Tutorial T1
Fundamentals of Music Processing:
An Introduction using Python and Jupyter Notebooks

Audio Decomposition

Meinard Müller, Frank Zalkow

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

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

Fraunhofer
IIS

Source Separation

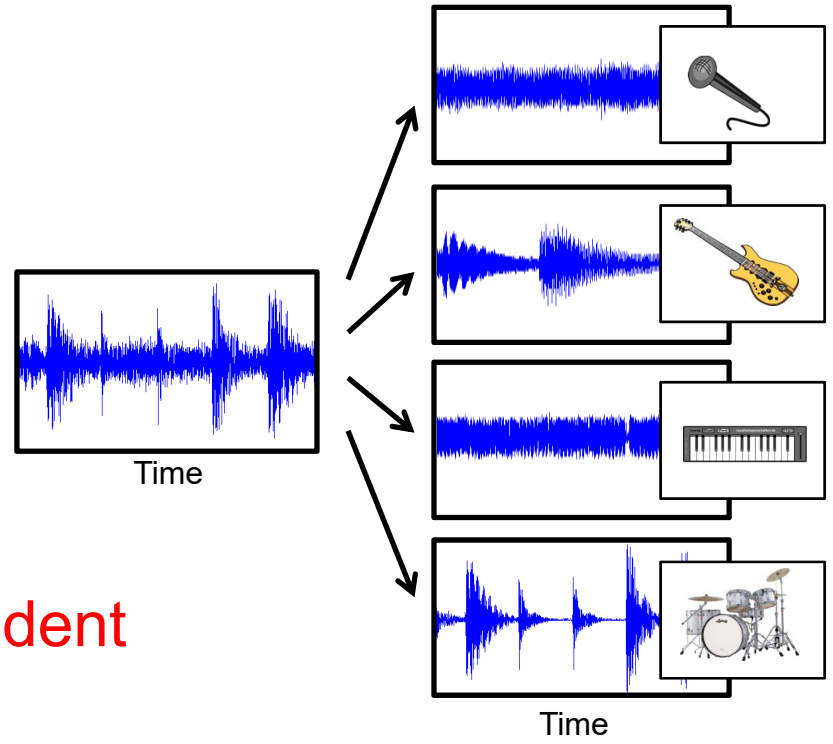
- Decomposition of audio stream into different sound sources
- Central task in digital signal processing
- “Cocktail party effect”

Source Separation

- Decomposition of audio stream into different sound sources
- Central task in digital signal processing
- “Cocktail party effect”
- Several input signals
- Sources are assumed to be statistically independent

Source Separation (Music)

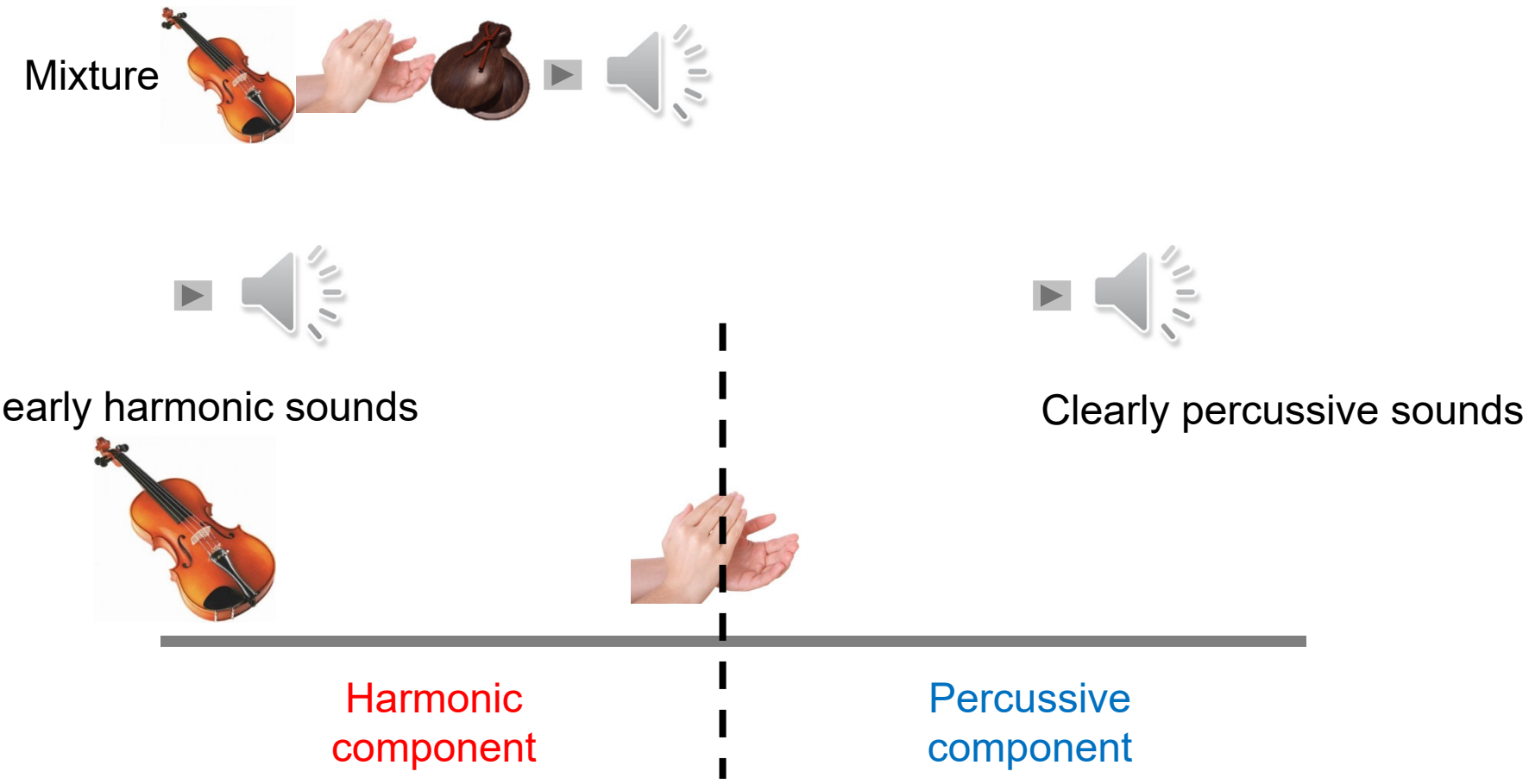
- Main melody, accompaniment, drum track
- Instrumental voices
- Individual note events
- Only mono or stereo
- Sources are often highly dependent



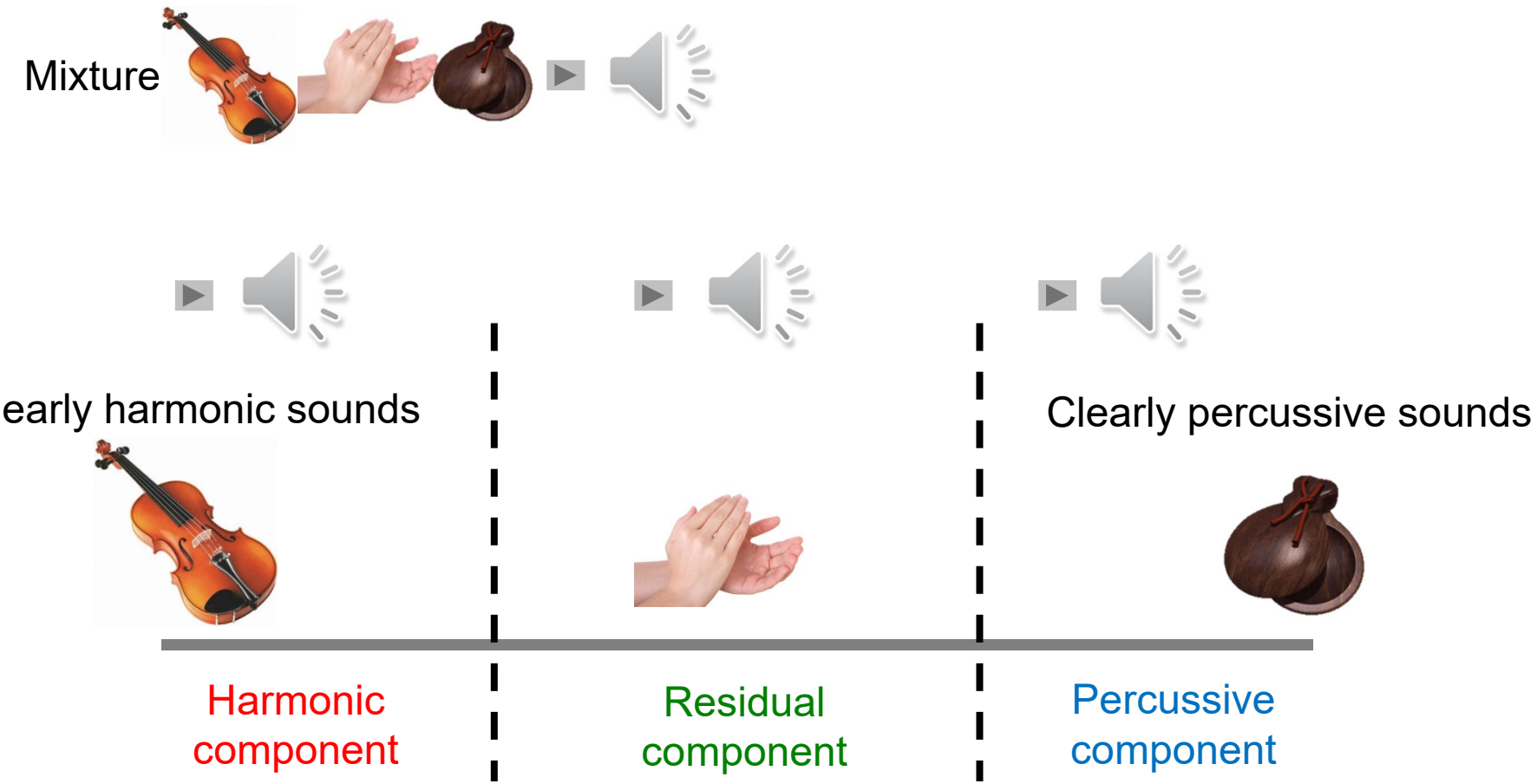
Harmonic-Percussive Decomposition

Mixture 

Harmonic-Percussive Decomposition

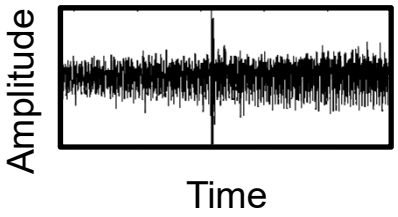


Harmonic-Percussive Decomposition



Harmonic-Percussive Decomposition

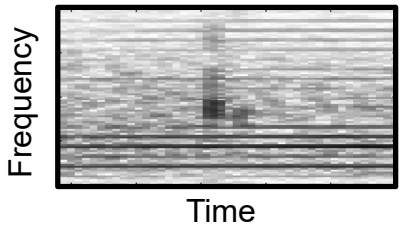
Audio



[Ono et al. ISMIR 2008, Fitzgerald DAFx 2010]

Harmonic-Percussive Decomposition

Spectrogram

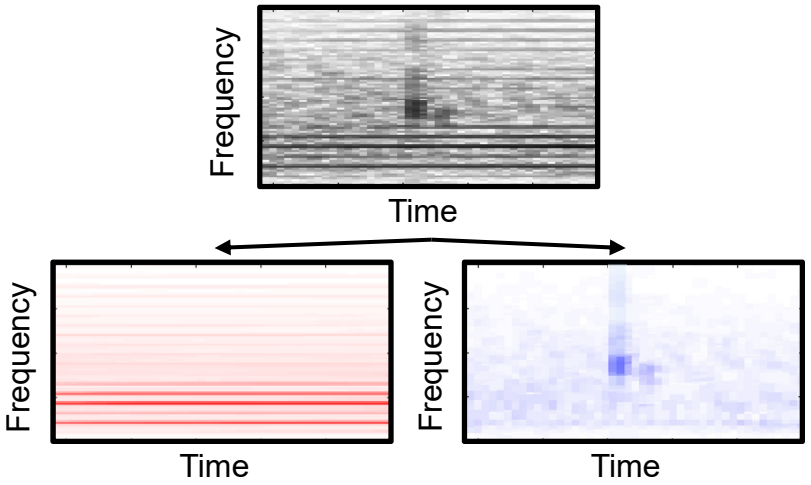


[Ono et al. ISMIR 2008, Fitzgerald DAFx 2010]

Harmonic-Percussive Decomposition

Spectrogram

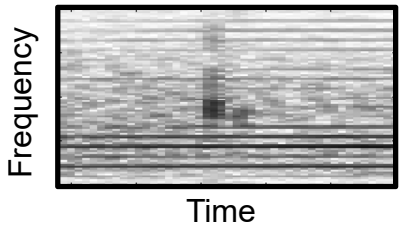
Horizontally and vertically enhanced spectrograms



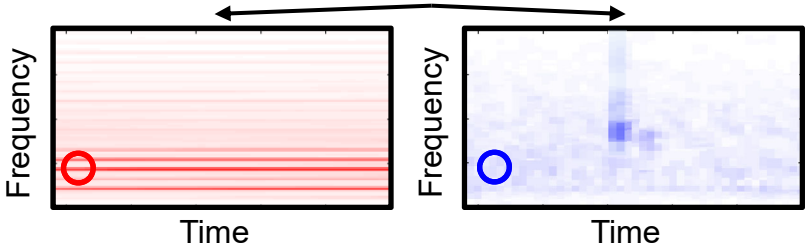
[Ono et al. ISMIR 2008, Fitzgerald DAFx 2010]

Harmonic-Percussive Decomposition

Spectrogram



Horizontally and vertically enhanced spectrograms

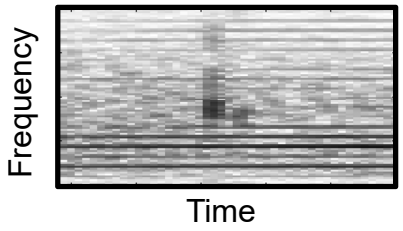


○ > ○ ➔ Harmonic component

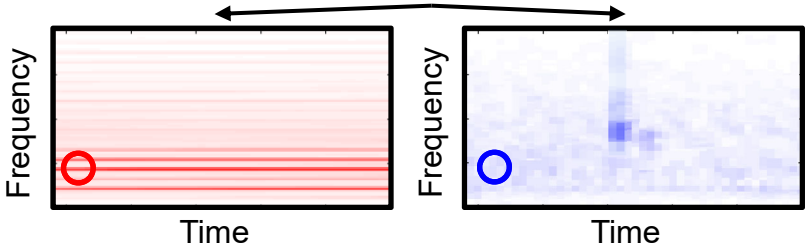
[Ono et al. ISMIR 2008, Fitzgerald DAFx 2010]

Harmonic-Percussive Decomposition

Spectrogram



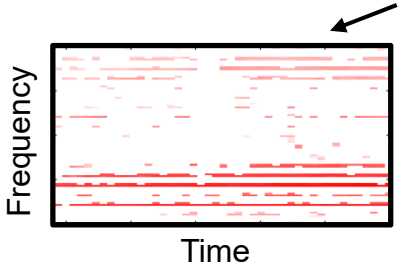
Horizontally and vertically enhanced spectrograms



Separation factor

$\beta \geq 1$

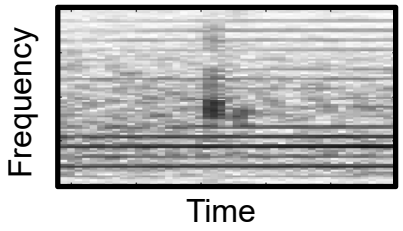
$\circ > \beta \cdot \circ$



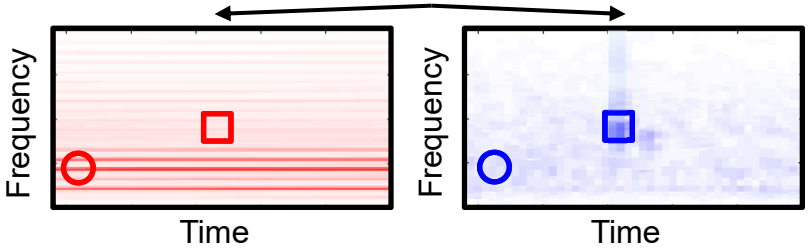
[Driedger et al. ISMIR 2014]

Harmonic-Percussive Decomposition

Spectrogram



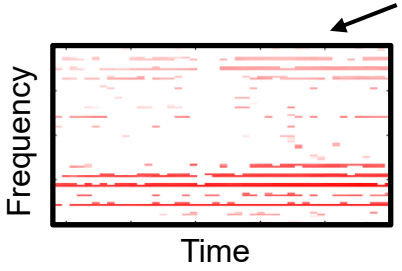
Horizontally and vertically enhanced spectrograms

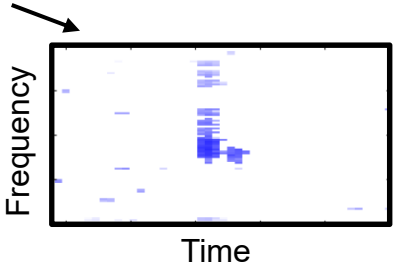


Separation factor
$\beta \geq 1$

$\circ > \beta \bullet \circ$ ➔ Harmonic component

$\beta \bullet \square < \square$ ➔ Percussive component

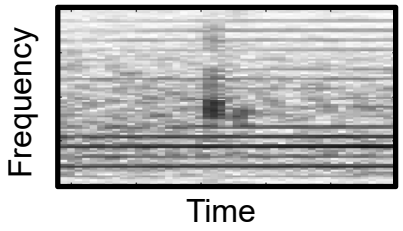




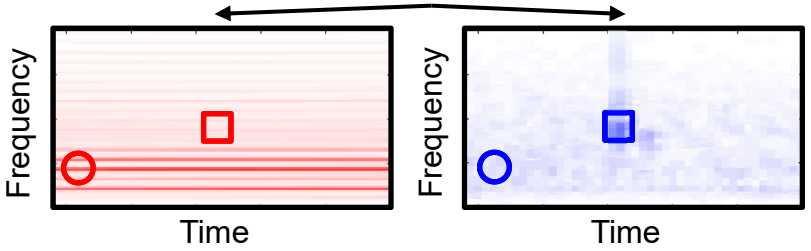
[Driedger et al. ISMIR 2014]

Harmonic-Percussive Decomposition

Spectrogram



Horizontally and vertically enhanced spectrograms



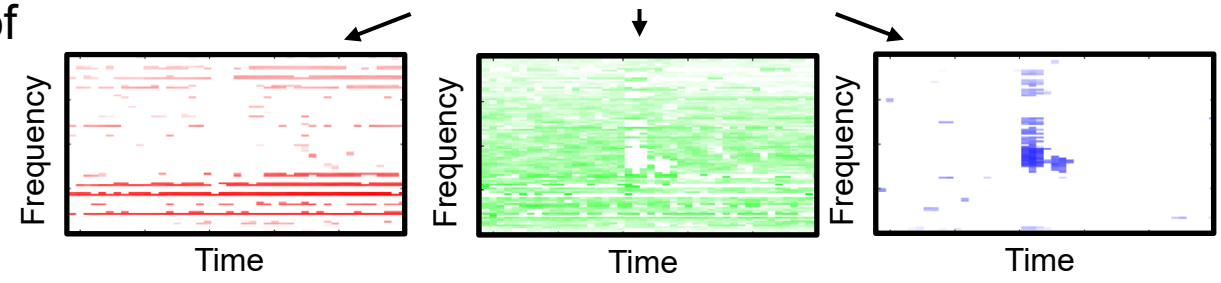
Separation factor
$\beta \geq 1$

$\bigcirc > \beta \cdot \bigcirc$ ➔ Harmonic component

$\beta \cdot \square < \square$ ➔ Percussive component

otherwise ➔ Residual component

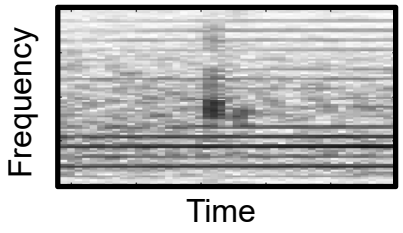
Spectrograms of the harmonic, residual and percussive components



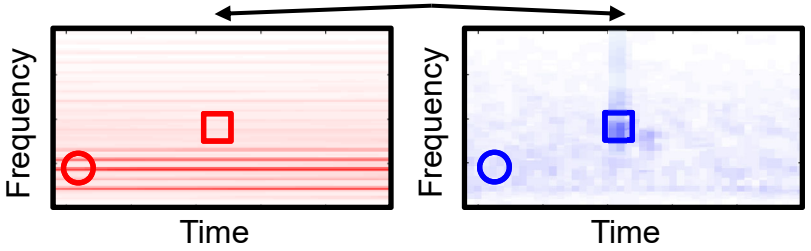
[Driedger et al. ISMIR 2014]

Harmonic-Percussive Decomposition

Spectrogram



Horizontally and vertically enhanced spectrograms



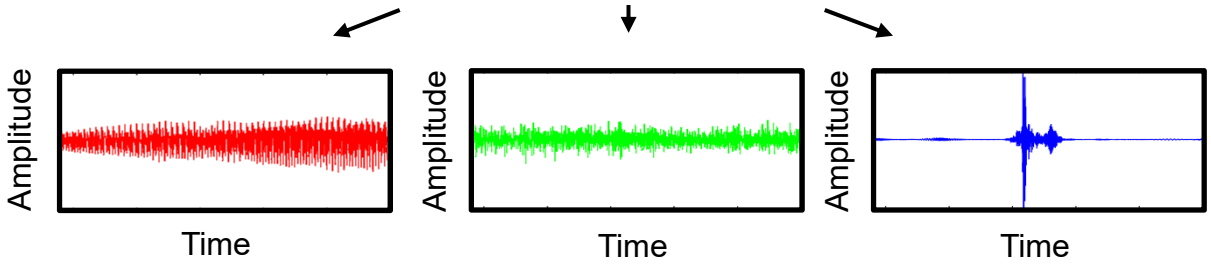
Separation factor
$\beta \geq 1$

$\bigcirc > \beta \bullet \bigcirc$ ➔ Harmonic component

$\beta \bullet \square < \square$ ➔ Percussive component

otherwise ➔ Residual component

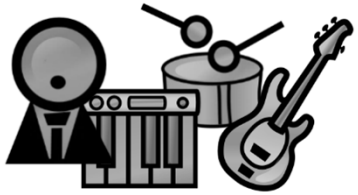
Harmonic, residual and percussive components



[Driedger et al. ISMIR 2014]

Harmonic-Percussive Decomposition

Mixture:



- Clearly harmonic sounds of singing voice and accompaniment

- Noise-like sounds
- Vibrato/glissando sounds

- Drum hits
- Fricatives & plosives in singing voice

Harmonic component

Residual component

Percussive component

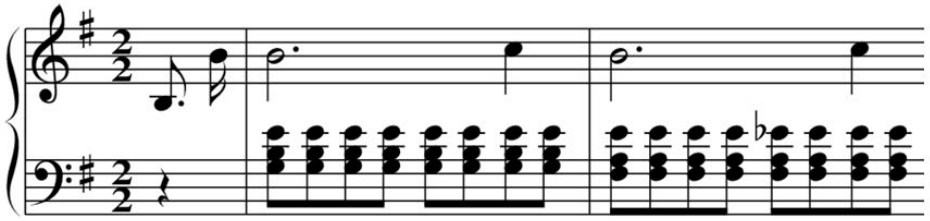
Demo: https://www.audiolabs-erlangen.de/resources/2014-ISMIR-ExtHPSep/

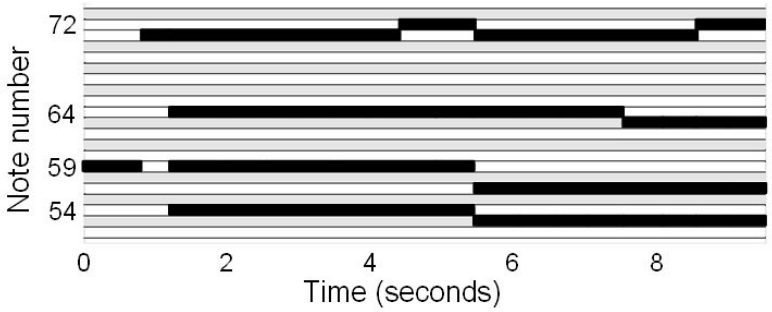
[Driedger et al. ISMIR 2014]

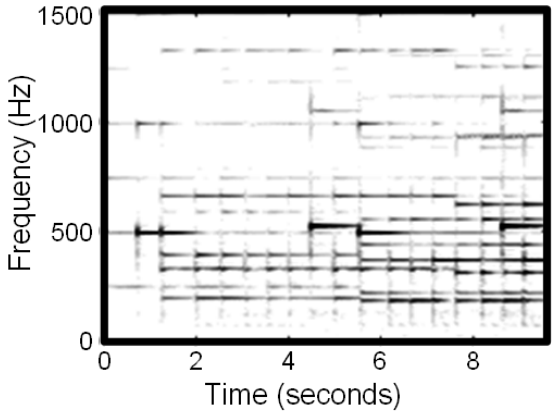
Score-Informed Audio Decomposition

Score-Informed Audio Decomposition

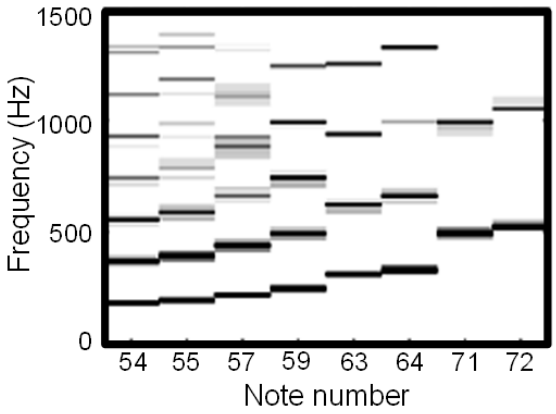
Exploit musical score to support separation process



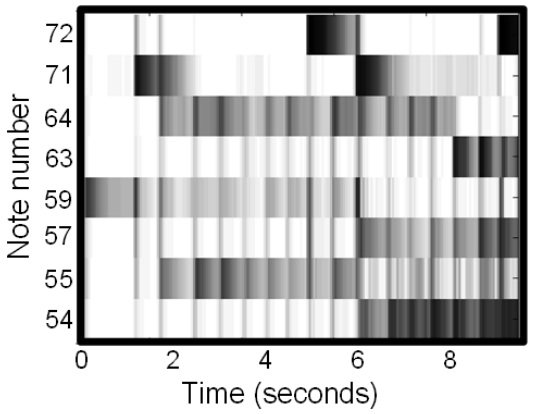




≈

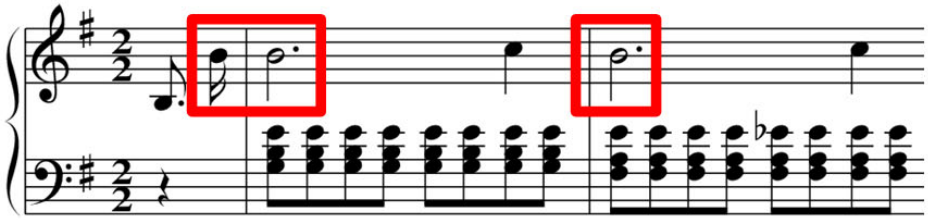


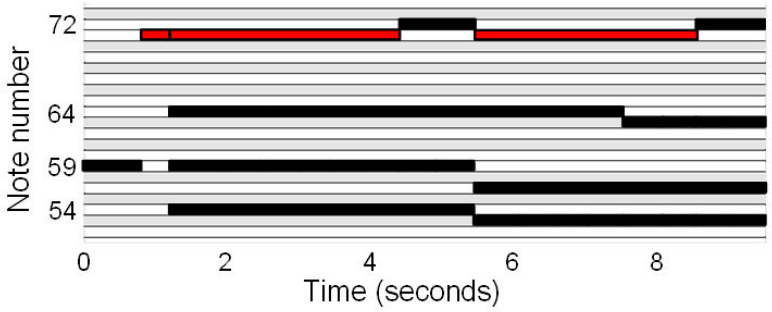
•

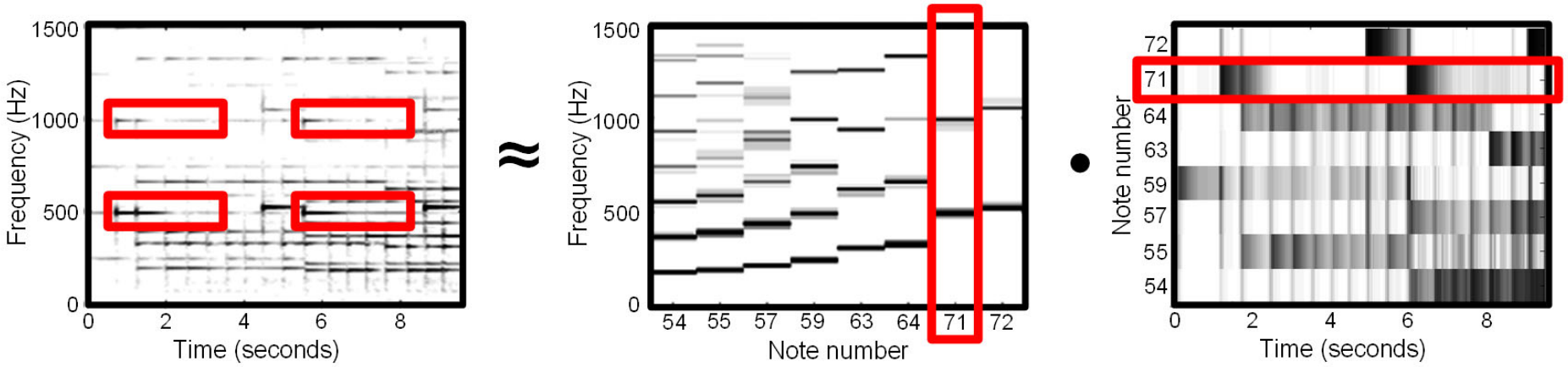


Score-Informed Audio Decomposition

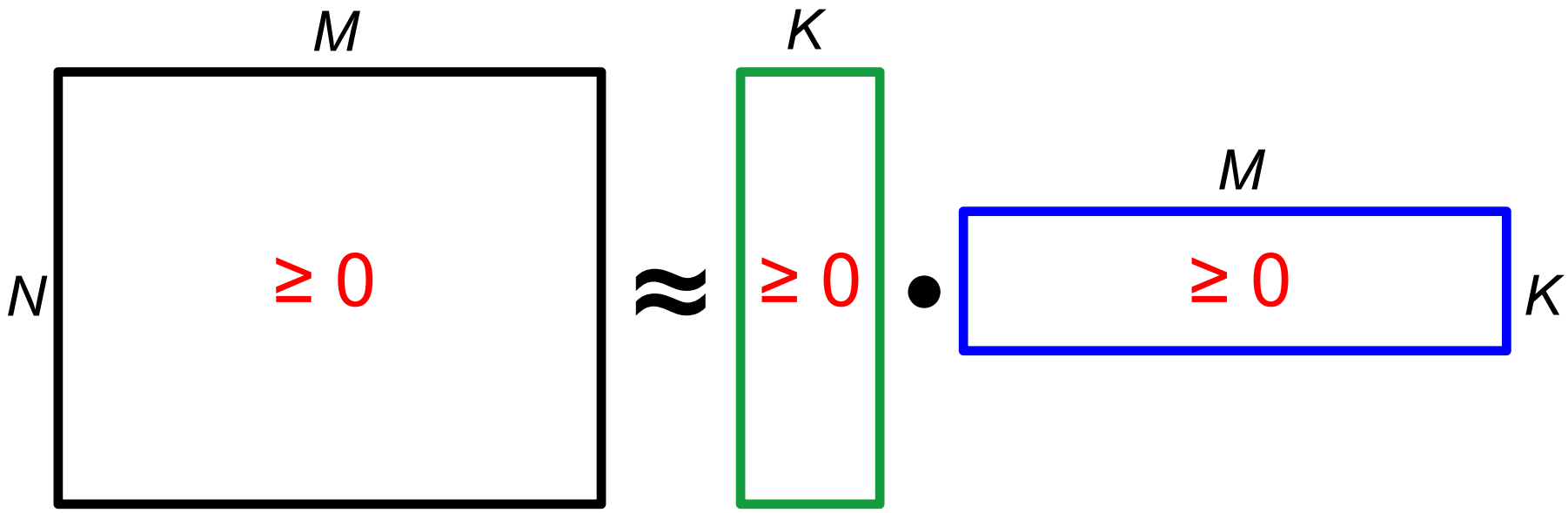
Exploit musical score to support separation process



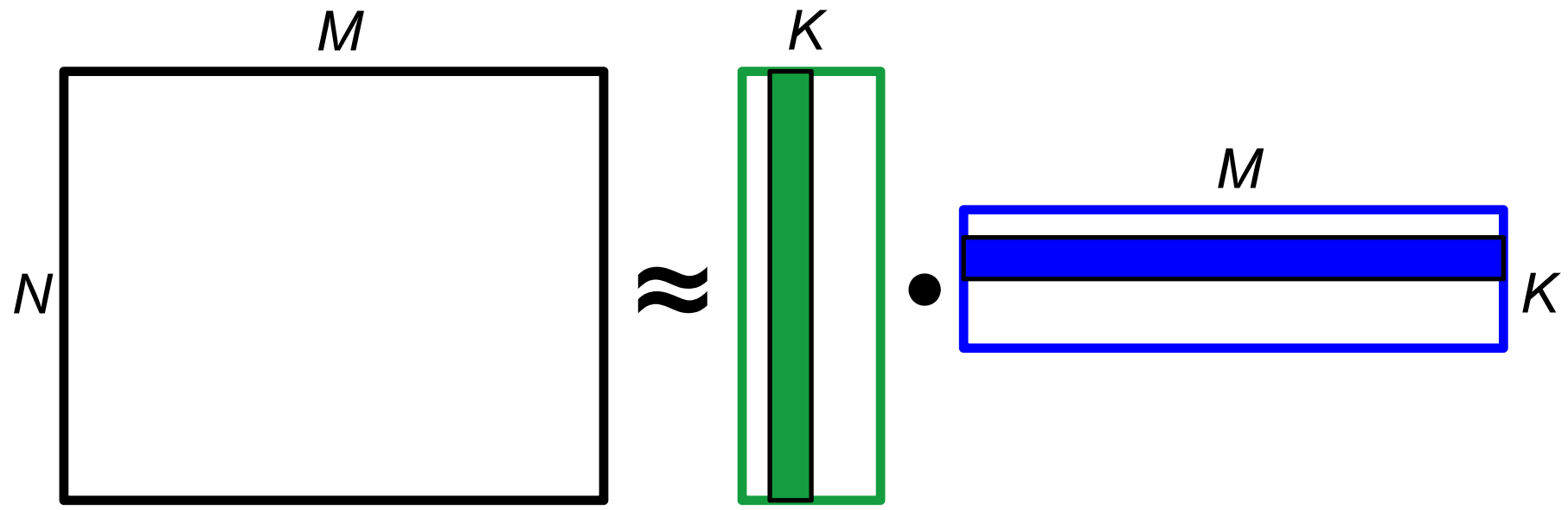




NMF (Nonnegative Matrix Factorization)



NMF (Nonnegative Matrix Factorization)

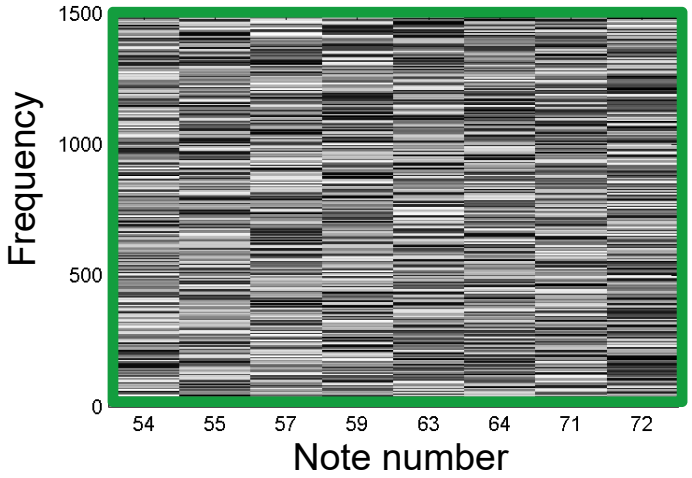


Templates: Pitch + Timbre — "How does it sound"

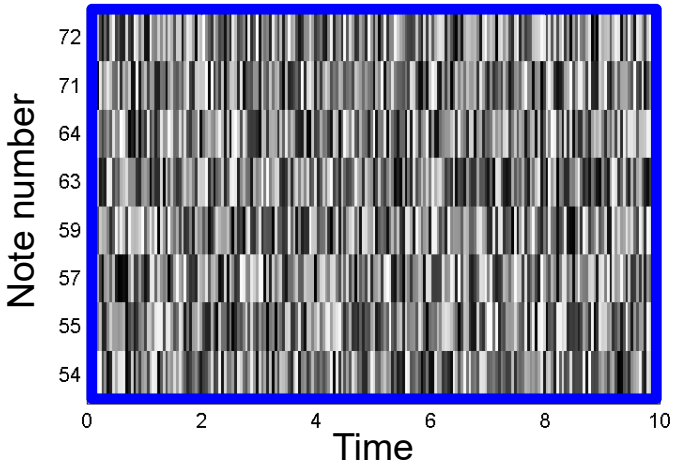
Activations: Onset time + Duration — "When does it sound"

NMF-Decomposition

Initialized template



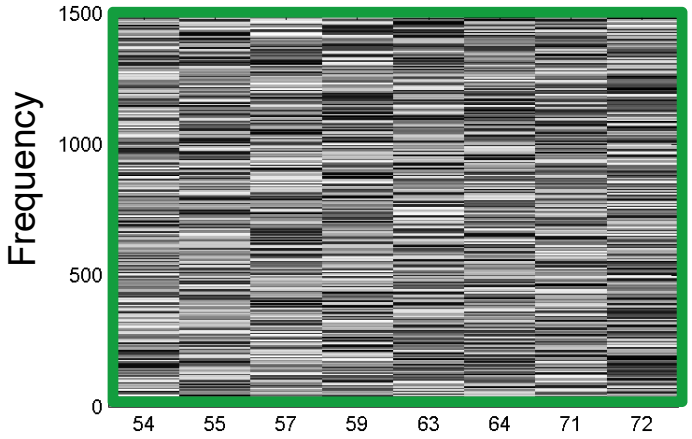
Initialized activations



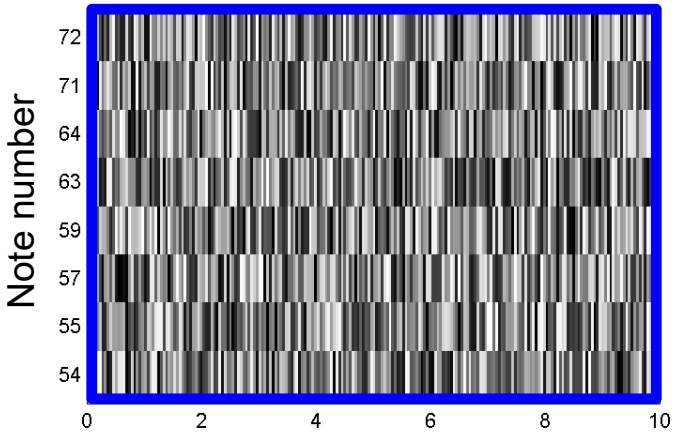
Random initialization

NMF-Decomposition

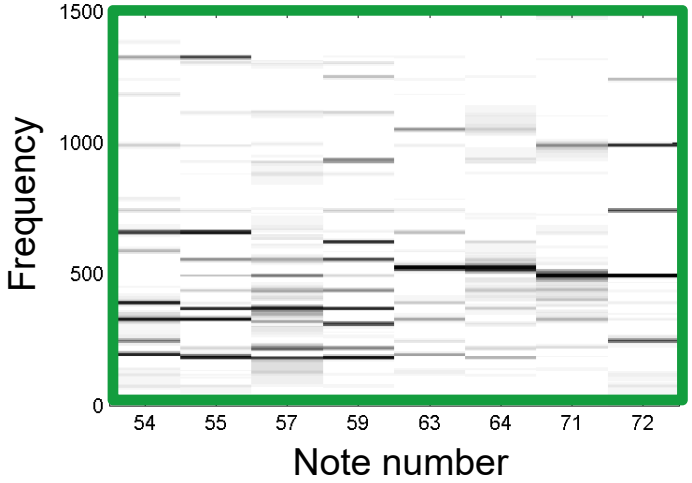
Initialized template



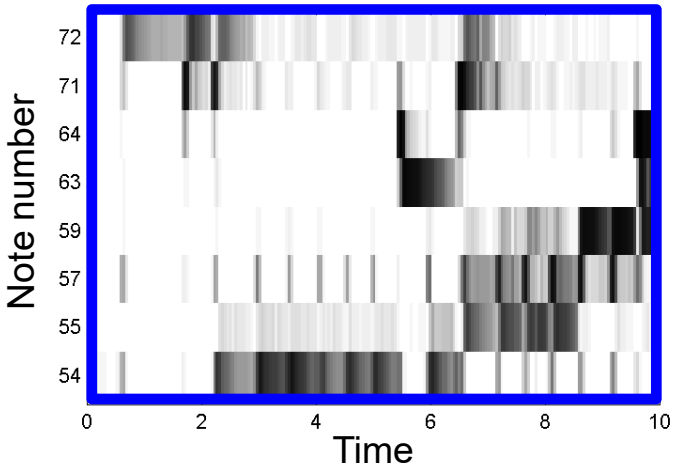
Initialized activations



Learnt templates



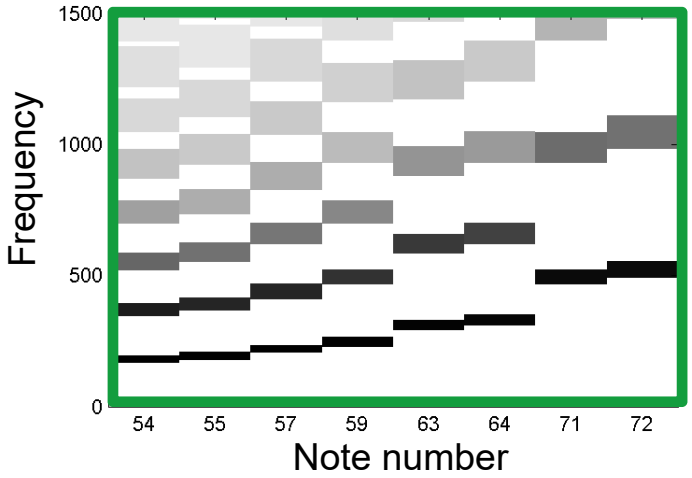
Learnt activations



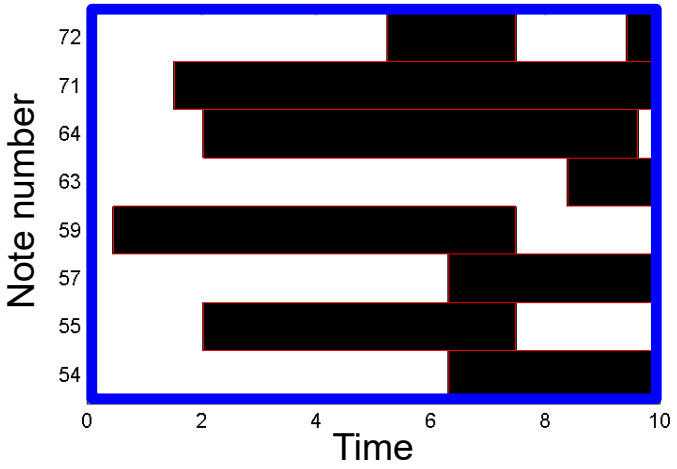
Random initialization → No semantic meaning

NMF-Decomposition

Initialized template



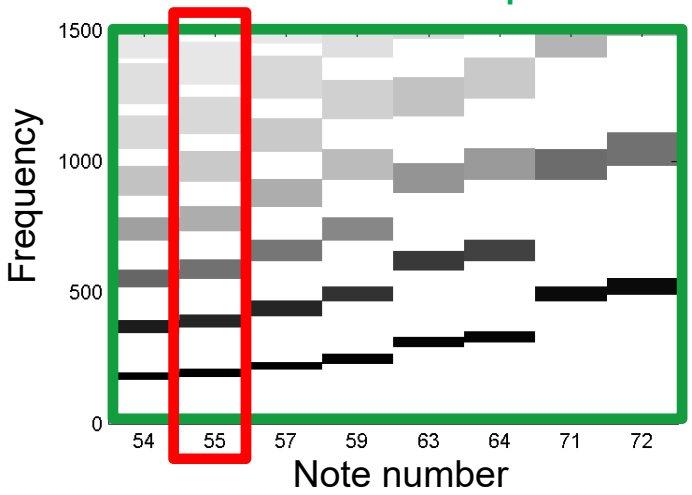
Initialized activations



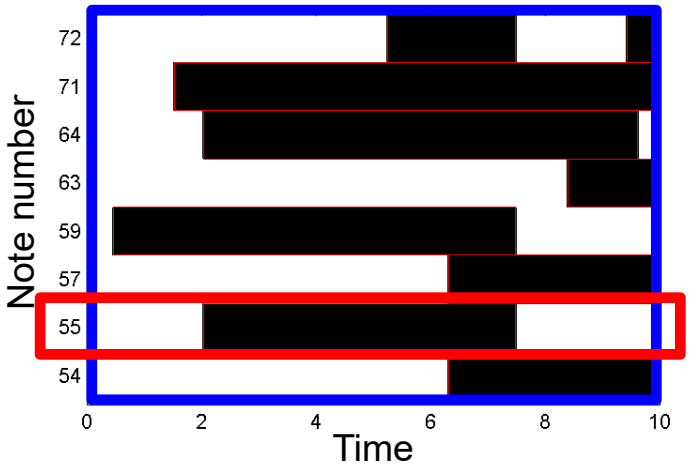
Constrained initialization

NMF-Decomposition

Initialized template



Initialized activations

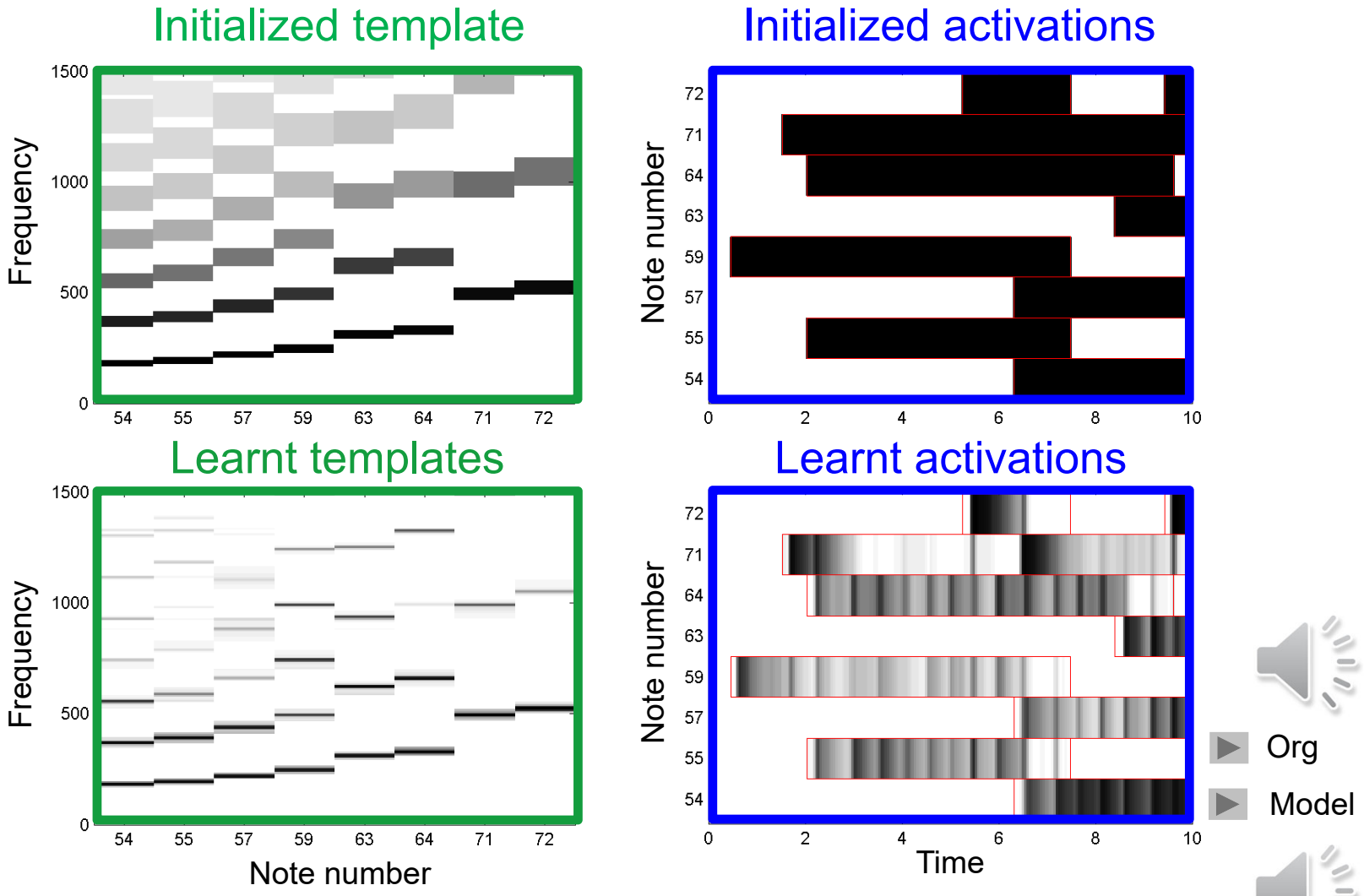


Template constraint for p=55

Activation constraints for p=55

Constrained initialization

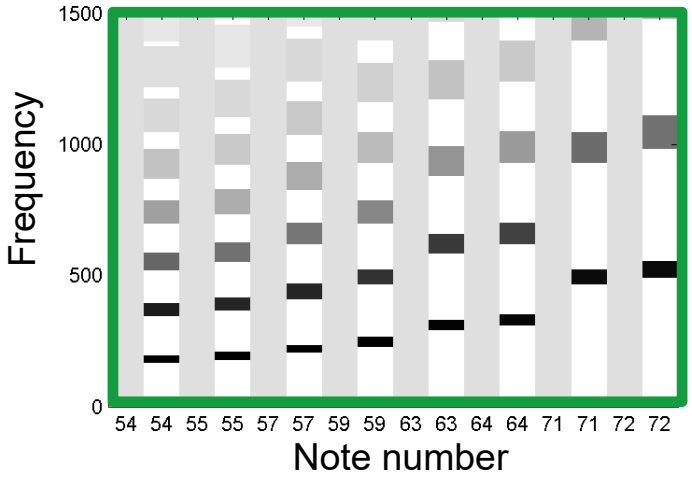
NMF-Decomposition



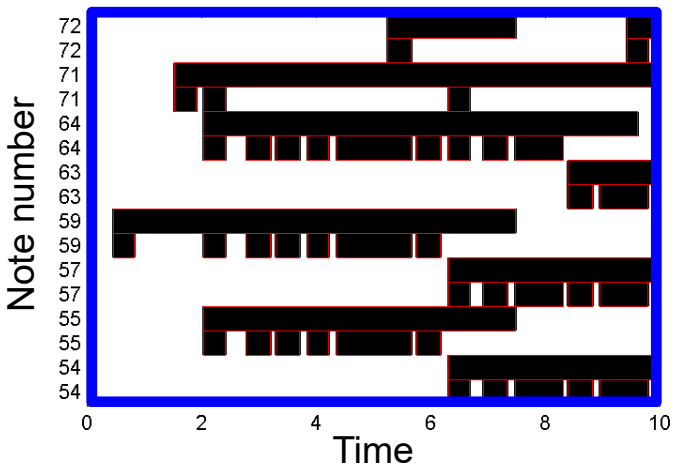
Constrained initialization → NMF as refinement

NMF-Decomposition

Initialized template



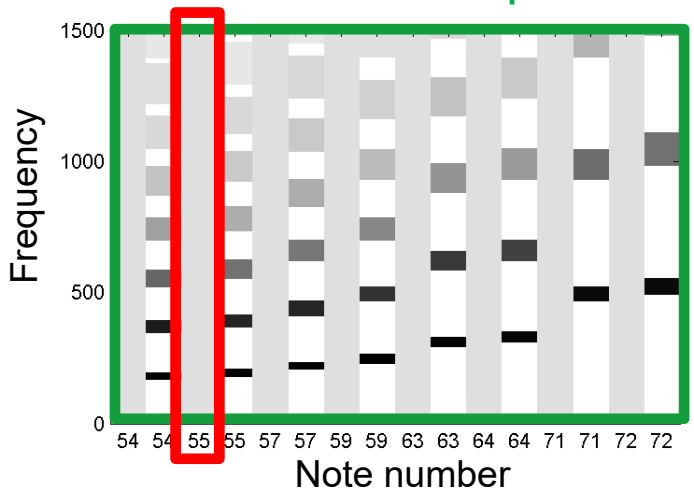
Initialized activations



Additional onset models → NMF as refinement

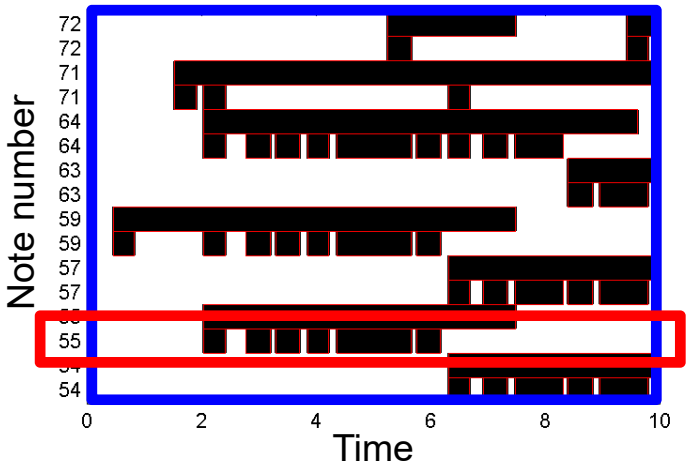
NMF-Decomposition

Initialized template



Onset template for p=55

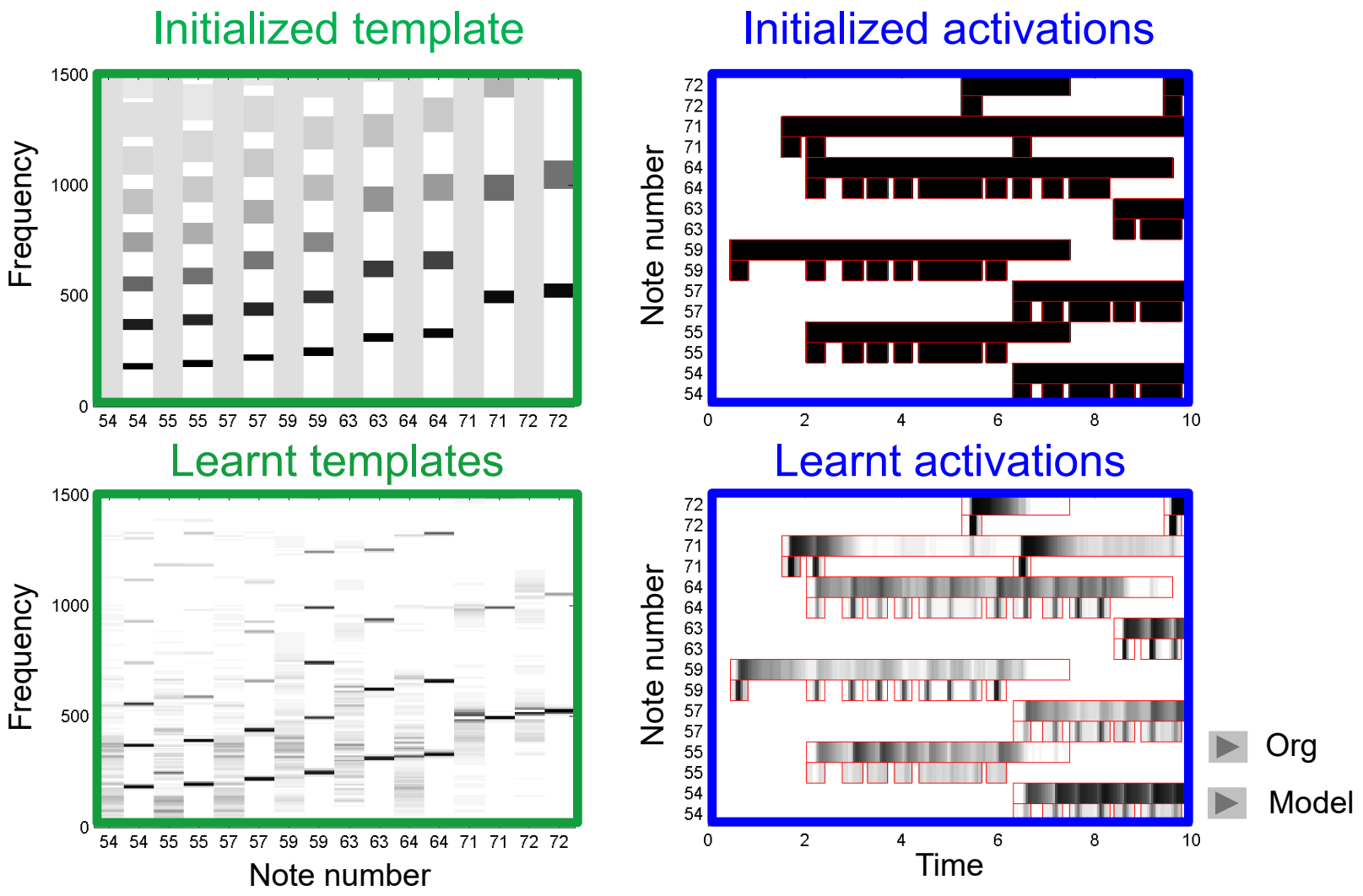
Initialized activations



Activation constraints for onset template

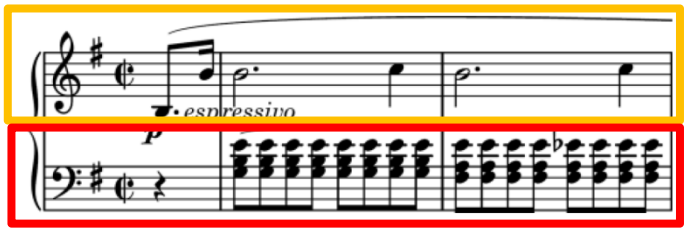
Additional onset models → NMF as refinement

NMF-Decomposition

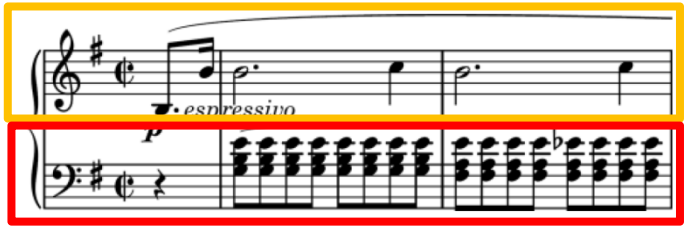


Additional onset models → NMF as refinement

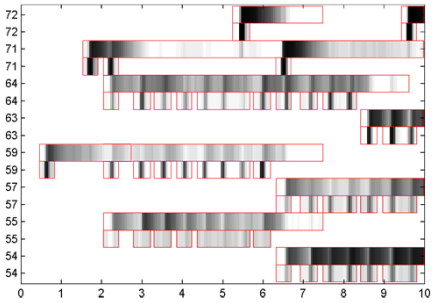
Score-Informed Audio Decomposition



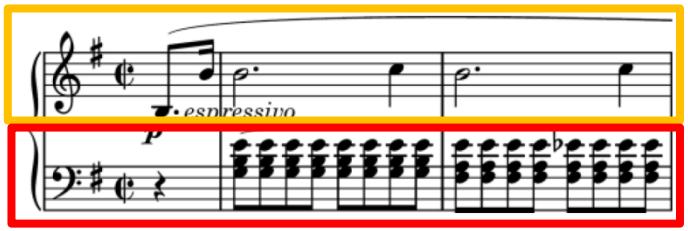
Score-Informed Audio Decomposition



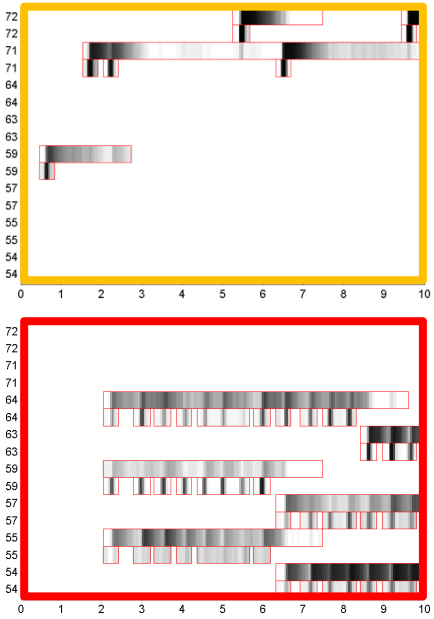
1. Split activation matrix



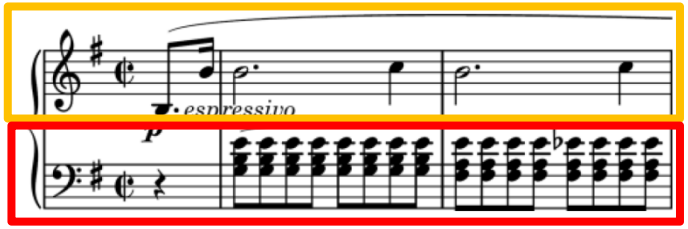
Score-Informed Audio Decomposition



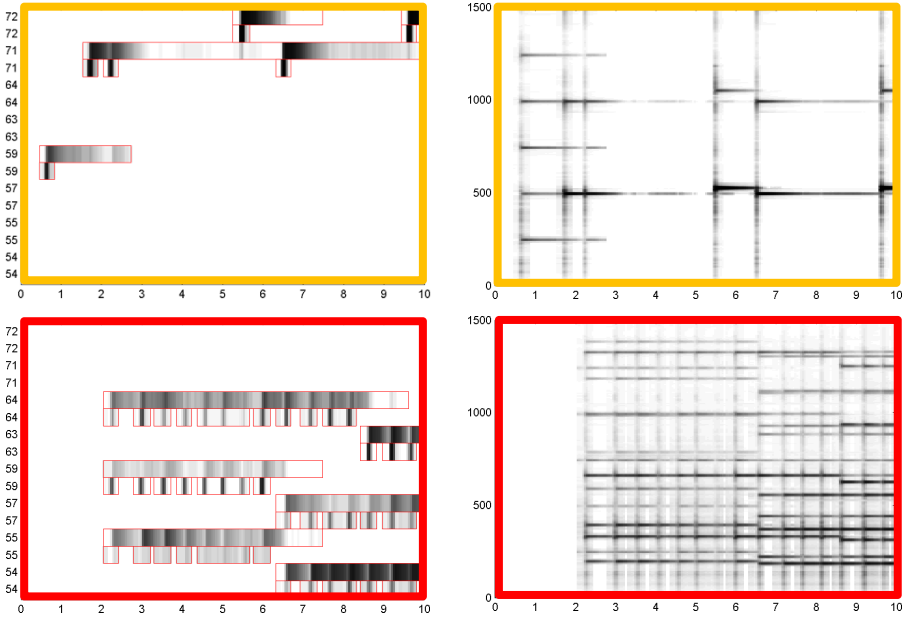
1. Split activation matrix



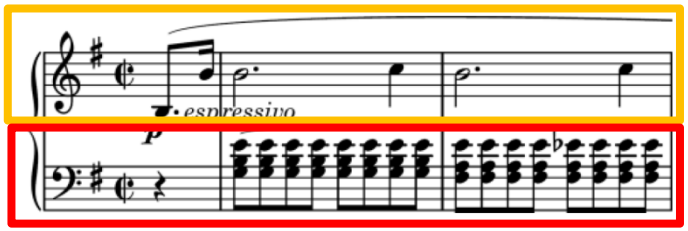
Score-Informed Audio Decomposition



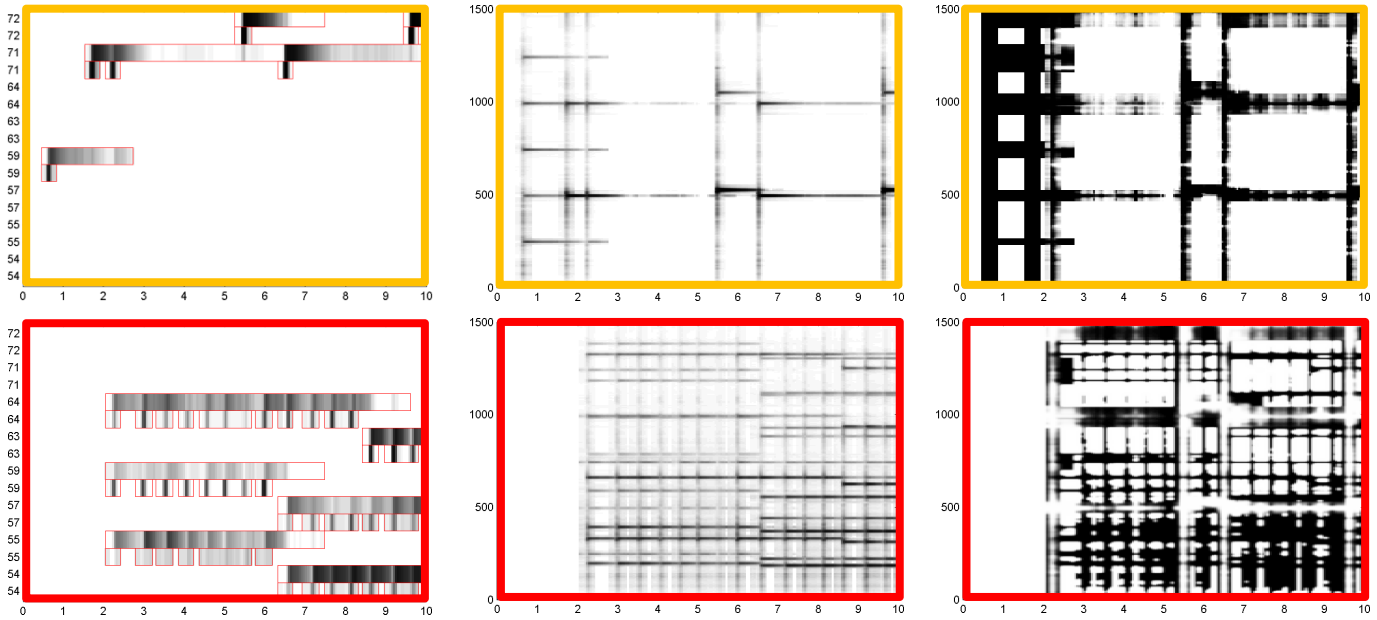
1. Split activation matrix
2. Model spectrogram for left/right



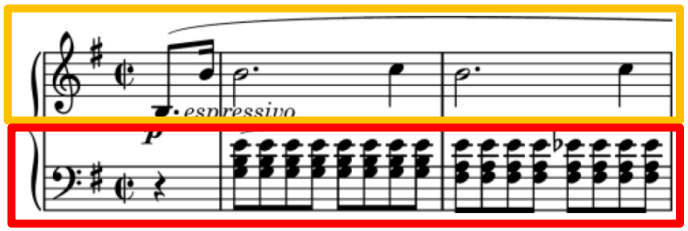
Score-Informed Audio Decomposition



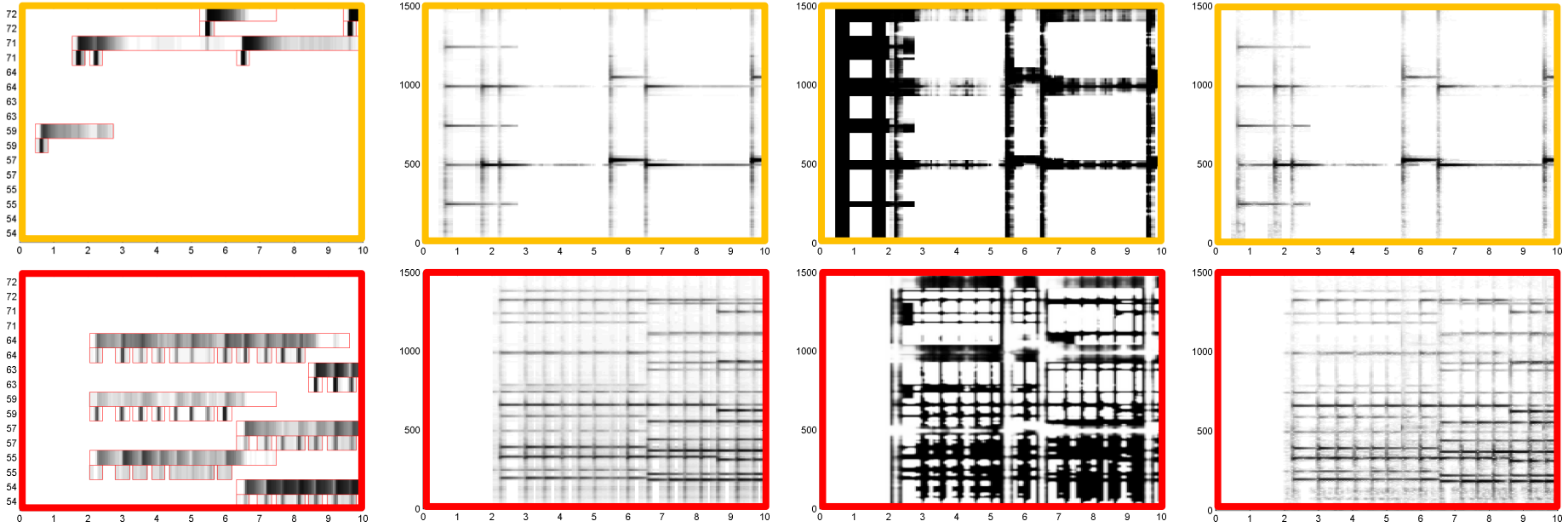
1. Split activation matrix
2. Model spectrogram for left/right
3. Separation masks for left/right



Score-Informed Audio Decomposition



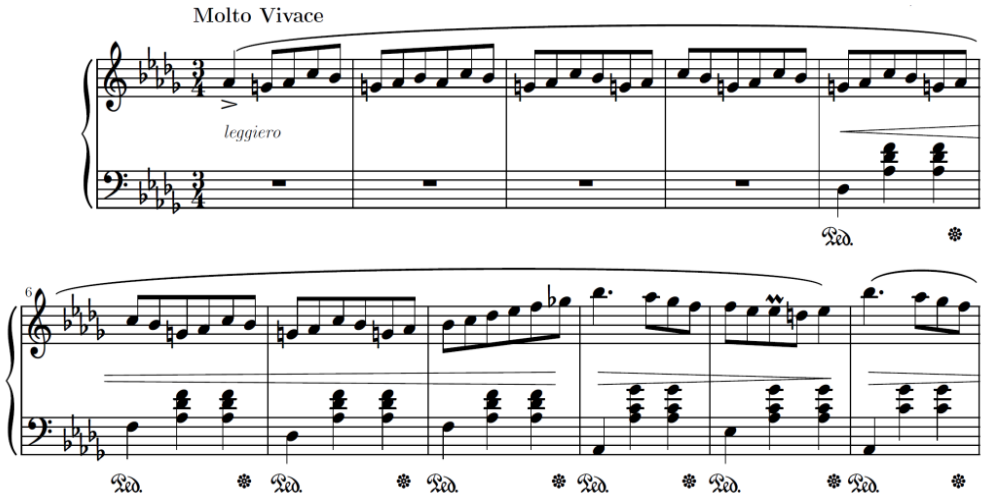
1. Split activation matrix
2. Model spectrogram for left/right
3. Separation masks for left/right
4. Estimated spectrograms for left/right



Score-Informed Audio Decomposition

Application: Separating left and right hands for piano

Chopin, Waltz Op. 64, No. 1

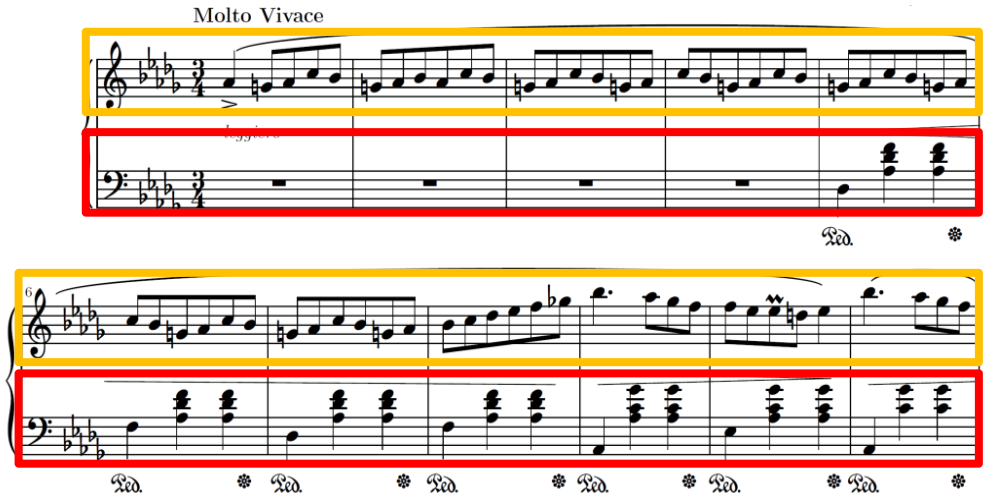


Original

Score-Informed Audio Decomposition

Application: Separating left and right hands for piano

Chopin, Waltz Op. 64, No. 1



Original

Left/right hand

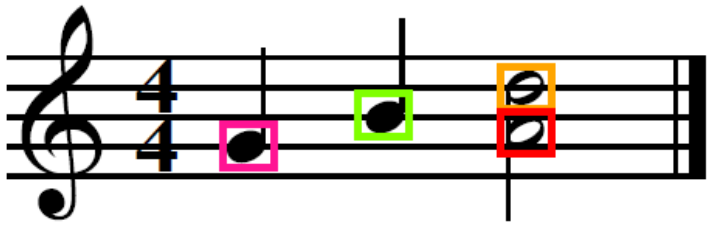
Right hand

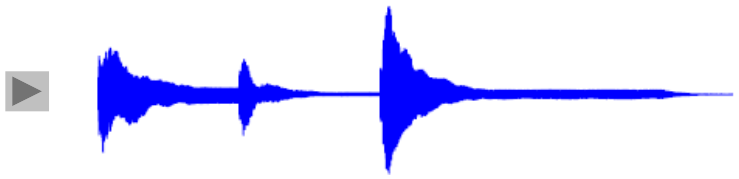
Left hand

Score-Informed Audio Decomposition

Parameterize audio signal using score's note events

Score

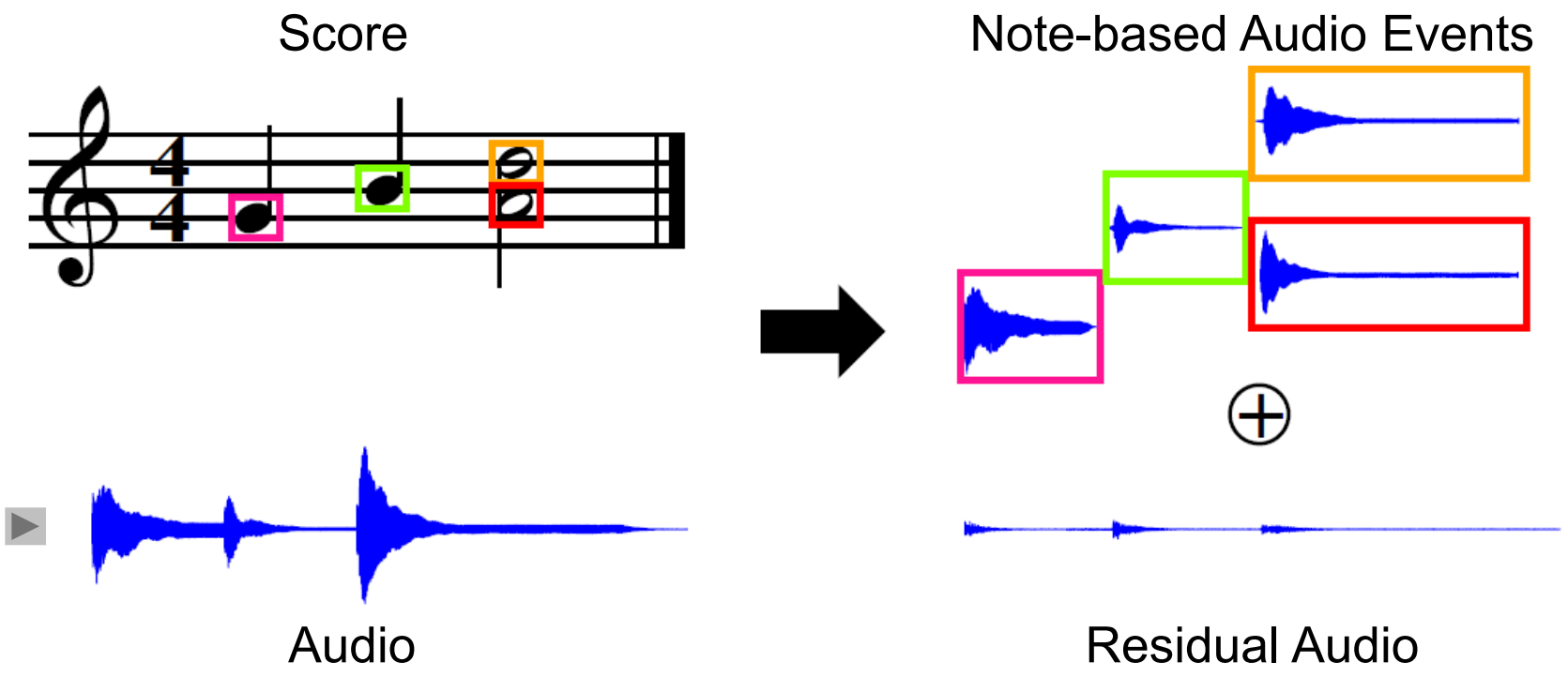




Audio

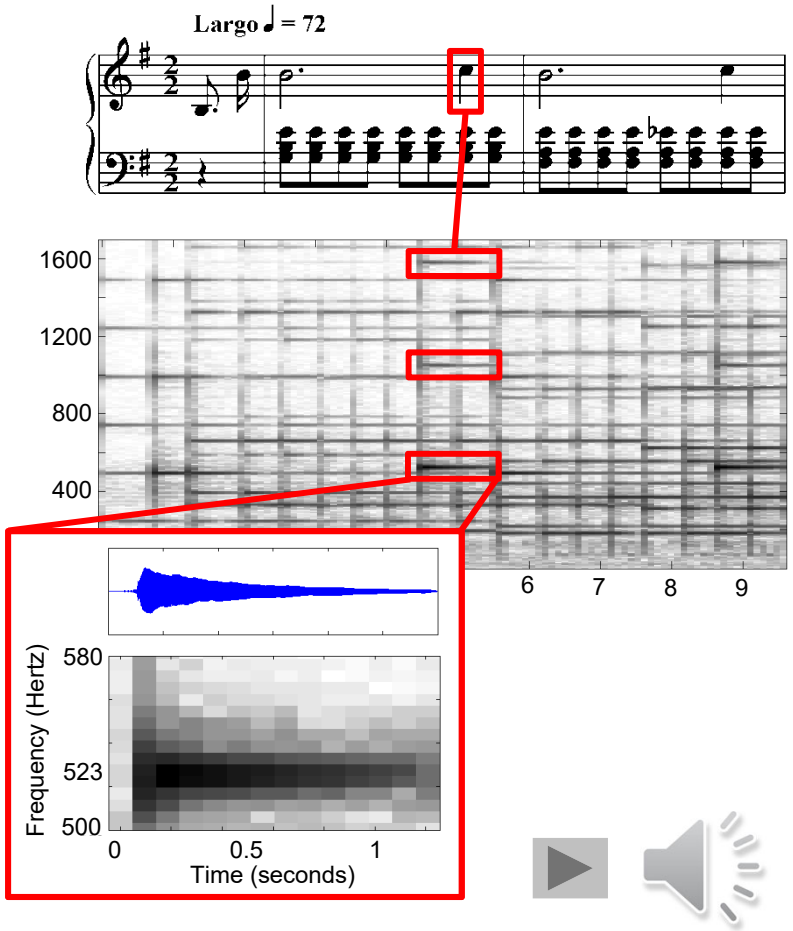
Score-Informed Audio Decomposition

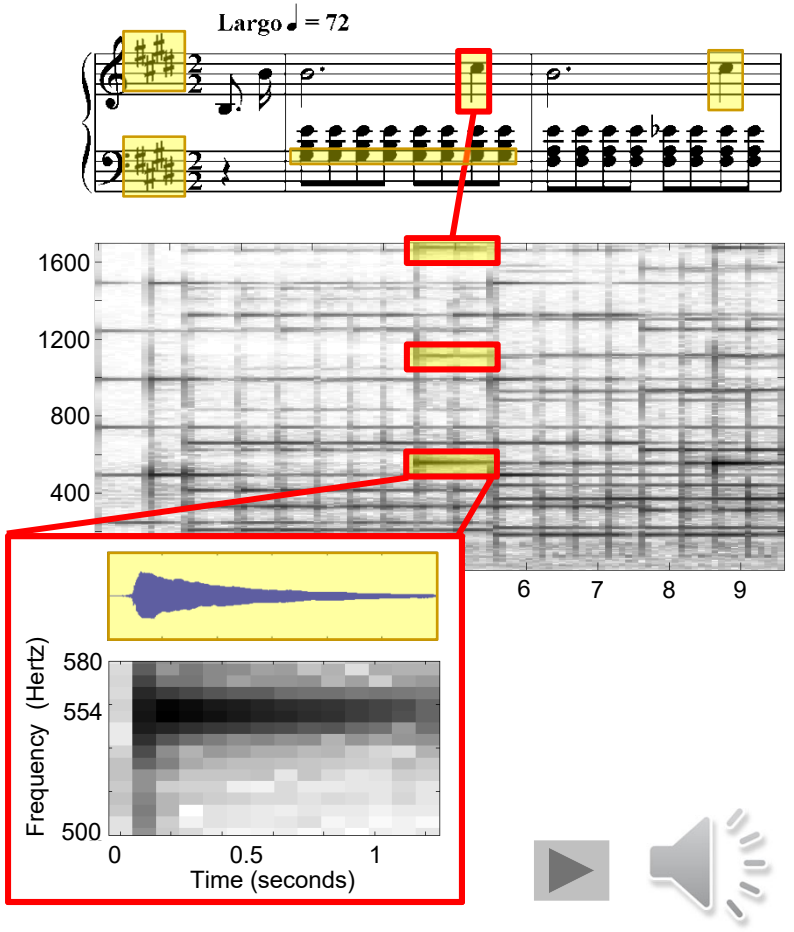
Parameterize audio signal using score's note events



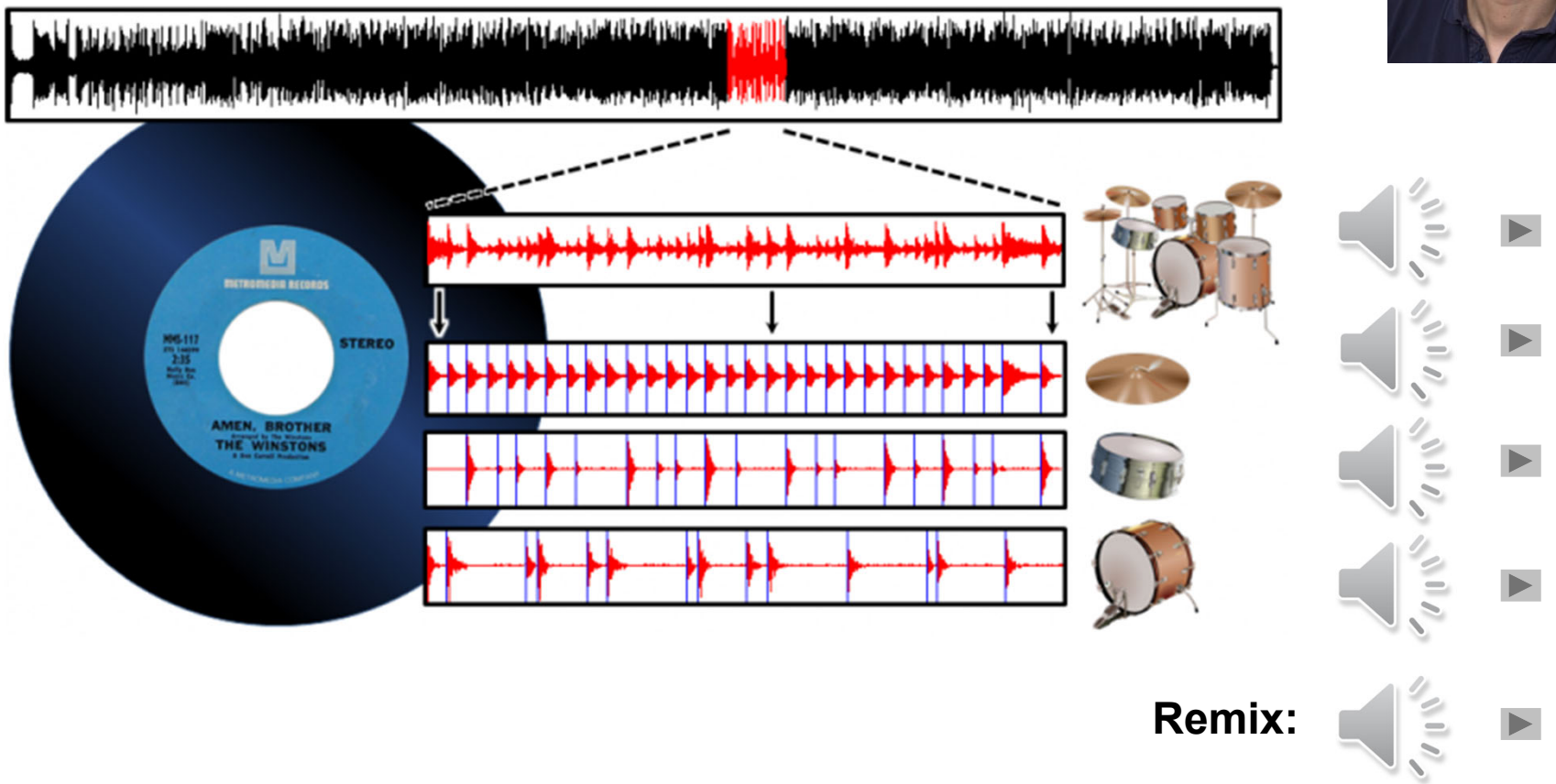
Score-Informed Audio Decomposition

Application: Audio editing





Informed Drum-Sound Decomposition

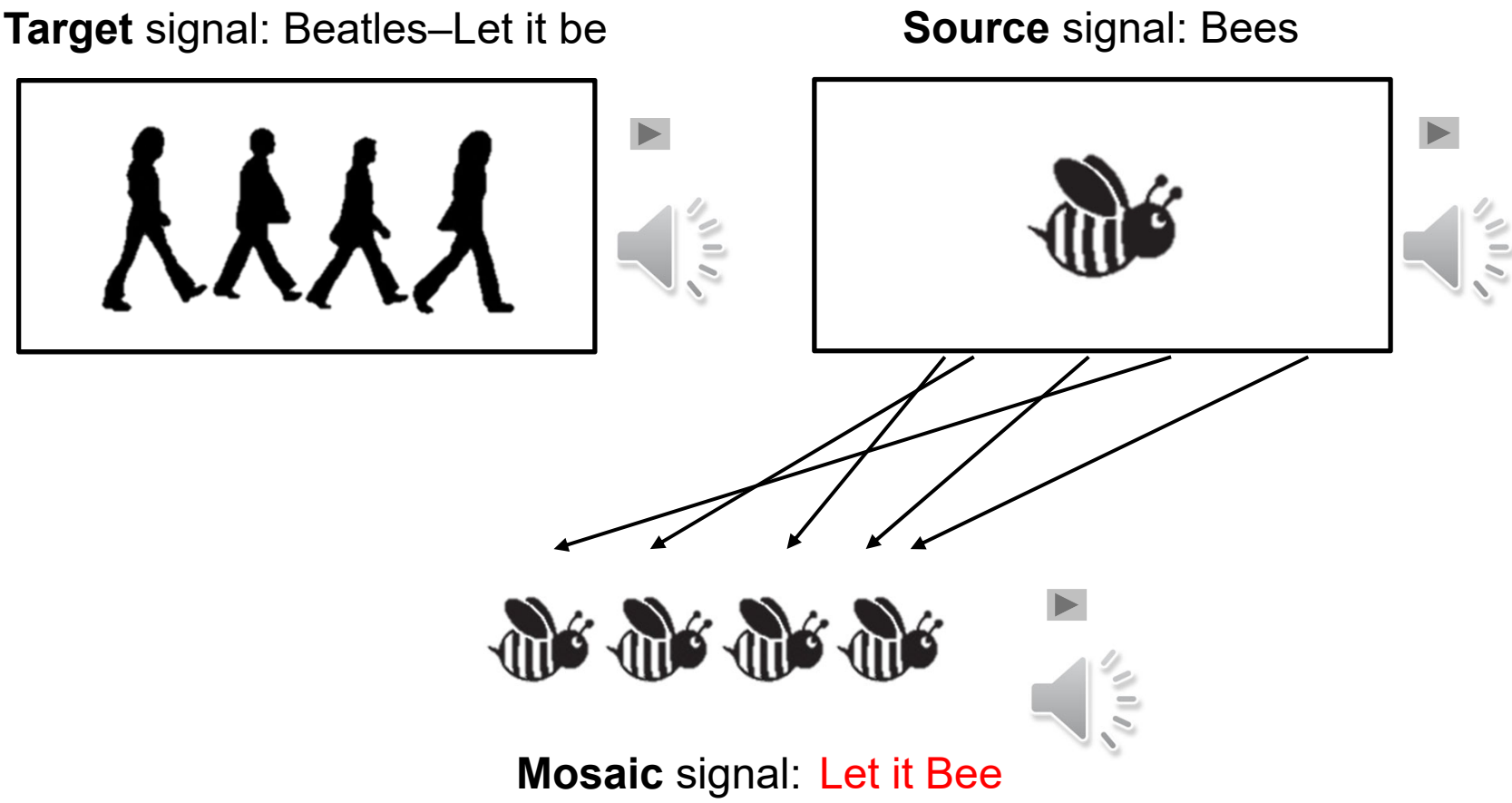


Remix:

Literature: [Dittmar/Müller, IEEE/ACM-TASLP 2016]

Demo: https://www.audiolabs-erlangen.de/resources/MIR/2016-IEEE-TASLP-DrumSeparation

Audio Mosaicing

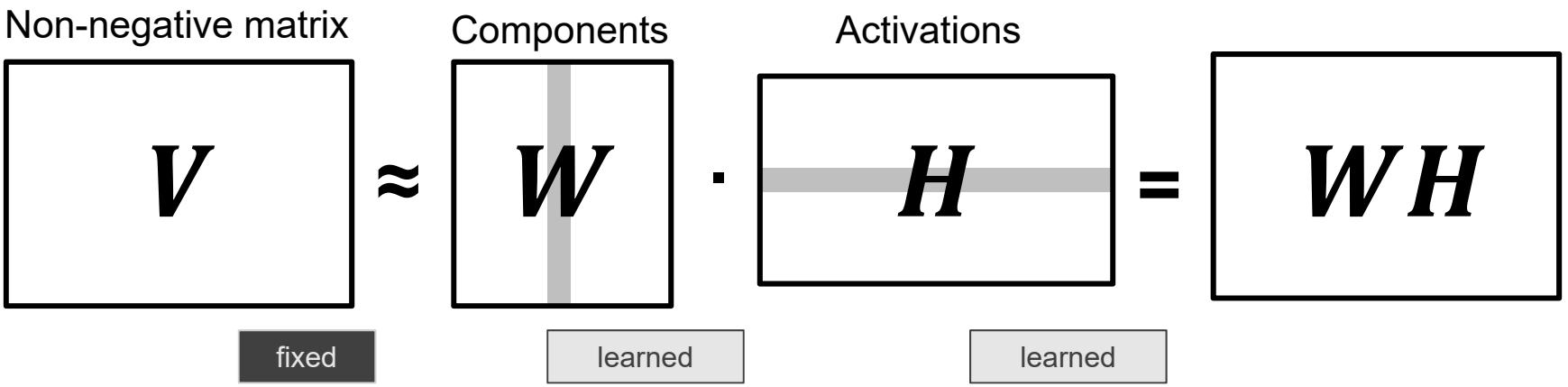


Literature: [Driedger/Müller, ISMIR 2015]

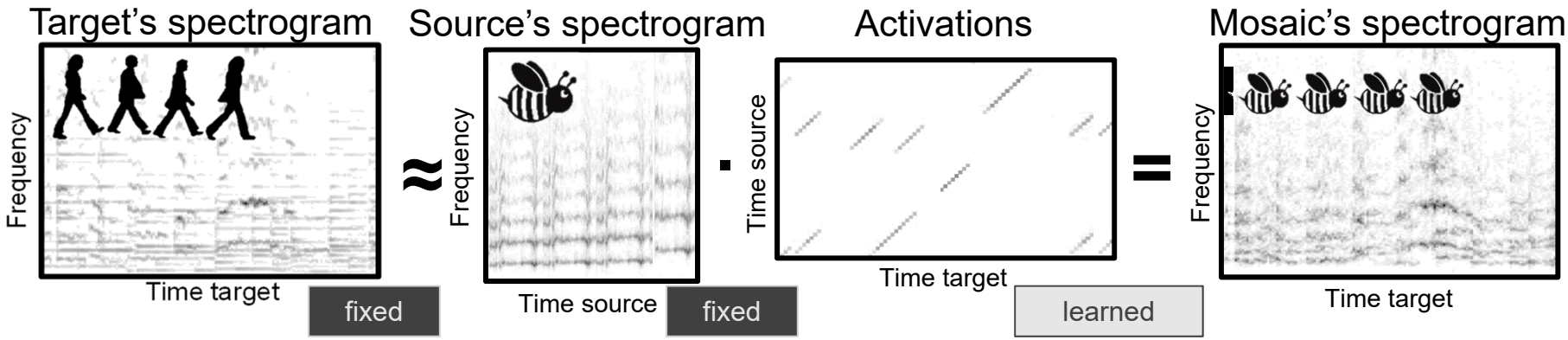
Demo: https://www.audiolabs-erlangen.de/resources/MIR/2015-ISMIR-LetItBee

NMF-Inspired Audio Mosaicing

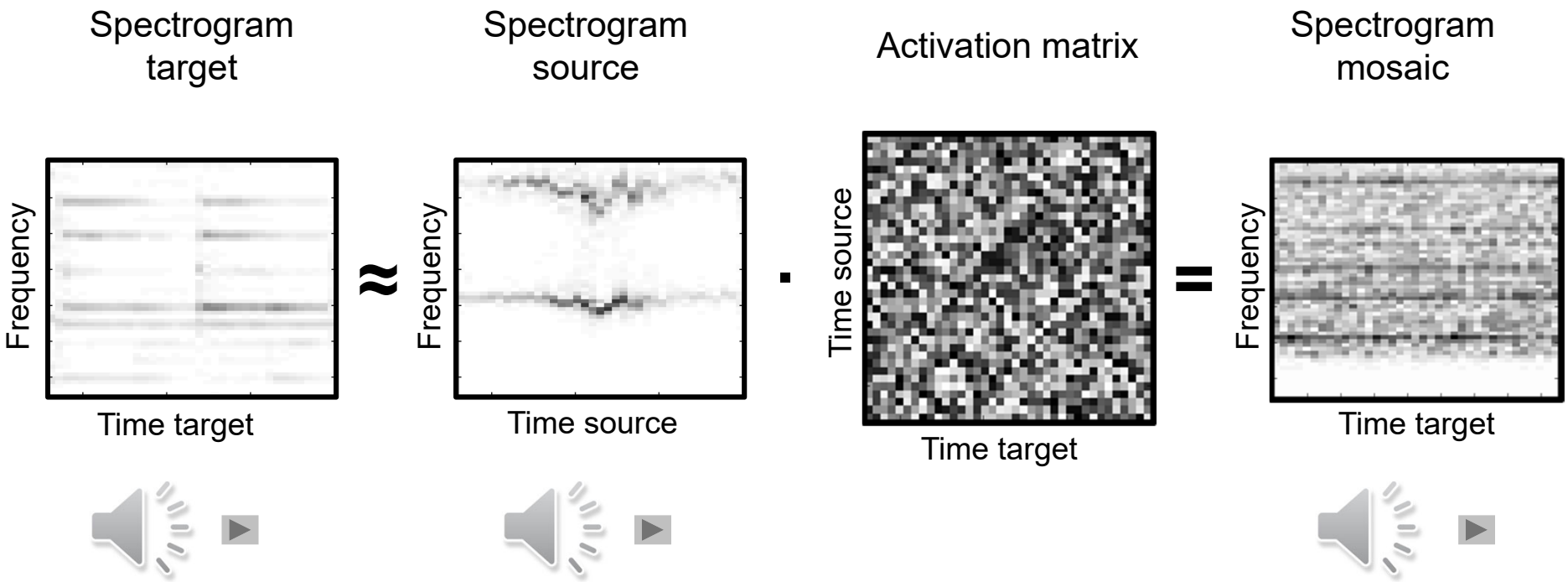
Non-negative matrix factorization (NMF)



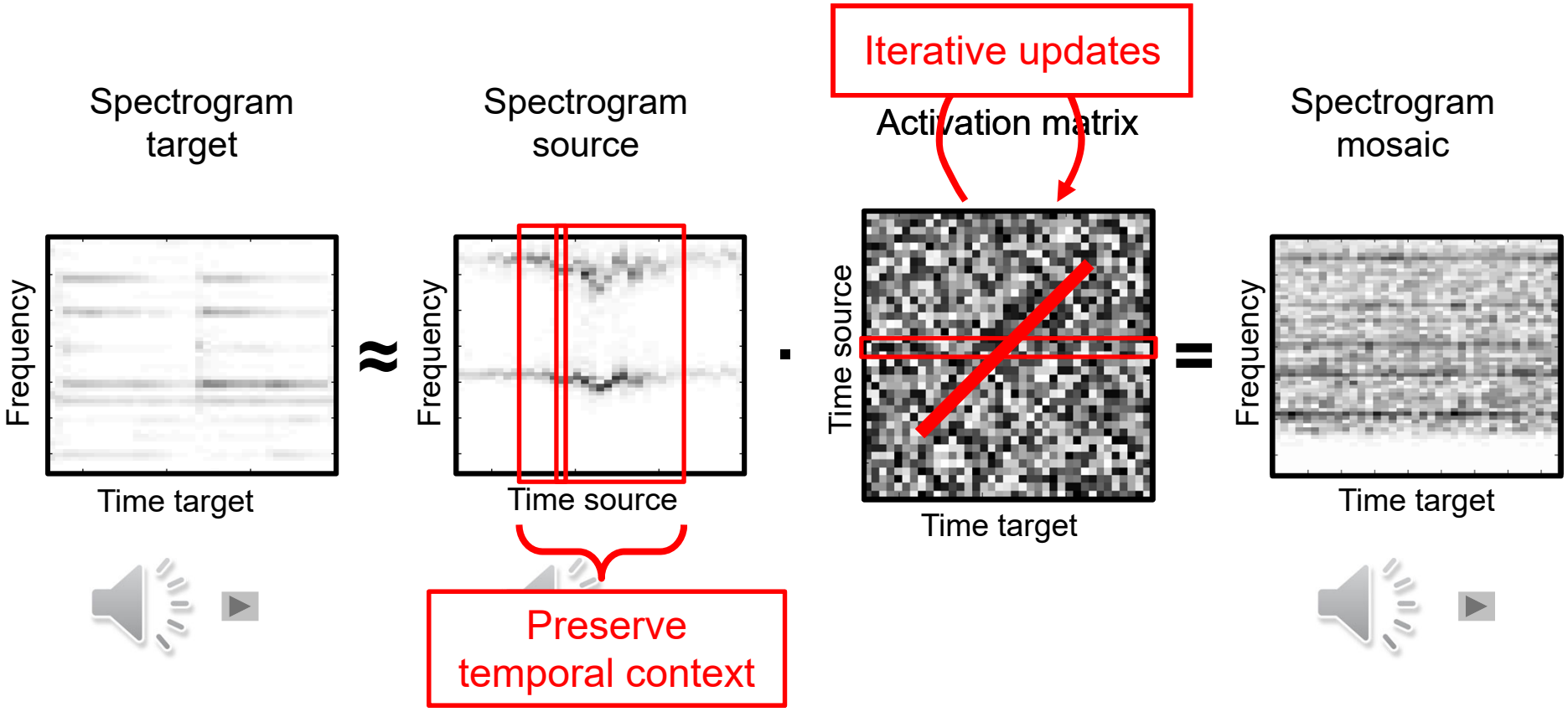
Proposed audio mosaicing approach



NMF-Inspired Audio Mosaicing

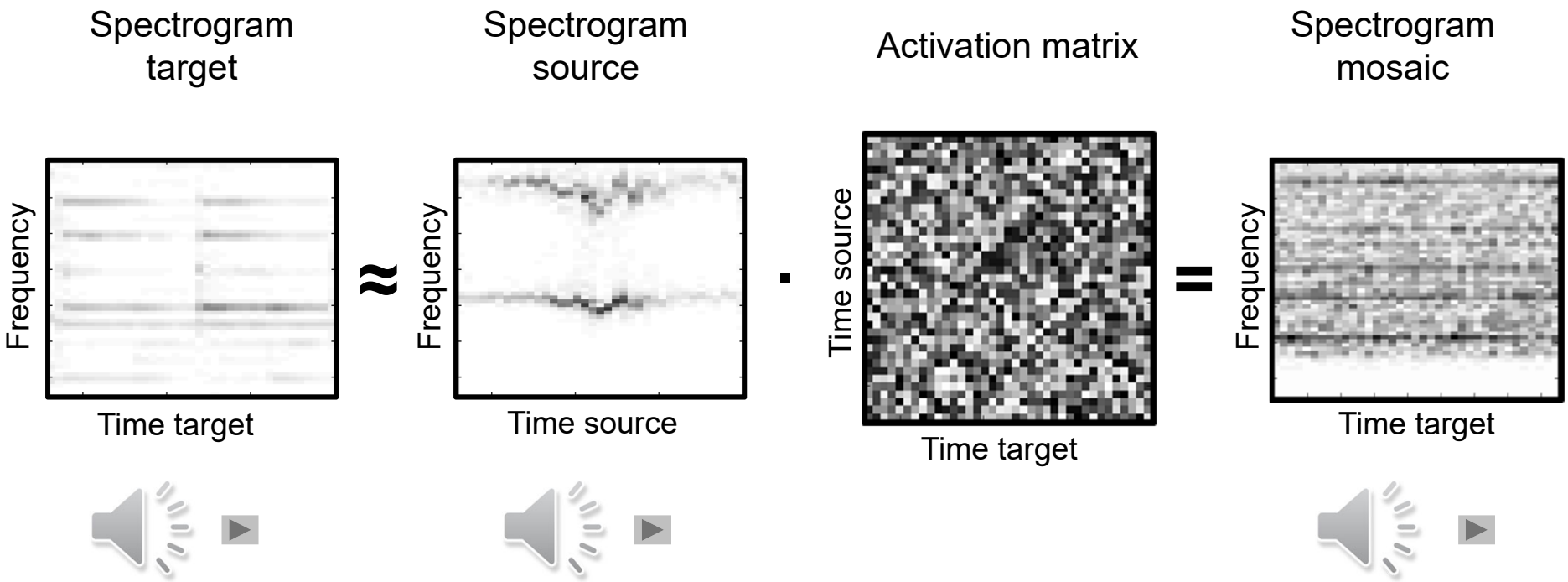


NMF-Inspired Audio Mosaicing

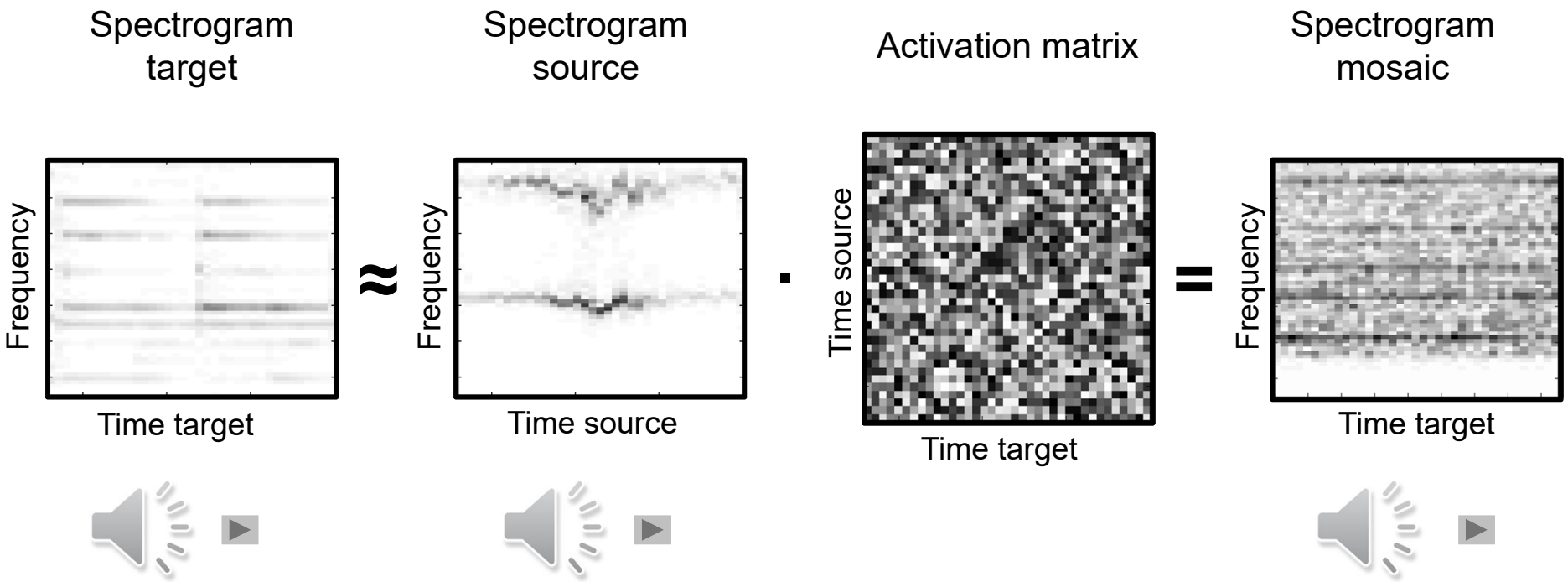


Core idea: support the development of sparse diagonal activation structures

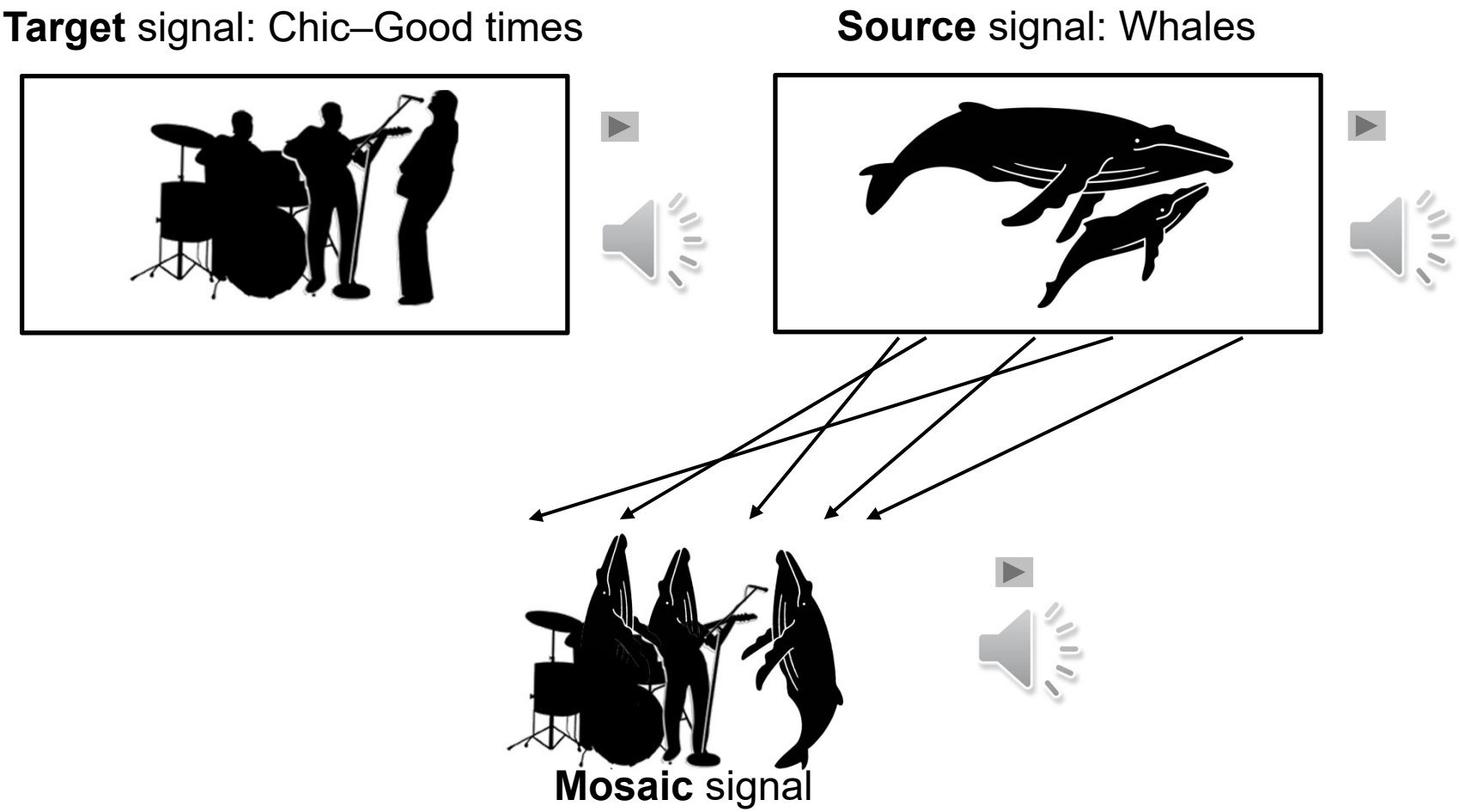
NMF-Inspired Audio Mosaicing



NMF-Inspired Audio Mosaicing



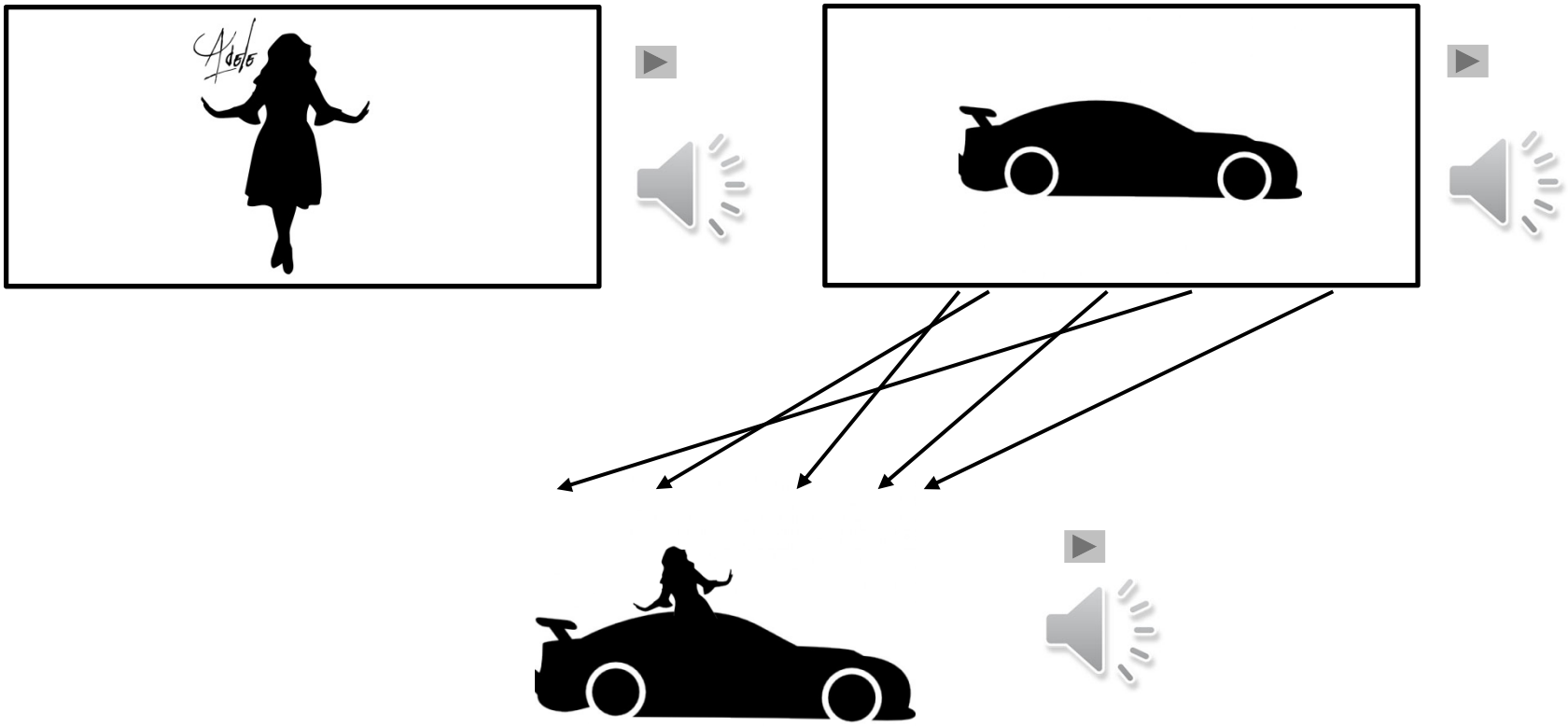
Audio Mosaicing



Audio Mosaicing

Target signal: Adele–Rolling in the Deep

Source signal: Race car

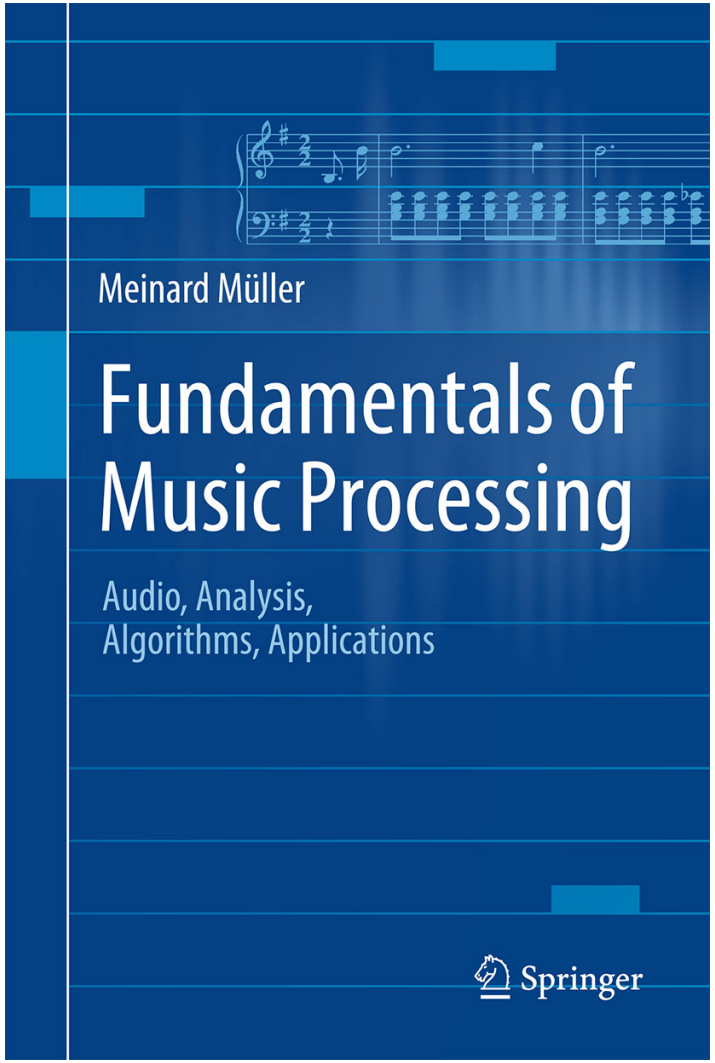


Mosaic signal

Links

- SiSEC: Signal Separation Evaluation Campaign
  https://www.sisec17.audiolabs-erlangen.de/
- MedleyDB: A Dataset of Multitrack Audio
  http://steinhardt.nyu.edu/marl/research/medleydb
- LibROSA (Python)
  https://librosa.github.io/librosa/

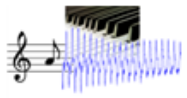
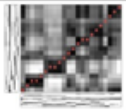
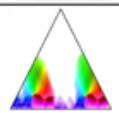
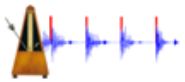
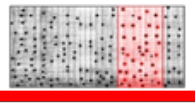
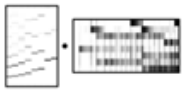
Book: Fundamentals of Music Processing



Meinard Müller
Fundamentals of Music Processing
Audio, Analysis, Algorithms, Applications
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de

Book: Fundamentals of Music Processing

Chapter		Music Processing Scenario
1		Music Represenations
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
483 p., 249 illus., hardcover
ISBN: 978-3-319-21944-8
Springer, 2015

Accompanying website:
www.music-processing.de