

INTERNATIONAL AUDIO LABORATORIES ERLANGEN
A joint institution of Fraunhofer IIS and Universität Erlangen-Nürnberg



Learning with Music Signals: Technology Meets Education

Meinard Müller

International Audio Laboratories Erlangen
meinard.mueller@audiolabs-erlangen.de

Jahrestreffen GIBU, Dagstuhl
03. April 2023



Meinard Müller



- Mathematics (Diplom/Master, 1997)
Computer Science (PhD, 2001)
Information Retrieval (Habilitation, 2007)
- Senior Researcher (2007-2012)
- Professor Semantic Audio Processing (since 2012)
- Former President of the International Society for Music Information Retrieval (MIR)
- IEEE Fellow for contributions to Music Signal Processing



Meinard Müller: Research Group

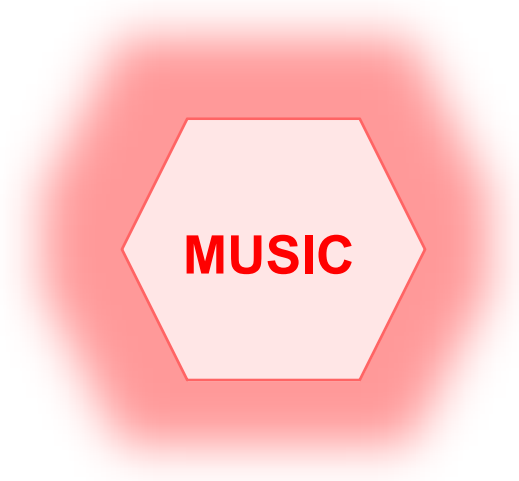
Semantic Audio Processing

- Michael Krause
- Yigitcan Özer
- Simon Schwär
- Johannes Zeitler
- Peter Meier (external)

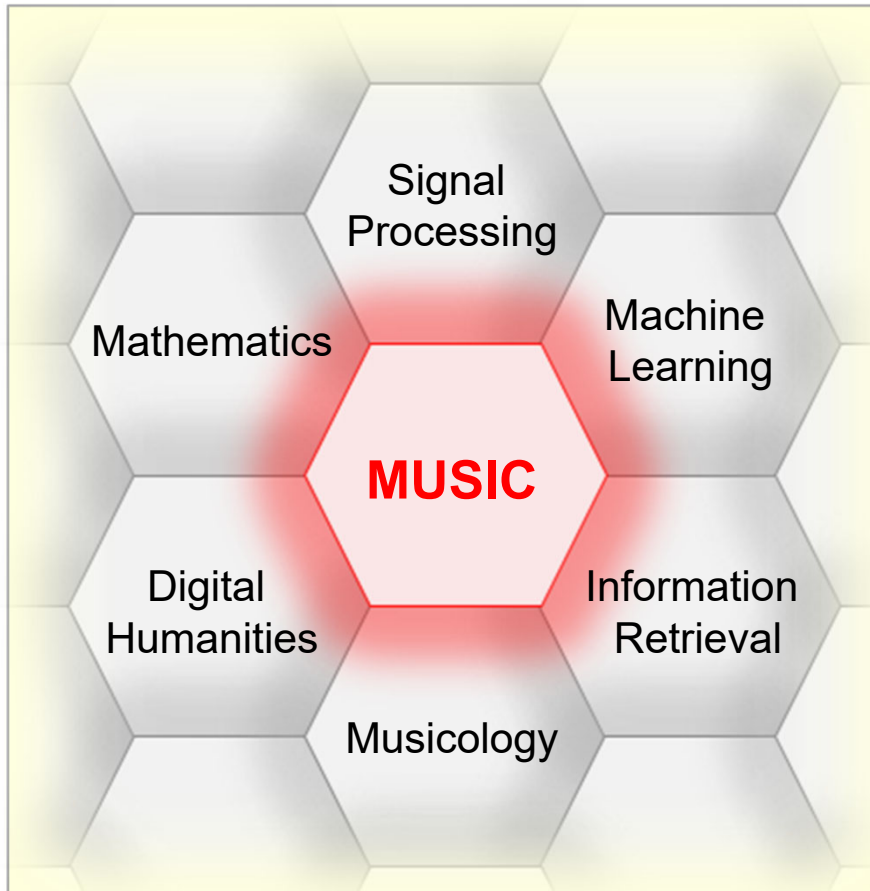
- Christof Weiß
- Sebastian Rosenzweig
- Frank Zalkow
- Christian Dittmar
- Stefan Balke
- Jonathan Driedger
- Thomas Prätzlich
- ...



Music Processing



Music Processing: A Multifaceted Research Area

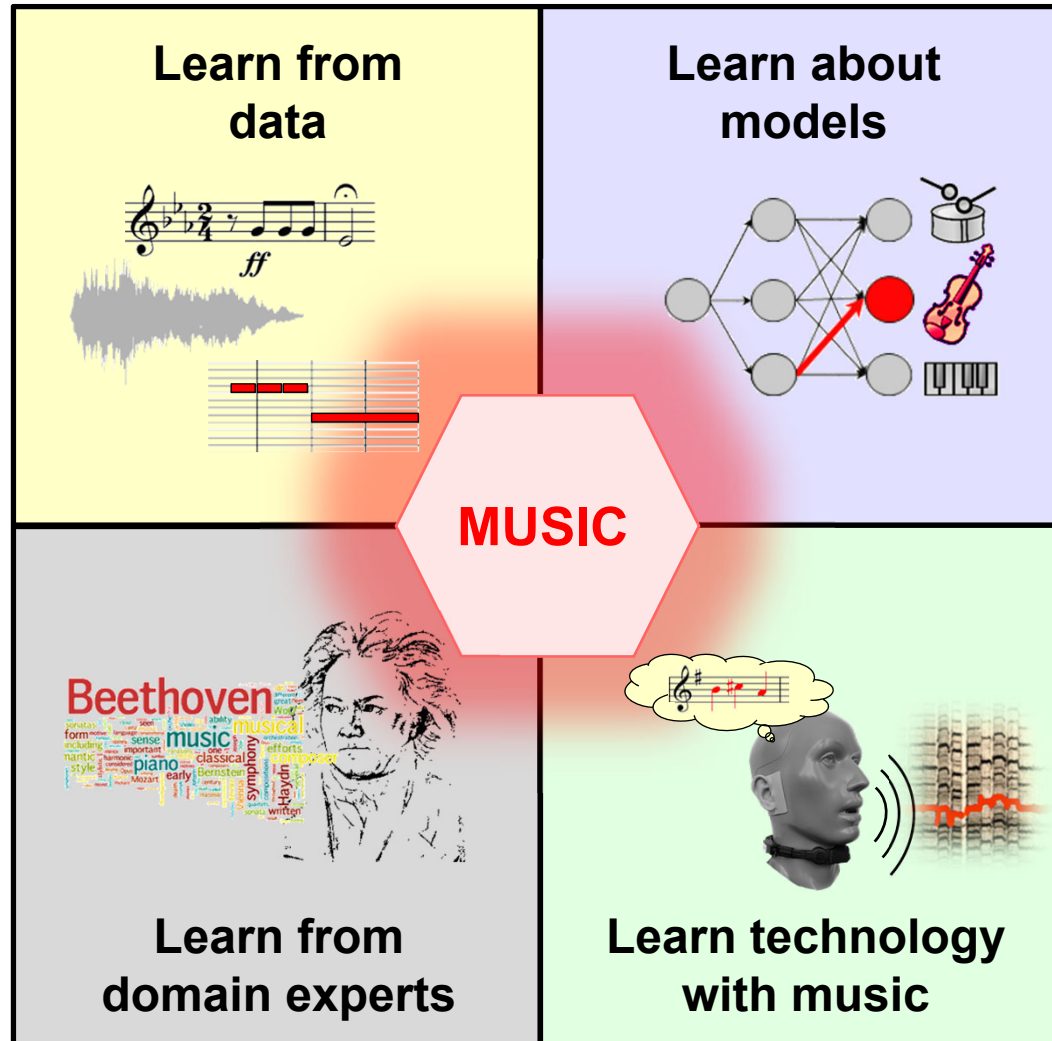


Music ...

- important part of our lives ...
- ... Spotify, Pandora, iTunes, ...
- interdisciplinary research
- intuitive entry point to education

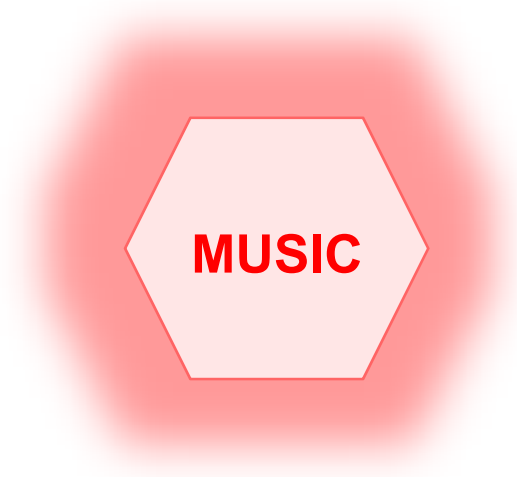
Reinhart Koselleck-Projekt: LEARN

Learning with Music Signals: Technology Meets Education



- Machine learning for music signal processing
- Interpretable models and knowledge integration
- Music understanding and applications
- Interactive learning in engineering through music

Multimodal Music Processing



Multimodal Music Processing

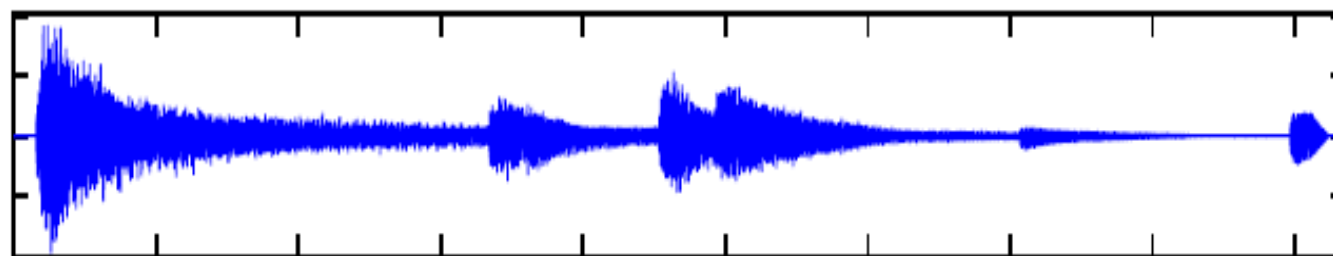


Multimodal Music Processing

Image



Audio



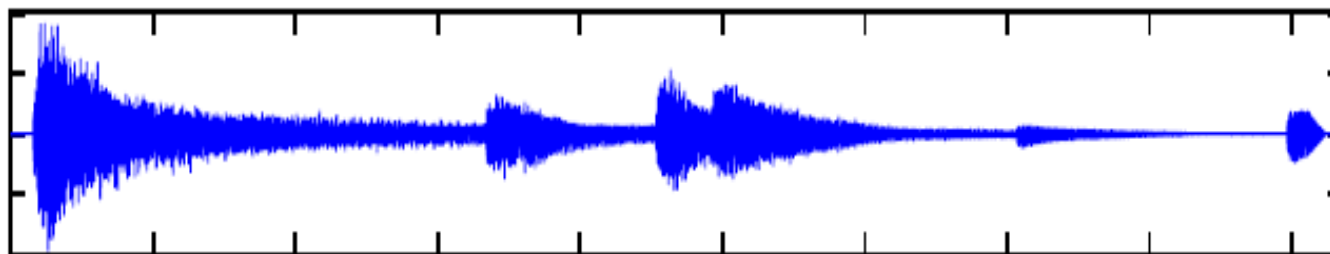
Multimodal Music Processing

Image Processing: Optical Music Recognition

Image



Audio



Multimodal Music Processing

Image Processing: Optical Music Recognition

Image



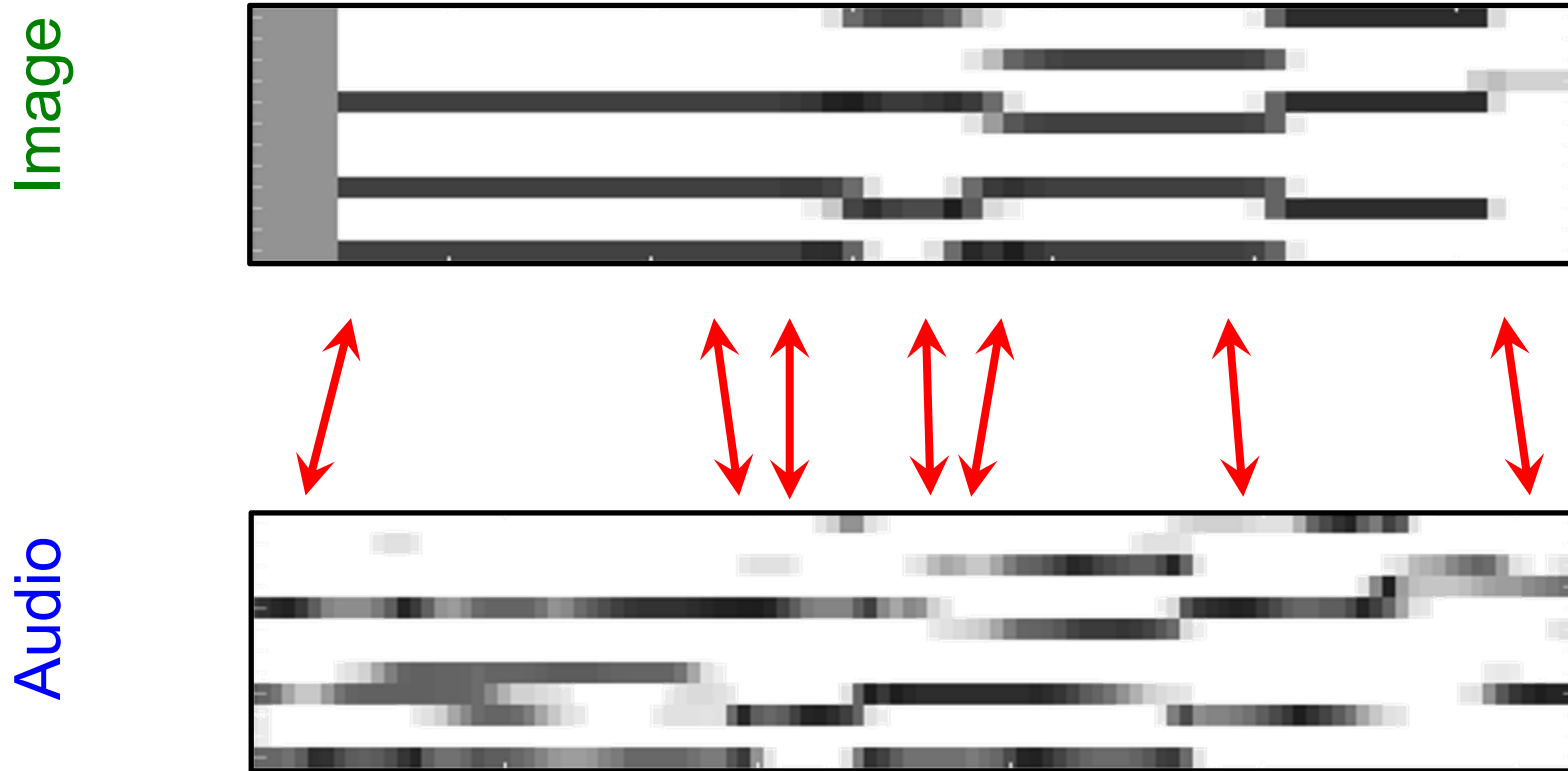
Audio



Audio Processing: Fourier Analysis

Multimodal Music Processing

Image Processing: Optical Music Recognition



Audio Processing: Fourier Analysis

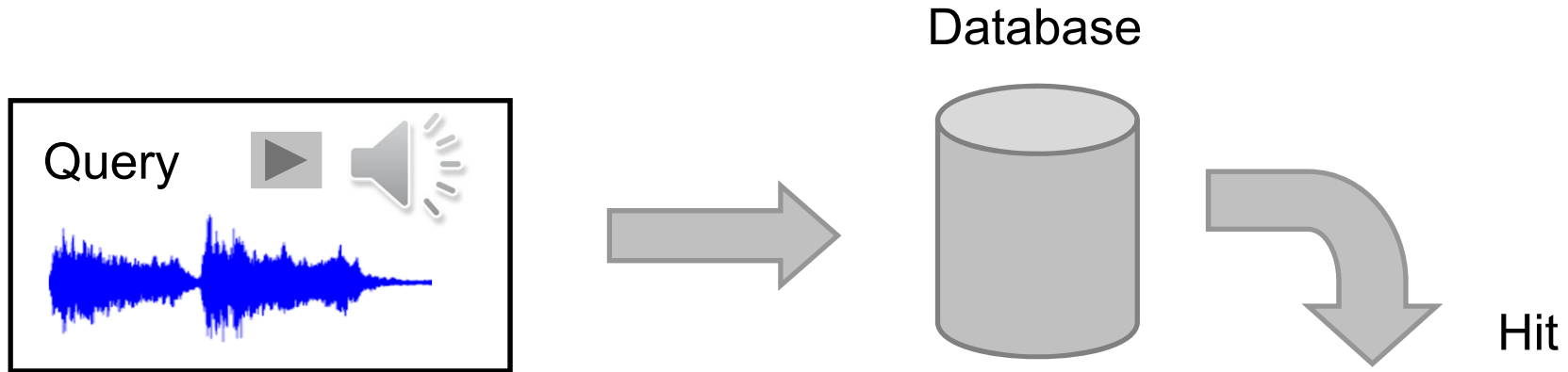


Multimodal Music Processing

The image displays two windows from a music processing application. The top window, titled "ScoreViewer", shows a musical score for "Beethoven - Klaviersonaten Band 1 - Henle". The score is for "Sonata no.8 in C minor, op.13 'Pathétique' / Rondo (Allegro)". The score is displayed in a multi-staff format, with the Rondo section highlighted in yellow. The interface includes navigation controls for Track (29 / 54), Bar (1 / 211), and Page (159 / 285). The bottom window, titled "AudioViewer", shows a track list for "Beethoven - Piano Sonatas-Alfred Brendel". The track list includes 11 tracks, with track 11, "Sonata no.8 in C minor, op.13 'Pathétique' / Rondo (Allegro)", selected. The interface includes navigation controls for Disc (1 / 11), Track (11 / 11), and Time (00:00.00 / 4:30.35). Both windows have "Play" and "Stop" buttons.



Multimodal Music Processing



Audio ID

Bernstein (1962)
Beethoven, Symphony No. 5

Version ID

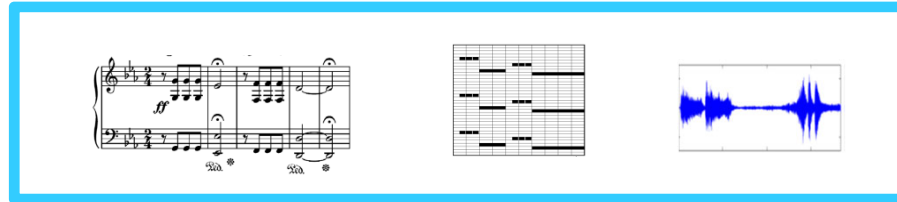
Beethoven, Symphony No. 5:
■ Bernstein (1962)
■ Karajan (1982)
■ Gould (1992)

Category ID

■ Beethoven, Symphony No. 9
■ Beethoven, Symphony No. 3
■ Haydn Symphony No. 94



Multimodal Music Processing Modalities



Retrieval tasks:

Specificity

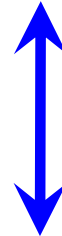
Granularity

Audio ID

High
specificity

Fragment-based
retrieval

Version ID

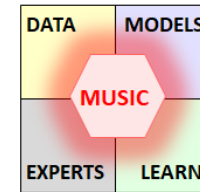


Category ID

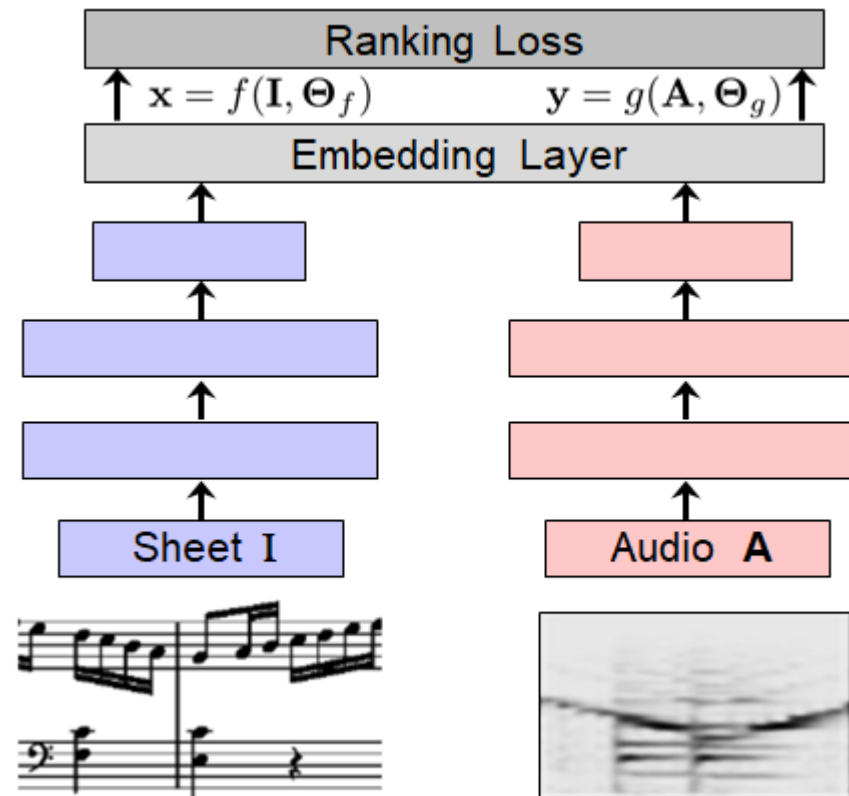
Low
specificity

Document-based
retrieval

Reinhart Koselleck-Projekt: LEARN



- Representation learning
- Embedding techniques
- Weak annotations
- Loss functions
- ...



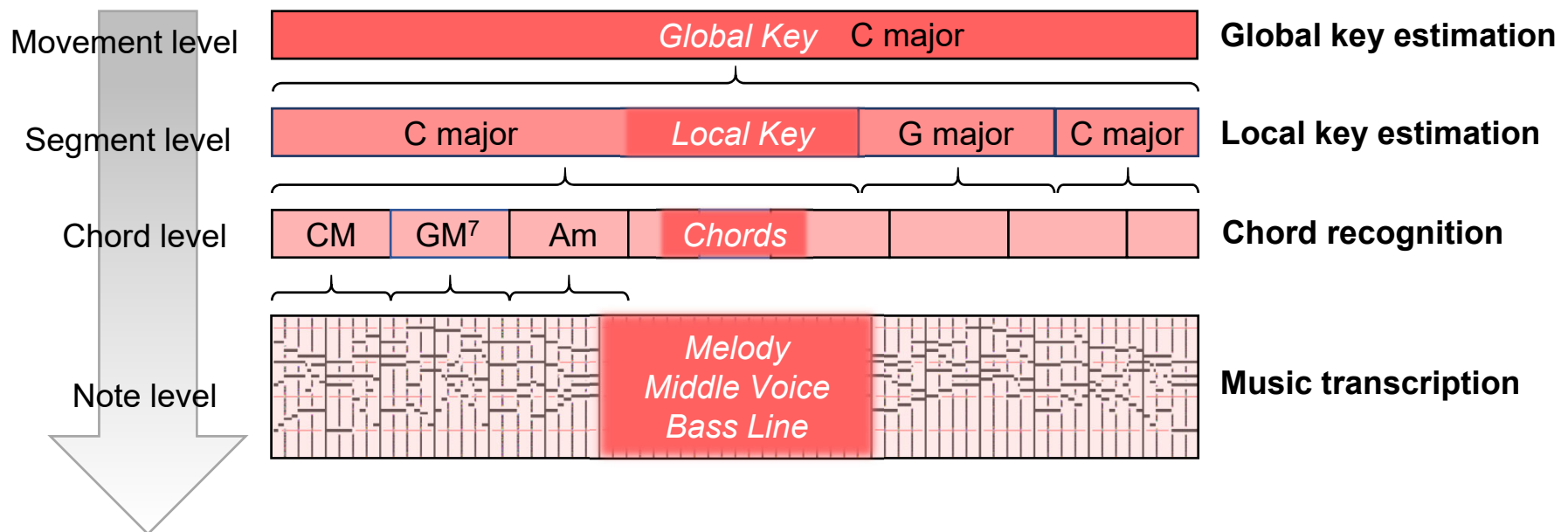
Computational Musicology

- Cooperation:
 - Rainer Kleinertz (Saarbrücken)
 - Stephanie Klauk (Saarbrücken)
 - Christof Weiß (Würzburg)
- Objectives
 - Harmony-based structural analysis
 - Beethoven Sonatas & Wagner's Ring
 - Interdisciplinary dialogue
- Since 2014: DFG-funded project



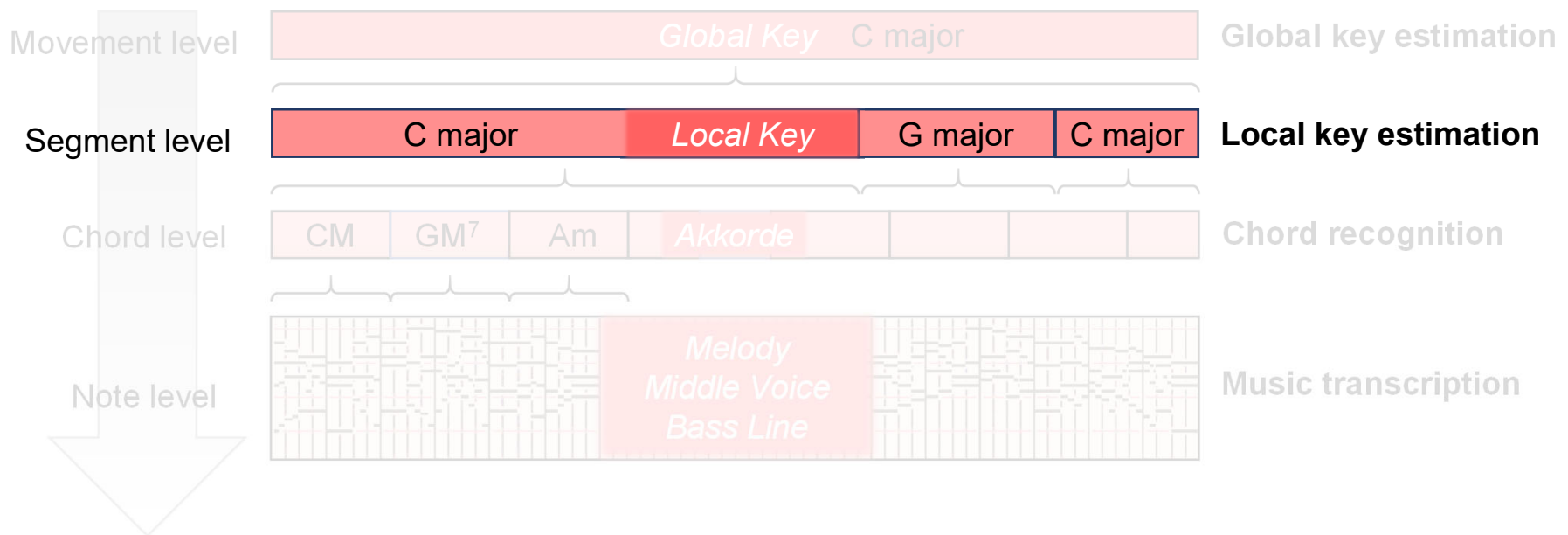
Computational Musicology: Harmony Analysis

- Different concepts
- Different temporal levels



Computational Musicology: Harmony Analysis

- Different concepts
- Different temporal levels

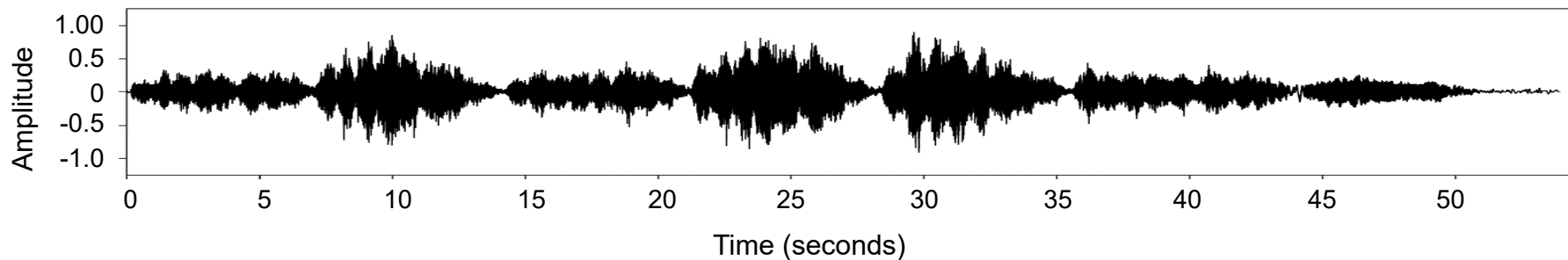


Local Key Estimation

Example: J.S. Bach, Choral "Durch Dein Gefängnis" (*Johannespassion*)



Musical score for the choral piece "Durch Dein Gefängnis" by J.S. Bach. The score is in G major (one sharp) and 4/4 time. It features a vocal line with German lyrics and a piano accompaniment. The lyrics are: "Durch dein Ge-fäng-nis, Got-tes Sohn, muß uns die Frei-heit kom-men; enn gingst du nicht die Knecht-schaft ein, müßt uns-re Knecht-schaft e-wig sein. Dein Ker-ker ist der Gna-den-thron, die Frei-statt al-ler From-men;".

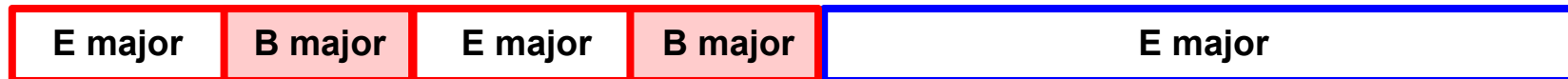
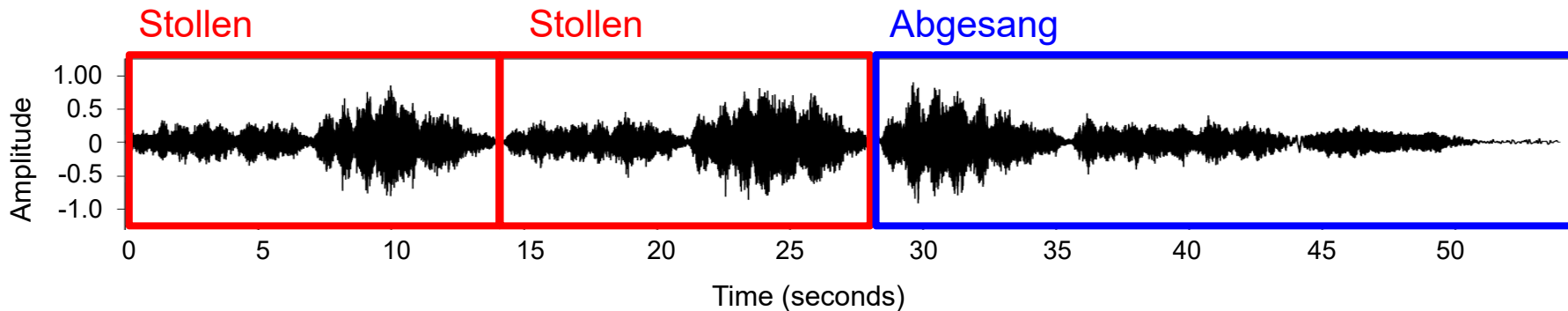


Local Key Estimation

Example: J.S. Bach, Choral "Durch Dein Gefängnis" (*Johannespassion*)

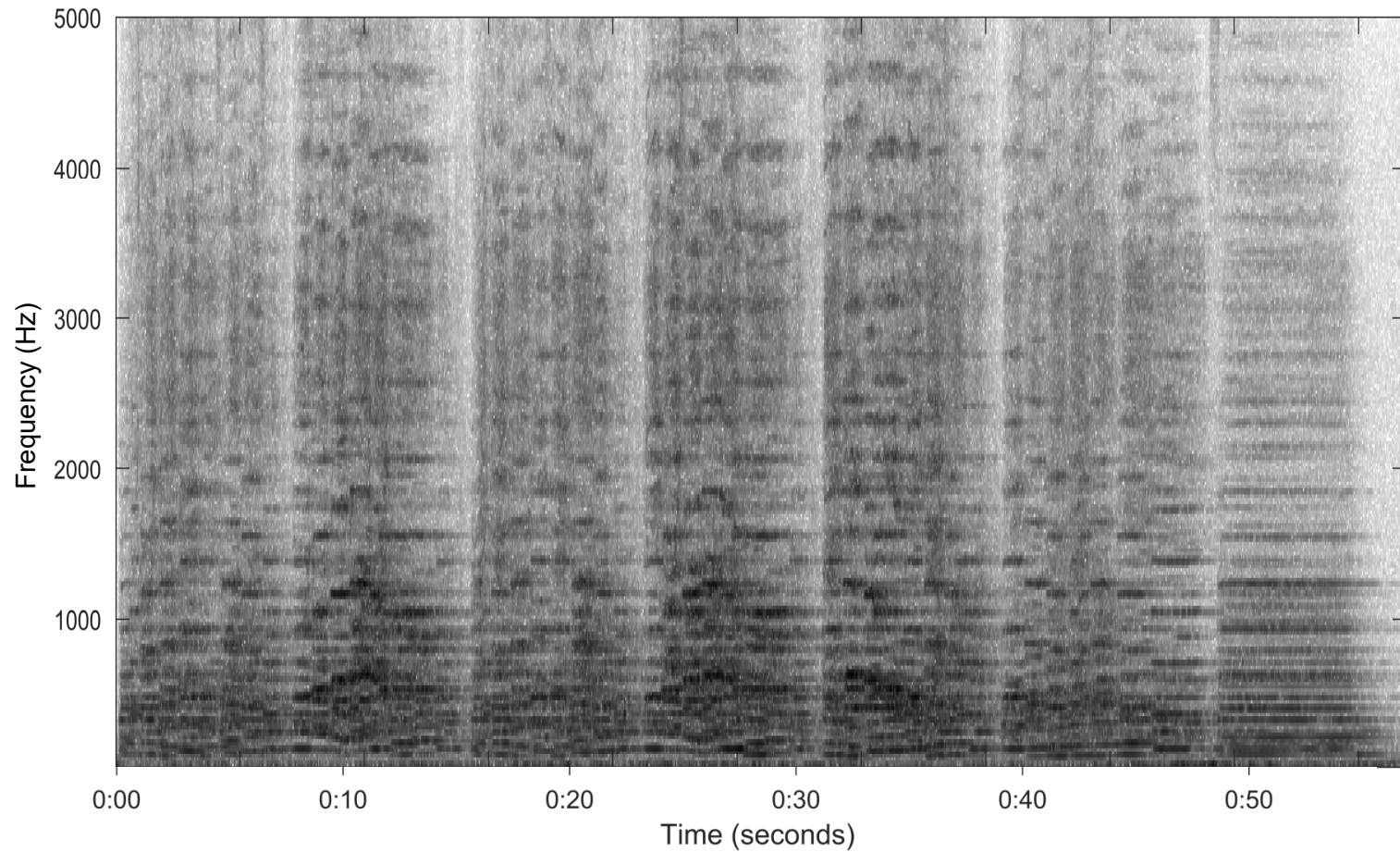
Stollen **Abgesang**

Durch dein Ge-fäng-nis, Got-tes Sohn, muß uns die Frei-heit kom-men; enn gingst du nicht die Knecht-schaft ein, müßt uns-re Knecht-schaft e-wig sein.
Dein Ker-ker ist der Gna-den-thron, die Frei-statt al-ler From-men;



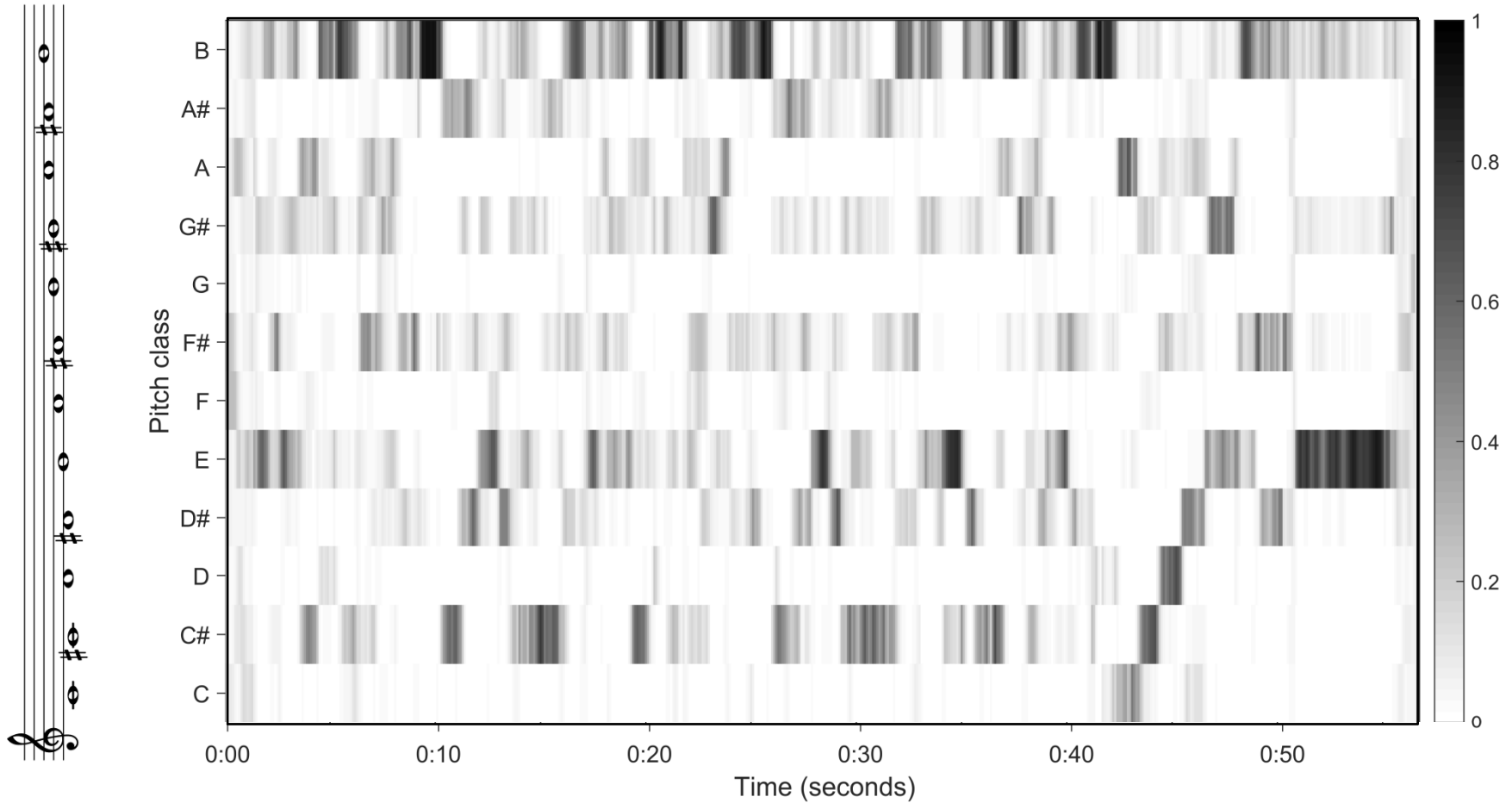
Local Key Estimation

Spectrogram



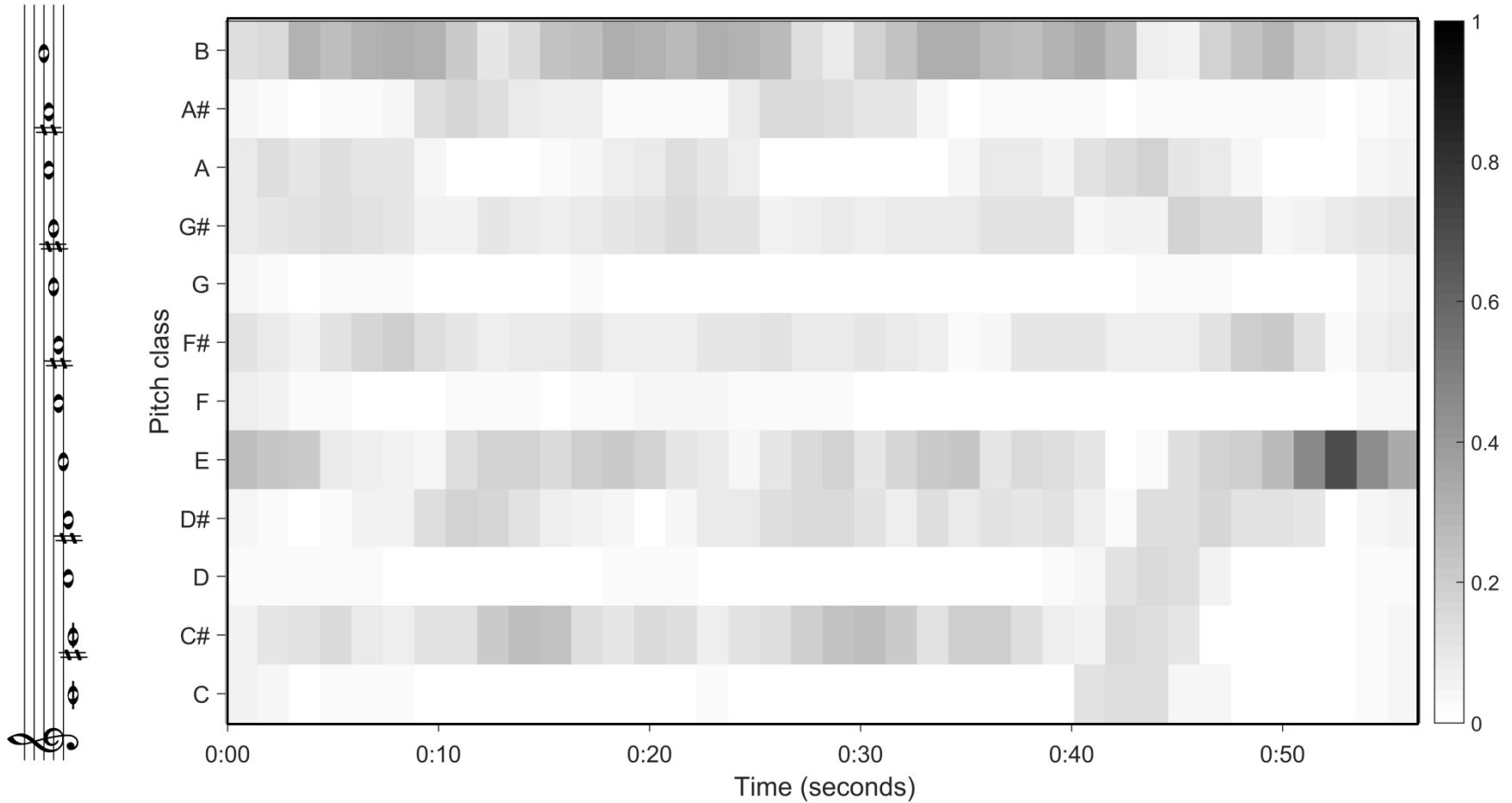
Local Key Estimation

Chromagram



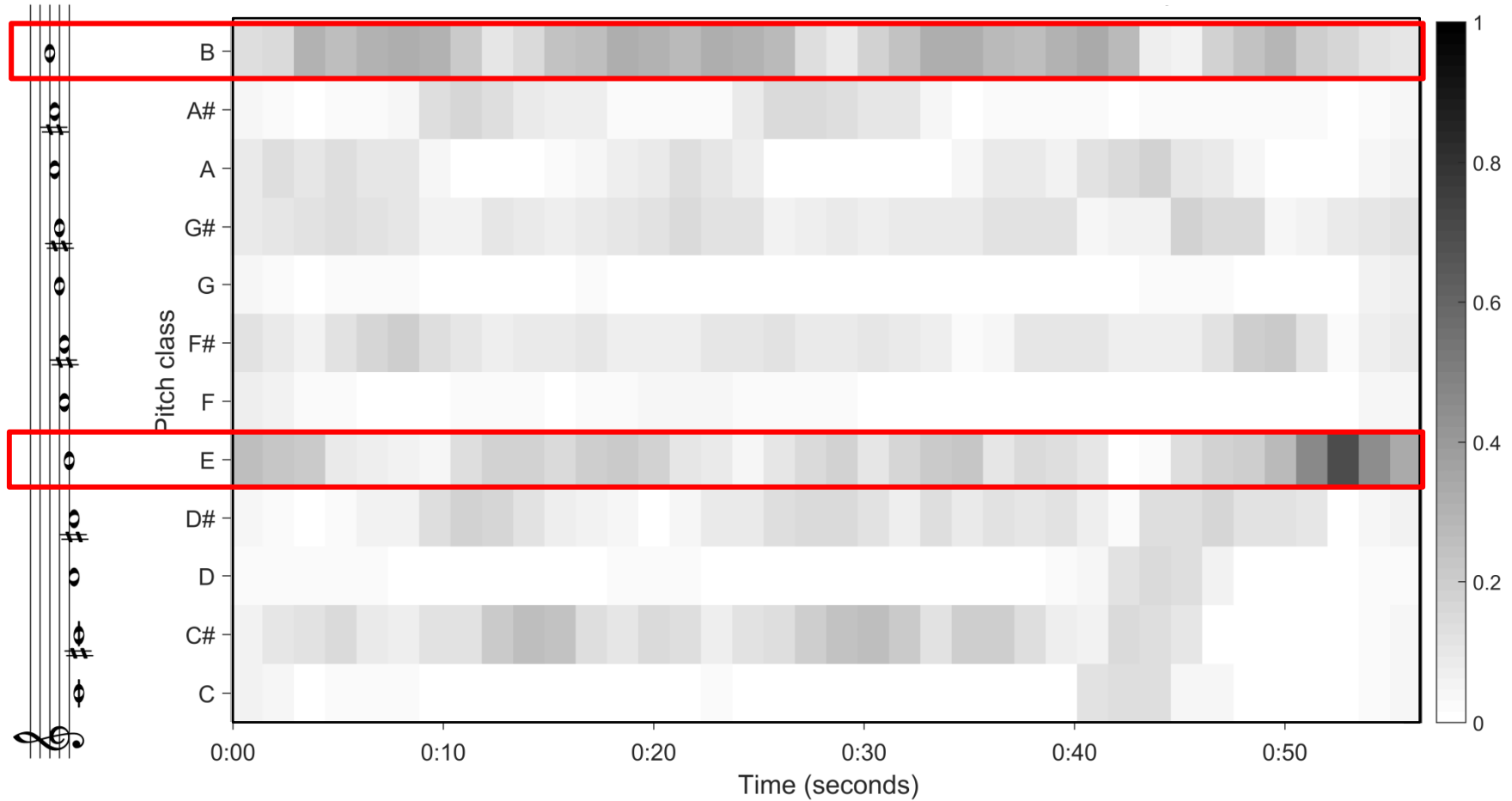
Local Key Estimation

Chromagram after smoothing



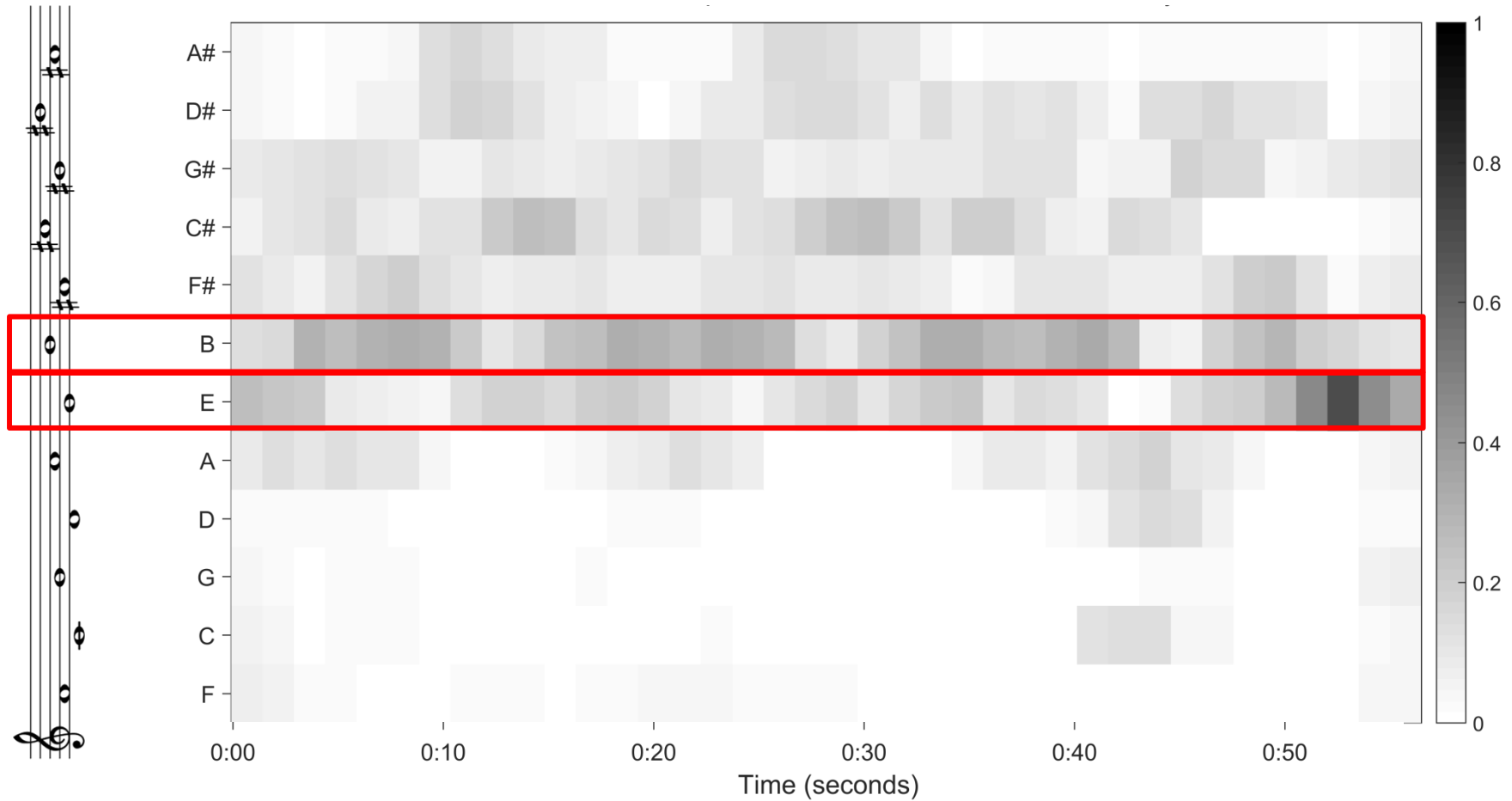
Local Key Estimation

Arrange pitch classes according to **perfect fifth series**



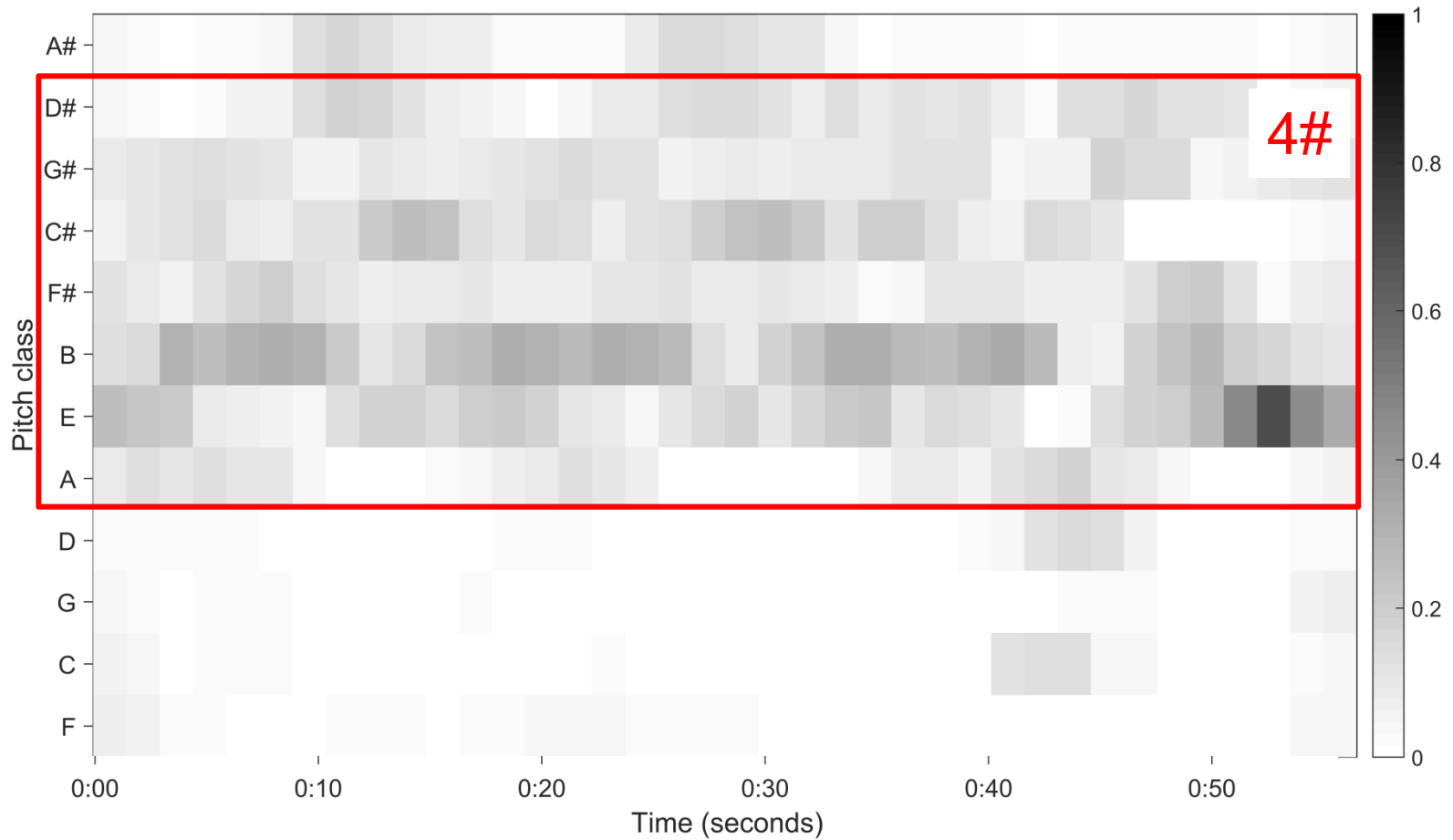
Local Key Estimation

Arrange pitch classes according to **perfect fifth series**



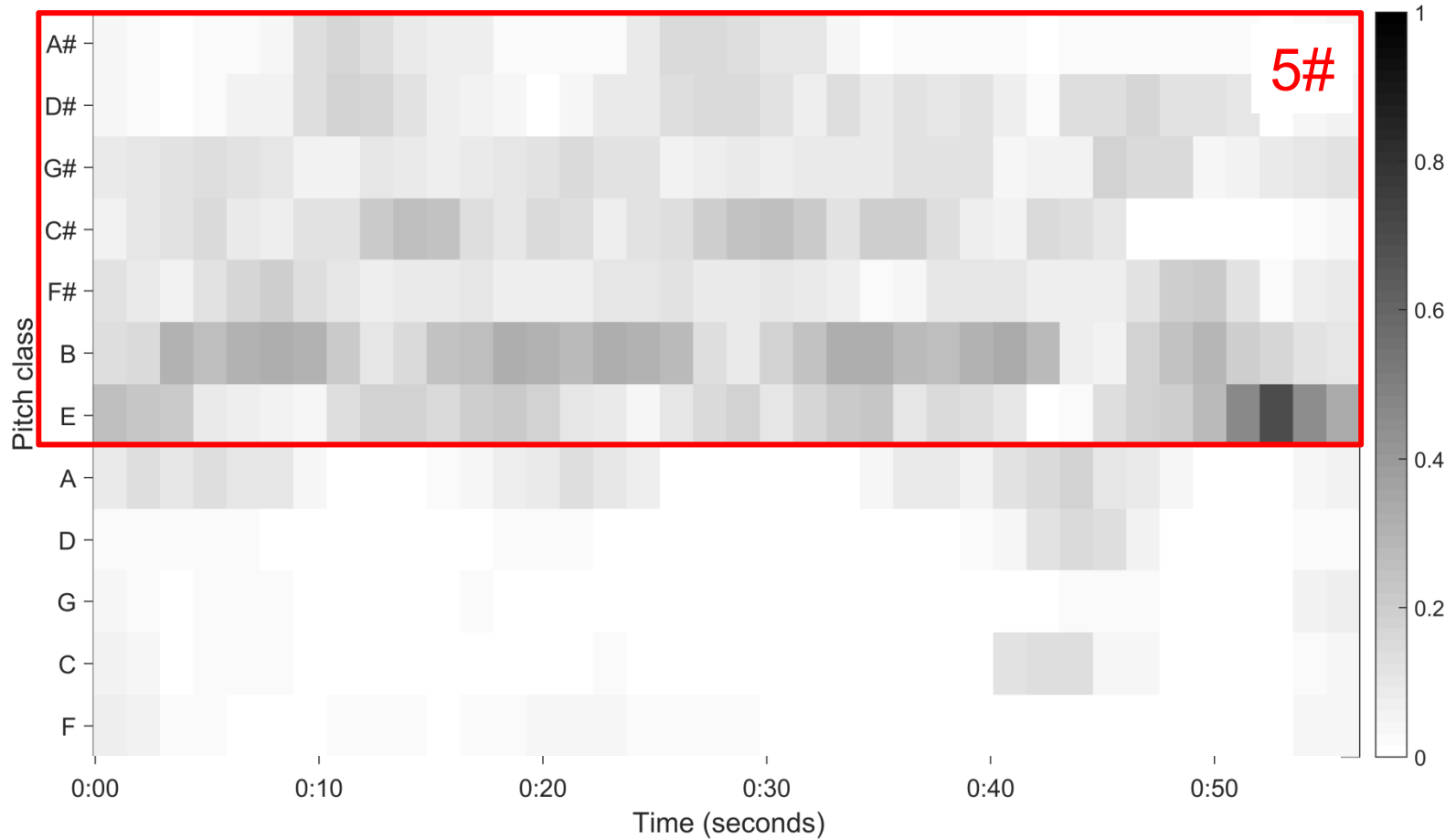
Local Key Estimation

Summarize pitch class content according to **diatonic scales**



Local Key Estimation

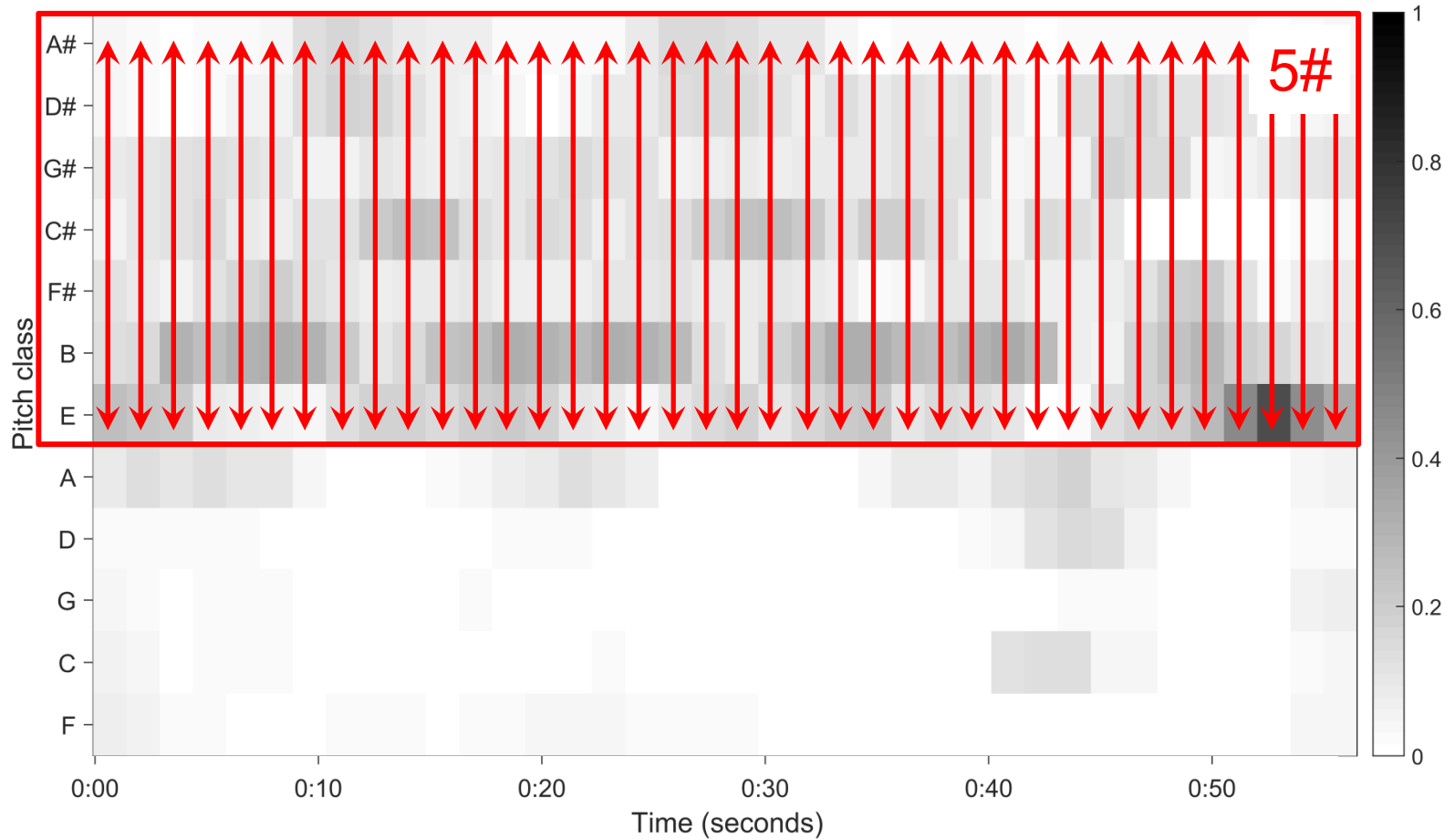
Summarize pitch class content according to **diatonic scales**



Local Key Estimation

Summarize pitch class content according to **diatonic scales**

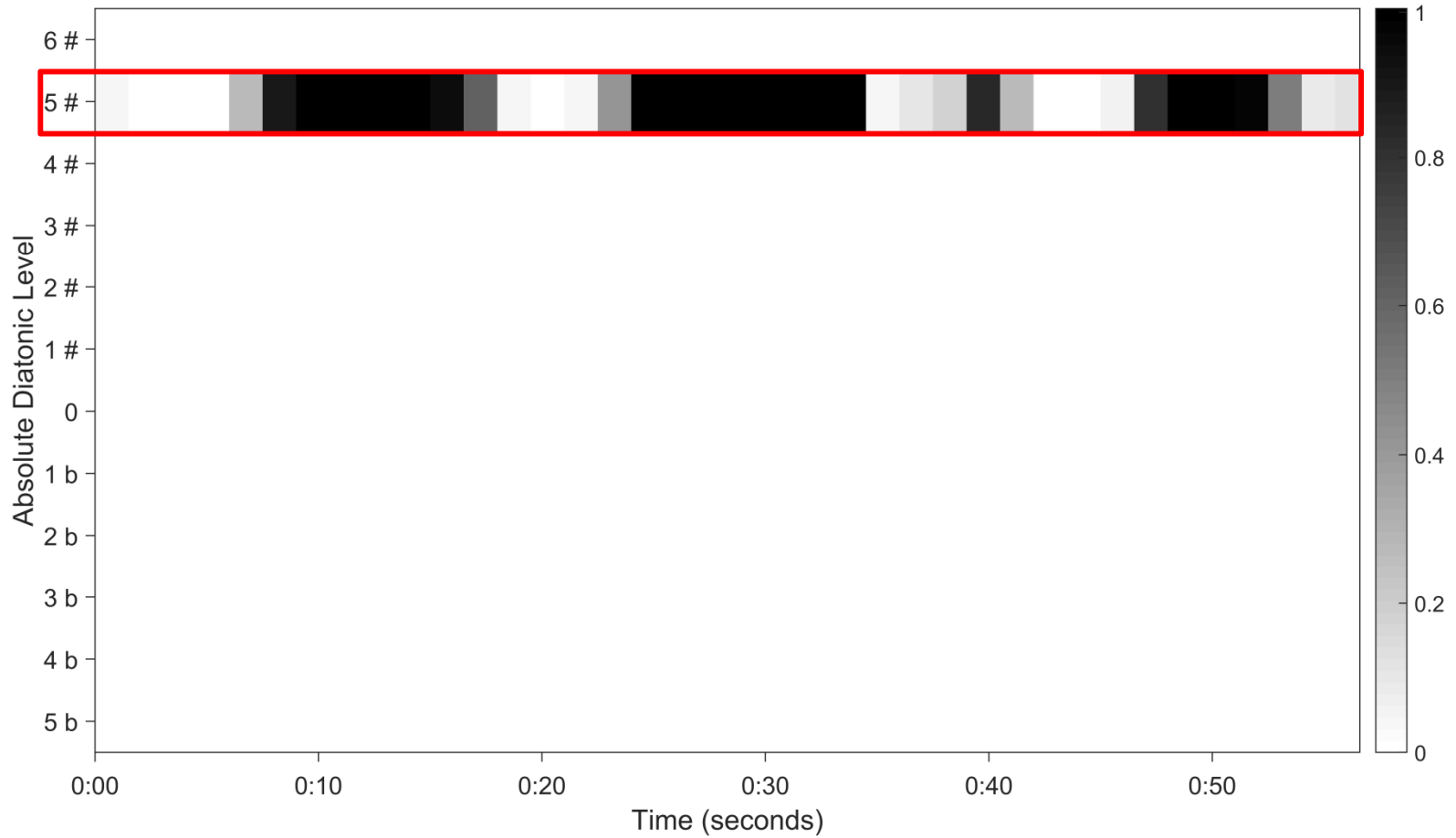
Multiply chroma values (in each column)



Local Key Estimation

Summarize pitch class content according to **diatonic scales**

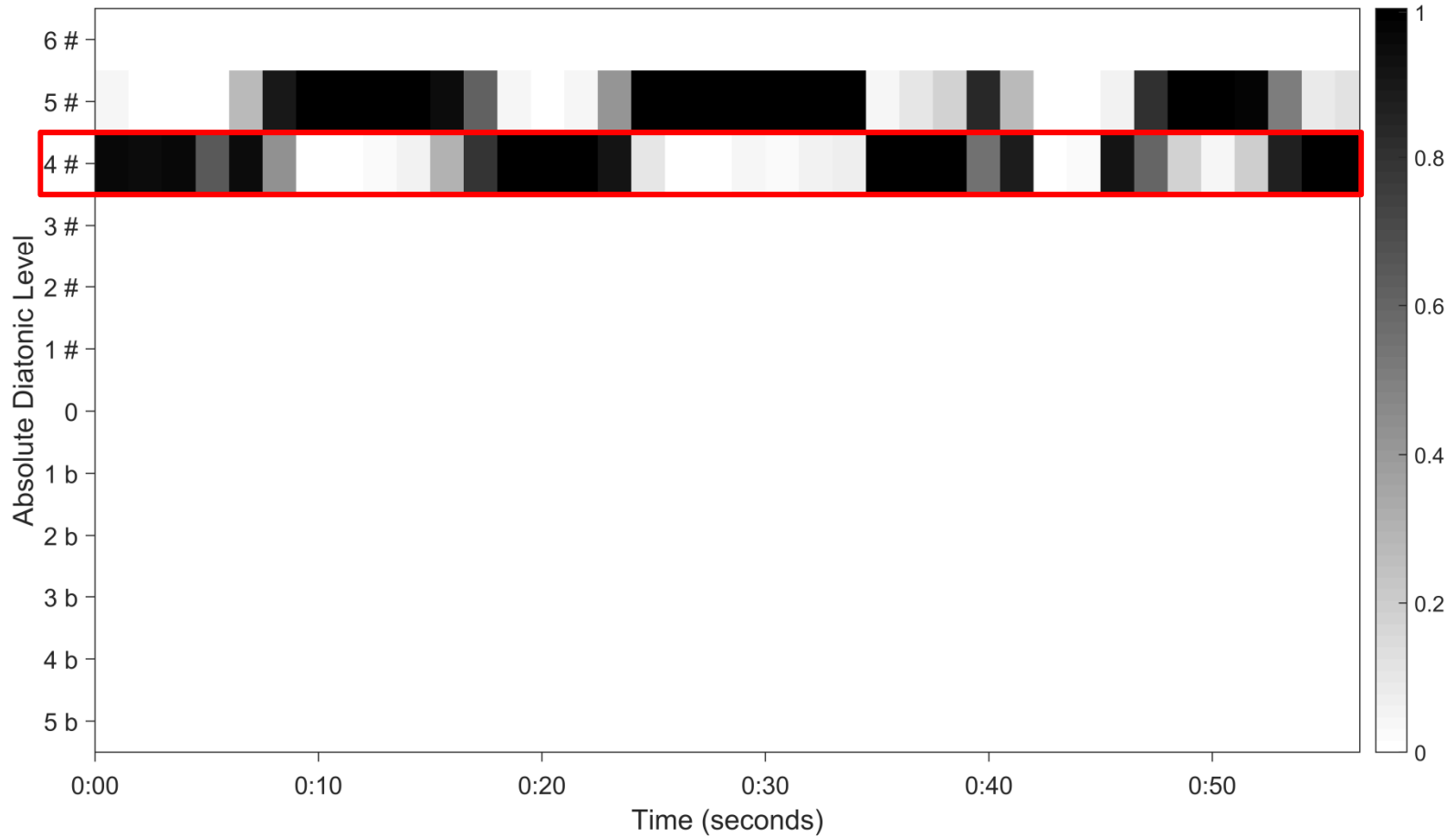
Multiply chroma values



Local Key Estimation

Summarize pitch class content according to **diatonic scales**

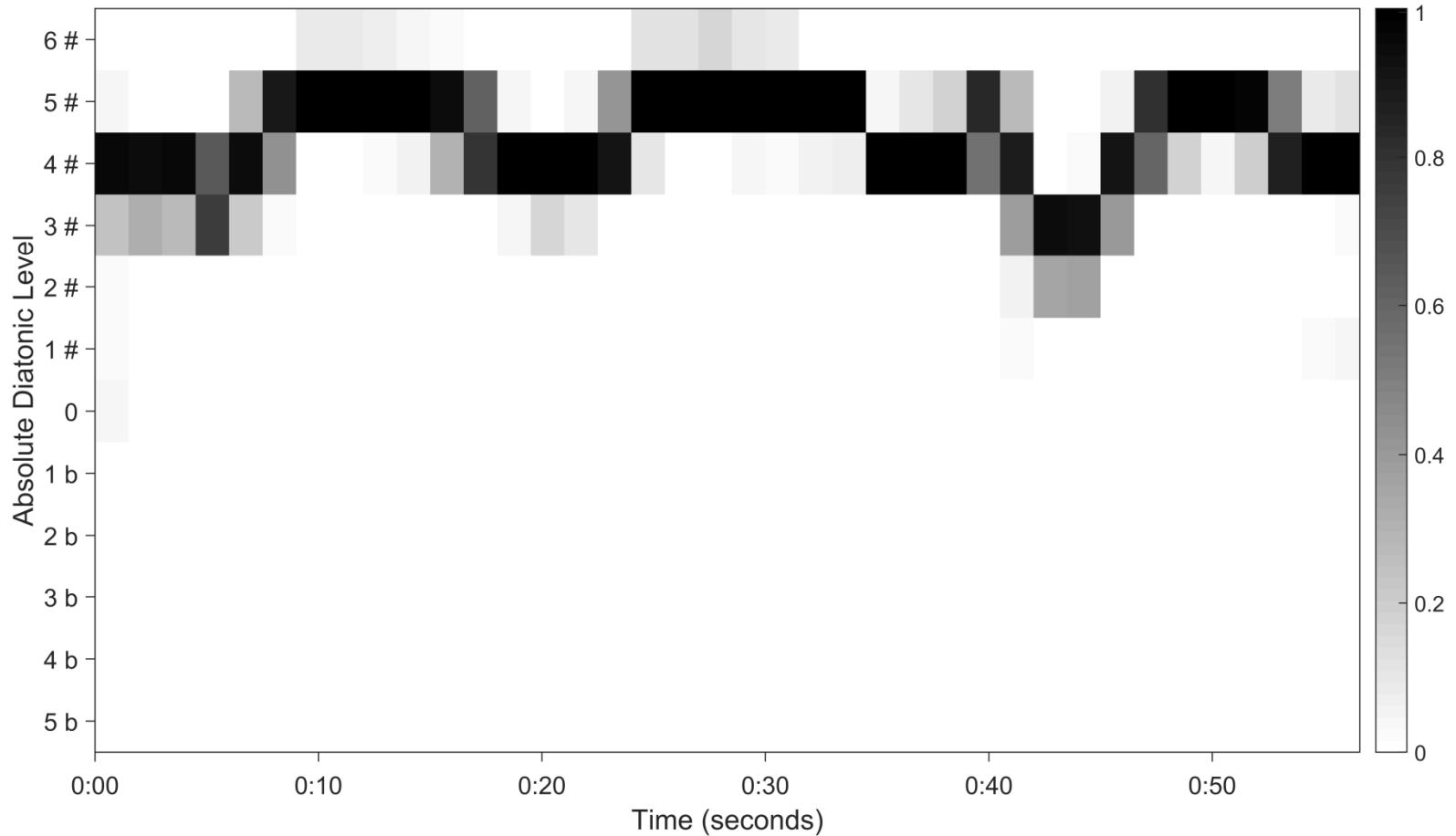
Multiply chroma values



Local Key Estimation

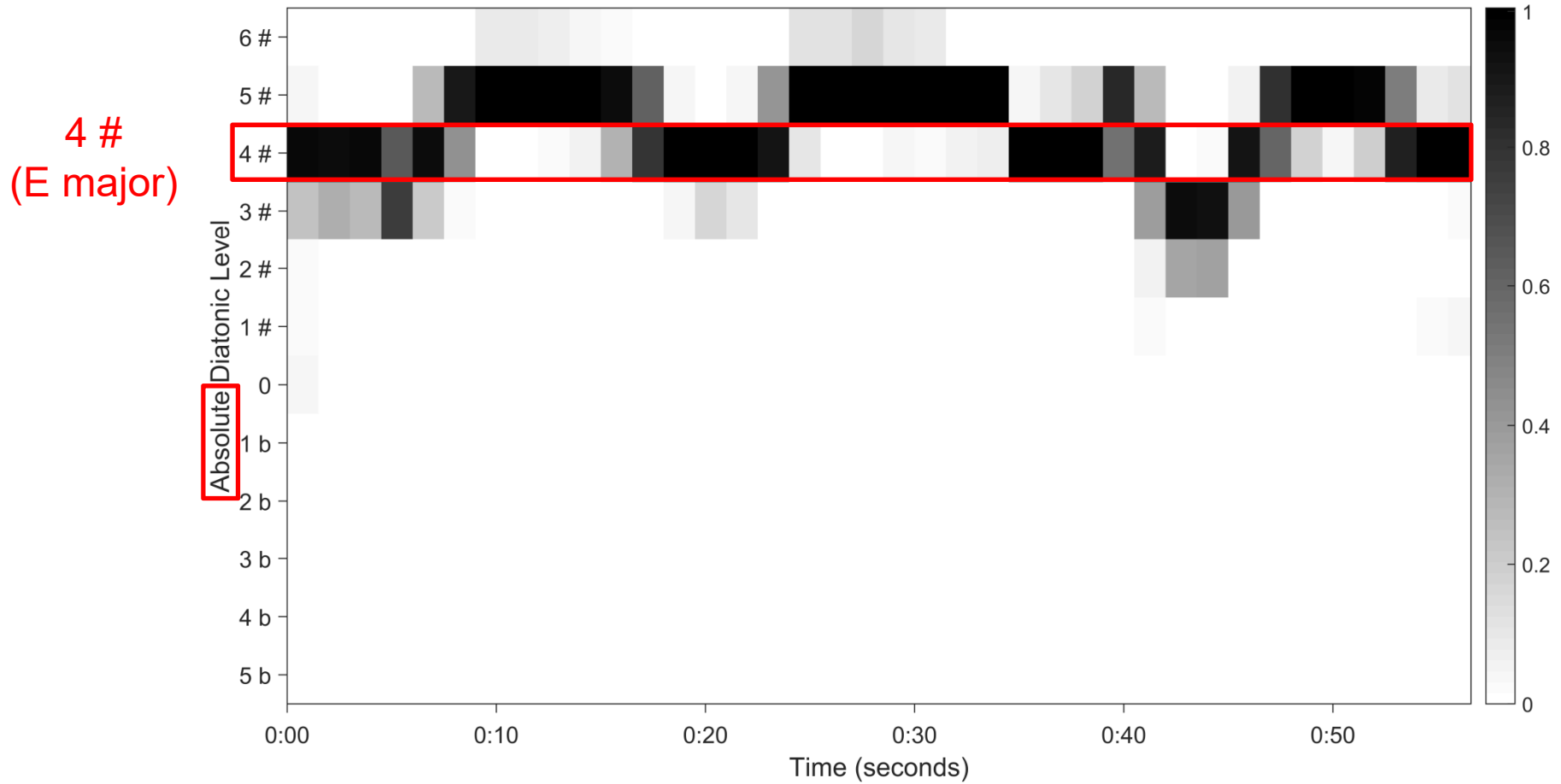
Summarize pitch class content according to **diatonic scales**

Multiply chroma values



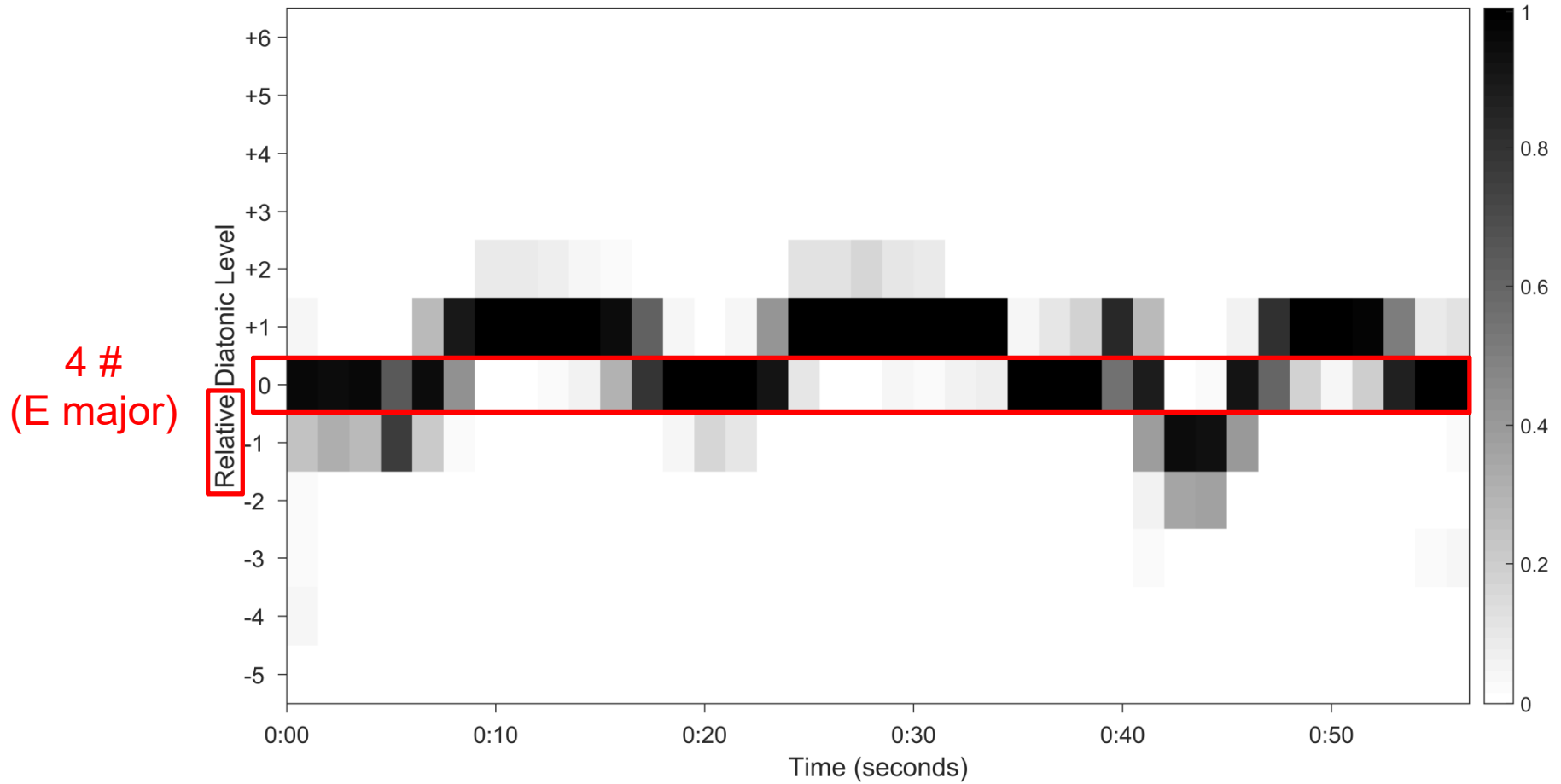
Local Key Estimation

Normalize representation relative to **global key**



Local Key Estimation

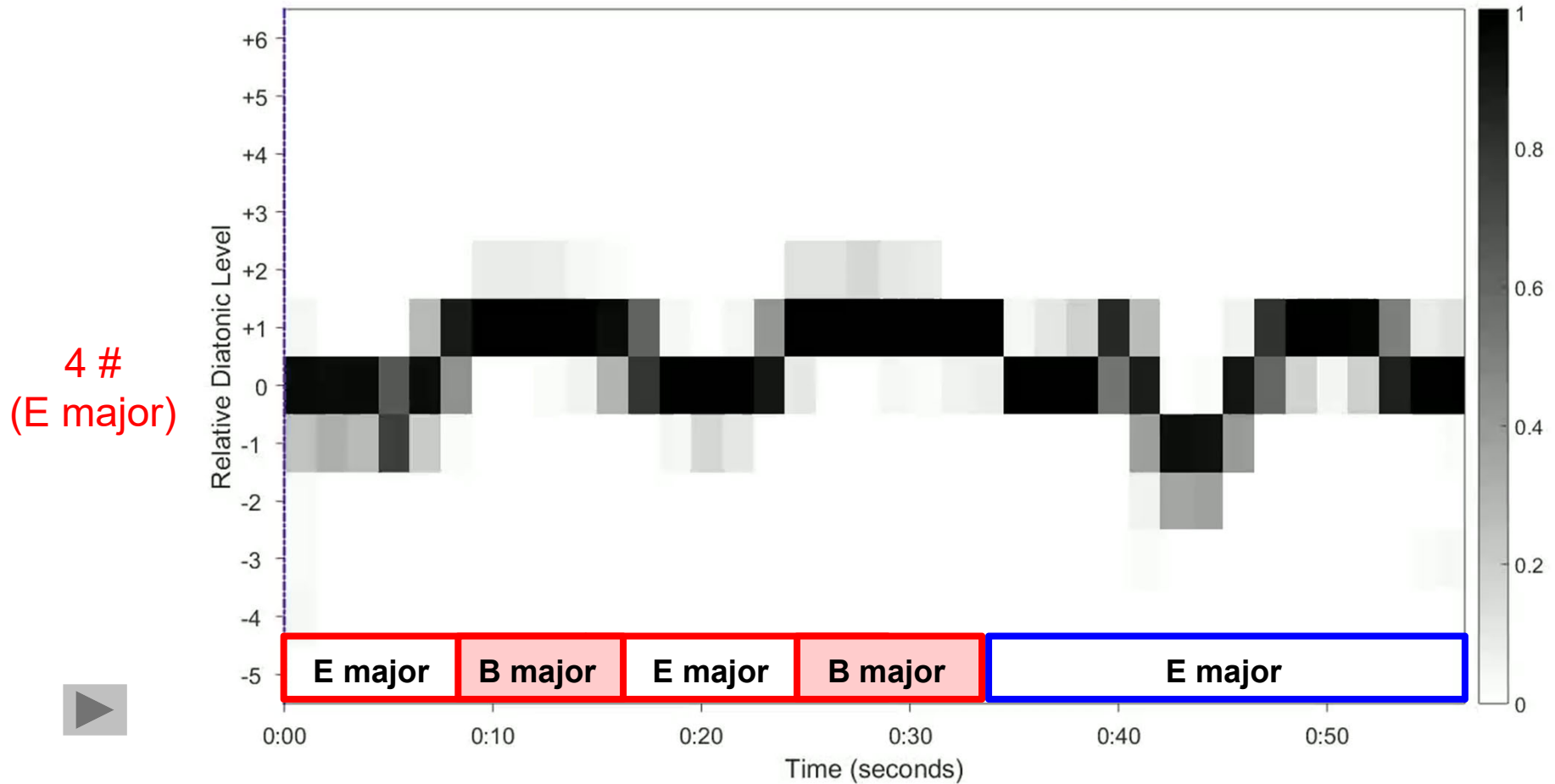
Normalize representation relative to **global key**



Local Key Estimation

J.S. Bach: Choral "Durch Dein Gefängnis" (*Johannespassion*)

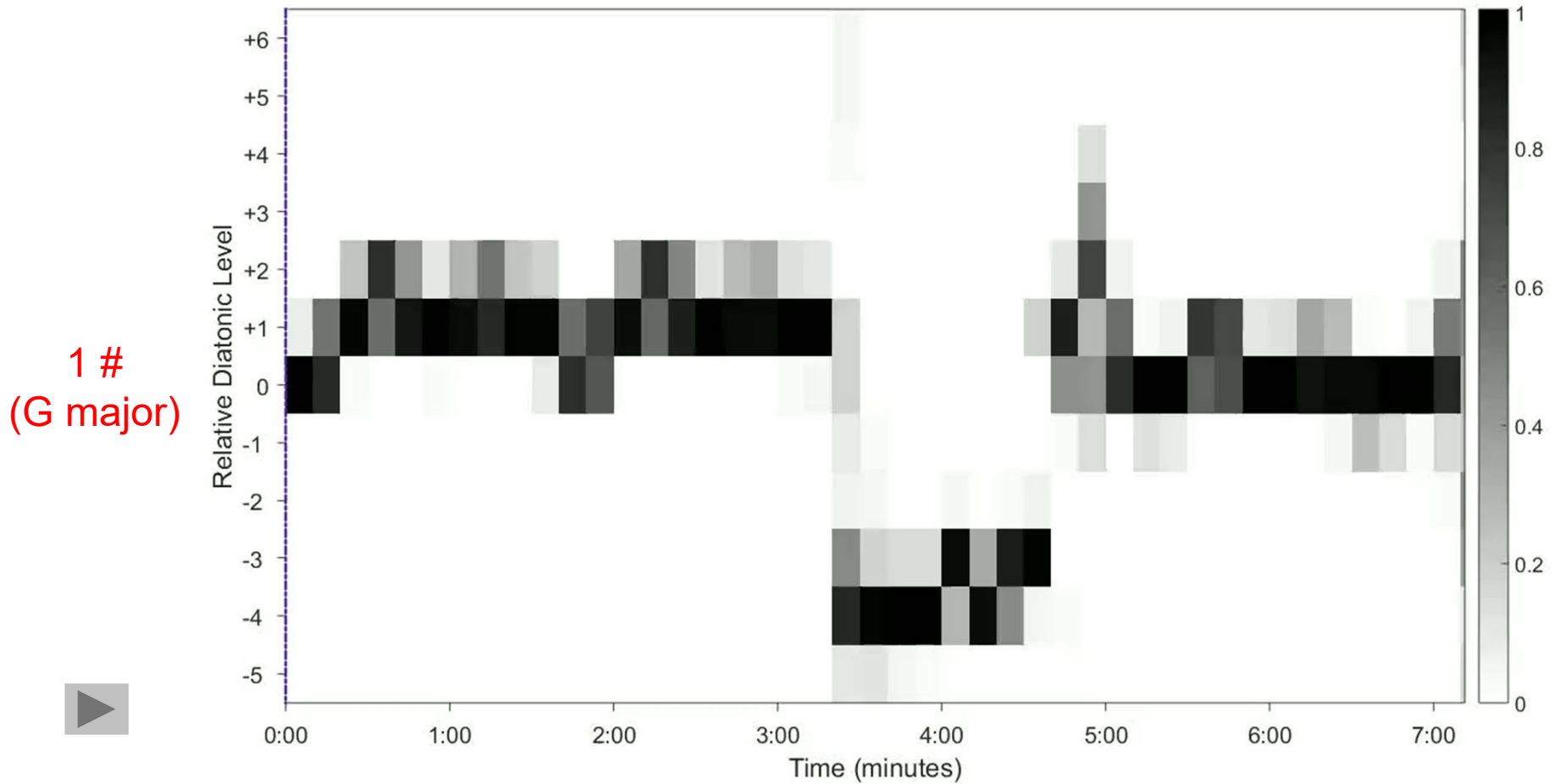
Recording: Scholars Baroque Ensemble, Naxos 1994



Local Key Estimation

L. v. Beethoven: Piano Sonata No. 10 (Op. 14 Nr. 2), 1. Allegro

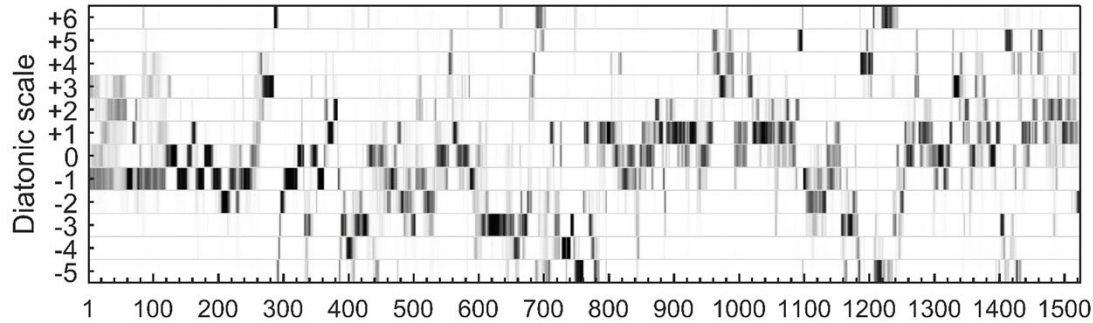
Recording: Barenboim, EMI 1998



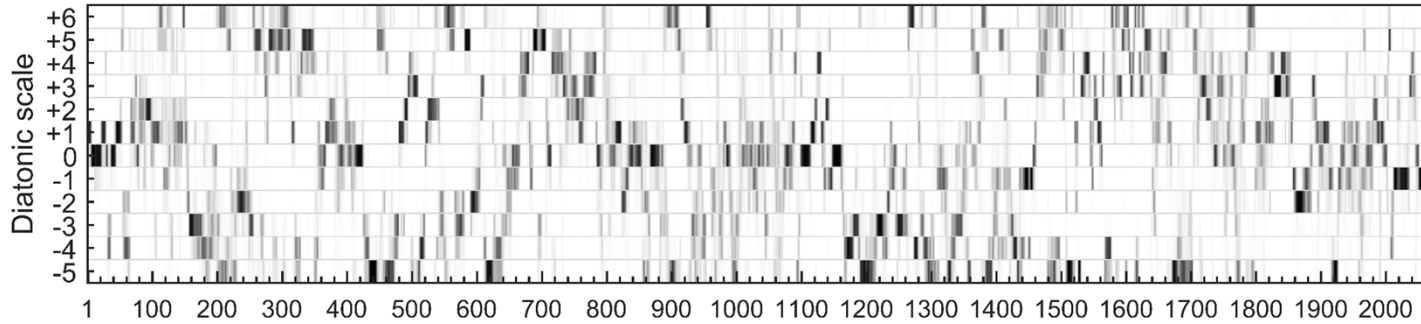
Local Key Estimation

R. Wagner: WWV 86 B (*Die Walküre*)

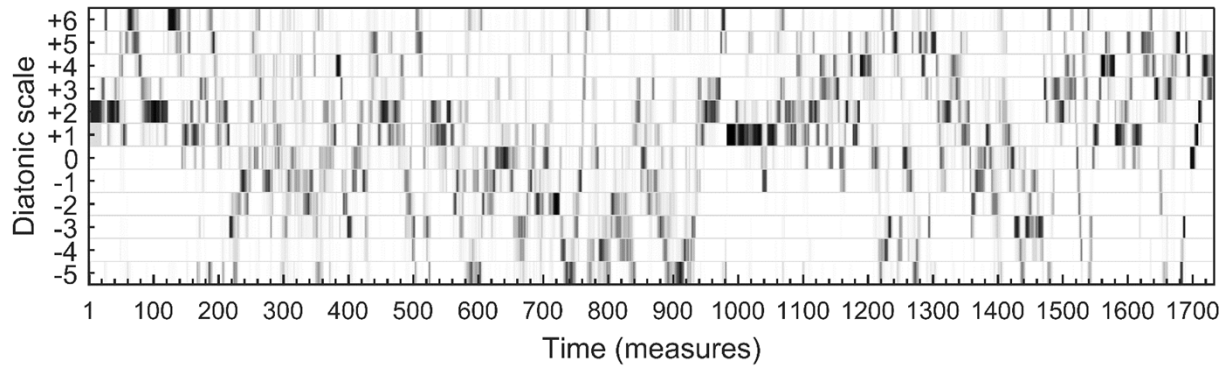
Act 1



Act 2



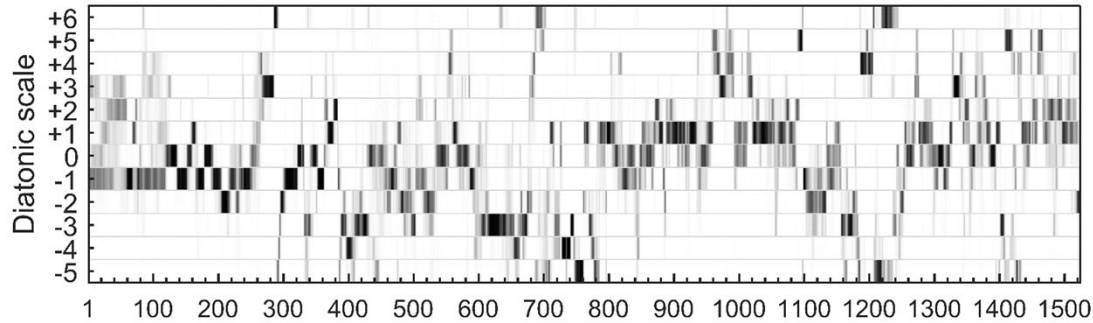
Act 3



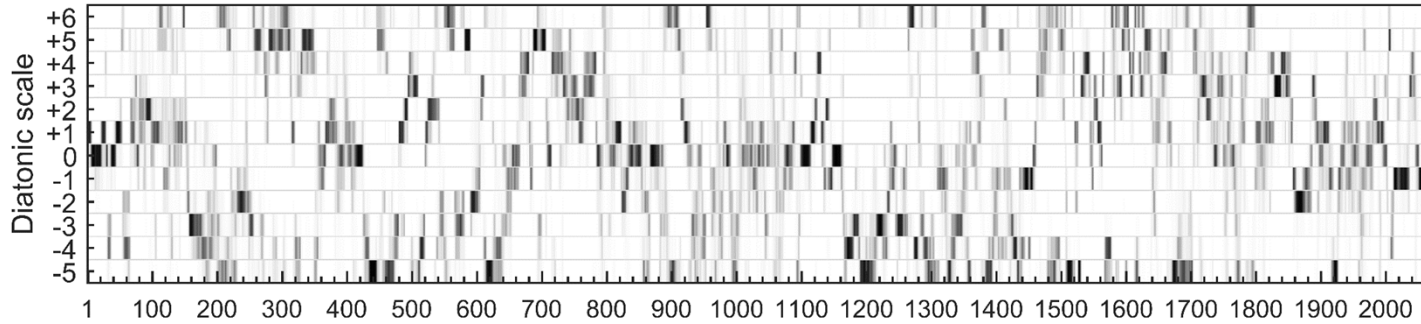
Local Key Estimation

R. Wagner: WWV 86 B (*Die Walküre*)

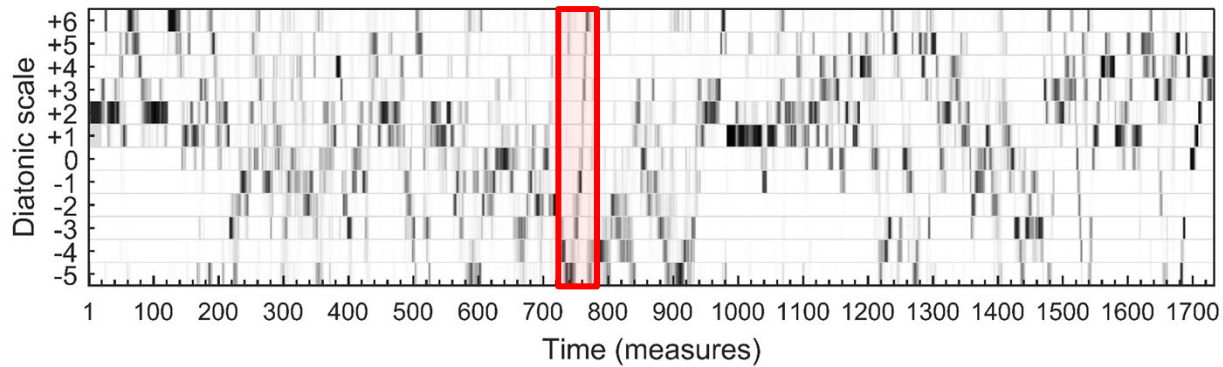
Act 1



Act 2



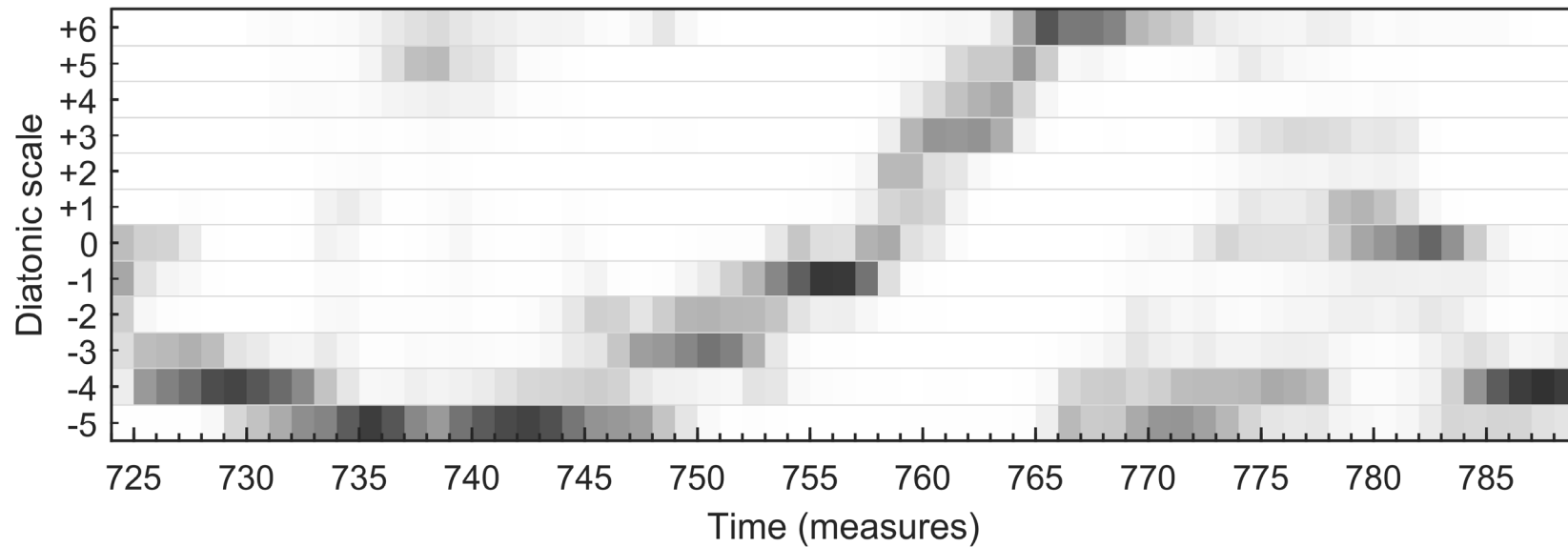
Act 3



Local Key Estimation

R. Wagner: WWV 86 B (*Die Walküre*)

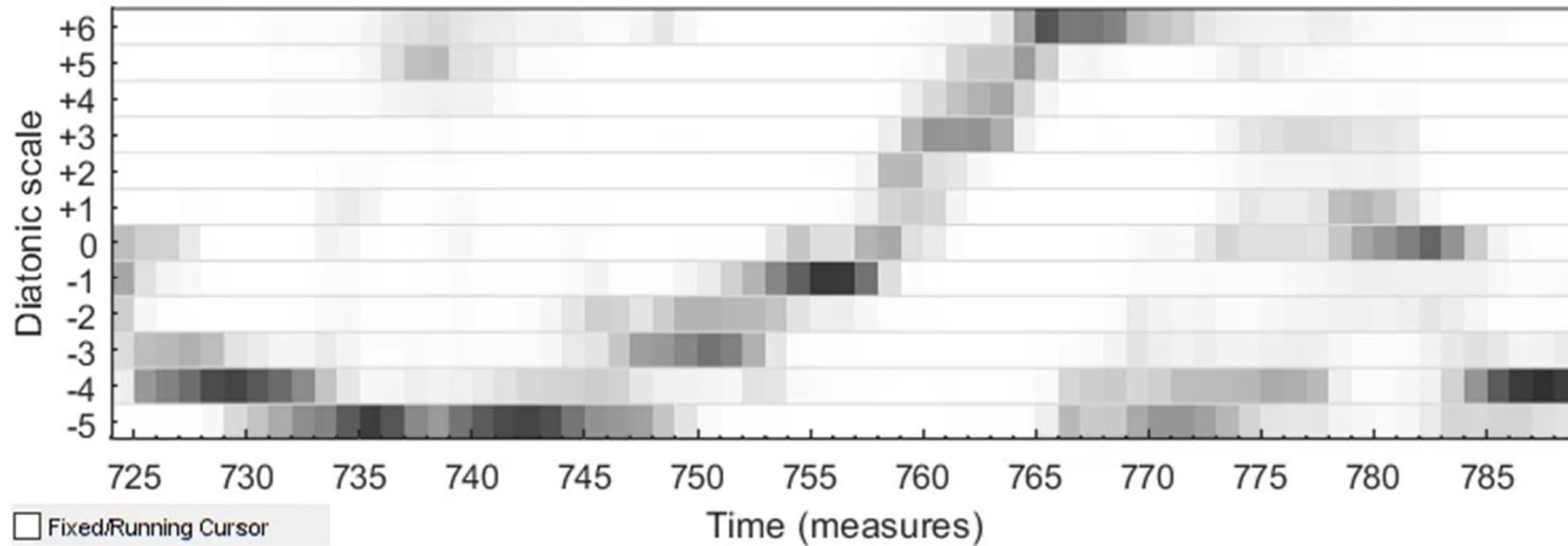
Act 3, measure 724–789 (*Wotan's punishment*)



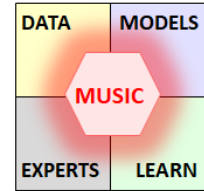
Local Key Estimation

R. Wagner: WWV 86 B (*Die Walküre*)

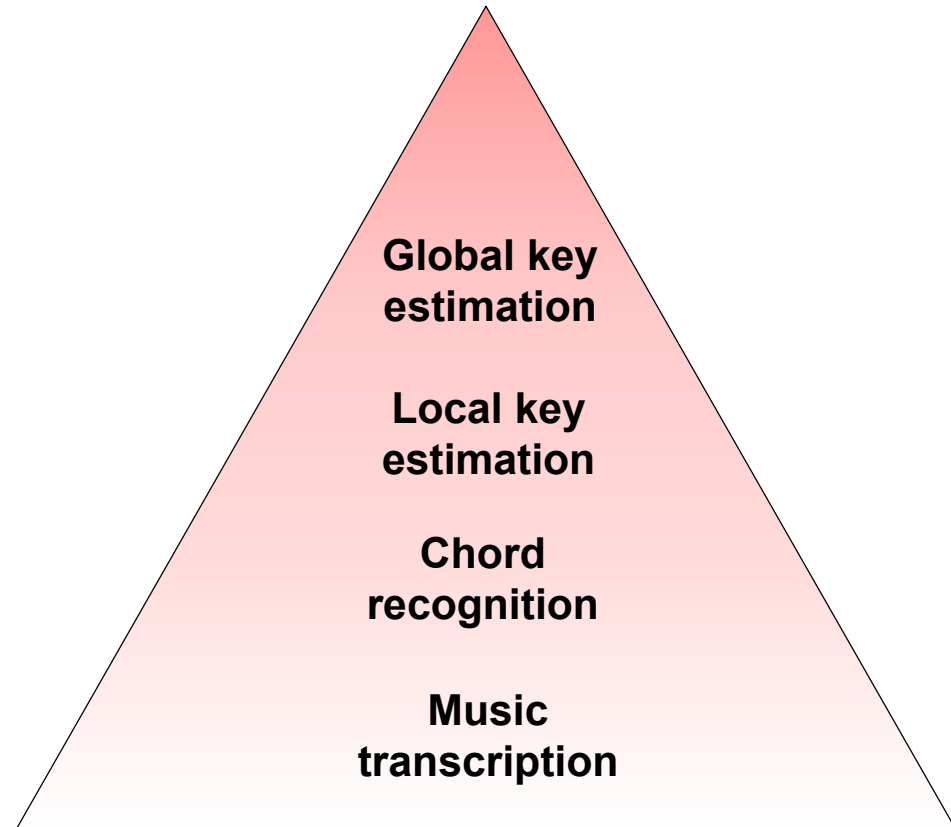
Act 3, measure 724–789 (*Wotan's punishment*)



Reinhart Koselleck-Projekt: LEARN

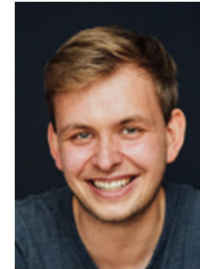


- Interpretability & explainability
- Knowledge integration
- Hybrid models
- Multitask learning
- Hierarchical approaches
- ...



Computational Ethnomusicology: Traditional Georgian Vocal Music

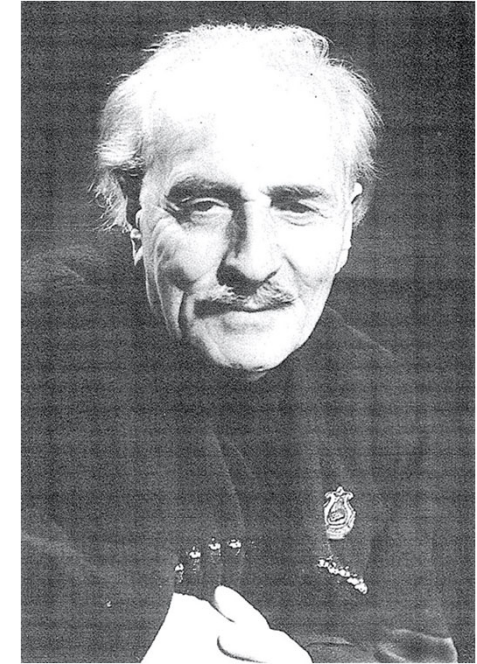
- Interdisciplinary research project
 - Prof. Dr. Frank Scherbaum (Potsdam)
 - Dr. Nana Mzhavanadze (Tbilisi)
 - Sebastian Rosenzweig (FAU)
- Objective: Tonal analysis
- 2018 – 2022: DFG-funded project



Traditional Georgian Vocal Music

Example: Erkomaishvili corpus

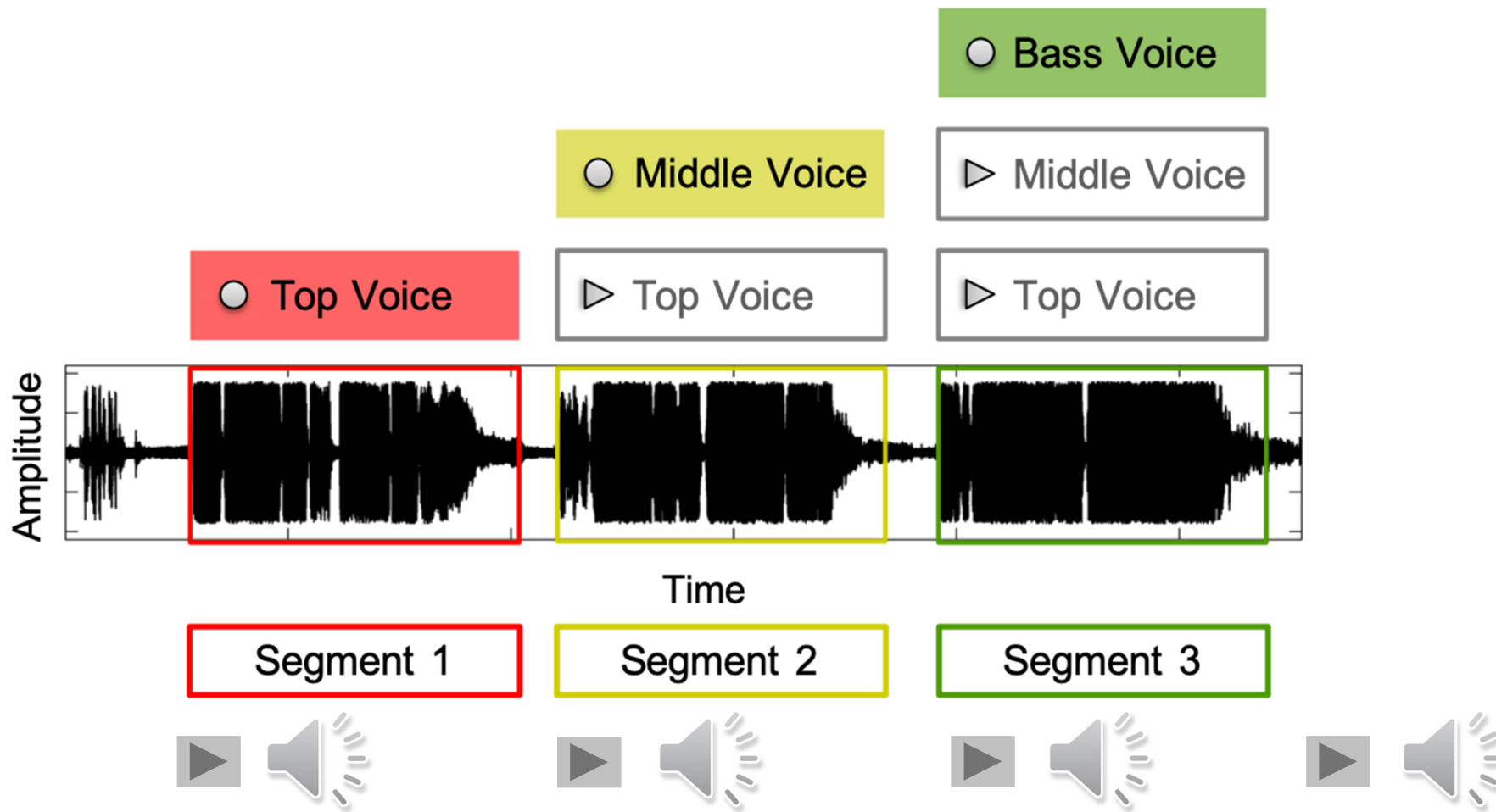
- Collection of traditional three-voice Georgian songs
- Performed by the former Georgian master chanter Artem Erkomaishvili (1887-1967)
- Recordings of 100 songs using tape recorders (1966)



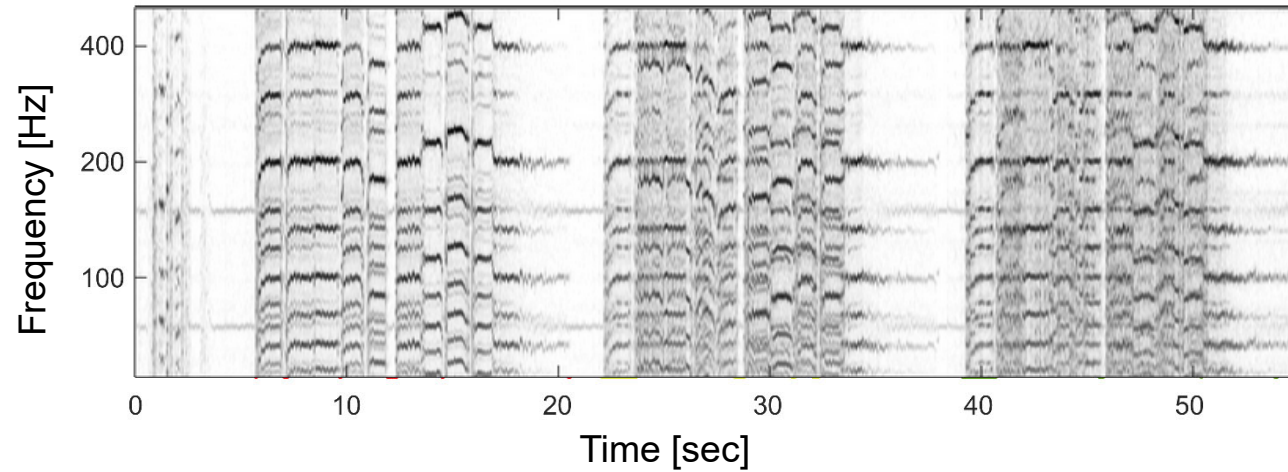
“Original masterpieces of Georgian musical thinking.” (Shugliashvili, 2014)

Traditional Georgian Vocal Music

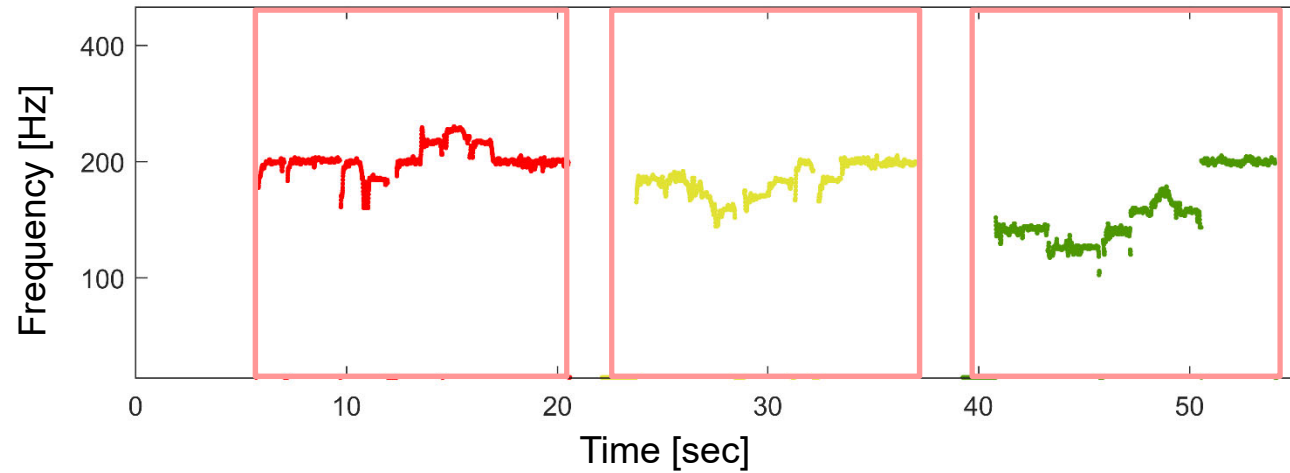
Example: Erkomaishvili corpus



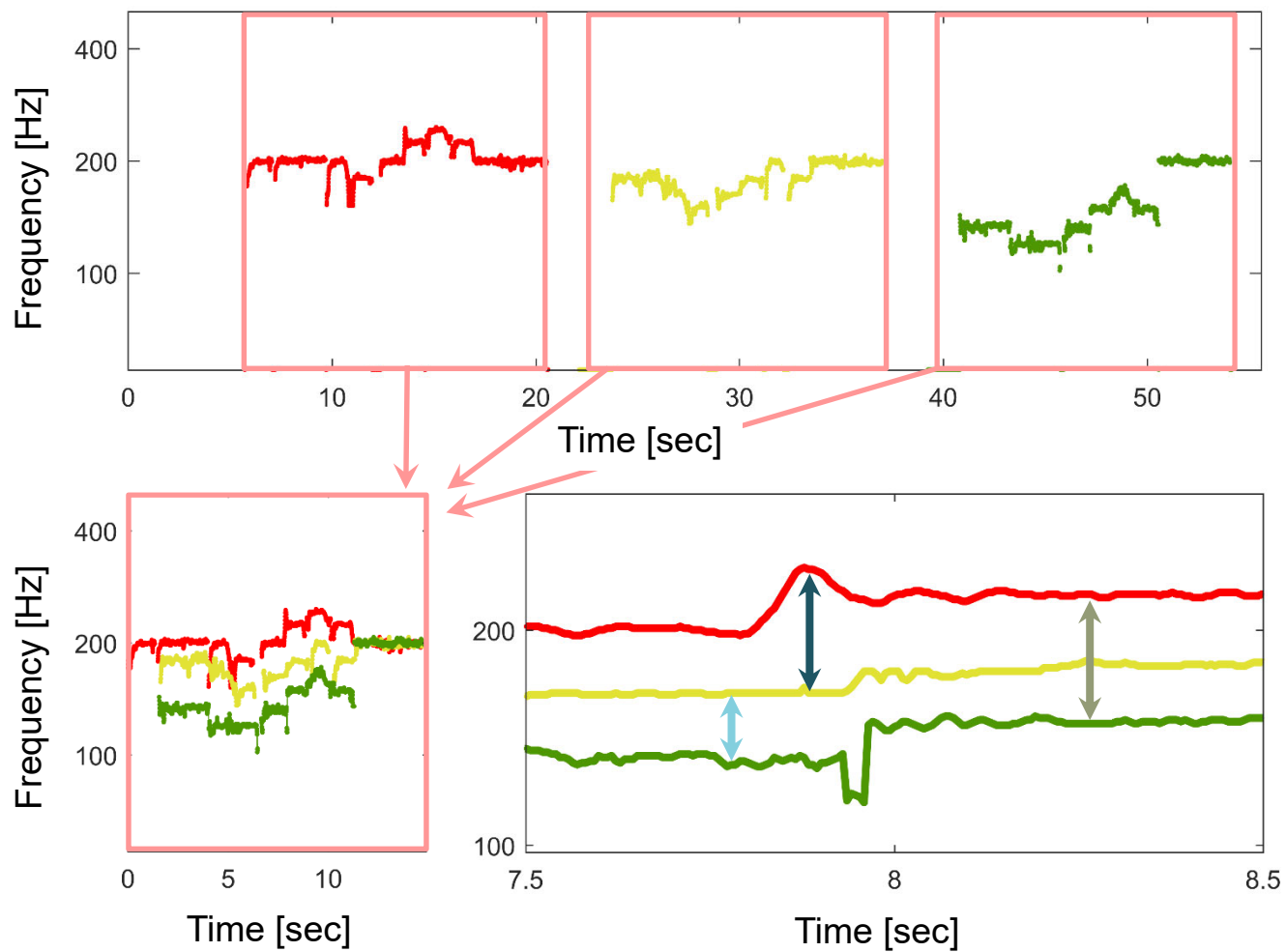
Traditional Georgian Vocal Music



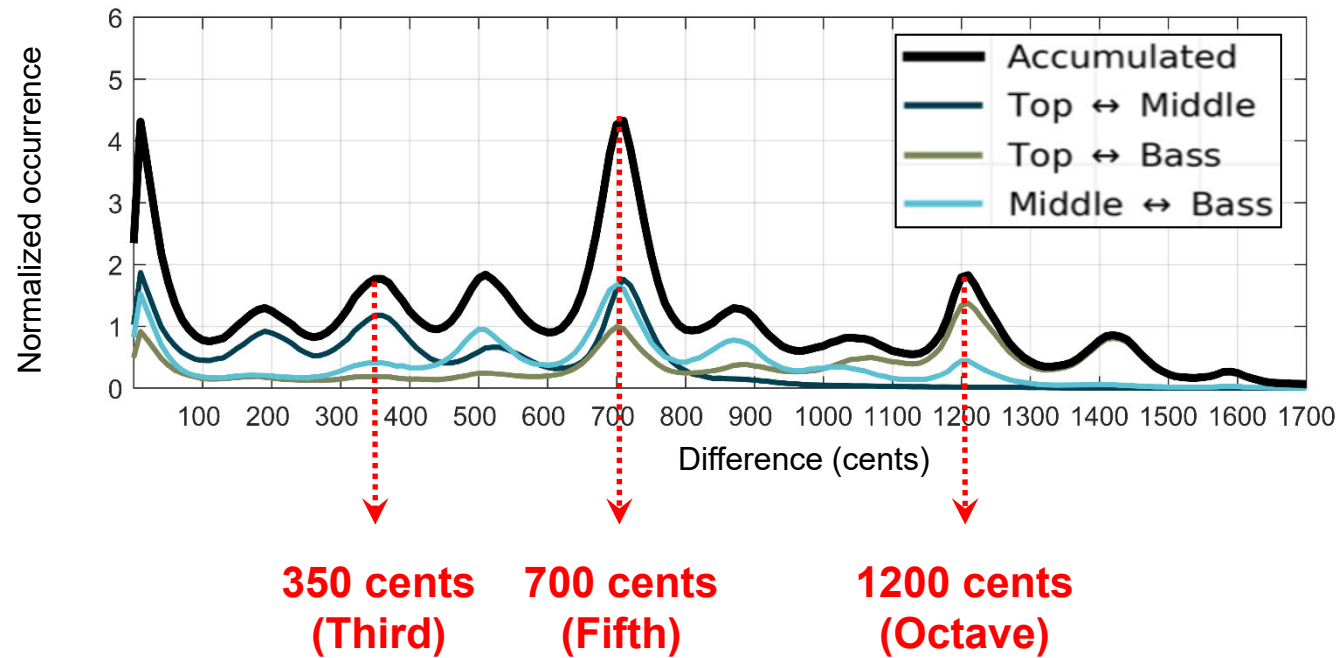
Traditional Georgian Vocal Music



Traditional Georgian Vocal Music



Traditional Georgian Vocal Music



- Peak at **350 cents** (between minor and major third)
- **Non-western temperament**

Traditional Georgian Vocal Music

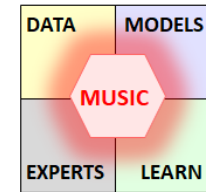


- Recordings from field expedition in 2016
- 216 performances
- Multitrack audio + video
 - Room, **HSM**, **LRX**
- Total duration: 6 h

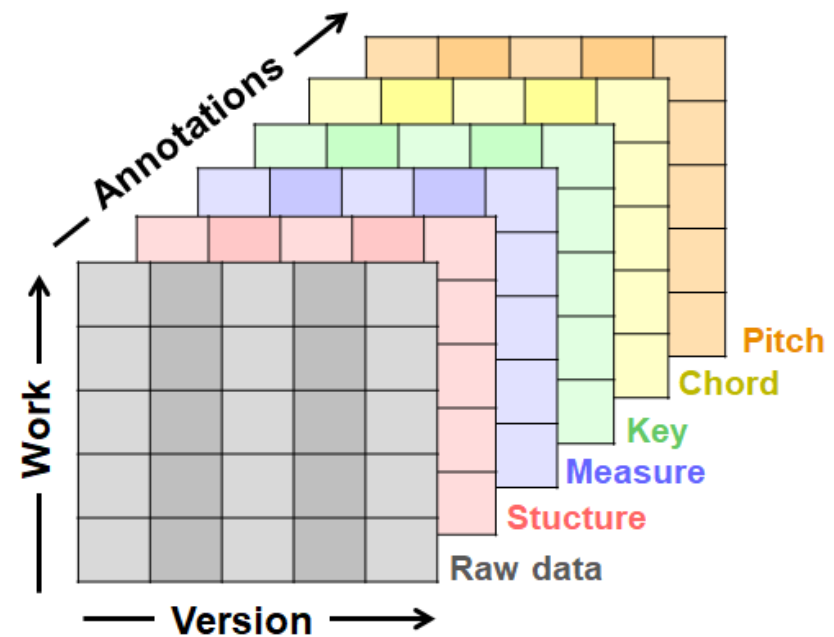


Room
Microphone

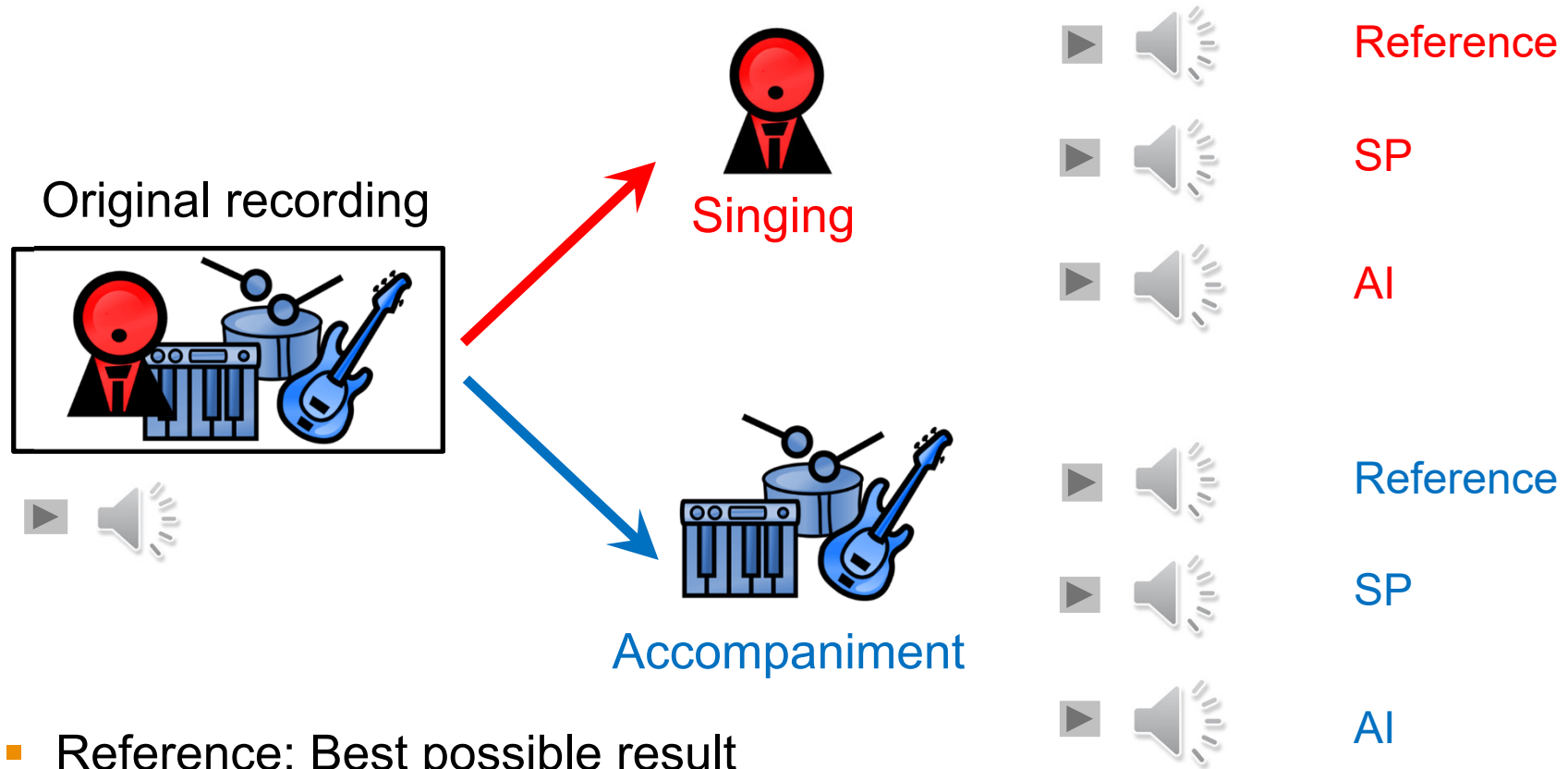
Reinhart Koselleck-Projekt: LEARN



- Non-standard datasets
 - Variety of music
 - Poor audio quality
 - Various sensor types
- Exploring DL models
 - Generalization
 - Overfitting
 - Data scarceness



Source Separation (Singing)



- Reference: Best possible result
- SP: Using traditional signal processing
- AI: Using data-driven approach

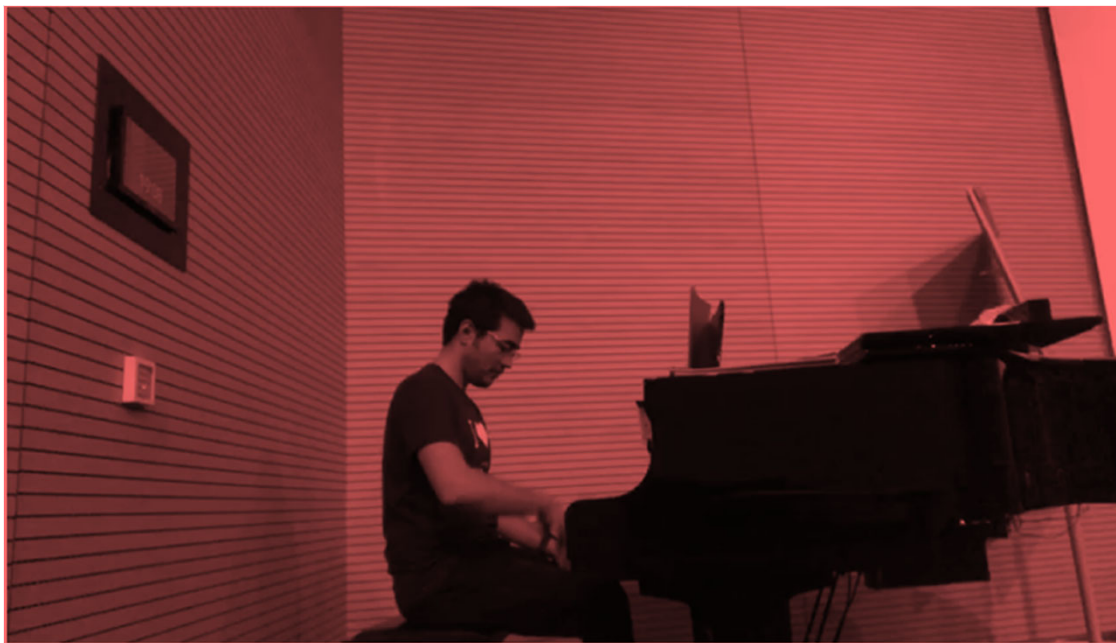
Source Separation (Piano Concerto)

- Yigitcan Özer
- PhD student in engineering
- Pianist



Source Separation (Piano Concerto)

- Yigitcan Özer
- PhD student in engineering
- Pianist



Only Piano!



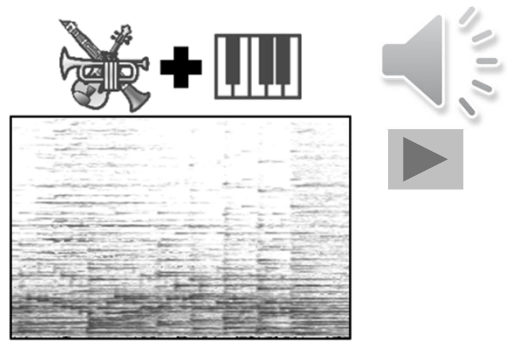
**Where is the
orchestra?**



Source Separation (Piano Concerto)



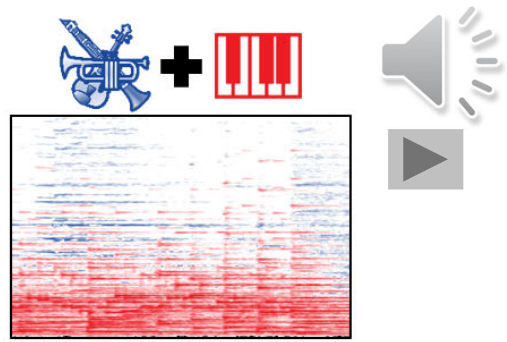
A musical score for a piano concerto, starting at measure 89. The score is written for a full orchestra and piano. The instruments are represented by icons on the left side of the score: woodwinds (flute, oboe, clarinet, bassoon), brass (trumpet, trombone, horn, tuba), strings (violin, viola, cello, double bass), and piano. The piano part is the most prominent, featuring a complex, fast-moving melody in the right hand and a steady accompaniment in the left hand. The score is in 2/4 time and the key signature has one flat.



A diagram illustrating the source separation process. It shows a plus sign between a trumpet icon and a piano keyboard icon, indicating the combination of these two sources. Below this, a spectrogram shows the frequency content of the combined signal. To the right of the spectrogram are a speaker icon and a play button icon, suggesting the output is an audio file that can be played.

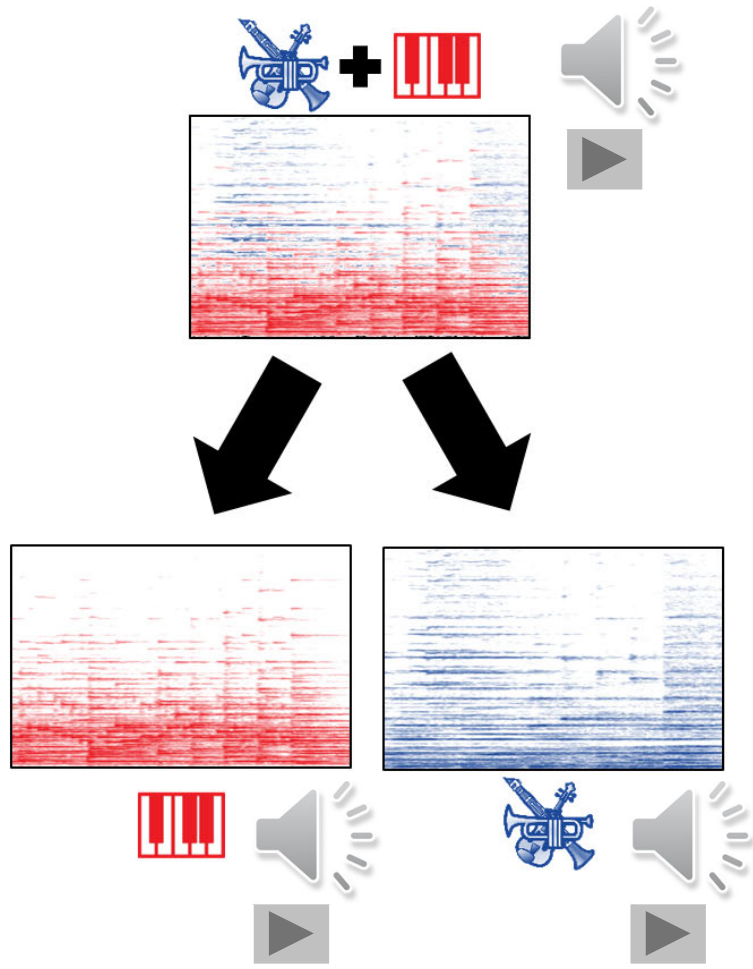
Source Separation (Piano Concerto)

A musical score for a piano concerto, starting at measure 89. The score is divided into two main sections. The upper section, in blue, includes staves for woodwinds (flute, oboe, clarinet, bassoon), brass (trumpet, trombone, horn, tuba), and strings (violin, viola, cello, double bass). The lower section, in red, is the piano part, featuring a piano keyboard icon to its left. The piano part shows a complex melodic line in the right hand and a supporting bass line in the left hand.

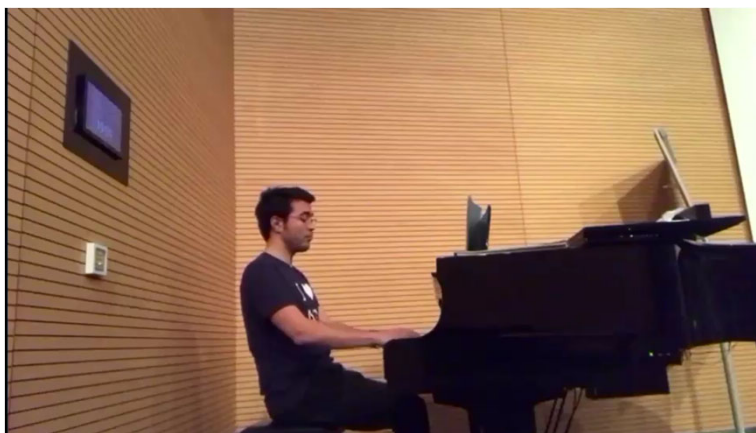
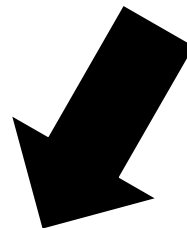
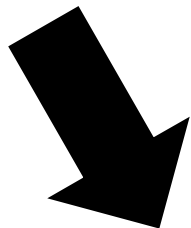
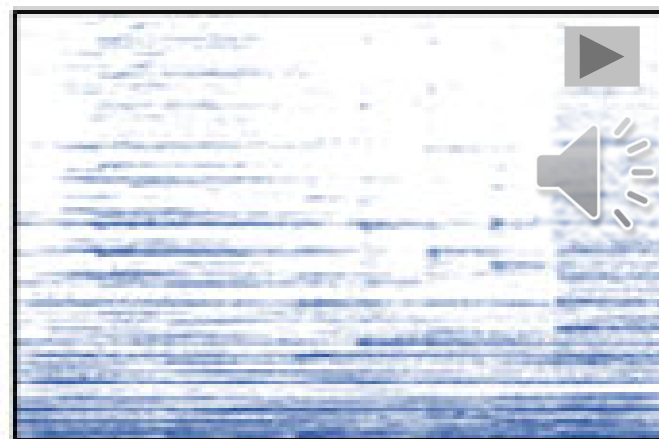
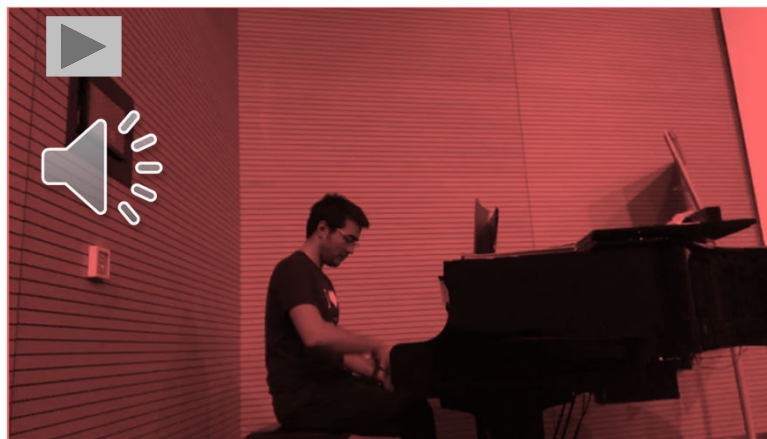


Source Separation (Piano Concerto)

A musical score for a piano concerto, starting at measure 89. The score includes staves for woodwinds (flute, oboe, clarinet, bassoon, horn, trumpet, trombone), strings (violin I, violin II, viola, cello, double bass), and piano. The piano part is highlighted in red, showing a complex melodic line. A red piano keyboard icon is placed to the left of the piano staves.

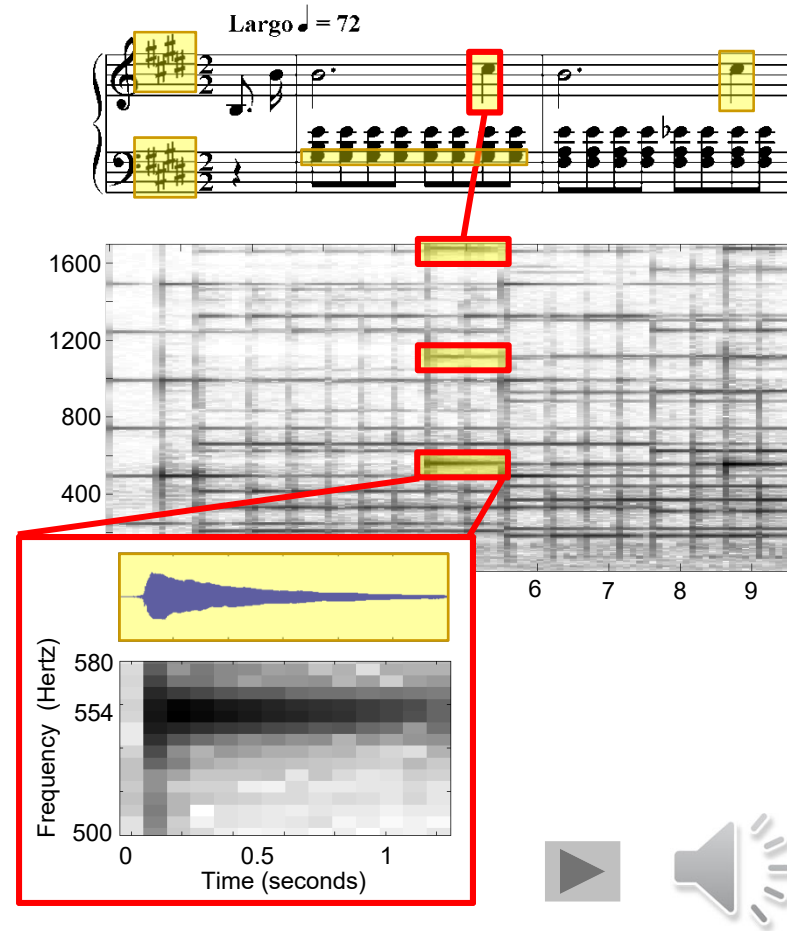
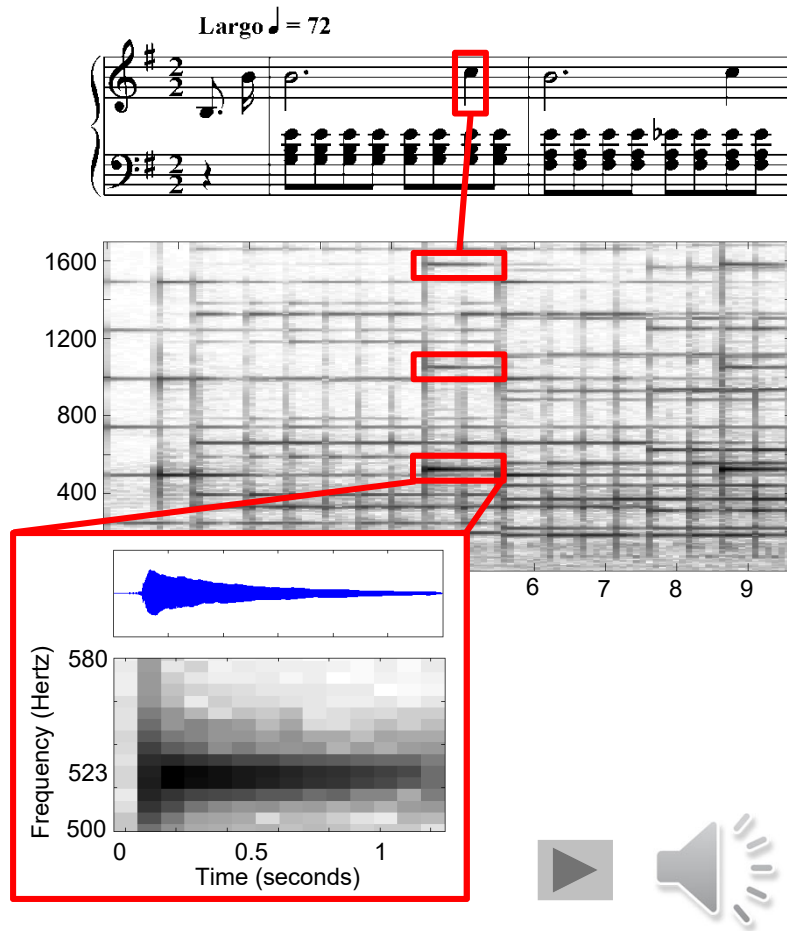


Source Separation (Piano Concerto)



Source Separation

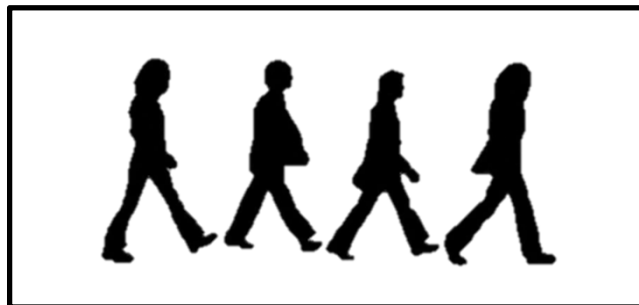
Score-informed audio editing



Source Separation

Audio mosaicing (style transfer)

Target signal: Beatles–Let it be



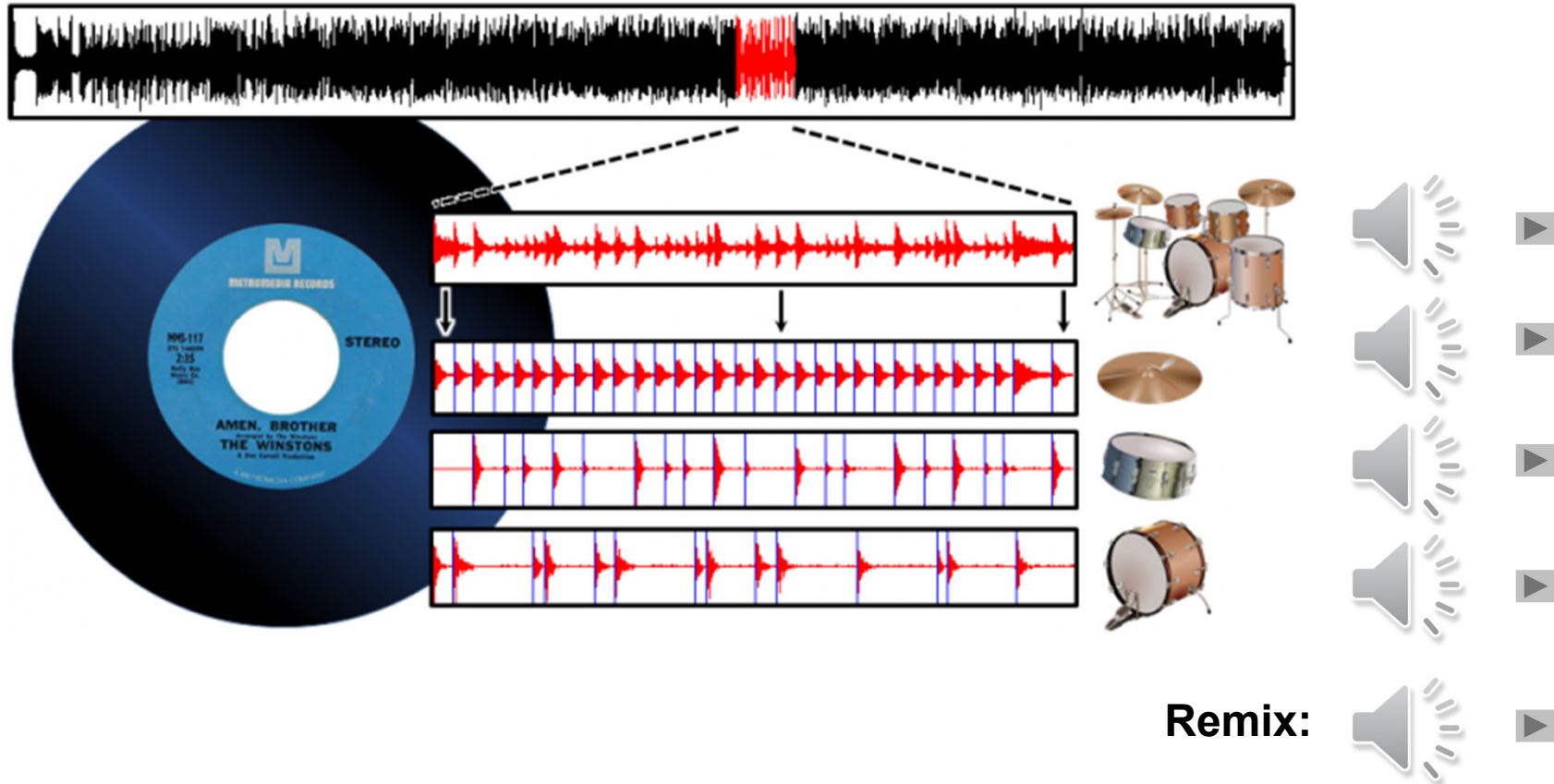
Source signal: Bees



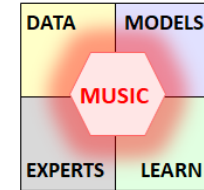
Mosaic signal: **Let it Bee**

Source Separation

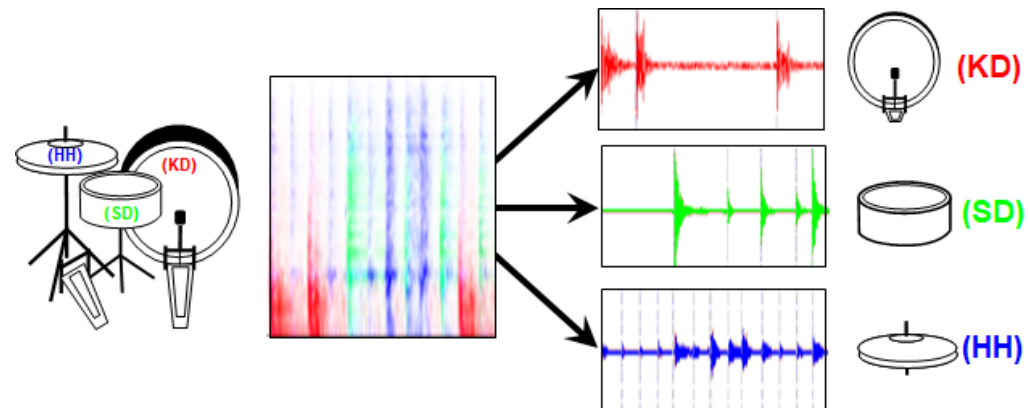
Informed Drum-Sound Decomposition



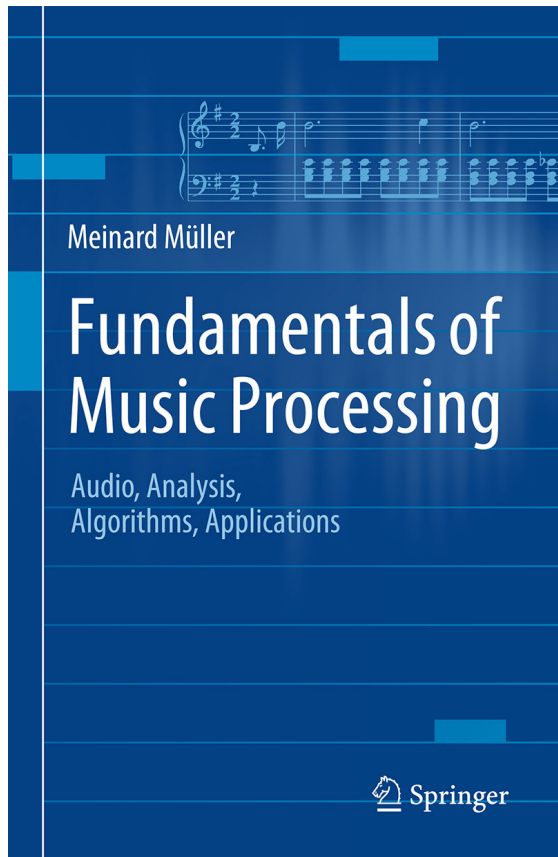
Reinhart Koselleck-Projekt: LEARN



- Reconstruction of sources
- Generative models
- Differentiable DSP
- Analysis by synthesis
- ...



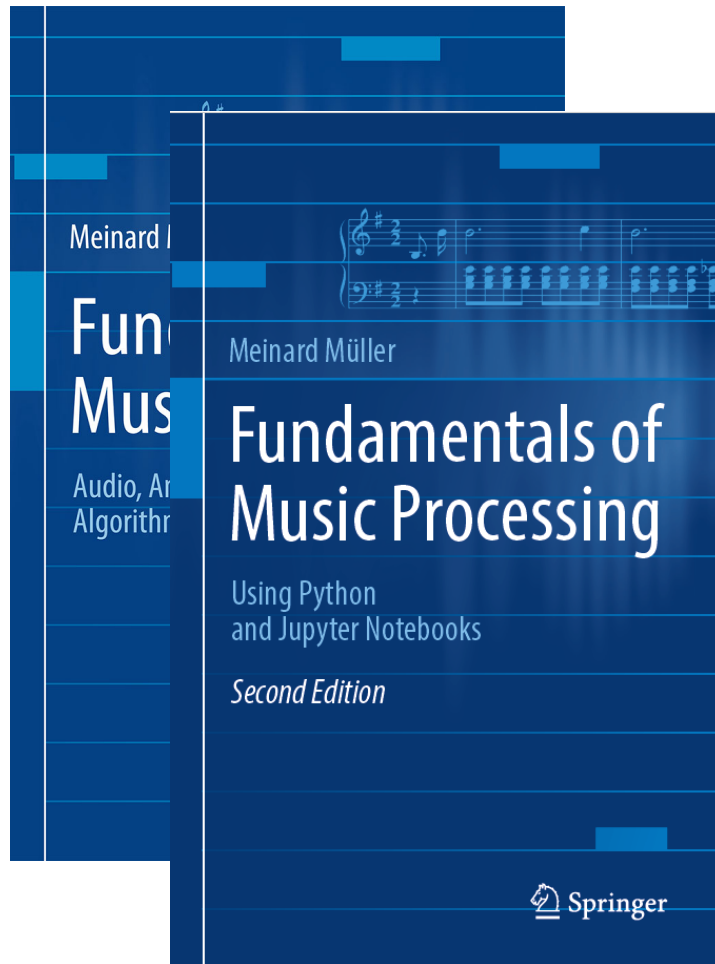
Fundamentals of Music Processing (FMP)



Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
Springer, 2015

Accompanying website:
www.music-processing.de

Fundamentals of Music Processing (FMP)

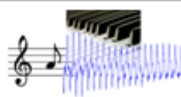

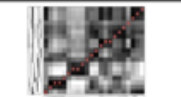
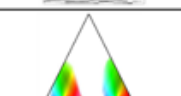

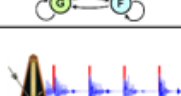




Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
Springer, 2015

Accompanying website:
www.music-processing.de

2nd edition
Meinard Müller
Fundamentals of Music Processing
Using Python and Jupyter Notebooks
Springer, 2021

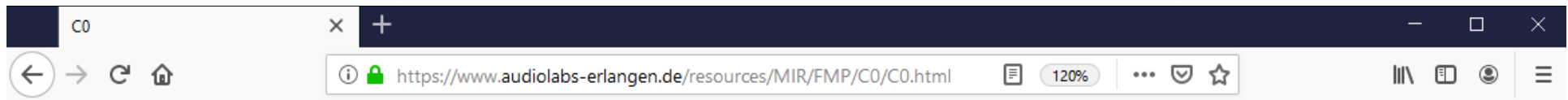
Fundamentals of Music Processing (FMP)

Chapter	Music Processing Scenario
1	 Music Representations
2	 Fourier Analysis of Signals
3	 Music Synchronization
4	 Music Structure Analysis
5	 Chord Recognition
6	 Tempo and Beat Tracking
7	 Content-Based Audio Retrieval
8	 Musically Informed Audio Decomposition

Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
Springer, 2015

Accompanying website:
www.music-processing.de

2nd edition
Meinard Müller
Fundamentals of Music Processing
Using Python and Jupyter Notebooks
Springer, 2021

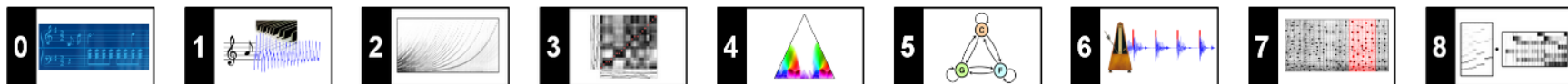


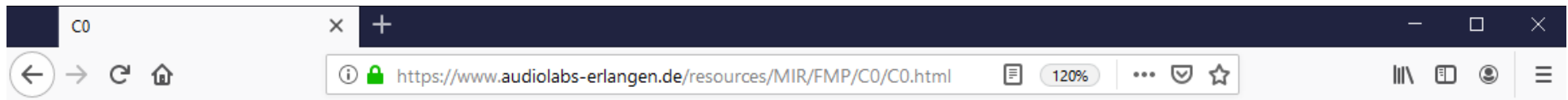
FMP Notebooks

Python Notebooks for Fundamentals of Music Processing



<https://www.audiolabs-erlangen.de/FMP>



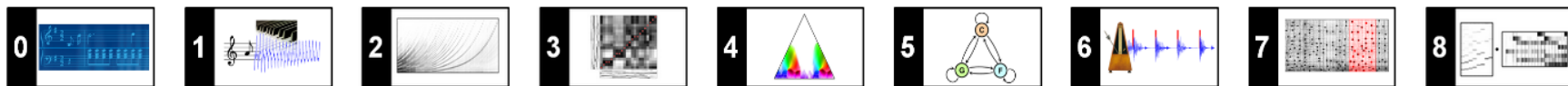


FMP Notebooks

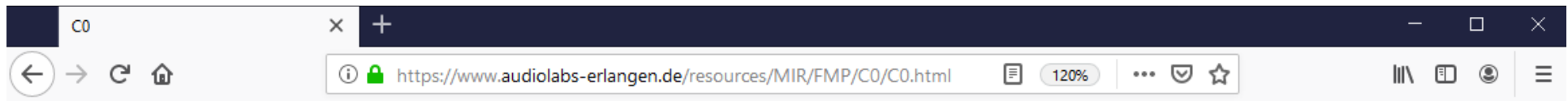
Python Notebooks for Fundamentals of Music Processing



<https://www.audiolabs-erlangen.de/FMP>



Basics + 8 Chapters

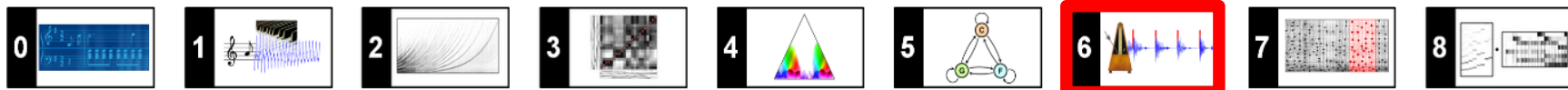


FMP Notebooks

Python Notebooks for Fundamentals of Music Processing

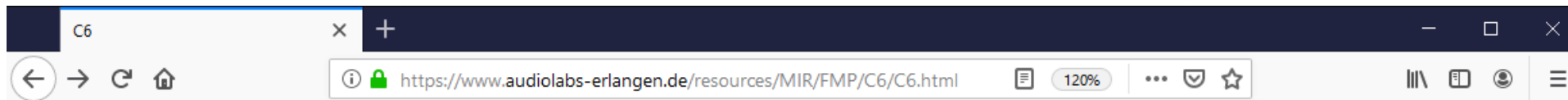


<https://www.audiolabs-erlangen.de/FMP>

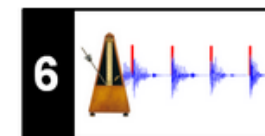


Basics + 8 Chapters

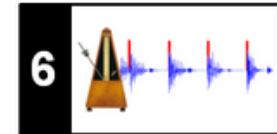
Tempo and Beat Tracking



Tempo and Beat Tracking



Tempo and Beat Tracking

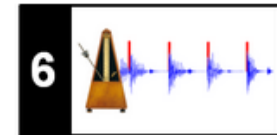


Definition

We assume that we are given a discrete-time novelty function $\Delta : \mathbb{Z} \rightarrow \mathbb{I}$ indicate note onset candidates. The idea of Fourier analysis is to detect local periodicity in novelty curve by comparing it with windowed sinusoids. A high correlation of Δ with a windowed sinusoid indicates a periodicity of the novelty curve (given a suitable phase). This correlation (along with the phase) can be computed via the short-time Fourier transform. To this end, we fix a window function $w : \mathbb{Z} \rightarrow \mathbb{R}$ of length centered at $n = 0$ (e.g., a sampled Hann window). Then, for a frequency parameter $\omega \in \mathbb{R}_{\geq 0}$ and time parameter $n \in \mathbb{Z}$, the complex Fourier coefficient is defined by

$$\mathcal{F}(n, \omega) := \sum_{m \in \mathbb{Z}} \Delta(m) \bar{w}(m - n) \exp(-2\pi i \omega m).$$

Tempo and Beat Tracking



Definition

We assume that we are given a discrete-time novelty function $\Delta : \mathbb{Z} \rightarrow \mathbb{I}$ indicate note onset and in novelty curve by computing the correlation of Δ with a window function w (given a suitable phase). This correlation (along with the phase) is analyzed using the short-time Fourier transform. To this end, we fix a window function w (e.g., a sampled Hann window). Then, for a time parameter $n \in \mathbb{Z}$, the complex Fourier coefficient is defined by

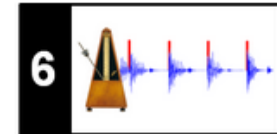
Explanations

Theory

Mathematics

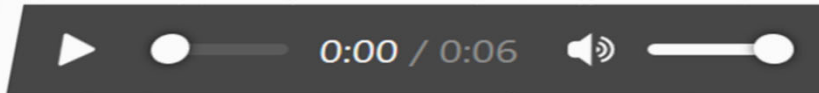
$$\mathcal{F}(n, \omega) := \sum_{m \in \mathbb{Z}} \Delta(m) \bar{w}(m - n) \exp(-2\pi i \omega m).$$

Tempo and Beat Tracking



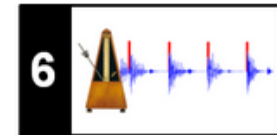
Example: Shostakovich

In the following example, we consider an excerpt of a recording of Dimitri Shostakovich's Suite for Variety Orchestra No. 1. The score version of the excerpt.



We start with a [spectral-based novelty function](#) resampled to F_s^Δ :
Furthermore, we use a window size corresponding to 5 seconds (λ)

Tempo and Beat Tracking

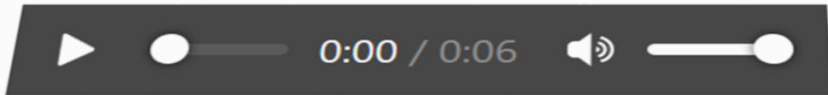


Example: Shostakovich

In the following example, we consider an excerpt from the Suite for Variety Orchestra.

Music Examples

Annotations

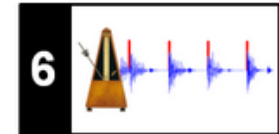


Audio

We start with a [spectral-based novelty function](#) resampled to F_s^Δ . Furthermore, we use a window size corresponding to 5 seconds (Δ).

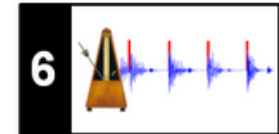
Links

Tempo and Beat Tracking



```
In [2]: def compute_sinusoid_optimal(c, tempo, n, Fs, N
        """Compute windowed sinusoid with optimal p
        Notebook: C6/C6S2_TempogramFourier.ipynb
        Args:
            c: Coefficient of tempogram (c=X(k,n))
            tempo: Tempo parameter corresponding to
            _coef_BPM[k])
            n: Frame parameter of c
            Fs: Sampling rate
            N: Window length
            H: Hop size
```


Tempo and Beat Tracking



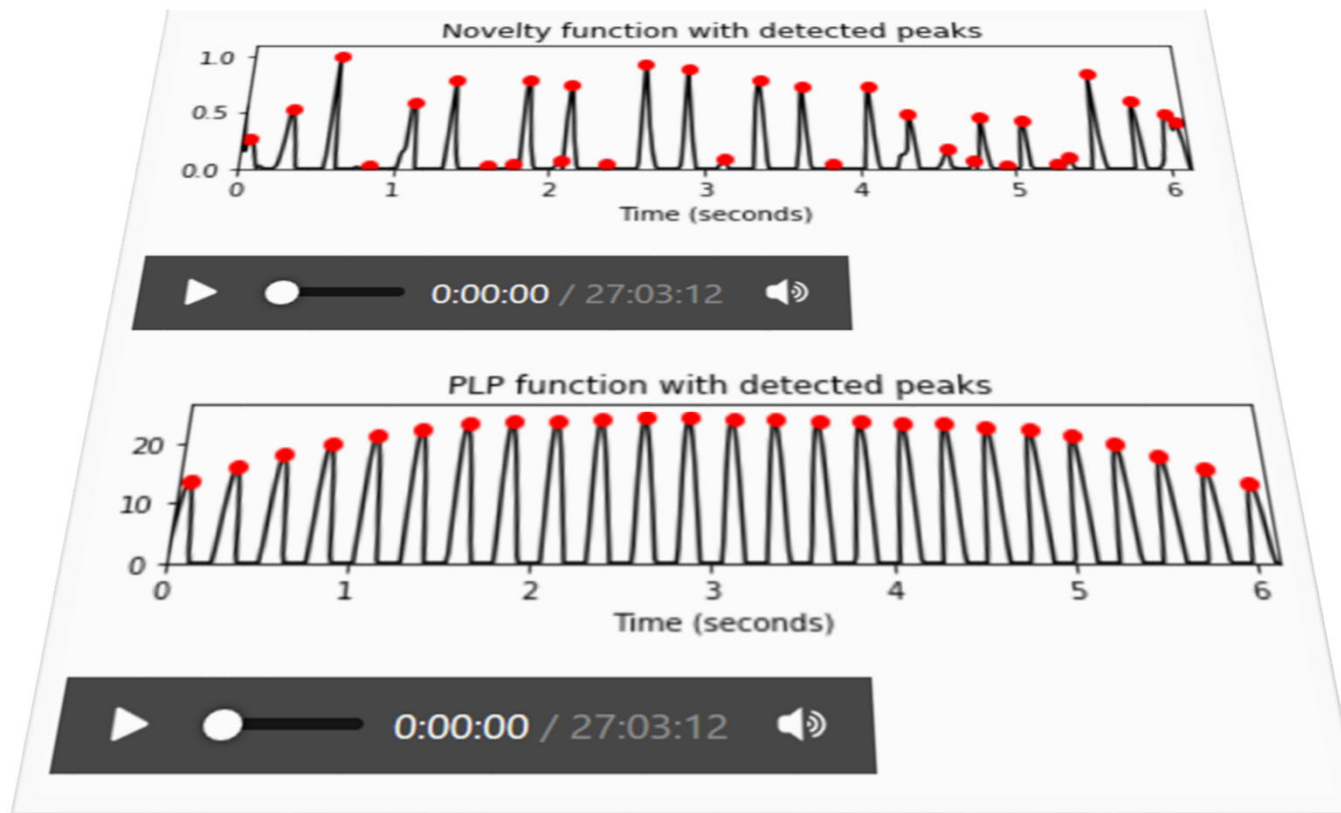
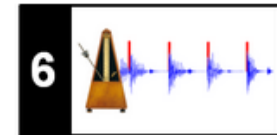
```
In [2]: def compute_sinusoid_optimal_p(n, Fs, N, H):  
        """Compute windowed sinusoid coefficients for optimal p  
        """  
        k = C6/C6S2... .ipynb  
  
        c: Coefficient of tempogram (c=X(k,n))  
        tempo: Tempo parameter corresponding to  
        _coef_BPM[k])  
        n: Frame parameter of c  
        Fs: Sampling rate  
        N: Window length  
        H: Hop size
```

Python Code

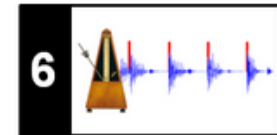
Algorithms

Functions

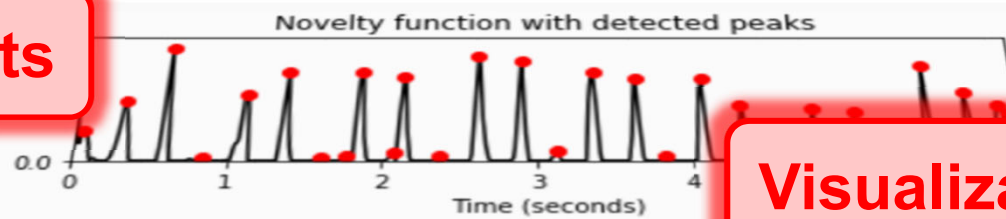
Tempo and Beat Tracking



Tempo and Beat Tracking

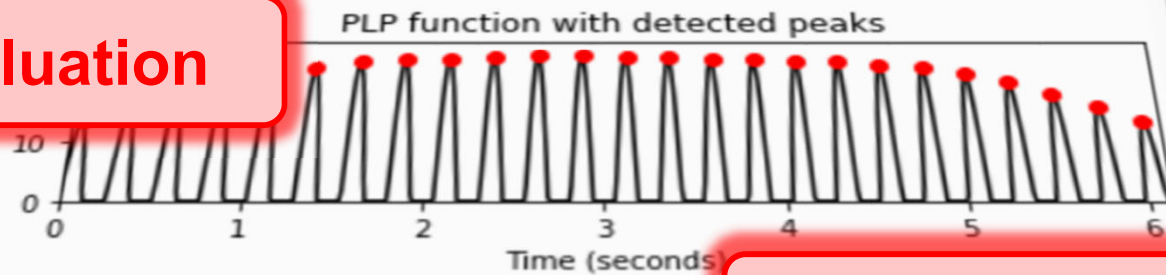


Results



Visualization

Evaluation



Sonification

FMP Notebooks

Teaching

Understanding

Programming

Baselines

Research

Multimedia

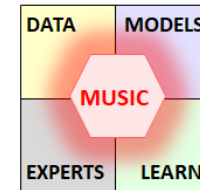
The screenshot displays a Jupyter Notebook titled 'C6' from the 'Fundamentals of Music Processing' series. The notebook content includes:

- Definition:** A discrete-time novelty function $\Delta : \mathbb{Z} \rightarrow \mathbb{R}$. The idea of Fourier analysis is to detect local sections of Δ with a windowed sinusoid. A high correlation between a section of Δ with a windowed sinusoid indicates a periodicity of the sinusoid (given a suitable phase). This correlation (along with the phase) can be captured by a short-time Fourier transform. To this end, we fix a window function $w : \mathbb{Z} \rightarrow \mathbb{R}$ of length centered at $n = 0$ (e.g., a sampled Hanning window). For a frequency parameter $\omega \in \mathbb{R}_{>0}$ and time parameter $n \in \mathbb{Z}$, the Fourier coefficients are defined by

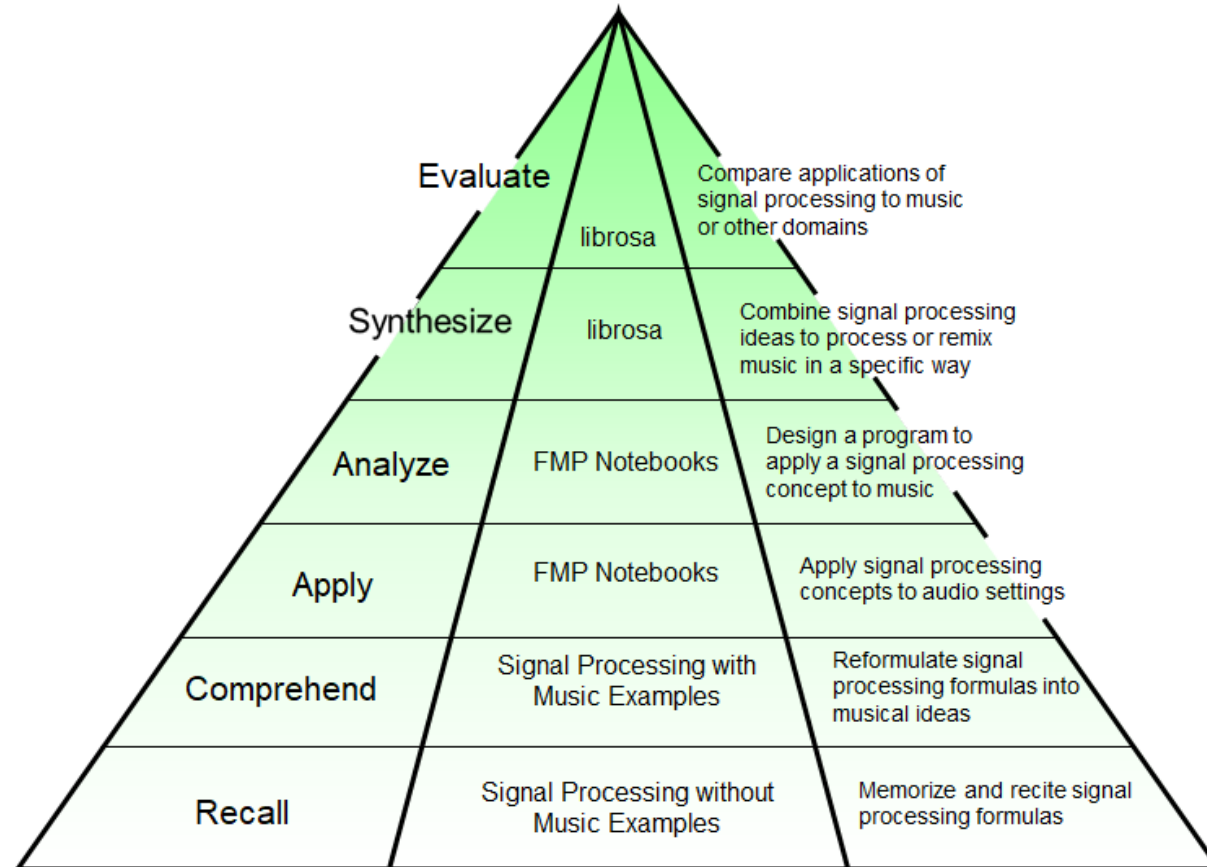
$$\mathcal{F}(n, \omega) := \sum_{m \in \mathbb{Z}} \Delta(m) \bar{w}(m - n) \exp(-2\pi i \omega m).$$
- Example: Shostakovich:** An excerpt of a recording of a Shostakovich symphony orchestra. The notebook includes a music score and an audio player showing 0:00 / 0:06. Annotations include 'Music Example', 'Annotations', and 'Audio'.
- Code:** A Python function `compute_sinusoid_optimal` is defined. Annotations include 'Python Code', 'Functions', and 'Algorithms'.
- Visualization:** A plot titled 'NOVELTY ESTIMATION WITH UNDETERMINED SPEAKERS' showing a signal over time (0 to 6 seconds). Annotations include 'Links', 'Visualization', and 'Results'.
- Evaluation:** A plot titled 'PLP function with detected peaks' showing a periodic signal over time (0 to 6 seconds). Annotations include 'Evaluation' and 'Sonification'.

At the bottom of the notebook interface, there is a navigation bar with icons for various components, numbered 0 through 8.

Reinhart Koselleck-Projekt: LEARN



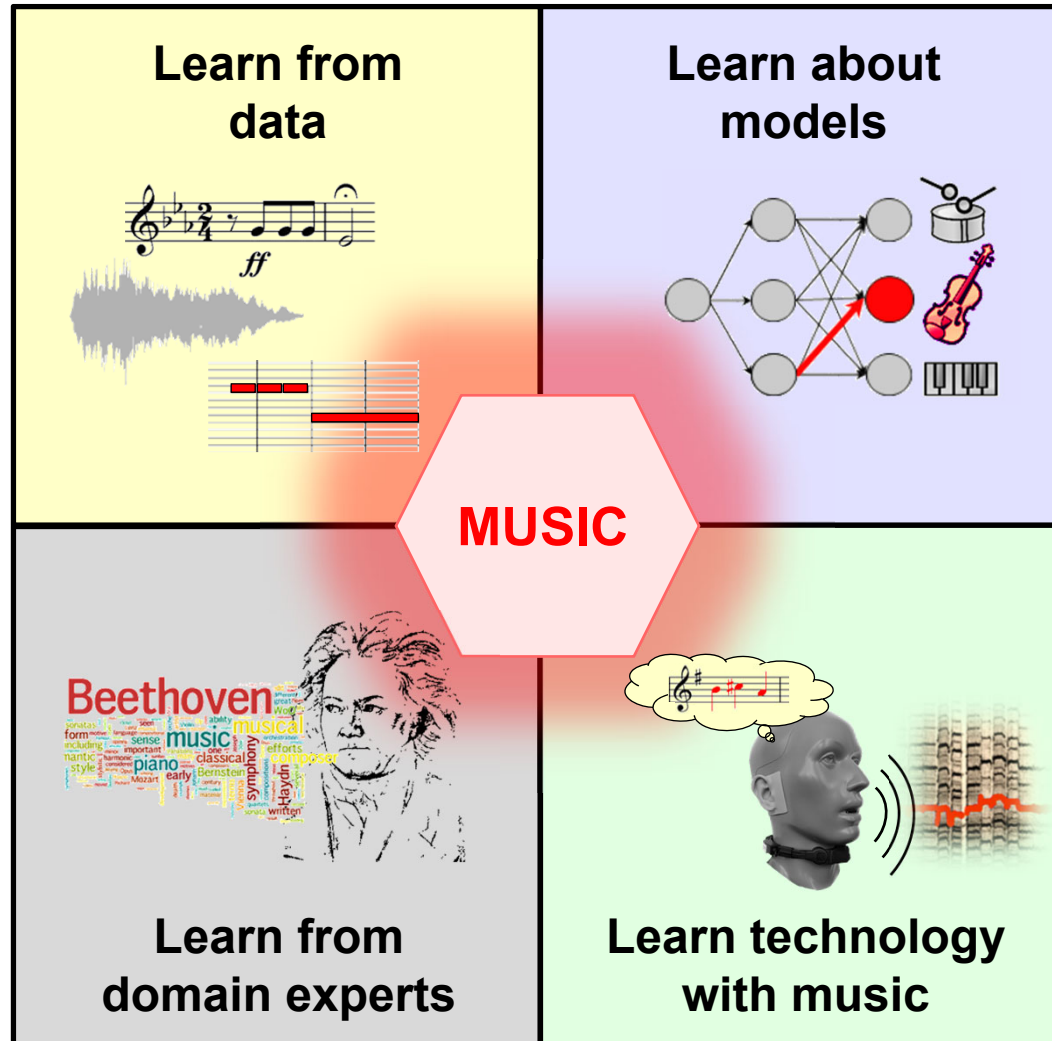
- Motivating examples
- Challenging domain
- Interdisciplinarity
- Structured learning
- ...



Bloom's taxonomy

Reinhart Koselleck-Projekt: LEARN

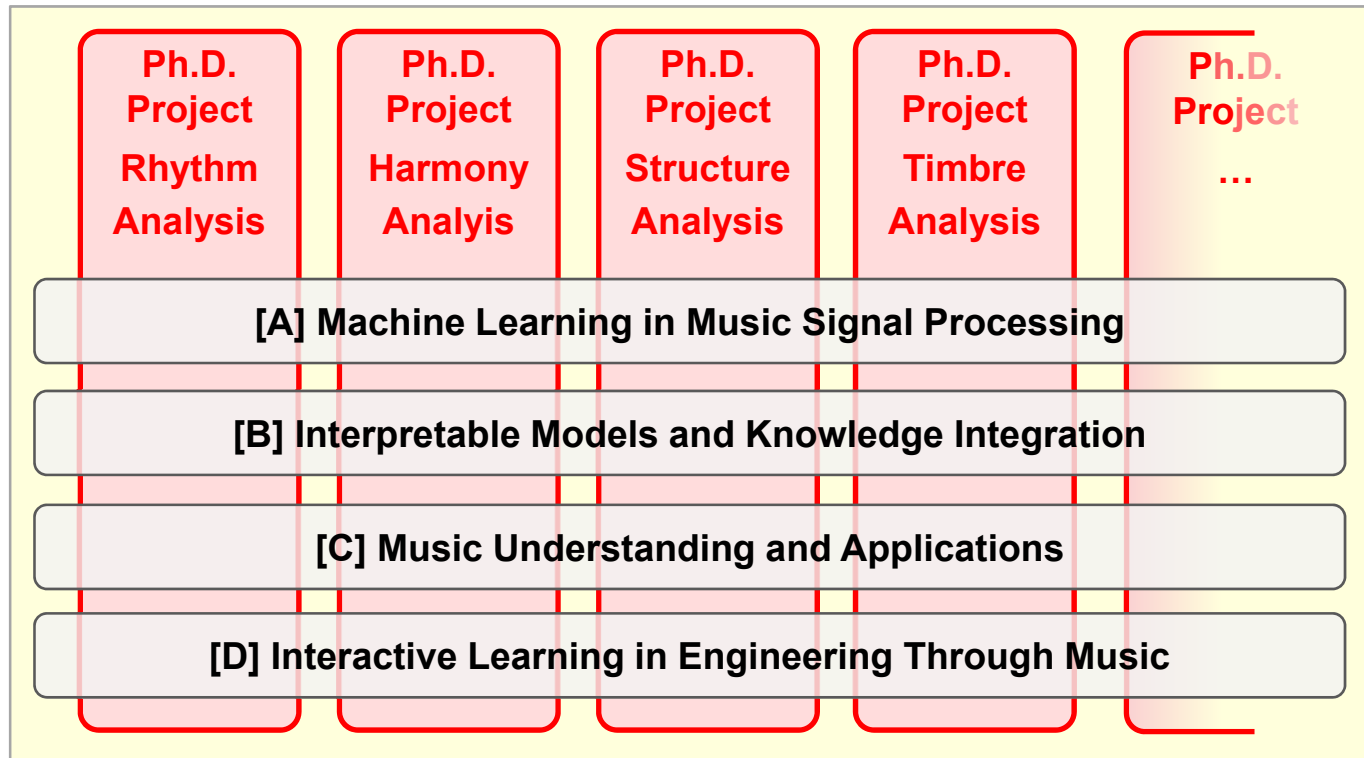
Learning with Music Signals: Technology Meets Education



- Machine learning for music signal processing
- Interpretable models and knowledge integration
- Music understanding and applications
- Interactive learning in engineering through music

Reinhart Koselleck-Projekt: LEARN

Learning with Music Signals: Technology Meets Education



References

- Meinard Müller: Fundamentals of Music Processing – Using Python and Jupyter Notebooks. 2nd Edition, Springer, 2021.
<https://www.springer.com/gp/book/9783030698072>
- Meinard Müller and Frank Zalkow: libfmp: A Python Package for Fundamentals of Music Processing. Journal of Open Source Software (JOSS), 6(63): 1–5, 2021.
<https://joss.theoj.org/papers/10.21105/joss.03326>
- Meinard Müller: An Educational Guide Through the FMP Notebooks for Teaching and Learning Fundamentals of Music Processing. Signals, 2(2): 245–285, 2021.
<https://www.mdpi.com/2624-6120/2/2/18>
- Meinard Müller and Frank Zalkow: FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing. Proc. International Society for Music Information Retrieval Conference (ISMIR): 573–580, 2019.
<https://zenodo.org/record/3527872#.YOhEQOgzaUk>
- Meinard Müller, Brian McFee, and Katherine Kinnaird: Interactive Learning of Signal Processing Through Music: Making Fourier Analysis Concrete for Students. IEEE Signal Processing Magazine, 38(3): 73–84, 2021.
<https://ieeexplore.ieee.org/document/9418542>

Resources (Group Meinard Müller)

- FMP Notebooks:

<https://www.audiolabs-erlangen.de/FMP>

- libfmp:

<https://github.com/meinardmueller/libfmp>

- synctoolbox:

<https://github.com/meinardmueller/synctoolbox>

- libtsm:

<https://github.com/meinardmueller/libtsm>

- Preparation Course Python (PCP) Notebooks:

<https://www.audiolabs-erlangen.de/resources/MIR/PCP/PCP.html>

<https://github.com/meinardmueller/PCP>