Dual Narrowband / Wideband Noise Reduction

C. Beaugeant, Nokia Siemens Networks

M. Schönle, Siemens AG

Introduction

- Difference between wideband noise reduction and narrowband noise reduction ?
- Dual systems narrowband / wideband
- Architecture proposal
- Example solutions
- Conclusion

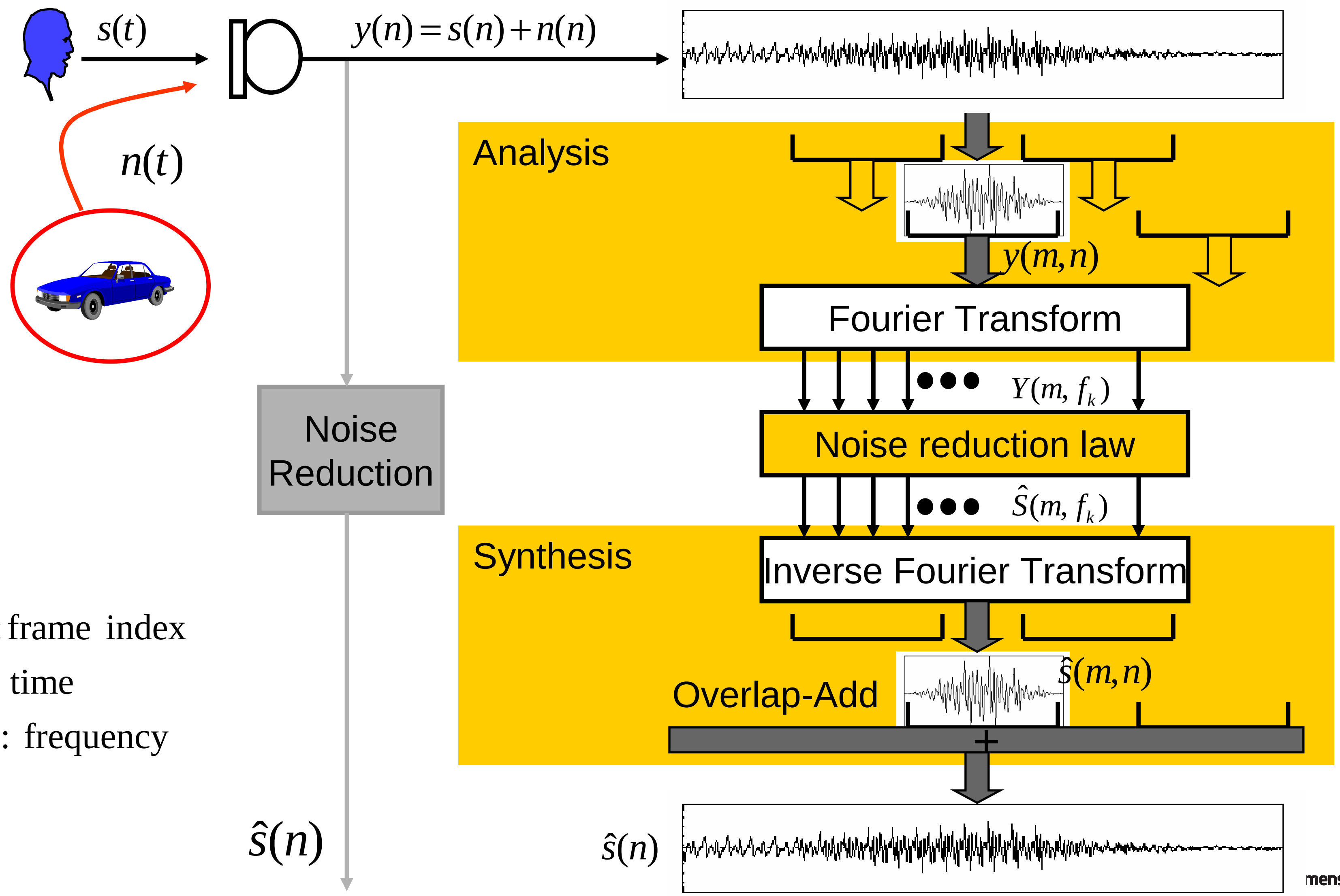
Wideband audio : What for?

- Historically, telecommunication uses limited bandwidth (300Hz-3400Hz)
- Wideband telephony: a simple solution to increase quality
- New applications propose wideband speech telecommunication
  - New services with VoIP (G.722, G.729.1)
  - Enhancement of quality for UMTS (AMR-WB)
- Front-end algorithms must follow codecs and applications development. Wideband solution needed for
  - Noise reduction
  - Echo cancellation
  - Automatic level control
  - Transducer enhancement
  - Microphone array

Overview on noise reduction

- Noise reduction mainly focused till now on the problematic for telephony (sampling frequency 8 kHz)
- Rare examples for other sampling rate than 8 kHz (speech recognition, combination of noise reduction with AMR-WB)
- First solutions at the beginning of the 80‘s
- Aim: reducing noise on the microphone signal to enhance quality, reduce tiredness for the far-end user
- Boost of industrial solution widely implemented when telephony get mobile:
  - Mobile terminals used in highly noisy environment
  - Carkit solutions
- Widely used solution based on short time frequency domain analysis
  - Analysis of successive windowed frames
  - Fourier transform
  - Applying a weighting gain on the amplitude of the frequency signal
  - Inverse Fourier transform
  - Overlap-add

Overview of Noise reduction in the frequency domain



m: frame index

n: time

f_k: frequency

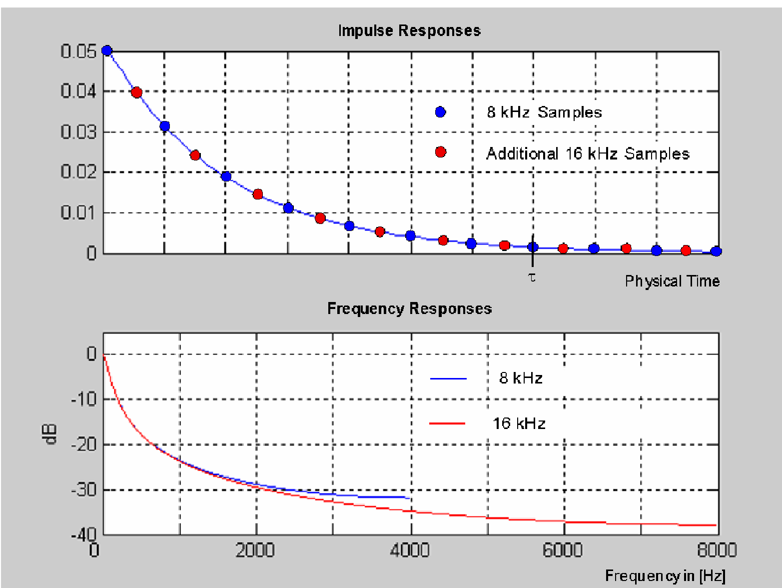
Difference wideband / narrowband noise reduction

- A priori same solution
  - FFT analysis/synthesis independent from sampling frequency
  - Weighting rule a priori independent from sampling frequency
  - Test on transposing narrowband solution to wideband solution already successful [1]
- But differences on many parameters!
  - FFT length different to get same resolution in narrowband and wideband
  - Smoothing factors highly depending on sampling frequency
  - Many variables depend on the sampling frequency: values of threshold, energy estimation, SNR estimation
- Important tuning needed to make the transposition of a narrowband noise reduction to a wideband noise reduction [2]
- Increase of computation load (up to twice more)

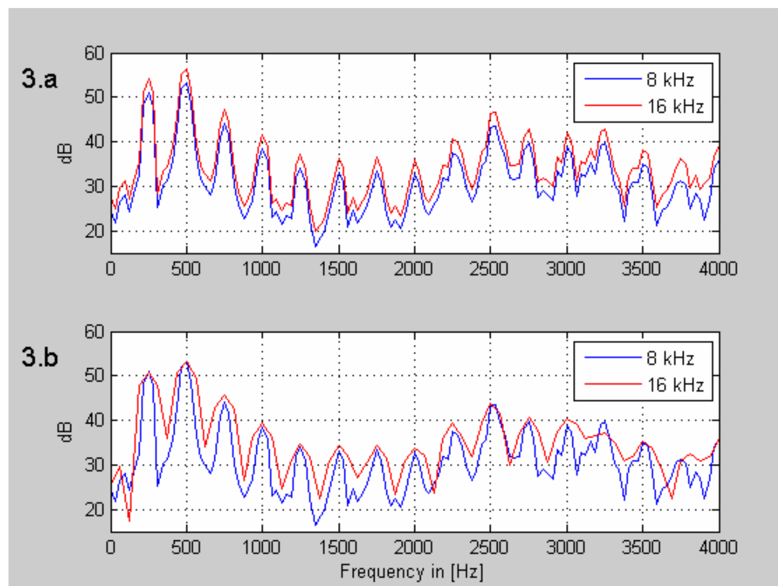
[1] C. Beaugeant, I. Varga, T. Lotter, P. Jax, P. Vary, Noise Reduction for AMR-WB, 2nd workshop on Wideband Speech Quality, 06.2005
[2] C. Beaugeant, M. Schönle, I. Varga, `Challenges of 16 kHz in Acoustic Pre- and Post-Processing for Terminals IEEE Com magazine, May 2006.

Challenge of dual systems

- Modern terminals to get both narrowband / wideband functionalities
- Example: UMTS phone
  - Wideband telephony in UMTS network
  - Narrowband telephony in GSM network
- Idea: Get a single front-end for both modes
- Problem: Many values to be tuned to get such universal front-end
- Two examples Filter design, FFT resolution



Filter Design

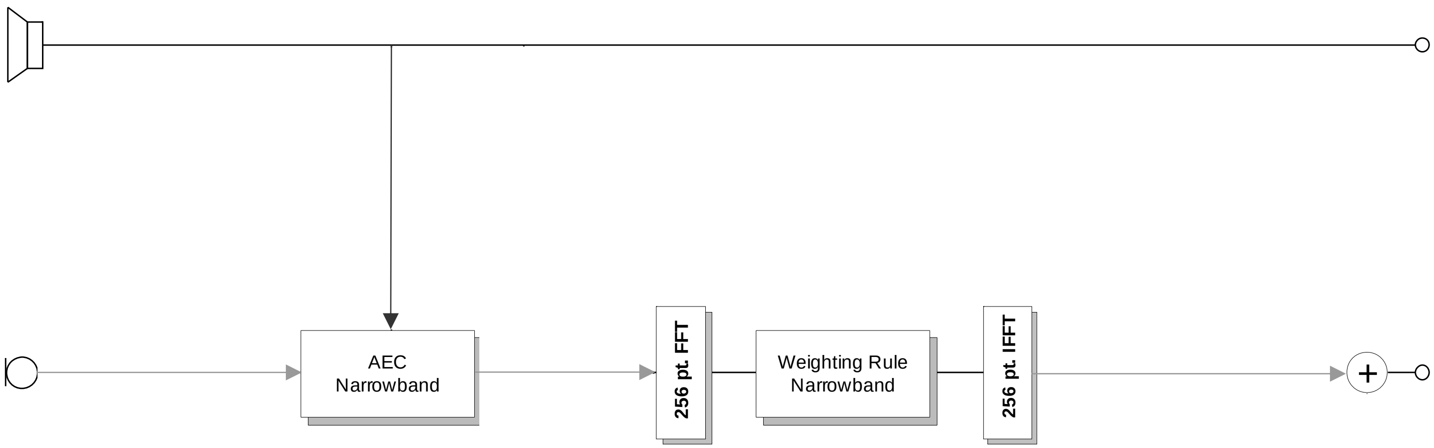


3.a: Speech PSD (FFT length doubled @ 16 kHz)

3.b: Frequency resolution with same FFT length

Architecture for dual system

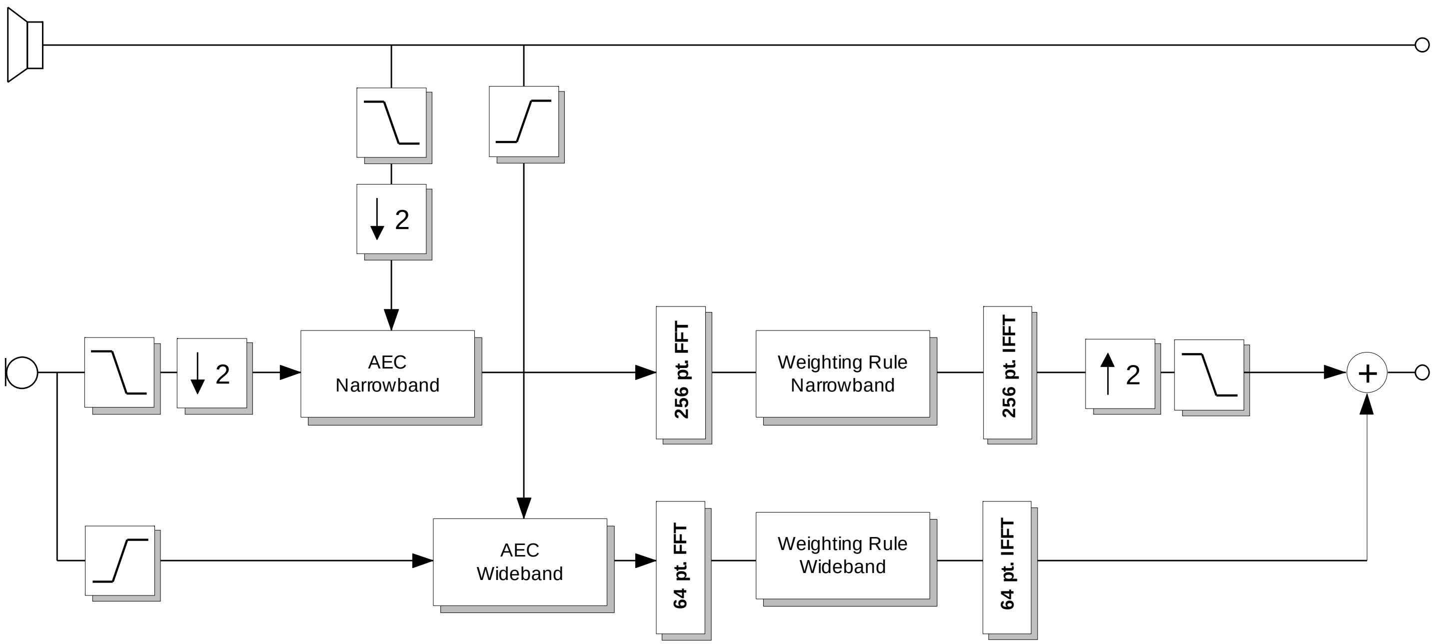
- Backward compatibility
- Reuse of 8 kHz system



Nokia Siemens Networks

Architecture for dual system

- Extension to wideband audio
- Plug-in of simple two-channels filter bank

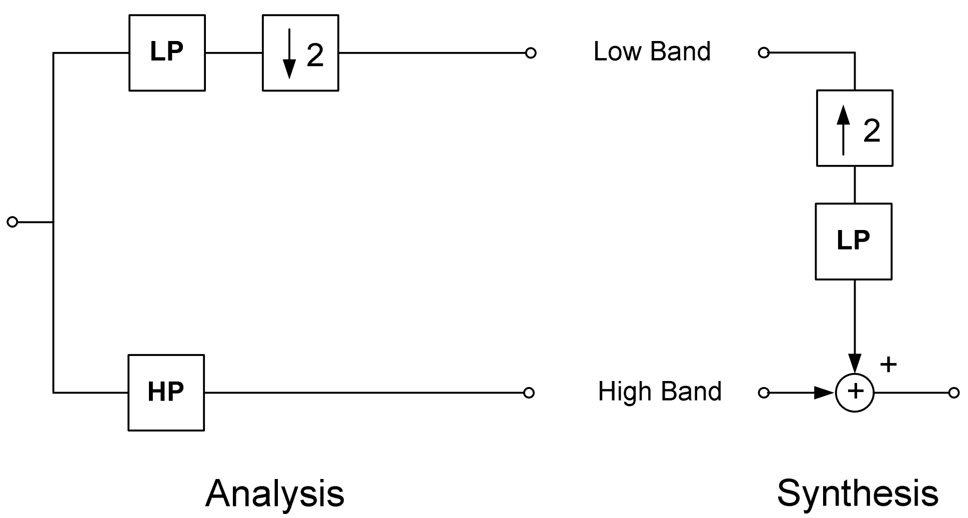


Requirements to filter bank

- High stopband attenuation
- Only small amount of aliasing
- Perfect Reconstruction

Trade-off

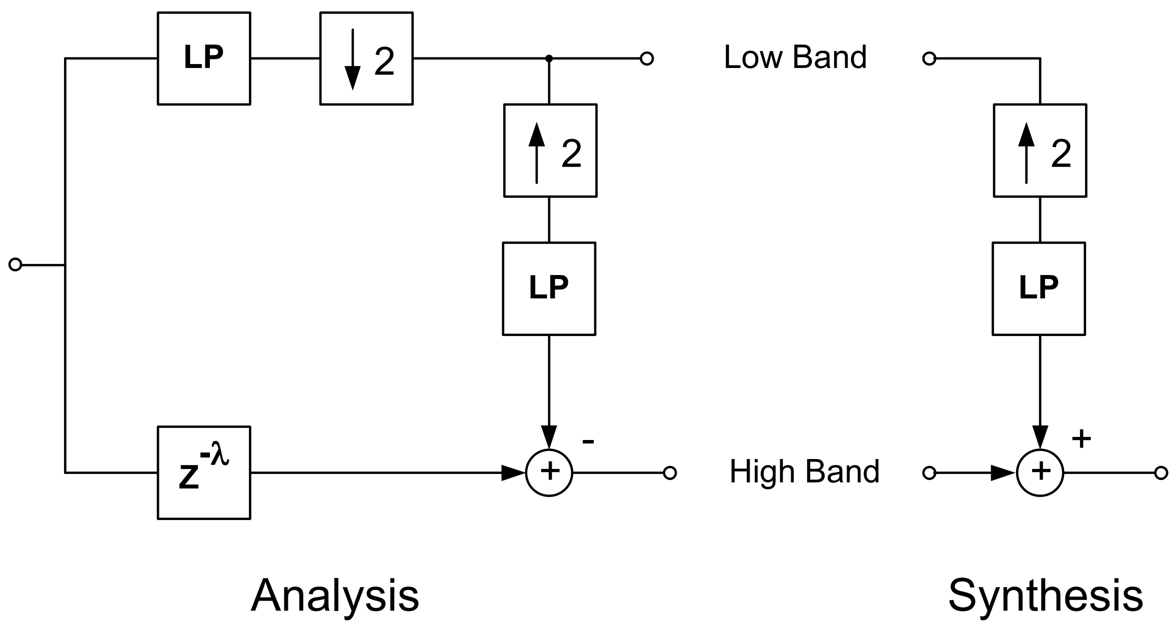
- Low complexity
- Low group delay
- Simple design



  C. Beaugeant, M. Schönle / Date

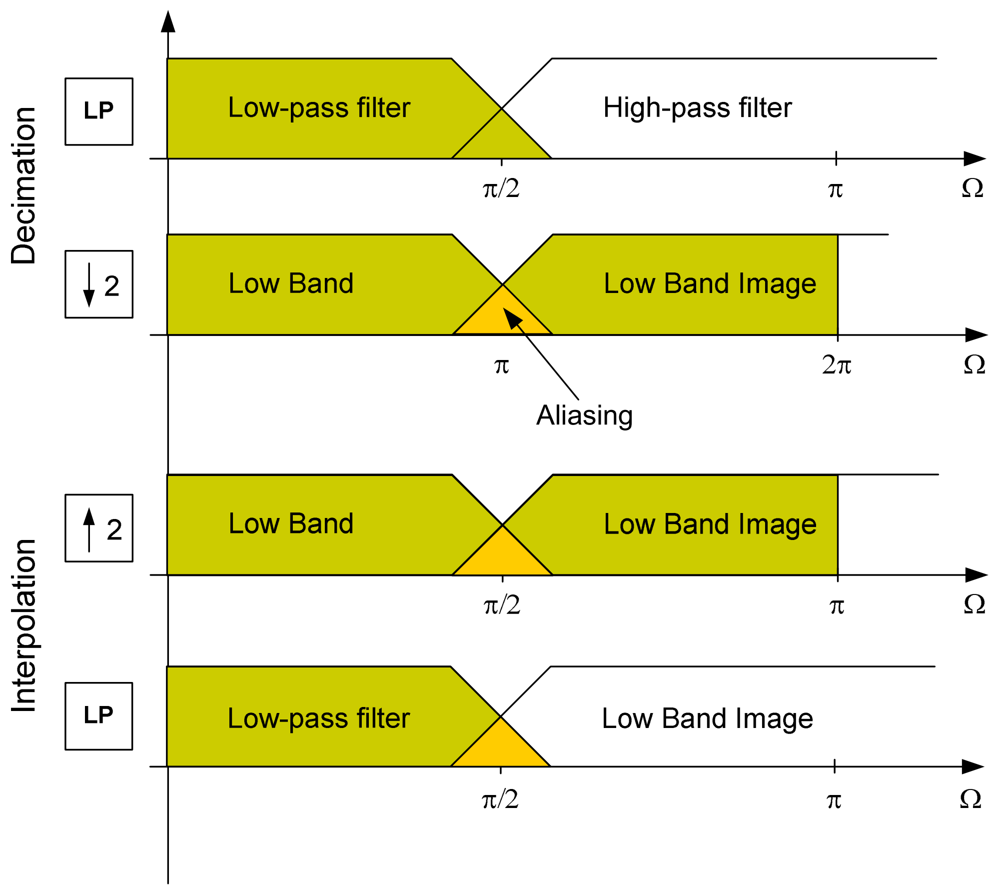
Example solution : filter bank design

- All-pass-Complementary Filter Bank
  - “Laplacian Pyramid“: High band adds details
  - Perfect reconstruction guaranteed by structure
  - No synthesis filter required in high band



Nokia Siemens Networks

Example solution : filter bank design



Anti-aliasing filtering by low-pass

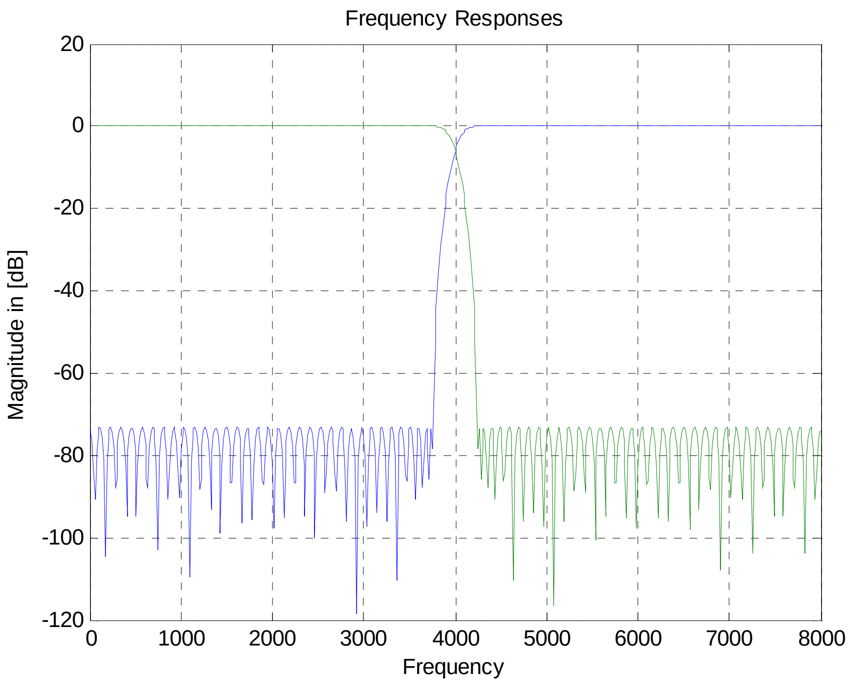
Decimation of sampling rate
→ aliasing components

Doubling of sampling rate

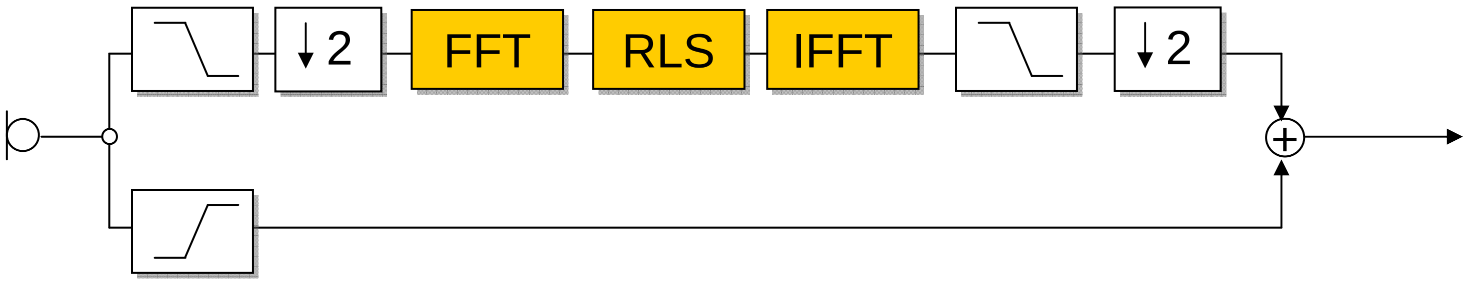
Same low-pass filter used for
anti-imaging filtering

Example solution : filter bank design

- Design of only one FIR low-pass filter
- High-pass filter is obtained by filter bank structure
- Simple design using Remez Algorithm
- Half-band filters can be used
- Example for N = 140
  - Stop-band attenuation 72 dB
  - No aliasing up to f = 3.75 kHz
  - Group Delay 4.4 ms @ 16kHz



Solution 1: noise reduction in low band



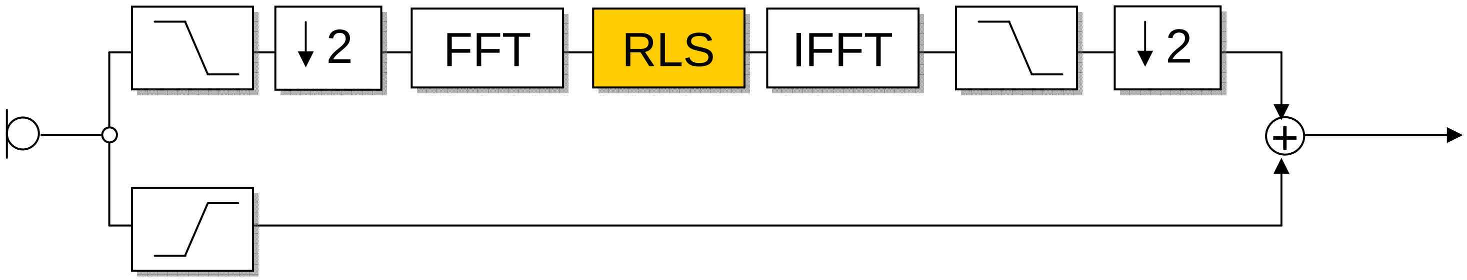
- Analysis / synthesis applied only in the narrowband part
- No noise reduction in the high band
  - Assumption that noise is less disturbing in higher band (psychoacoustic)
  - Assumption than SNR is higher above 4 kHz (Plosive speech sound)

- Advantage:
  - small computation load increase compared to narrowband system

- Drawback:
  - No noise reduction in higher band

Nokia Siemens Networks

Recursive Least Square (RLS) in low band



- Minimization of the cost function $J_m(e(f_k)) = \sum_{l=0}^{m} \lambda^{m-l} e^2(l, f_k)$

with $e(l, f_k) = S(l, f_k) - \hat{S}(l, f_k)$

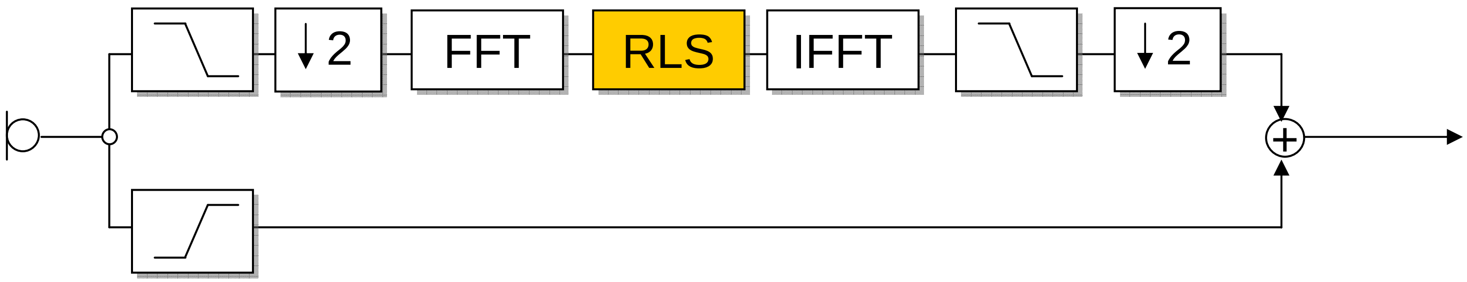
λ forgetting factor

- Leads to the following weighting rule (RLS)

$$\boxed{G_{LS}(m, f_k) = \frac{\sum_{l=0}^{m} \lambda^{m-l} S^2(l, f_k)}{\sum_{l=0}^{m} \lambda^{m-l} S^2(l, f_k) + \sum_{l=0}^{m} \lambda^{m-l} N^2(l, f_k)}}$$

So that $\hat{S}(m, f_k) = G_{LS}(m, f_k).Y(m, f_k)$

Recursive Least Square (RLS) in low band



Simplified version (no need of estimation of the speech signal)

$$G_{LS}(m, f_k) = \frac{\sum_{l=0}^{m} \lambda_Y^{\,m-l} Y^2(l, f_k)}{\sum_{l=0}^{m} \lambda_Y^{\,m-l} Y^2(l, f_k) + \sum_{l=0}^{m} \lambda_N^{\,m-l} \hat{N}^2(l, f_k)}$$

λ ➡ λ_Y λ_N

Tuning parameters

$G_{LS}(m, f_k)$ can be seen as a combination of two filters

➡ $G_{LS}(m, f_k) = G_1(m, f_k) + \beta_2 . G_2(m, f_k)$

$$\beta_2 = \frac{1-\lambda_N}{2-\lambda_N-\lambda_Y}$$

Noise reduction filter:

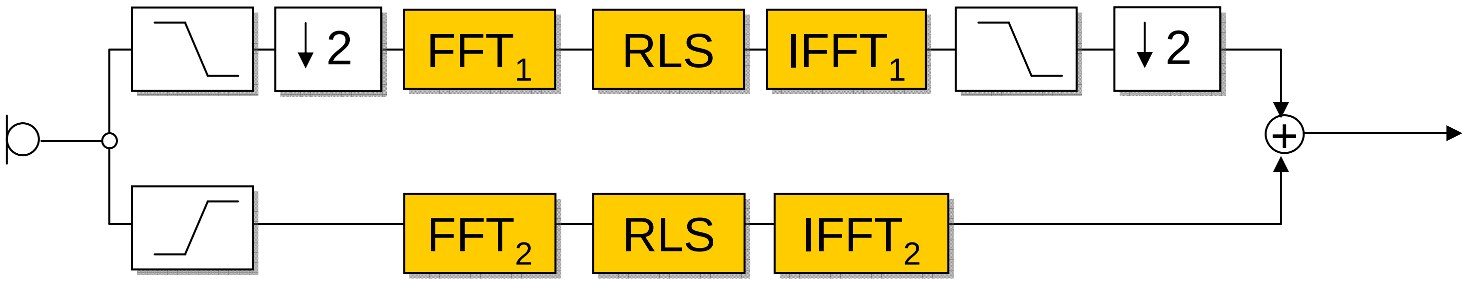
Output : ‚clean speech‘

Noise estimation filter:

Output : ‚clean noise‘

$$G_1 = G_1(\lambda_Y, \lambda_N)$$

Solution 2 : RLS in both bands

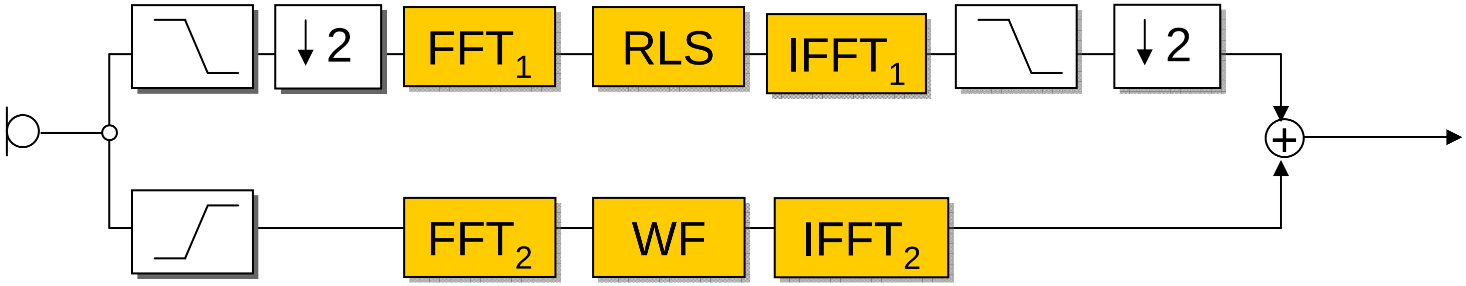


- Same algorithm in both bands: RLS

- Analysis / synthesis different on the bands
  - FFT_1 : 5 ms overlap, 256 pts
  - FFT_2 : 5 ms overlap, 32 pts

- Good compromise for ROM:
  - Partial re-use of the same functions for the higher and lower band

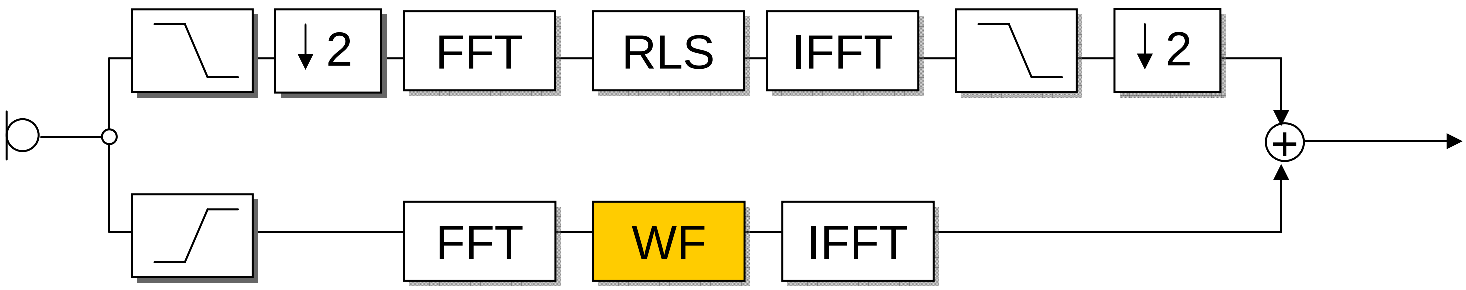
Solution 3 : different kind of noise reduction in the bands



- Study of the robustness of the architecture when different solutions are built on the two bands
- Analysis / synthesis different on the bands
  - FFT_1 : 5 ms overlap, 256 pts
  - FFT_2 : 5 ms overlap, 32 pts
  - Wiener filter in the higher band

- Idea: Possibility of special tuning, special algorithm for the higher band
  - Taking into account special SNR
  - Plosive characteristic of speech signal
  - Psycho-acoustic

Solution 3 : algorithm for the high band



- Wiener filter in the frequency domain:

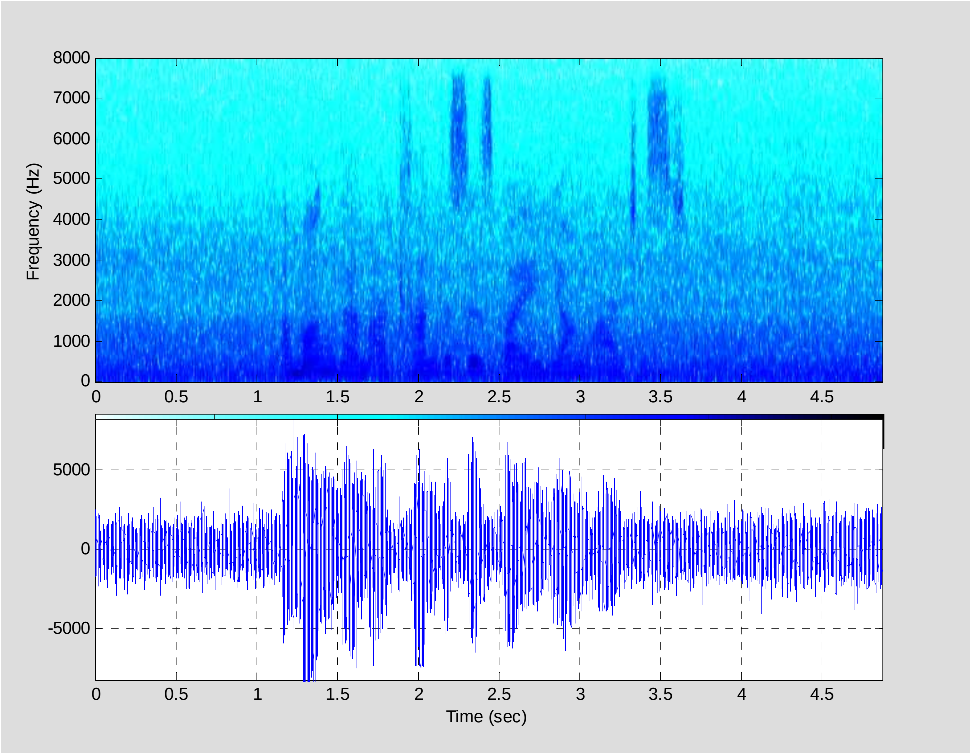
$$G(m, f_k) = \frac{SNR(m, f_k)}{1 + SNR(m, f_k)} \text{ for } k = 0,1,\cdots, NFFT - 1$$

- Ephraim & Malah SNR estimation [3]:

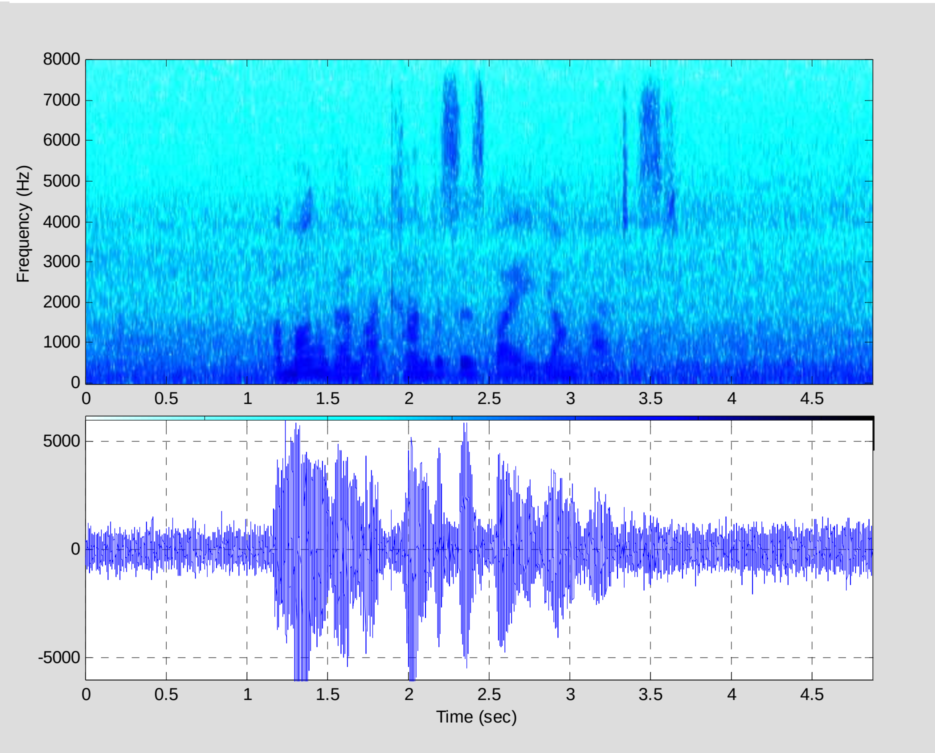
$$SNR(m, f_k) = \beta_1 \cdot \frac{|Y(m, f_k)|^2 - \hat{\gamma}_n(m, f_k)}{\hat{\gamma}_n(f_k)} + (1 - \beta_1) \cdot \frac{G^2(m-1, f_k)|Y(m-1, f_k)|^2}{\hat{\gamma}_n(f_k)}$$

[3] Y. Ephraim und D. Malah, "`Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator"'. IEEE Trans. On ASSP, ASSP-32, Nr. 6, Dec. 1984

Demos



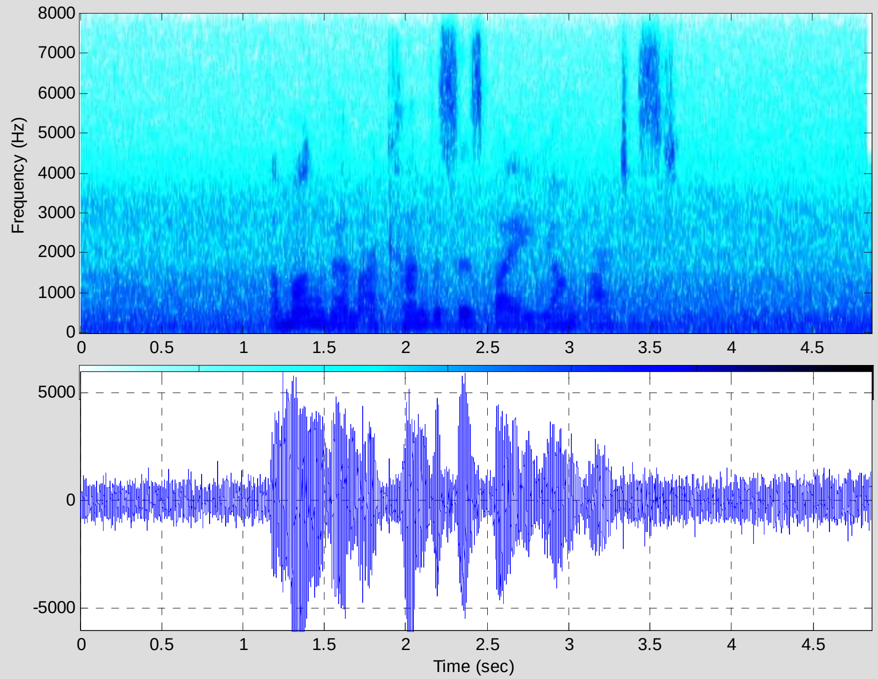
Noisy signal (SNR = 5 dB)



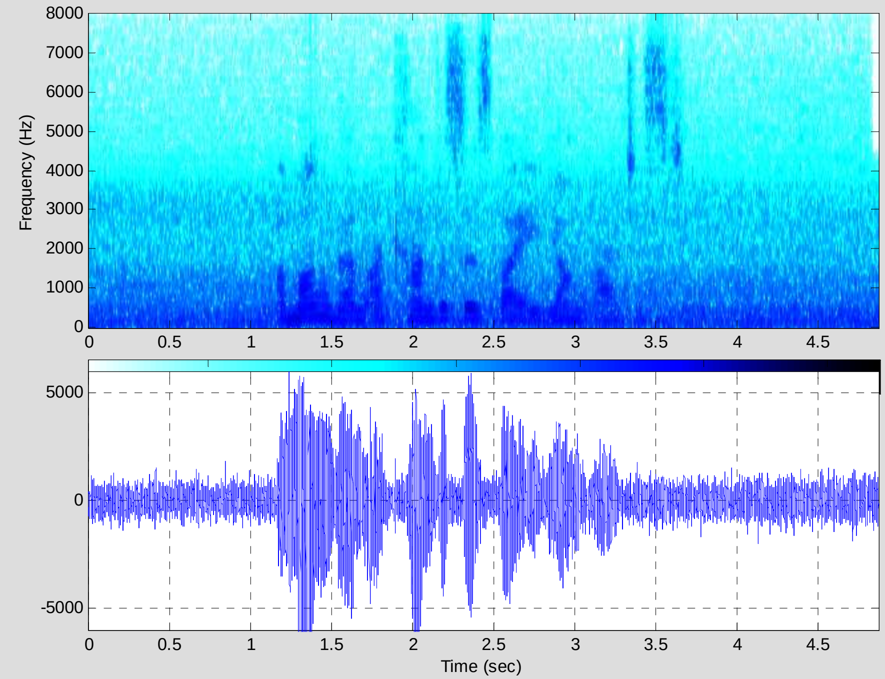
Noise reduction in the lowest band only (RLS)

Nokia Siemens Networks

Demos



RLS Noise reduction in both bands



RLS Noise reduction in the low band

Wiener noise reduction in the high band

Conclusion

- Proposed architecture gives an answer for designing a dual narrowband wideband noise reduction

- Study of particular dual noise reductions
  - No modification of already existing narrowband noise reduction
  - Scalable solution
  - 3 schemes proposed showing the feasibility and the potential of the architecture

- Perspectives:
  - Open questions on the optimal resolution needed in the high band, on the best compromise computation load vs. quality
  - Open field of research for better solution of noise reduction dedicated to the high frequency band [4 kHz, 8 kHz]
  - Easy scalability to other sampling rates (32-48kHz)
  - Same principle can be used for other front-end algorithms (echo cancellation, automatic level control…)
  - Interaction and optimization together with scalable speech codecs (G.729.1)

Nokia Siemens
Networks