On the Use of Artificial Bandwidth Extension Techniques in Wideband Speech Communications

Peter Jax and Peter Vary

RWTH Aachen University, Institute of Communications Systems (IND)
{jax,vary}@ind.rwth-aachen.de

2nd Workshop on Wideband Speech Quality
in Terminals and Networks: Assessment and Prediction

22nd and 23rd June 2005 – Mainz, Germany

ind **Institut für Nachrichtengeräte**
und Datenverarbeitung
Prof. Dr.-Ing. P. Vary

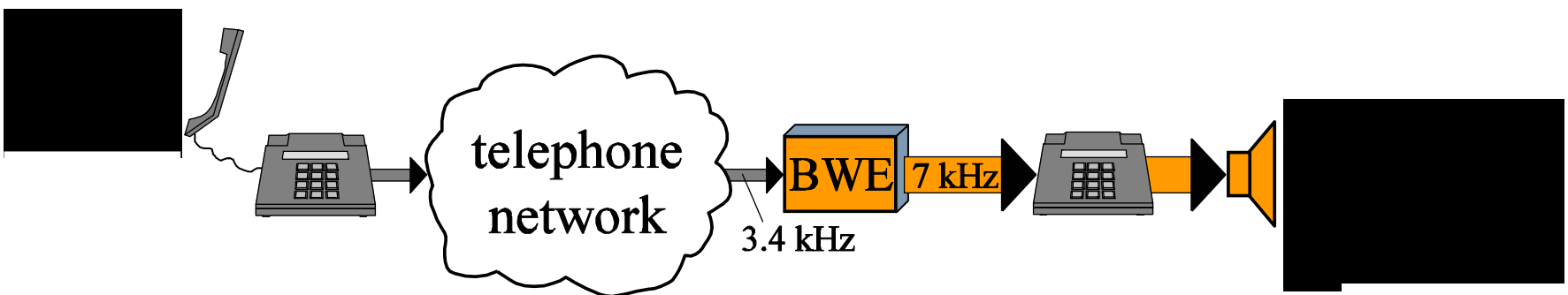
RWTH AACHEN

Outline

- Motivation & Bandwidth Extension Principle
- Exemplary Bandwidth Extension Algorithm
- Applications in Wideband Speech Communication
- Conclusions

Motivation

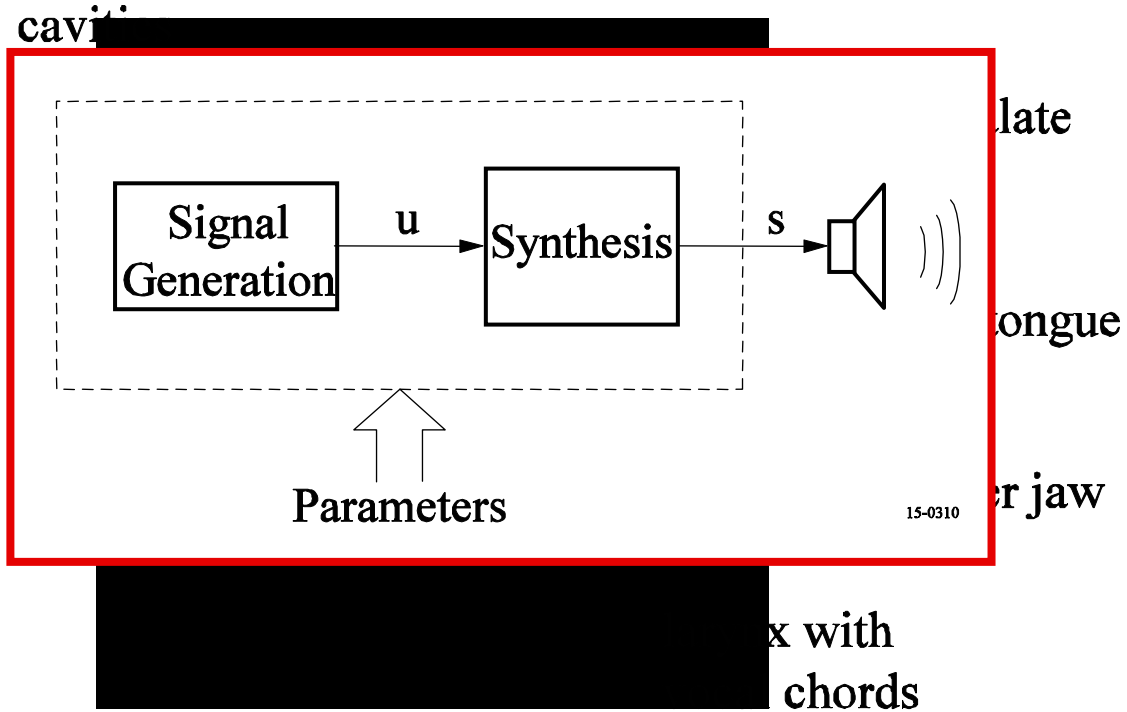
- Transition to wideband telephony
  - wideband-capable terminals needed on both sides
  - for mobile radio systems: changes in the network required
  - long transition period to be expected
- Artificial bandwidth extension (BWE) at the receiver
  - no modification of transmission link necessary
  - improvement of subjective speech quality



BWE Principle: Speech Production Model

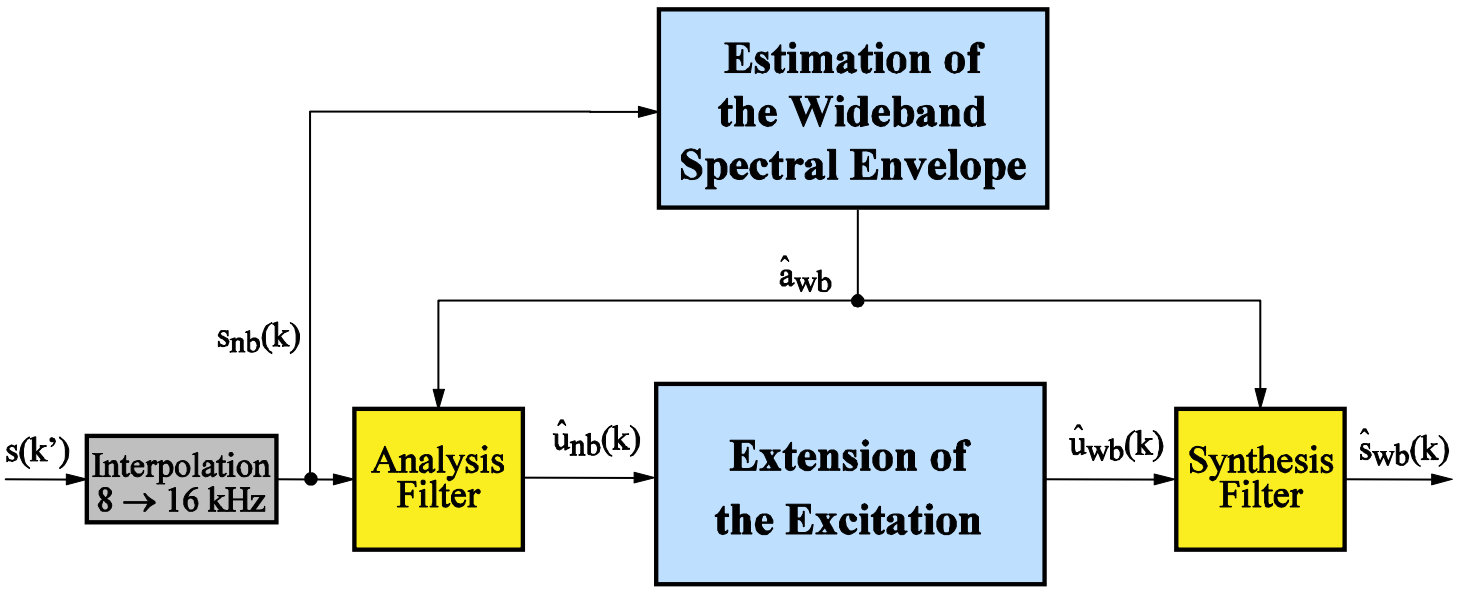
1st Step:

Use pattern recognition to estimate parameters of the wideband speech production model from the available narrowband speech signal



BWE Principle: Algorithm

2nd Step: Adaptive digital analysis-synthesis filtering with extension of the excitation

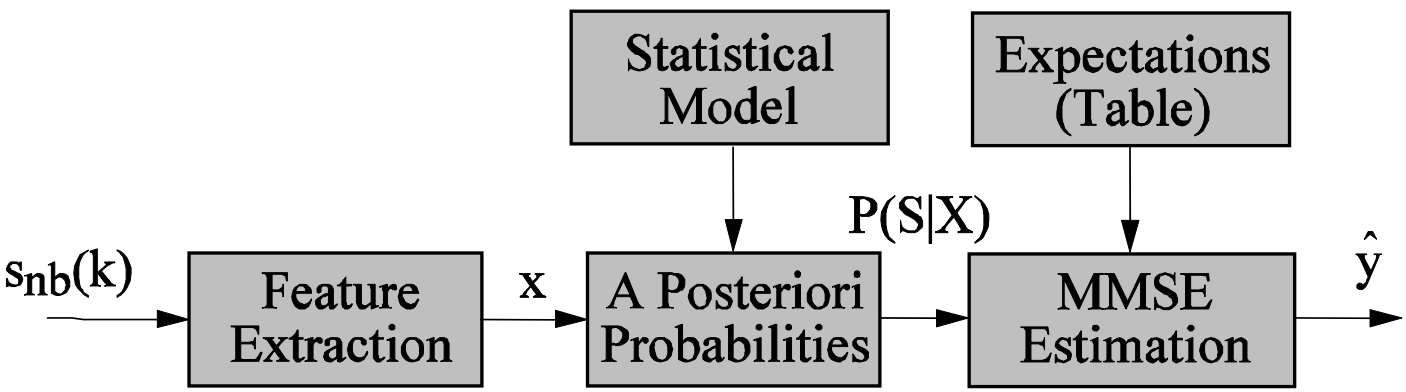


- Separate Extension of
  - wideband spectral envelope (filter coefficients $\hat{a}_{wb}$)
  - excitation signal u(k)

Estimation of Wideband Spectral Envelope

▶ Pattern recognition with Hidden Markov Model

- based on pre-trained statistical model
- each state represents a typical speech sound
- MMSE estimation from a posteriori state probabilities



MMSE Estimation with Hidden Markov Model

$$\hat{\mathbf{y}}_{\mathrm{wb}} = \sum_{i=1}^{N_S} \mathrm{E}\{\mathbf{y}(m) \mid S_i(m)\} \cdot P(S_i(m) \mid \mathbf{X}(m_k))$$

$$\mathrm{E}\{\mathbf{y}(m) \mid S_i(m), \mathbf{x}(m)\} = \sum_{l=1}^{L} \rho_{\mathbf{y}|\mathbf{x},il} \left(\boldsymbol{\mu}_{\mathbf{y},il} - \left((\mathbf{x} - \boldsymbol{\mu}_{\mathbf{x},il})^{\mathrm{T}} \mathbf{A}_{\mathbf{yx},il} \mathbf{A}_{\mathbf{yy},il}^{-1} \right)^{\mathrm{T}} \right)$$

$$P(S_i(m) \mid \mathbf{X}(m_k)) = \frac{\alpha_i(m)\, p(\mathbf{x}(m) \mid S_i(m))}{\sum_{i=1}^{N_S} \alpha_i(m)\, p(\mathbf{x}(m) \mid S_i(m))}$$

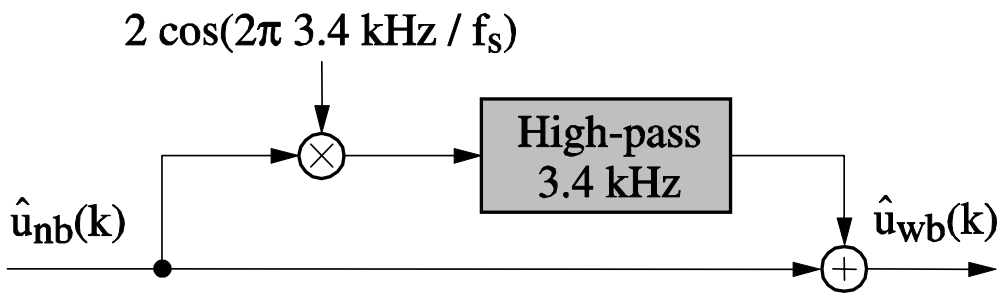
$$p(\mathbf{x}(m) \mid S_i(m)) = \sum_{l=1}^{L} \rho_{\mathbf{x},il} \frac{\sqrt{\det \mathbf{A}_{\mathbf{x},il}}}{(2\pi)^{d/2}} \exp\left(-\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{x},il})^{\mathrm{T}} \mathbf{A}_{\mathbf{x},il} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{x},il}) \right)$$

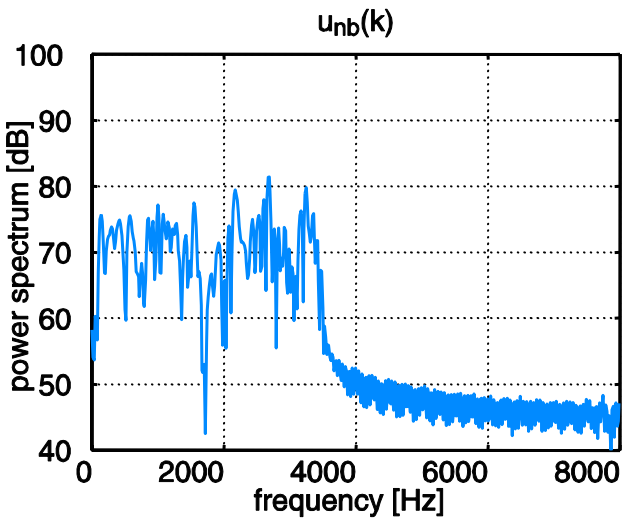
$$\alpha_i(m+1) = \sum_{j=1}^{N_S} \alpha_j(m)\, p(\mathbf{x}(m) \mid S_j(m)) P(S_i(m+1) \mid S_j(m)), \quad \alpha_i(1) = P(S_i)$$

.
.
.

Extension of the Excitation Signal

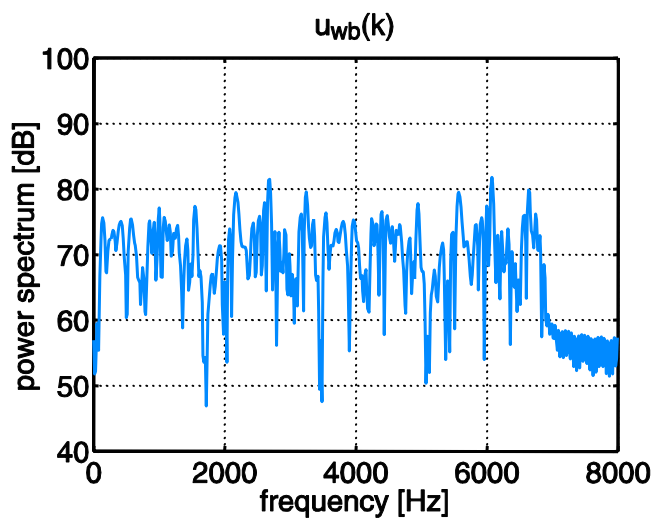
▶ Spectral shift



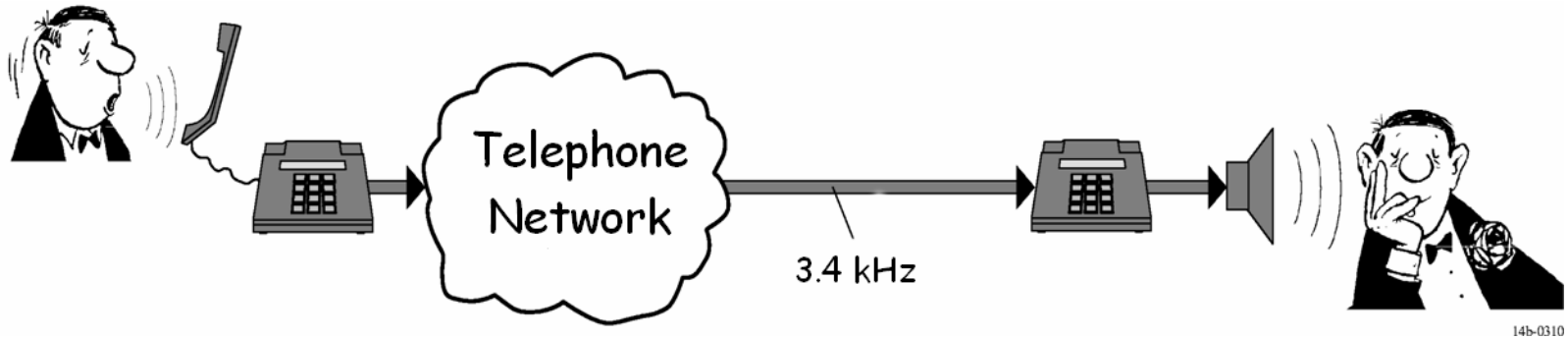


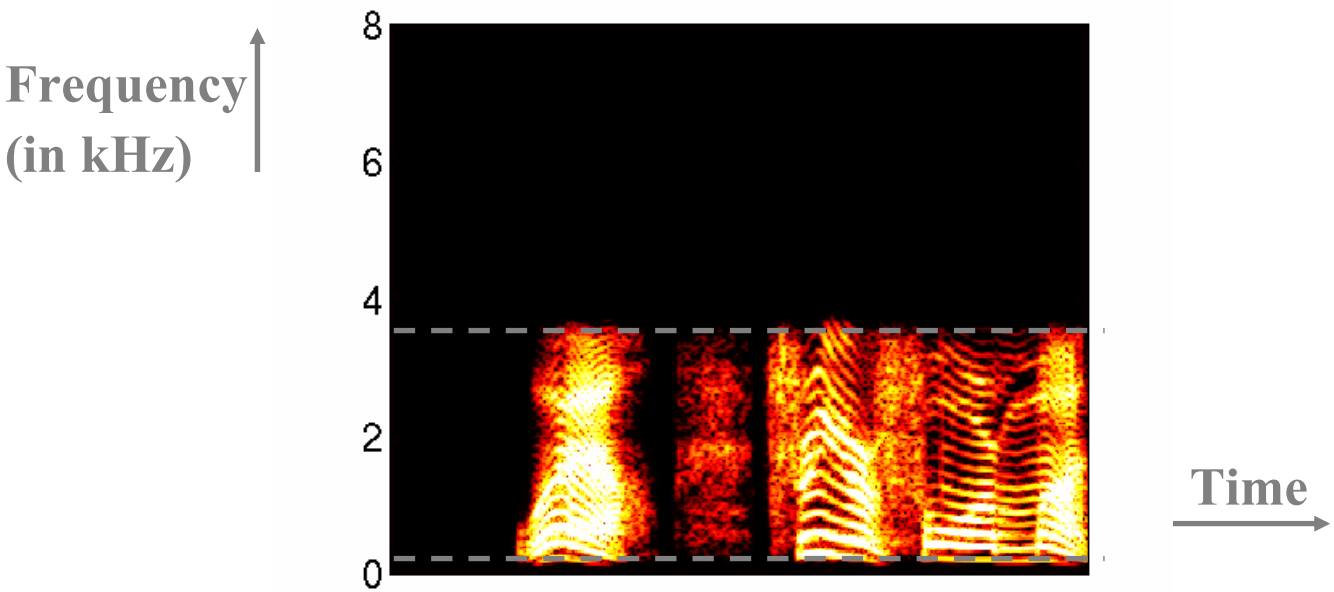
▶ Control of modulation freq.

- fixed spectral shift, e.g. 3.4 kHz
- optional: adaptive w.r.t. pitch frequency



Audio-Demo: Telephone Speech with BWE





GO

Applications in Speech Communication

▶ Stand-alone bandwidth extension, e.g.

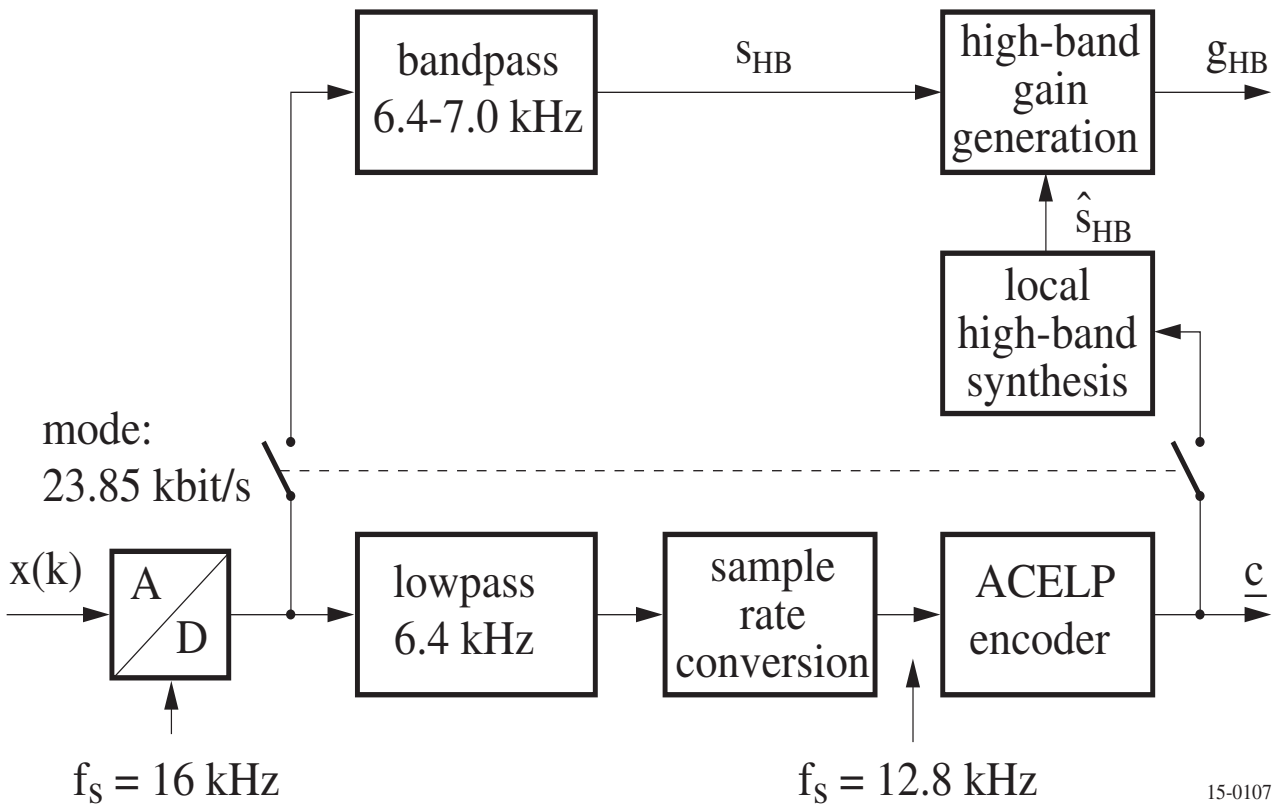
- integration in wideband terminals for transition period
- postprocessing of wideband speech decoding
  → super-wideband speech quality (50 – 16000 Hz)

▶ Use of BWE techniques in speech codecs

- conventional (e.g. CELP) coding at low frequencies
- coding with side information at high frequencies
- high quality obtainable with low additional rate
- standards: GSM-FR, AMR-WB/G.722.2, Extended AMR-WB+, Enhanced aacPlus

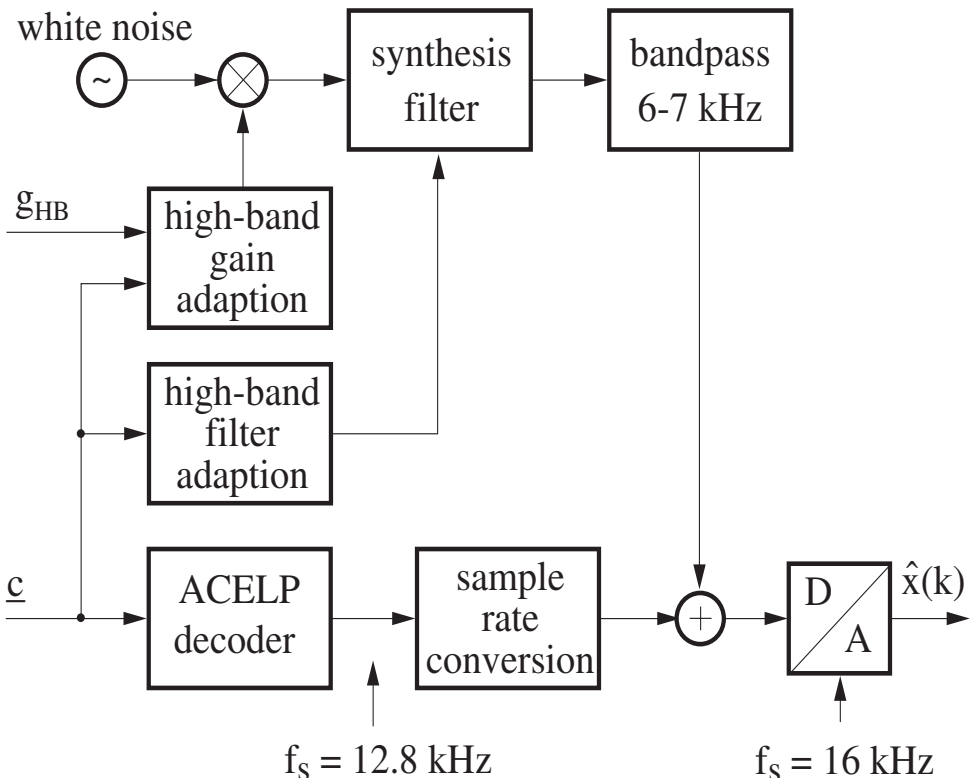
Example: AMR Wideband Speech Encoder

- Effective baseband sampling rate of 12.8 kHz
- Side info. for BWE above 6.4 kHz only in 23.85 kbit/s mode



Example: AMR Wideband Speech Decoder

- Artificial high frequency band signal (6.4 - 7 kHz):
  - white noise excitation
  - gain derived from power of low band and voicing information
  - synthesis filter derived from baseband linear prediction filter



Conclusions

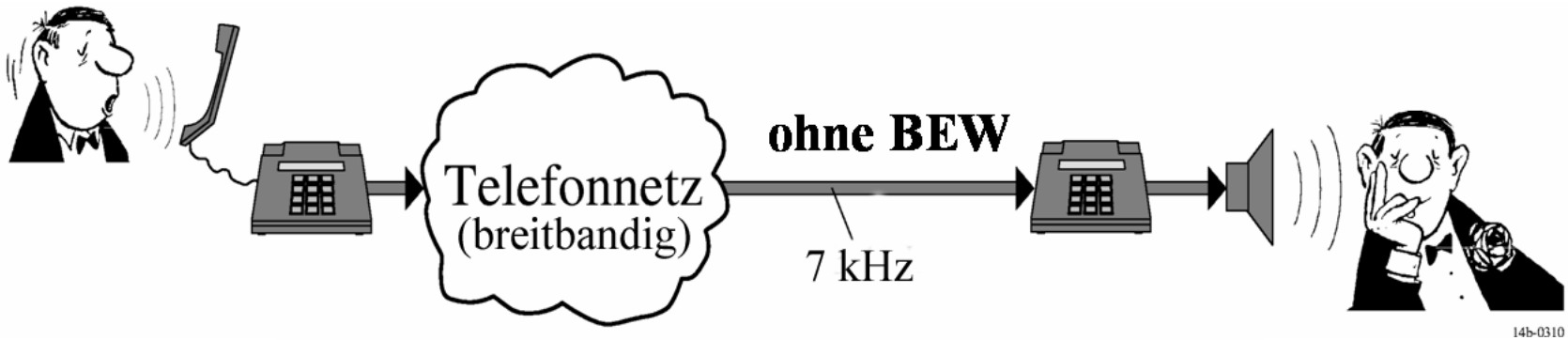
▶ BWE techniques in speech codecs

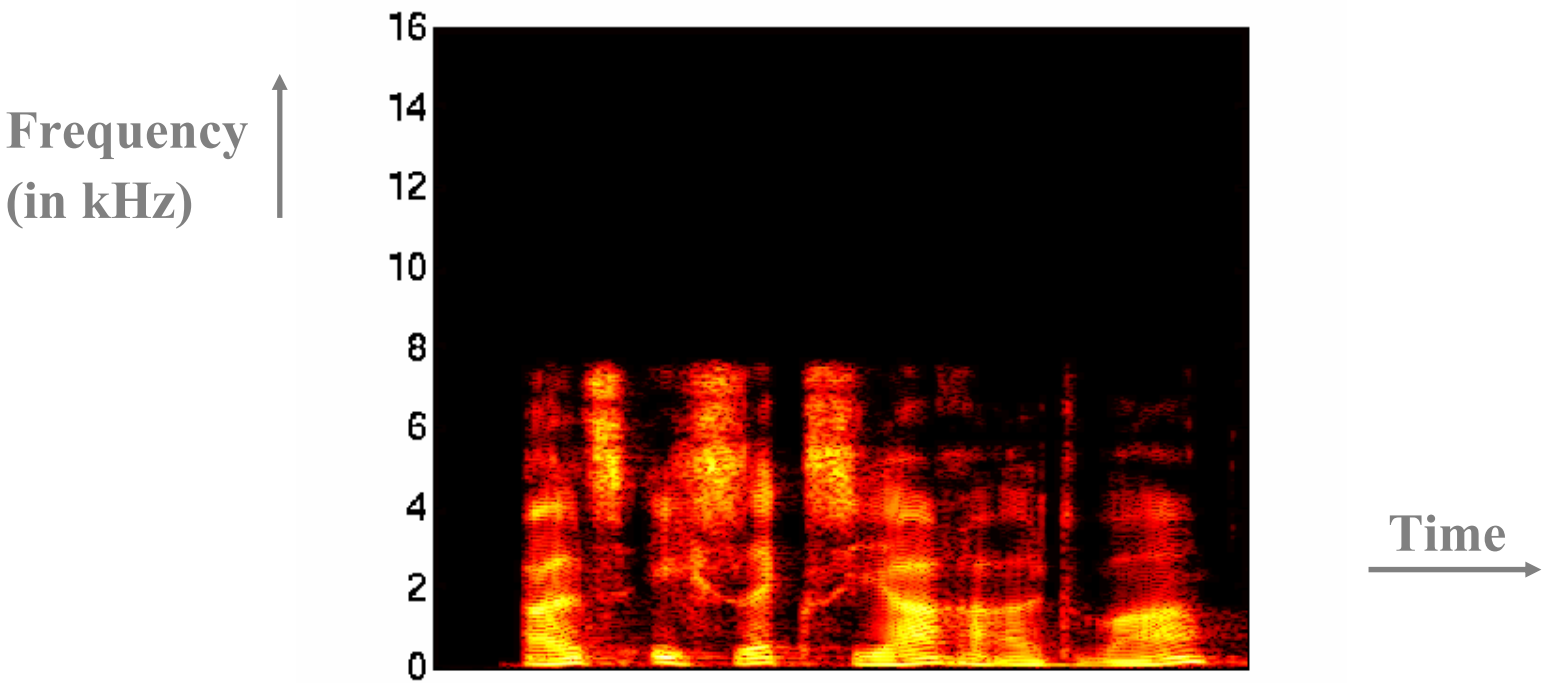
- found in established standards
- more sophisticated BWE techniques to be expected in future codecs

▶ Stand-alone bandwidth extension

- can be placed at the receiver or in the network
- no modification of sending terminal or network
- improvement of subjective speech quality
- *complement* to wideband speech coding

Audio-Demo: Wideband Speech plus BWE





GO

Audio-Demo: AMR-WB with BWE


Original „Wideband“ Speech (7 kHz)


AMR-WB Mode 12.65 kbit/s


AMR-WB 12.65 kbit/s plus BWE


BWE without AMR-WB


Original „CD-Quality“ Speech (16 kHz)