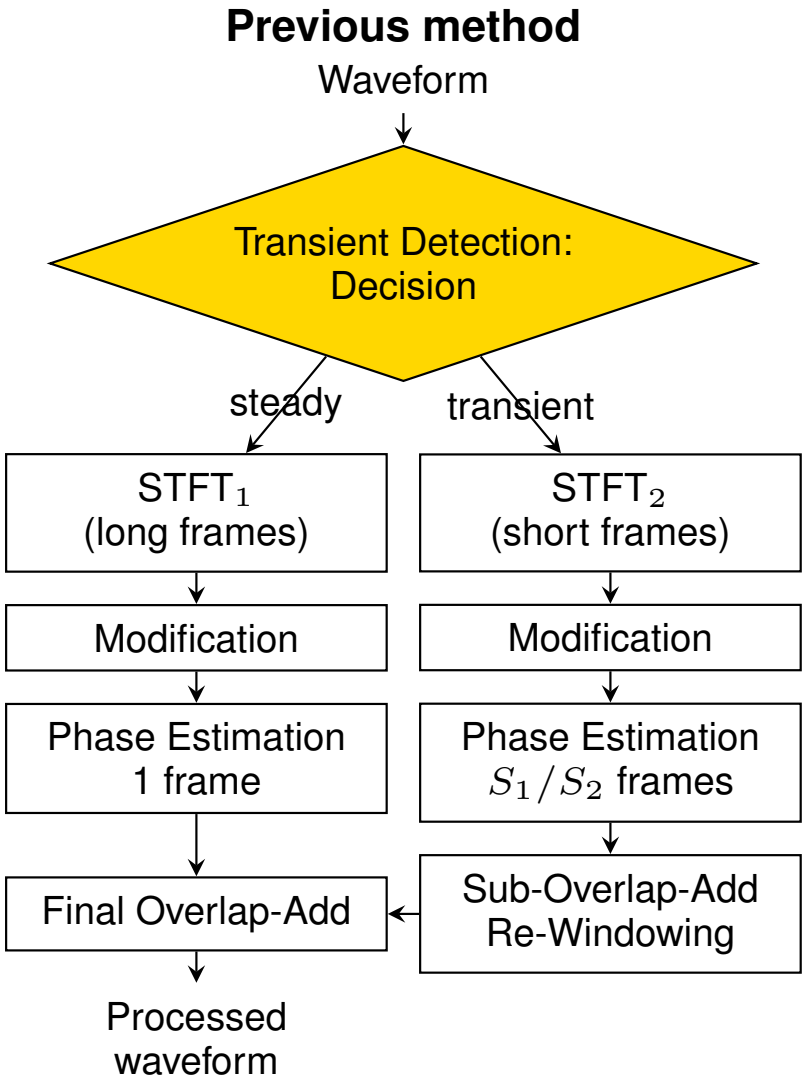


Order

- **Standard:** first the last frame, then second-last, ...
- **Improvement:** first loudest frame, then second-loudest, ...
- one iteration after another.

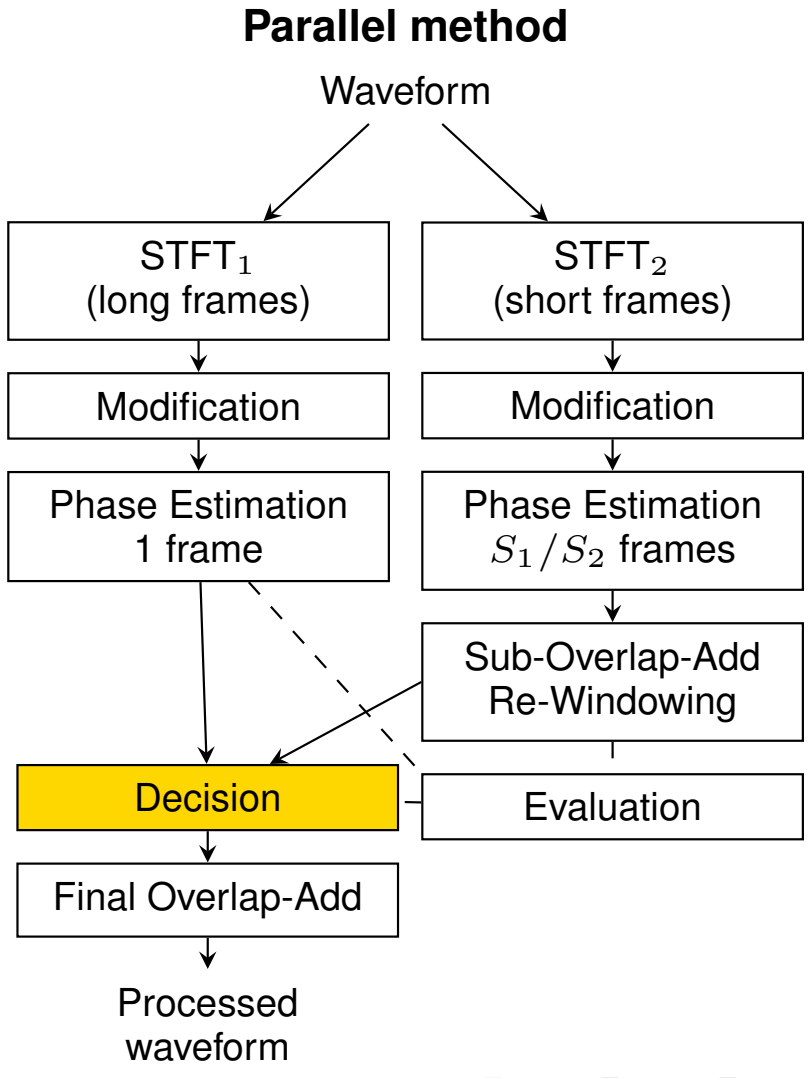
Volker Gnann Institut für Nachrichtentechnik IENT RWTH Aachen University

Parallel Multiresolution Processing (I)



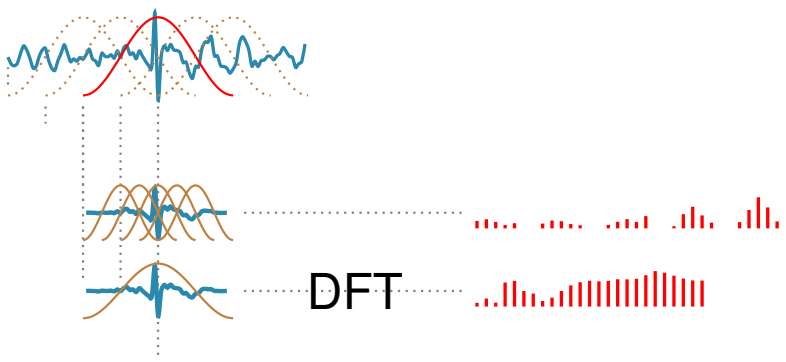
Volker Gnann Institut für Nachrichtentechnik IENT RWTH Aachen University

Parallel Multiresolution Processing (I)



Volker Gnann Institut für Nachrichtentechnik IENT RWTH Aachen University

Parallel Multiresolution Processing (II)



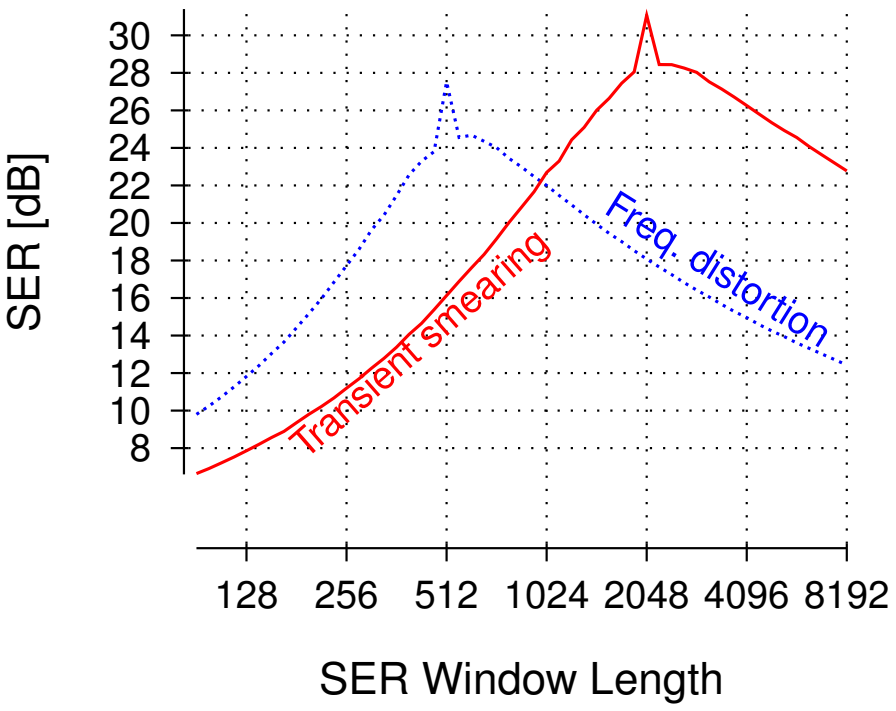
Properties

- Standard window: Hamming
- Parallel generation of long-w.len **spectrum** and short-w.len **spectrogram**

Advantages of parallel processing

- Decision is made when phase estimation is known.
  - No previous transient detection needed.
  - Better decisions can be expected.
- Modification algorithms know **the whole** magnitude spectrogram.
- Extension to more than two buffers straightforward.

Minimax Decision



Listening Examples

- Frequency distortion: ♫
- Transient smearing: ♫

Minimax Assumption

- The lower the SER, the better audible the error (the worse the result.)
- Choose phase estimation which maximizes the lowest SER. → Decide which error is the lesser evil.

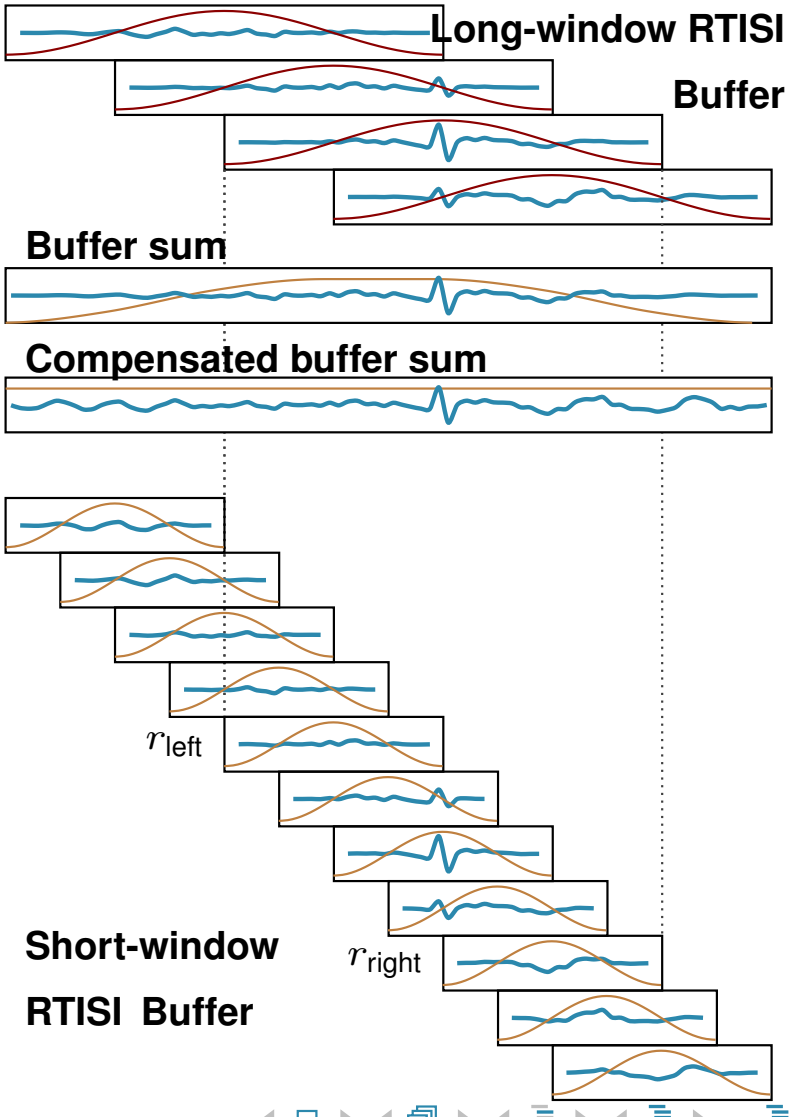
Synchronization between RTISI buffers

Long and short blocks

- Standard RTISI

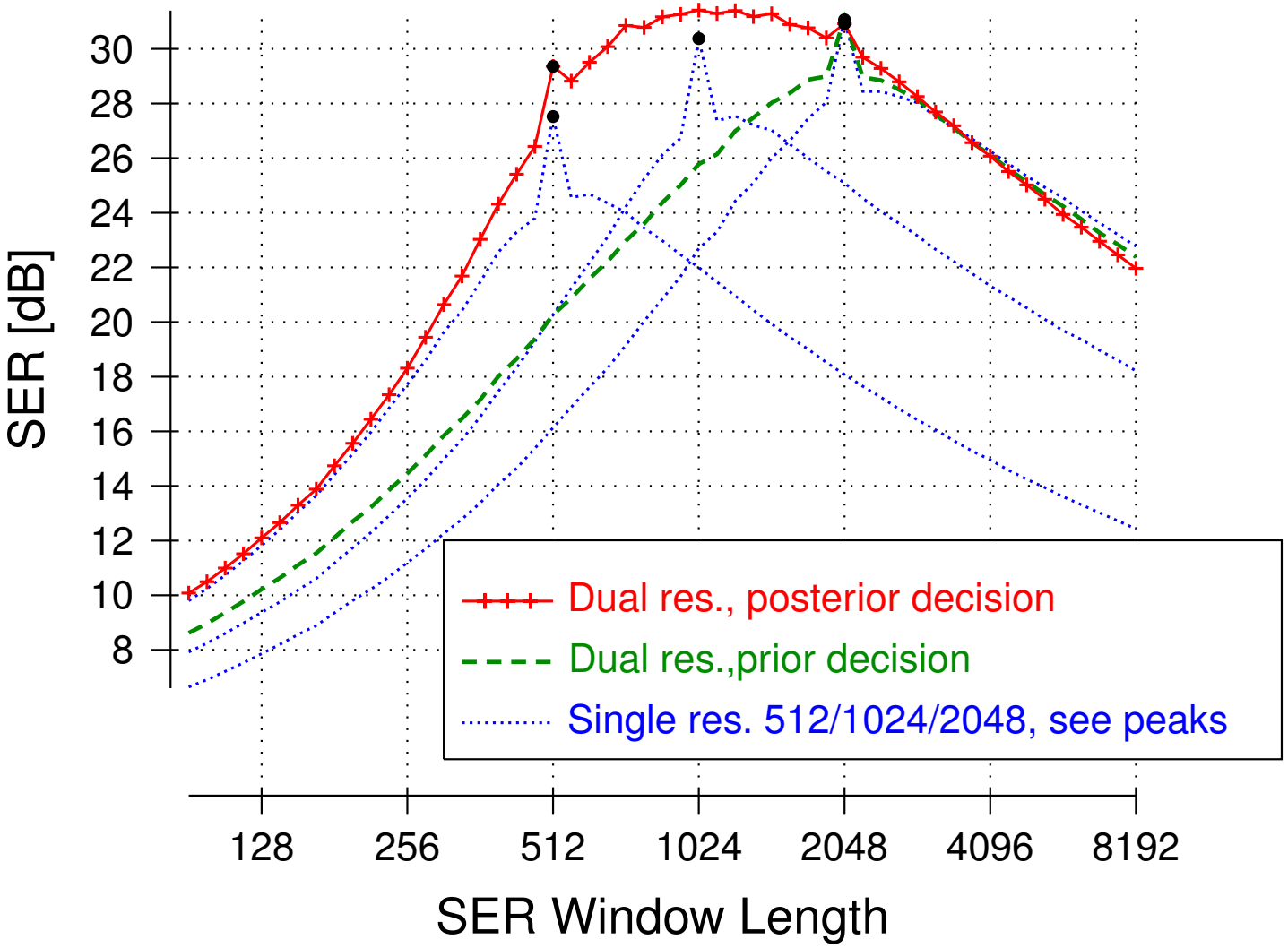
Synchronization

- Get source buffer sum.
- Compensate for the window sum.
- Fill target buffer with windowed sum segments.



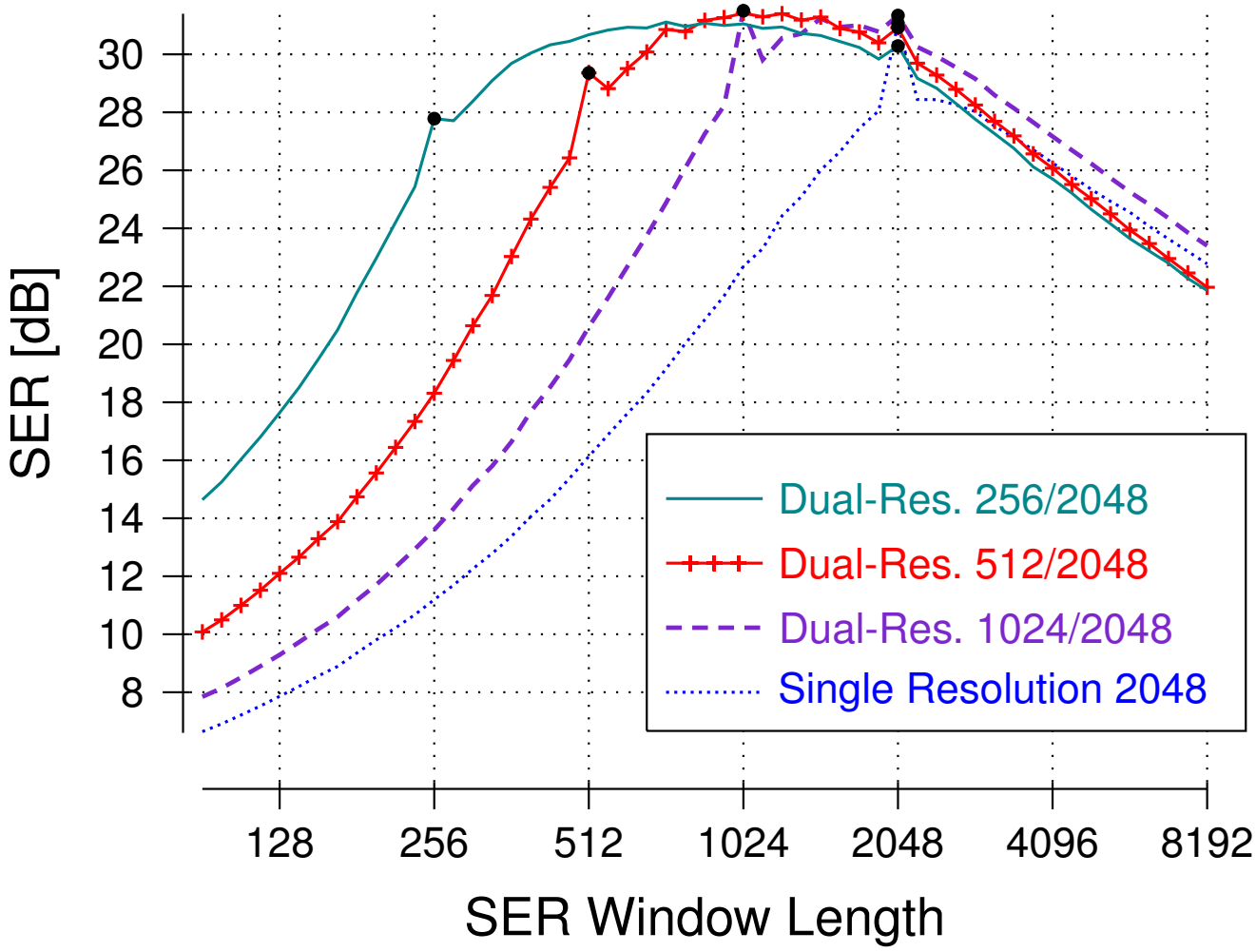
Volker Gnann Institut für Nachrichtentechnik IENT RWTH Aachen University

Experiments and Results (I)



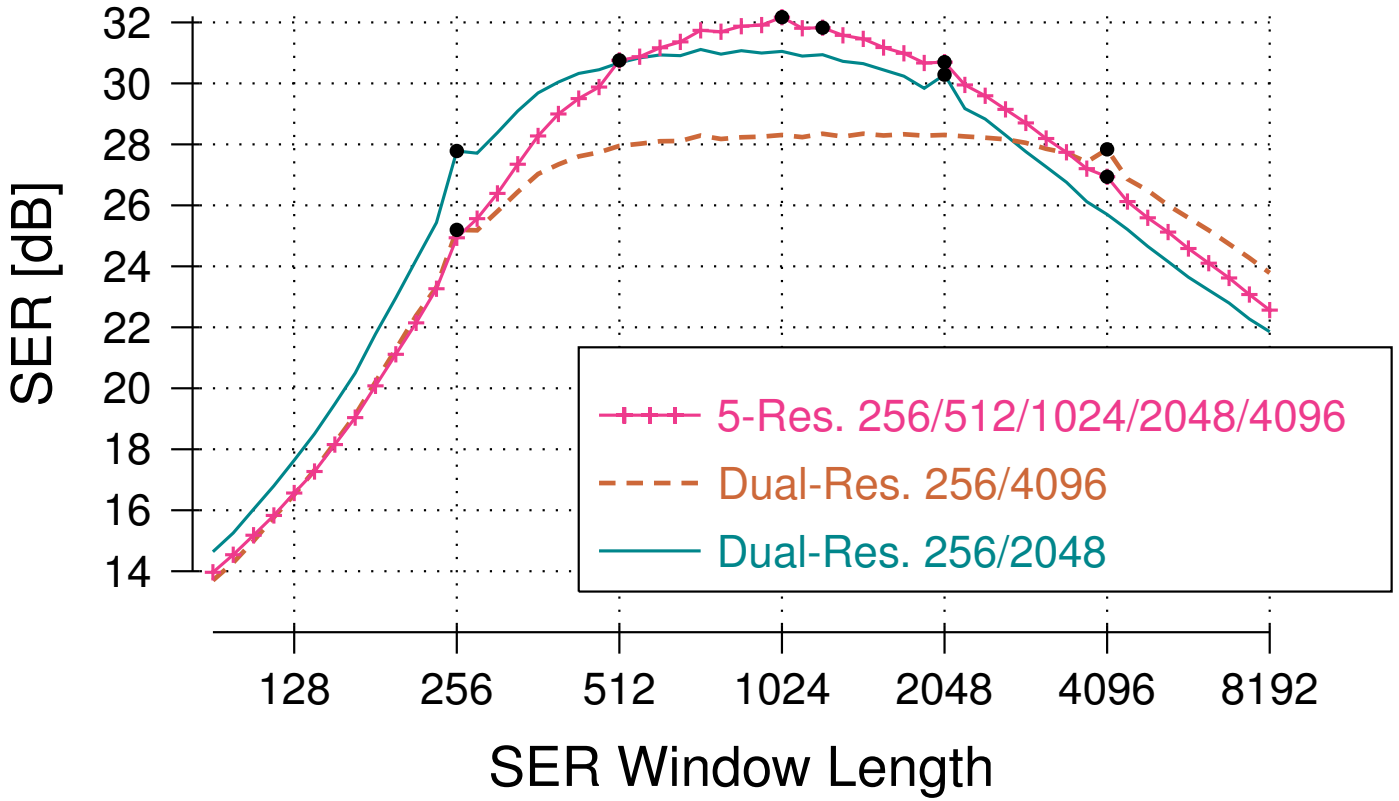
Volker Gnann Institut für Nachrichtentechnik IENT RWTH Aachen University

Experiments and Results (II)



Volker Gnann Institut für Nachrichtentechnik IENT RWTH Aachen University

Experiments and Results (III): Multiresolution



Volker Gnann Institut für Nachrichtentechnik IENT RWTH Aachen University

Listening Examples

- Original: ♫
- Single resolution:
  - 512 Samples: ♫
  - 1024 Samples: ♫
  - 4096 Samples: ♫
- Dual-Resolution 512/4096:
  - A-priori estimation: ♫
  - A-posteriori estimation: ♫
- Multiresolution 512/1024/2048/4096 ♫:

Volker Gnann | Institut für Nachrichtentechnik IENT | RWTH Aachen University

Thank you for your attention!