AN AMPLITUDE- AND FREQUENCY MODULATION VOCODER FOR AUDIO SIGNAL PROCESSING

Sascha Disch and Bernd Edler

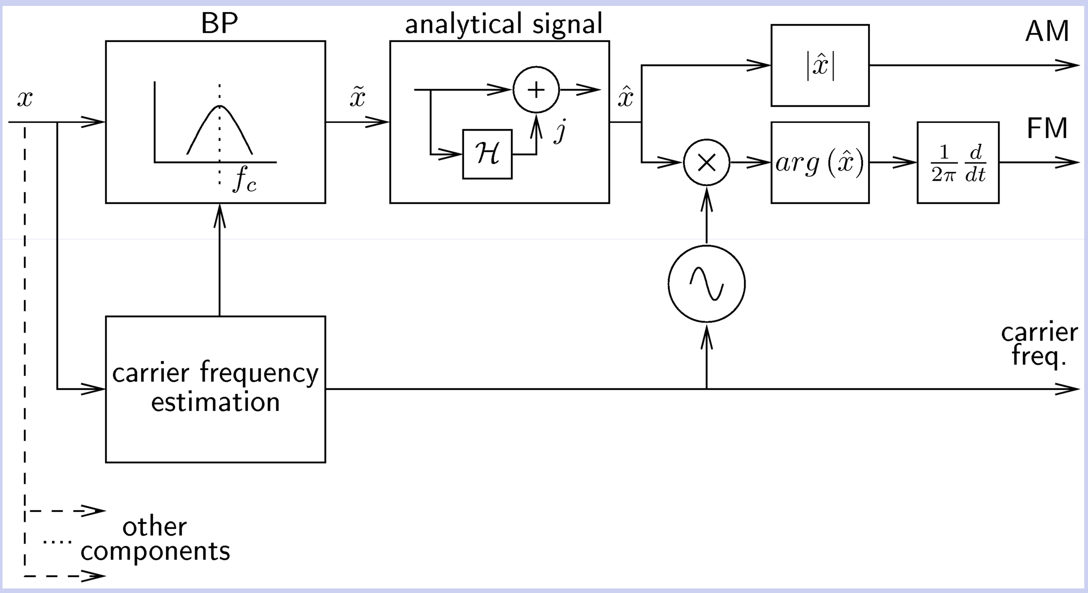
1. Introduction

The **decomposition of audio signals** into **perceptually meaningful modulation components** is highly desirable for e.g. efficient audio compression algorithms and new musical audio effects.

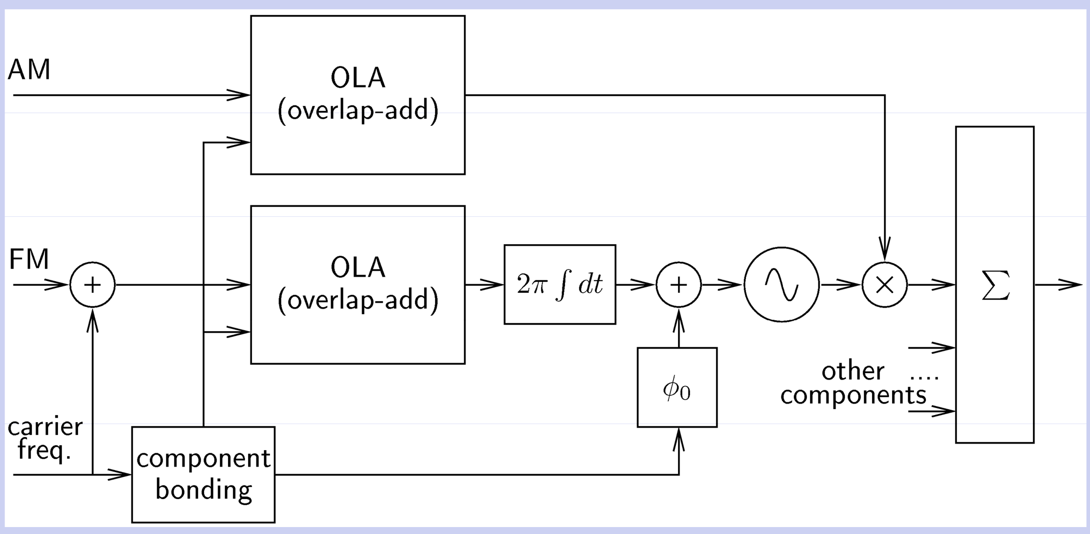
- Decomposition into sets of carriers, amplitude modulation (AM) and frequency modulation (FM) components
- Ill-defined problem, infinity number of possible solutions - thus need for border conditions
  - Adaption to human perception
  - Interpretability of modulation parameters
  - Scalability of audio texture detail
  - Minimal side artifacts for all types of modulation processing
- Lowest possible processing delay

2. Proposed System

- **Analysis**



- **Synthesis**



3. Spectral Segmentation

Temporal blocks
- 340ms duration
- 75% overlap

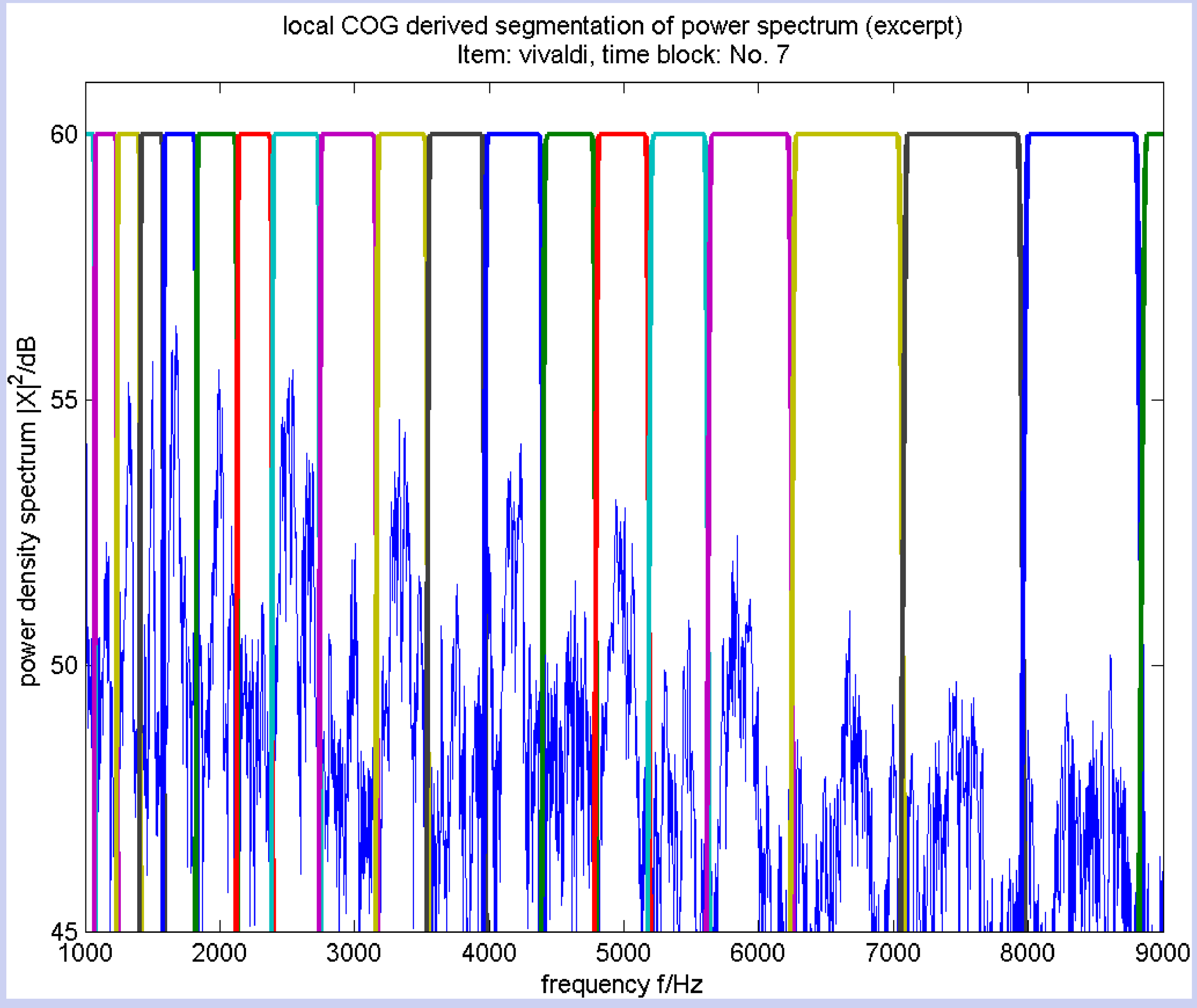
Spectral Segments
- Seamless spectral coverage
- Adaptive vs. critical bands
- Local **c**entre **o**f **g**ravity (cog)
- Post-selection

$$CogPos(k,m)=\frac{nom(k,m)}{denom(k,m)}$$

$$nom(k,m)=\alpha\sum_{i=-B(k)/2}^{+B(k)/2}\left(iw(i)\left|X(k+i,m)\right|^2\right)+(1-\alpha)\,nom(k,m-1)$$

$$denom(k,m)=\alpha\sum_{i=-B(k)/2}^{+B(k)/2}\left(w(i)\left|X(k+i,m)\right|^2\right)+(1-\alpha)\,denom(k,m-1)$$

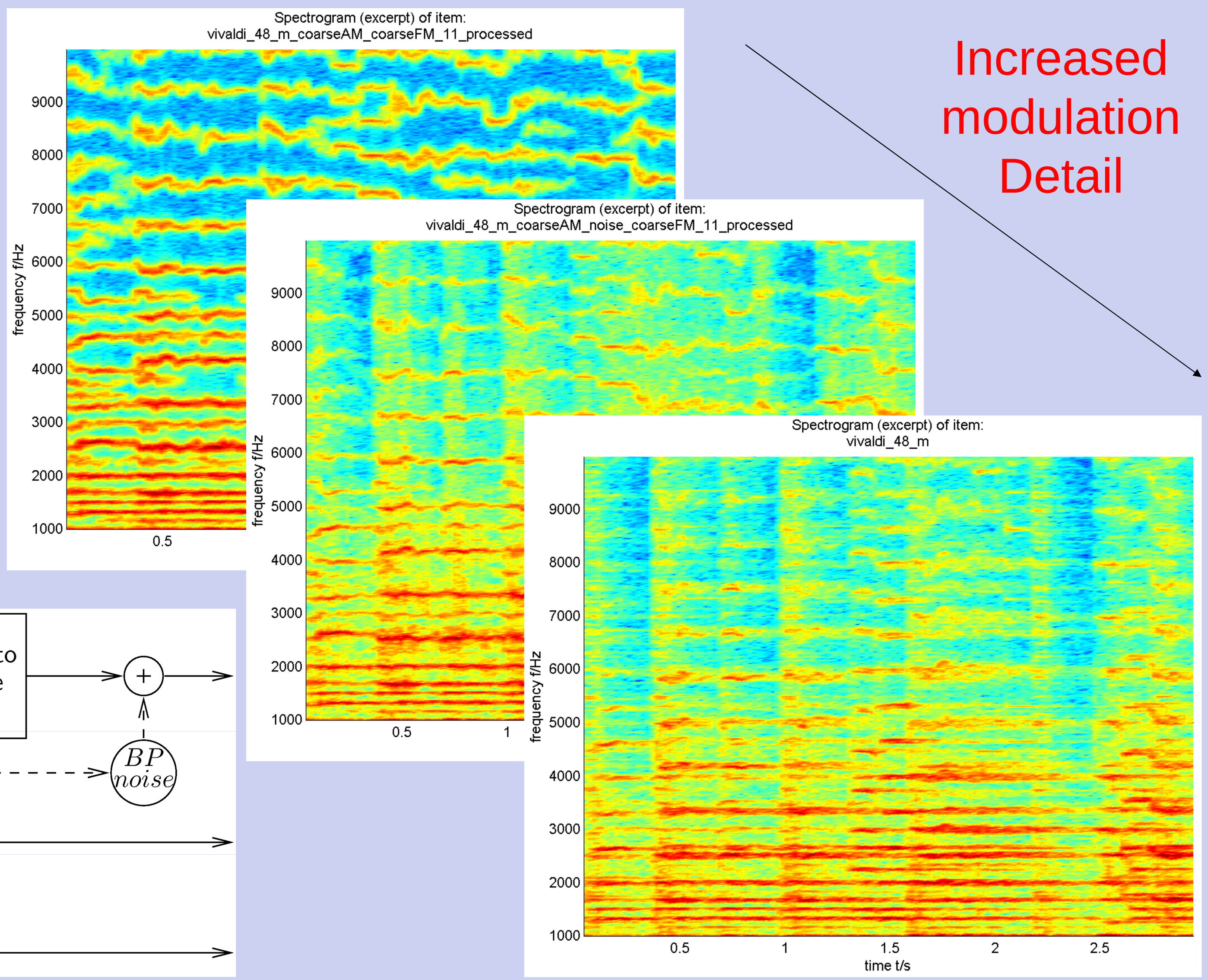
$$\alpha=\frac{1}{\tau F_s}$$



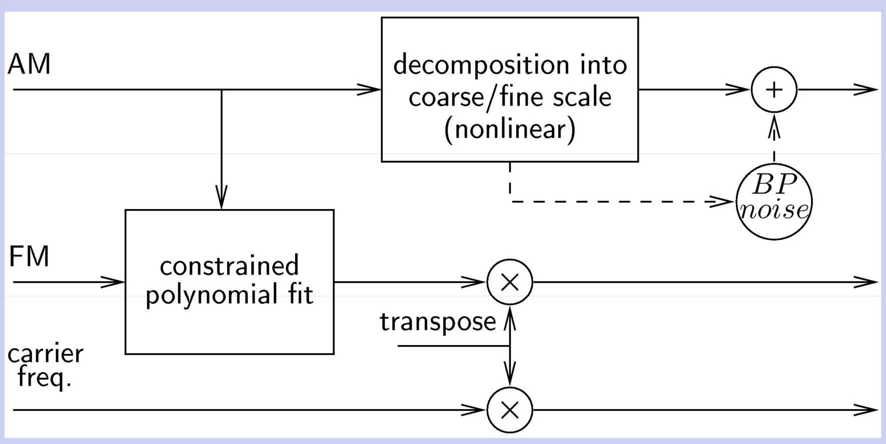
4. Modulation Processing

Modification of **AM/FM**
- Roughness control

Modification of **carrier**
- Pitch Transposition
- Local *vertical coherence* is preserved in AM ("envelope")

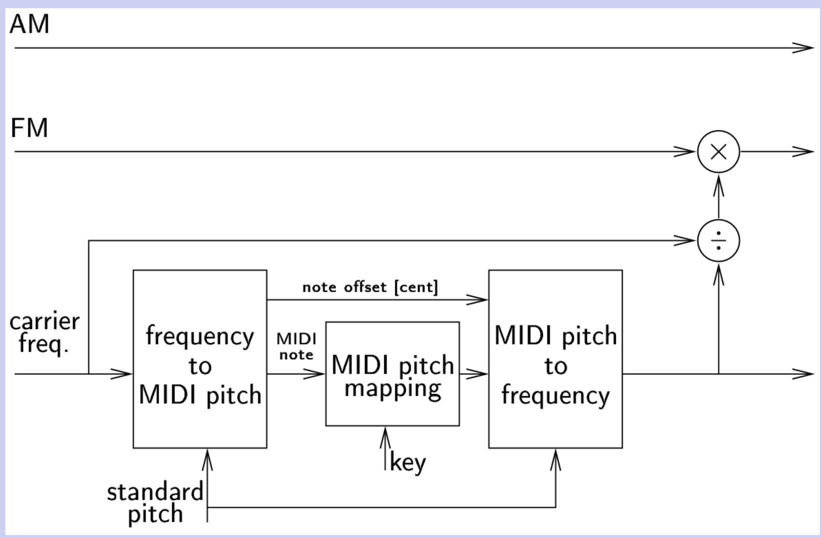


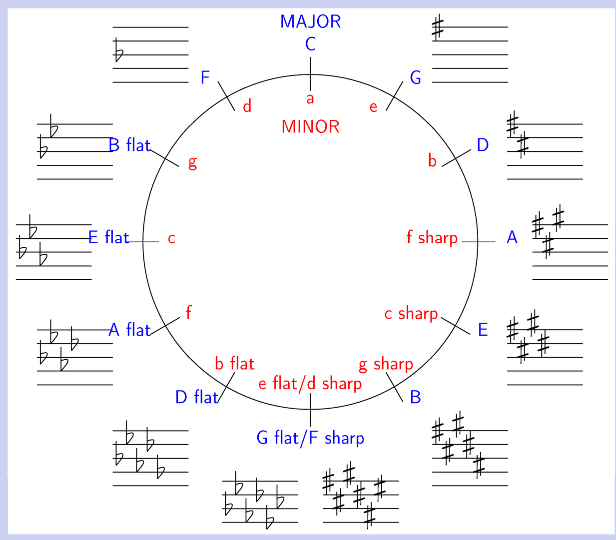
Increased modulation Detail



5. Key Mode Change

Change of the key mode of polyphonic music content
- **Carrier modification via MIDI pitch mapping**
- Circle of Fifth
  - Major to minor: 3 steps counterclockwise
  - minor to Major: 3 steps clockwise
- FM is adapted to new carrier frequency value
- Note onset detection not needed since ADSR info is contained in AM ("envelope")

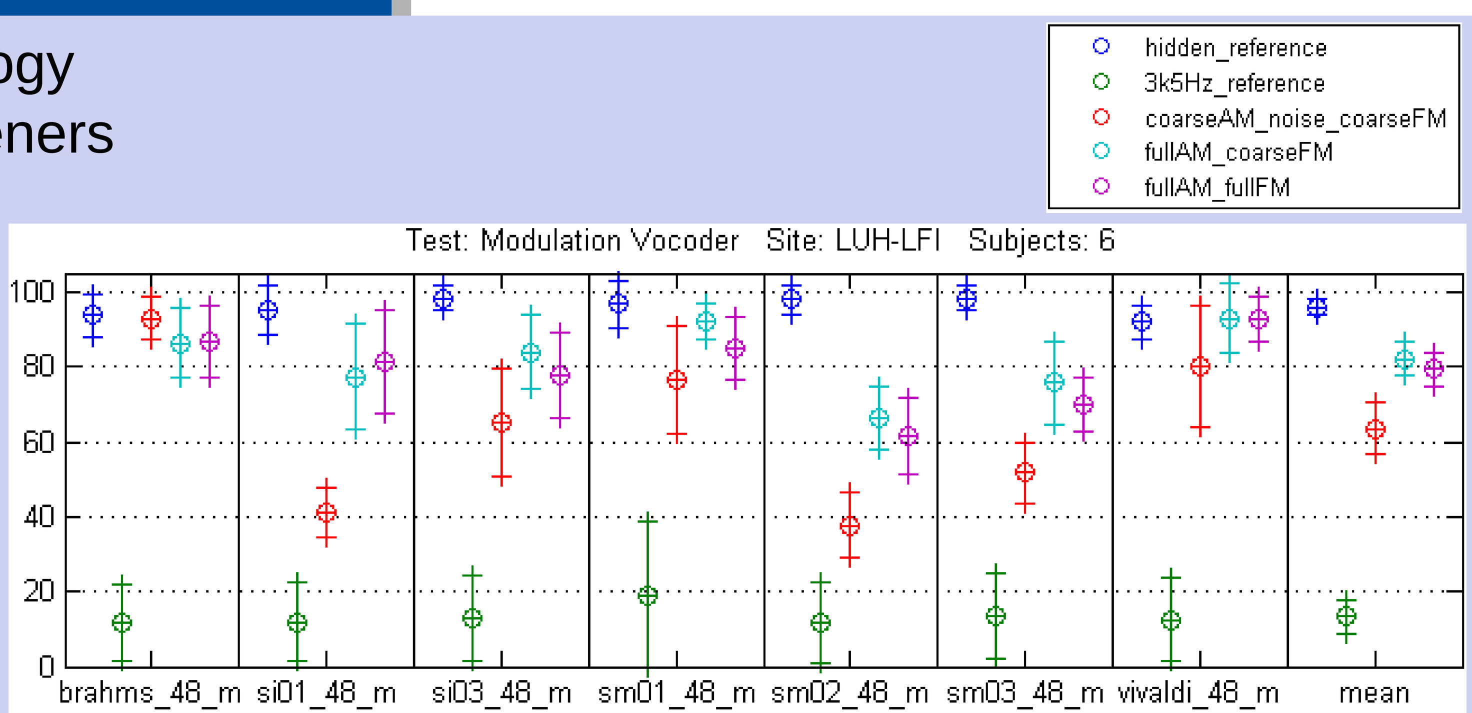




6. Listening Test Results

- MUSHRA methodology
- Six experienced listeners
- Tonal items

Item	Description
vivaldi_48_m	Classical orchestral music
brahms_48_m	Classical orchestral music
si01_48_m	Harpsichord/MPEG
si03_48_m	Pitch pipe/MPEG
sm01_48_m	Bagpipe/MPEG
sm02_48_m	Glockenspiel/MPEG
sm03_48_m	Plucked string/MPEG



7. Conclusion

- A promising novel method for audio decomposition into modulation components
- High quality synthesis
- Scalable modulation detail
- Link between waveform and parametric coding
- Applications scenarios
  - Audio codec bandwidth extension
  - Research tool for auditory perception
  - New audio effects

Leibniz Universität Hannover

Laboratorium für Informationstechnologie