PAT 498/598 (Winter 2025)

Music & AI

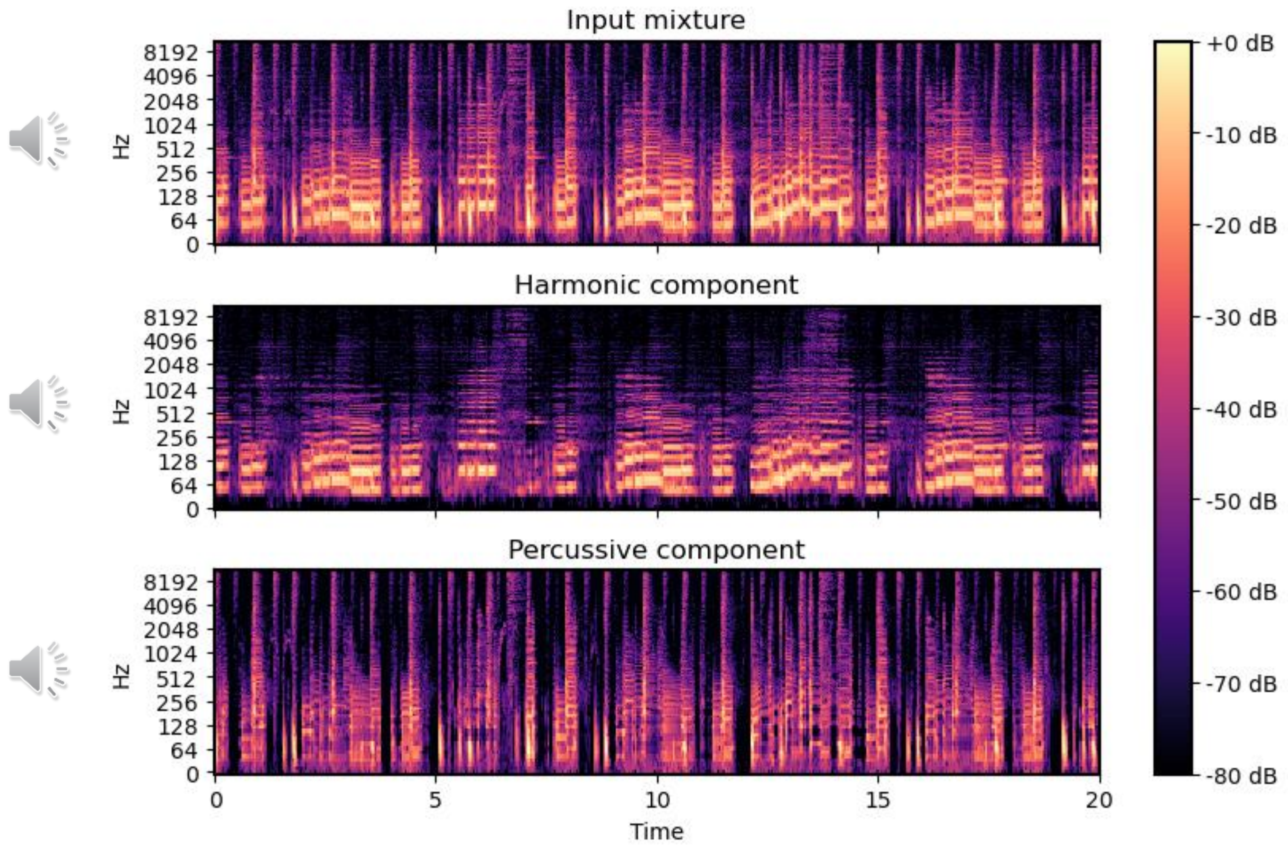
Lecture 13: Music Analysis

Instructor: Hao-Wen Dong

SCHOOL OF MUSIC, THEATRE & DANCE
PERFORMING ARTS TECHNOLOGY
UNIVERSITY OF MICHIGAN

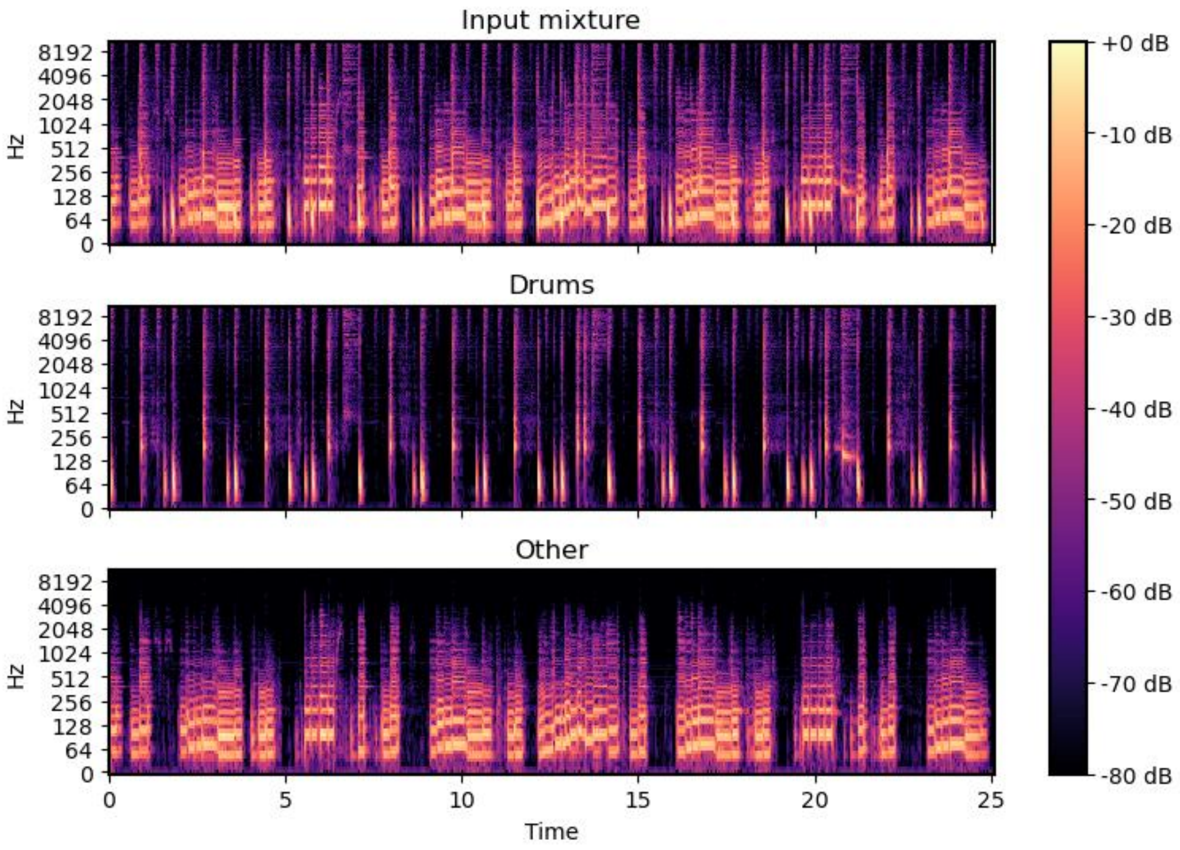
Homework 4: Source Separation

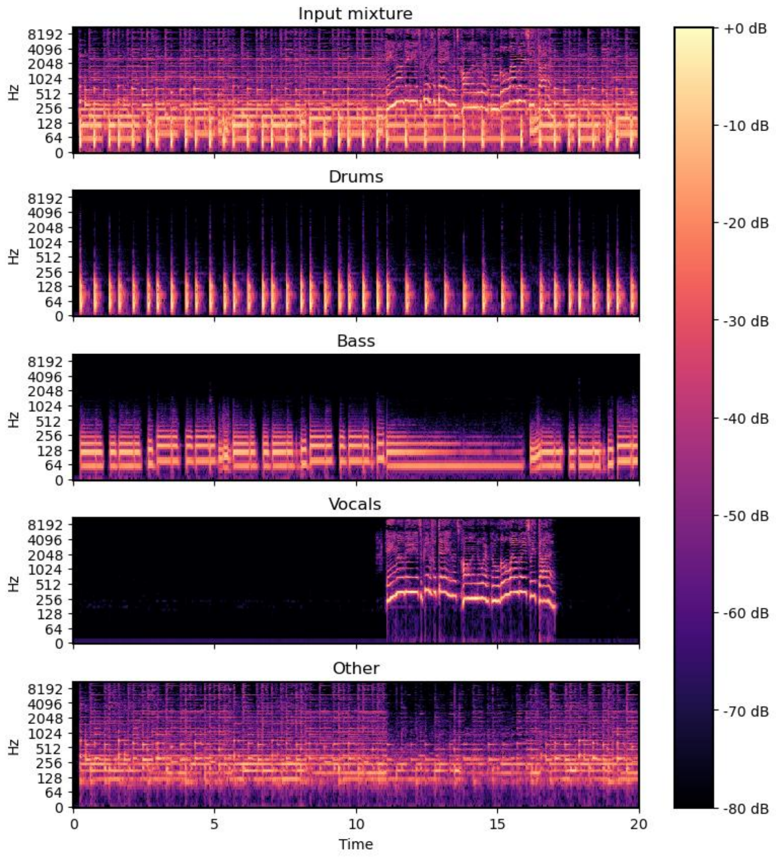
- **Part 1**: Harmonic-Percussive Source Separation (HPSS) using **librosa**



Homework 4: Source Separation

- **Part 2**: Music Source Separation using **Demucs**

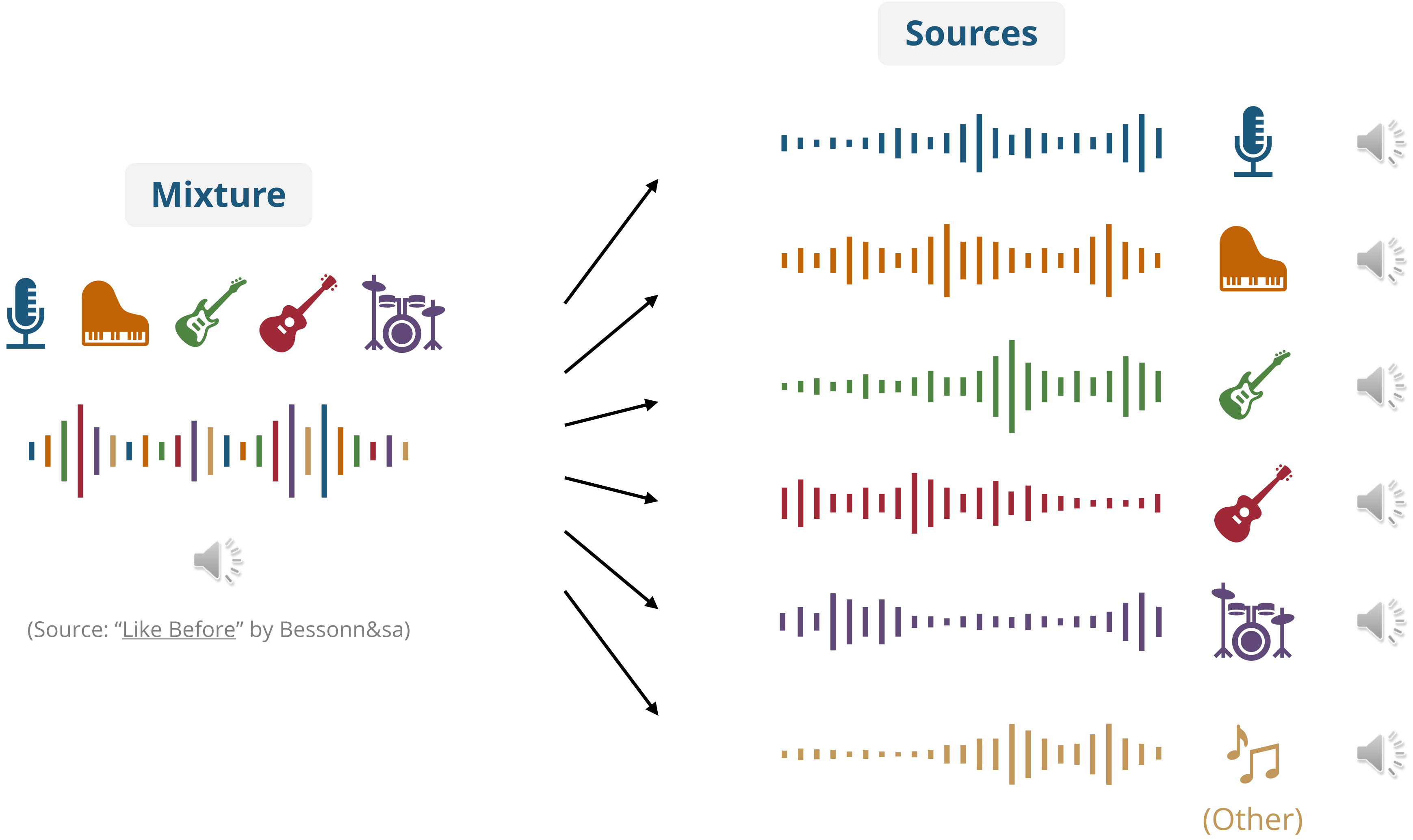




Homework 4: Source Separation

- Instructions will be released on the course website
- Please submit your work to Gradescope
- Due at **11:59pm ET** on **February 28**
- Late submissions: **1 point deducted per day**
- No late submission is allowed a week after the due date

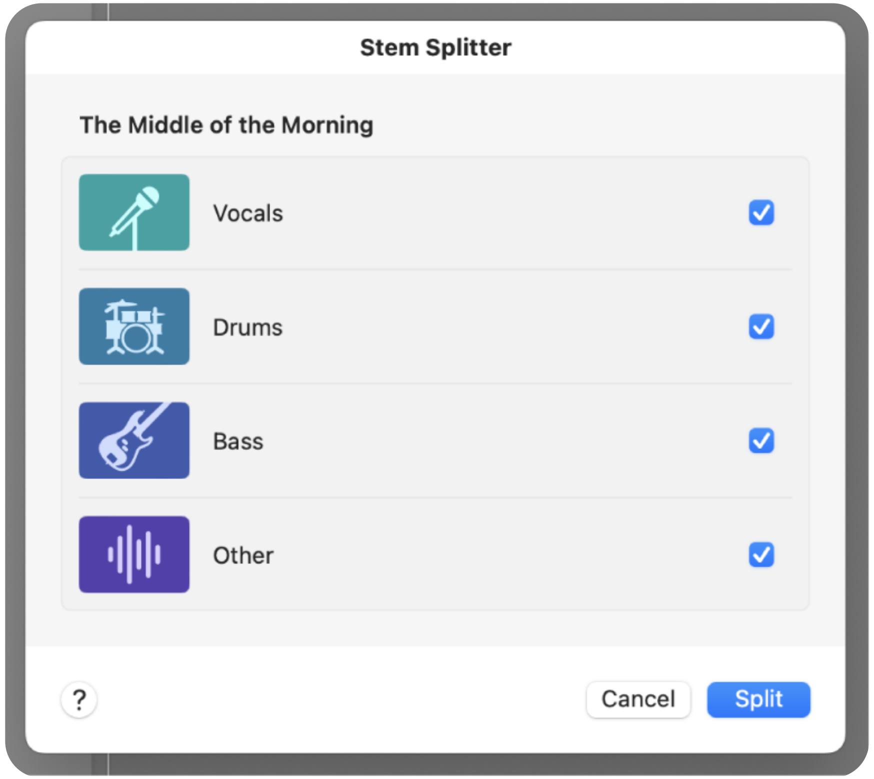
(Recap) Source Separation



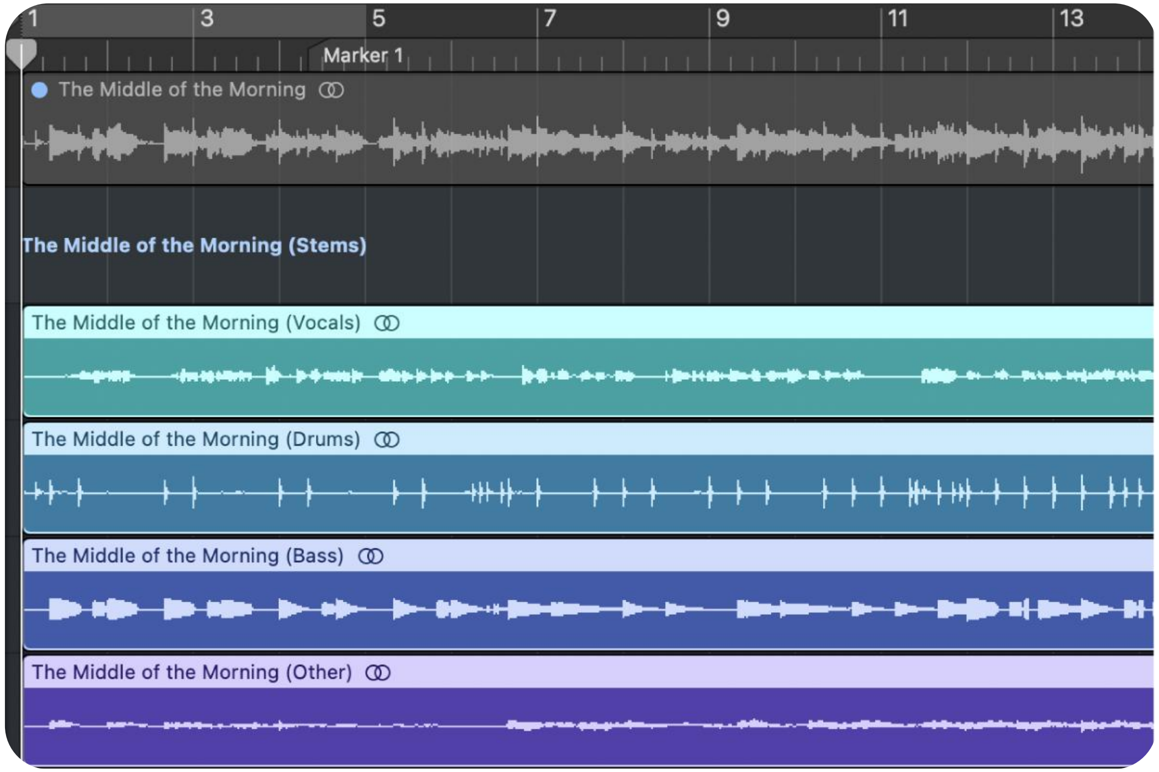
(Source: "Like Before" by Bessonn&sa)

Bessonn&sa, CC BY-NC-SA, via Jamendo
Illustration inspired by Ethan Manilow, Prem Seetharman, and Justin Salamon's Tutorial on "Open Source Tools & Data for Music Source Separation" at ISMIR 2020.

(Recap) Stem Splitter in Logic Pro



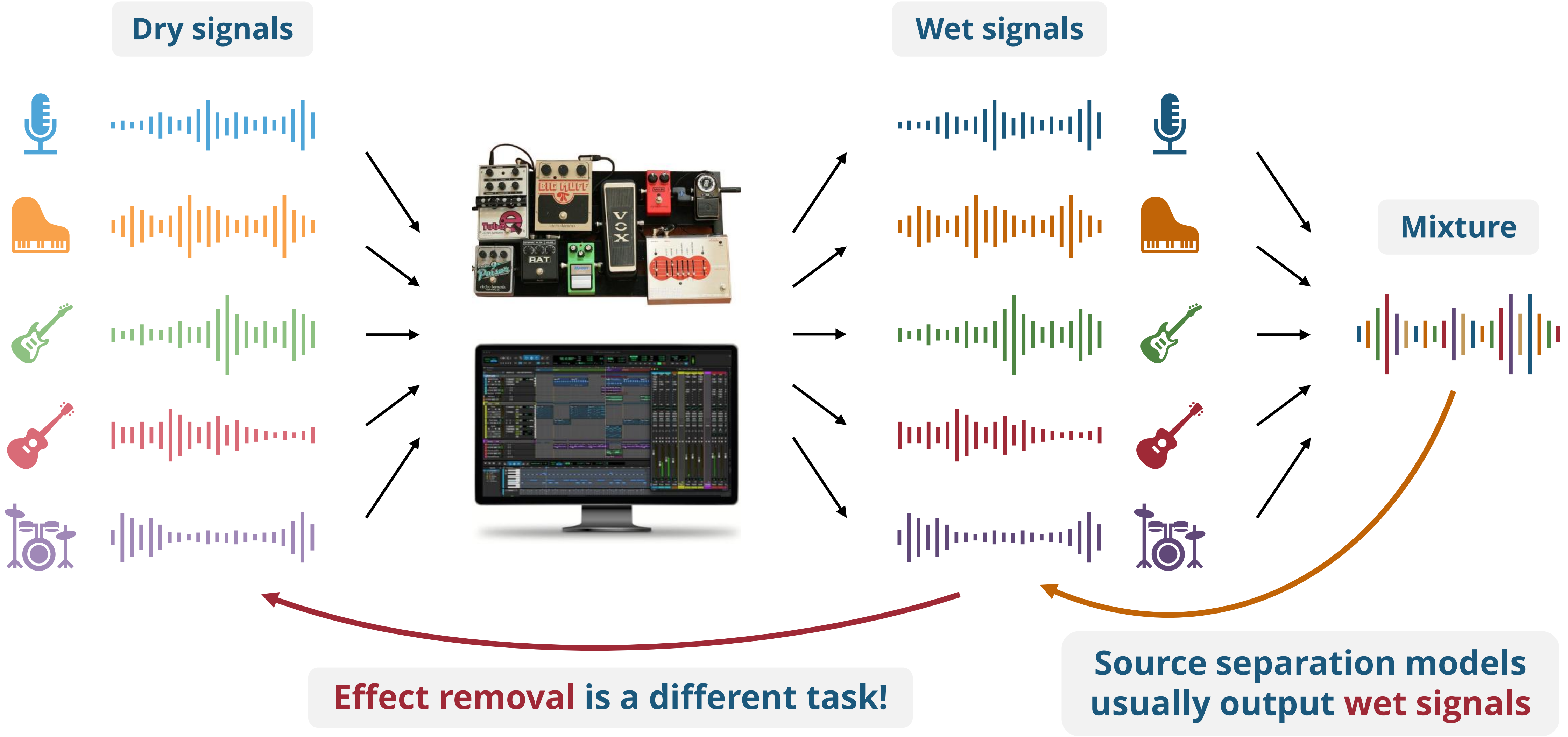
(Source: Logic Pro User Guide)



(Source: Logic Pro User Guide)

support.apple.com/guide/logicpro/extract-vocal-instrumental-stems-stem-lgcp61bae908/mac

(Recap) Source Separation does NOT Remove Effects

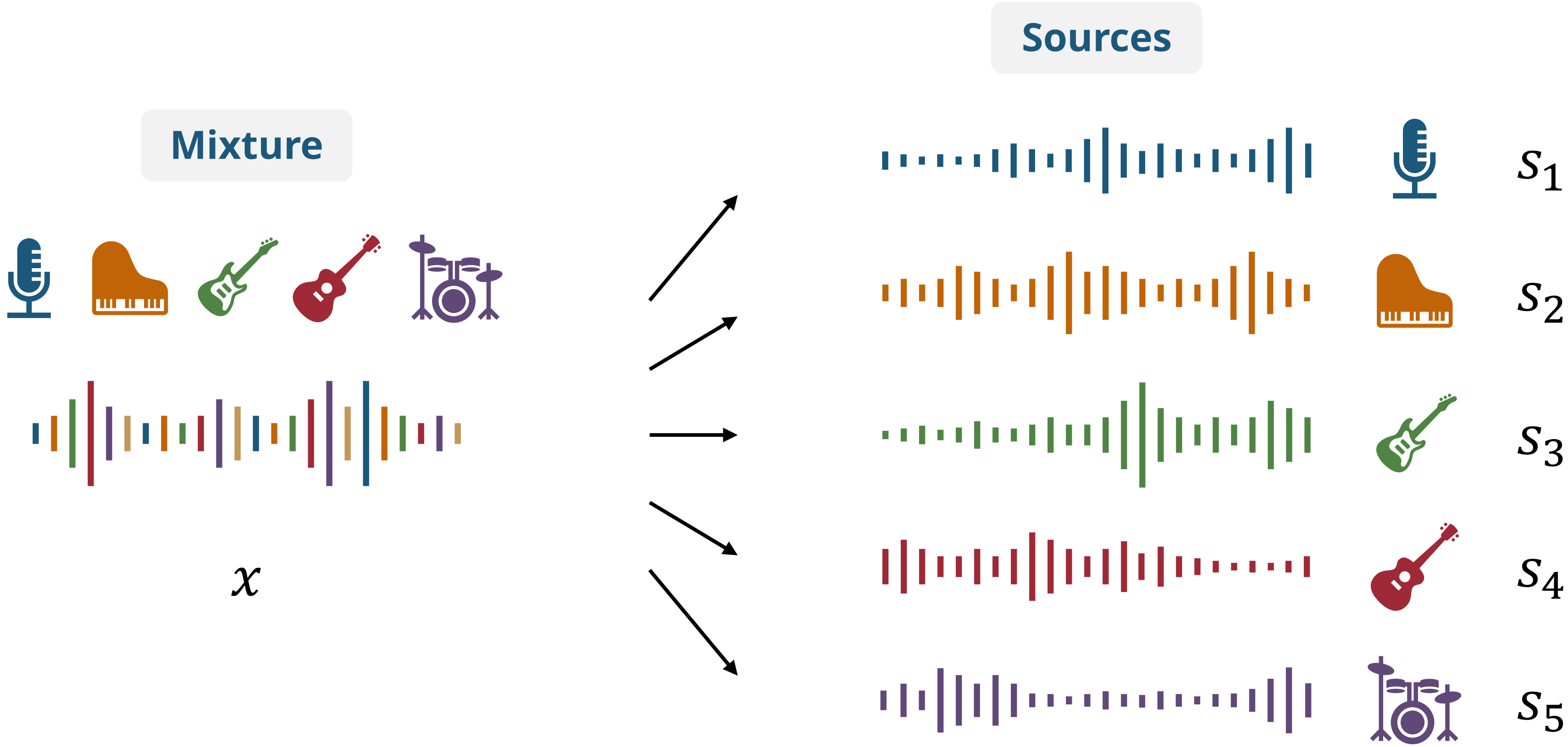


(Recap) Moises

- 🤩 **Free Moises Pro license** until Summer 2025
- Register at studio.moises.ai/claim-trial/UMichFree/monthly/
  - Use your **U-M email** (@umich.edu)
  - Sign up in your **desktop browser**
  - Ignore the prompt to upgrade your account
  - **Deadline to sign up: March 14**

moises

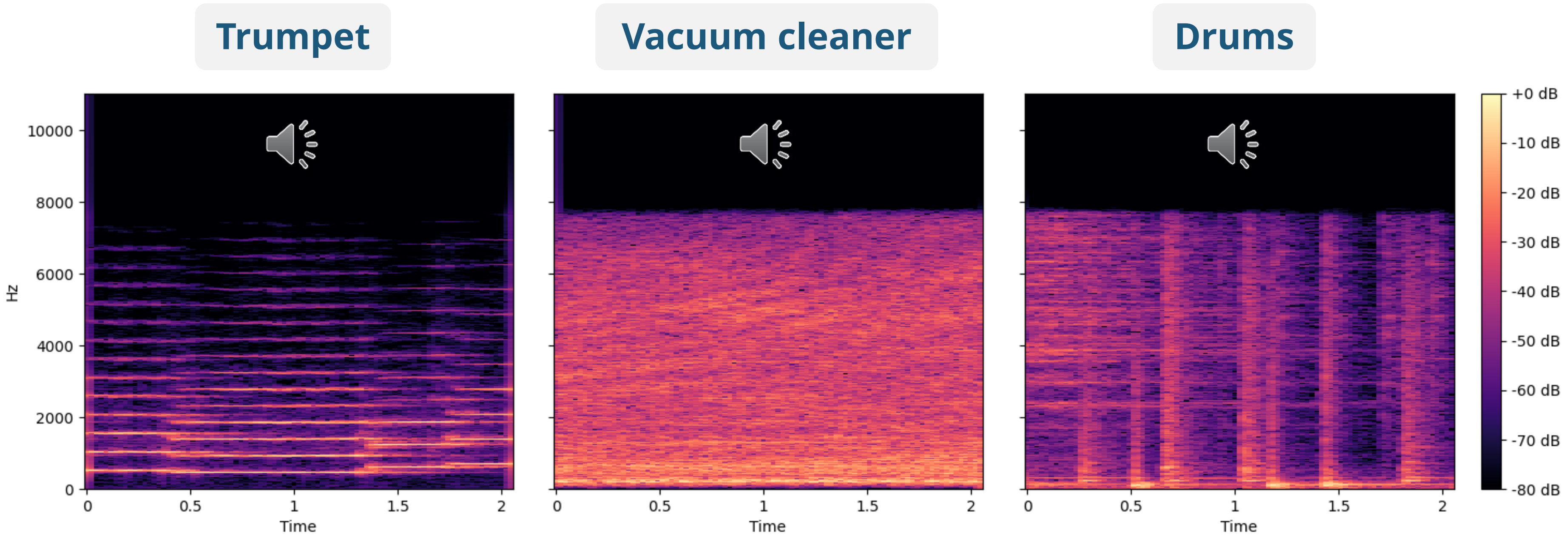
(Recap) Mathematical Formulation



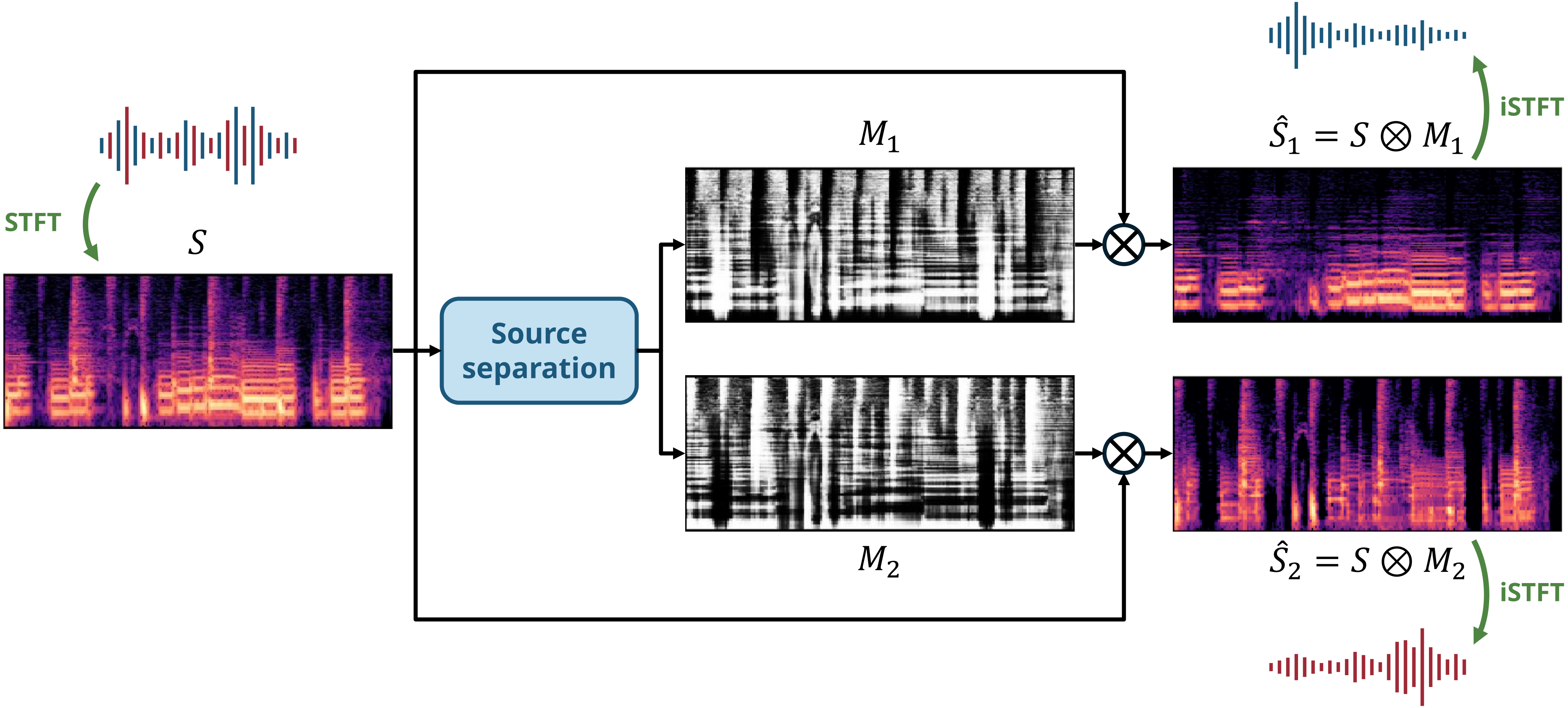
$$x = s_1 + s_2 + s_3 + s_4 + s_5 = \sum_i s_i$$

(Recap) Source Separation is an Ill-posed Problem

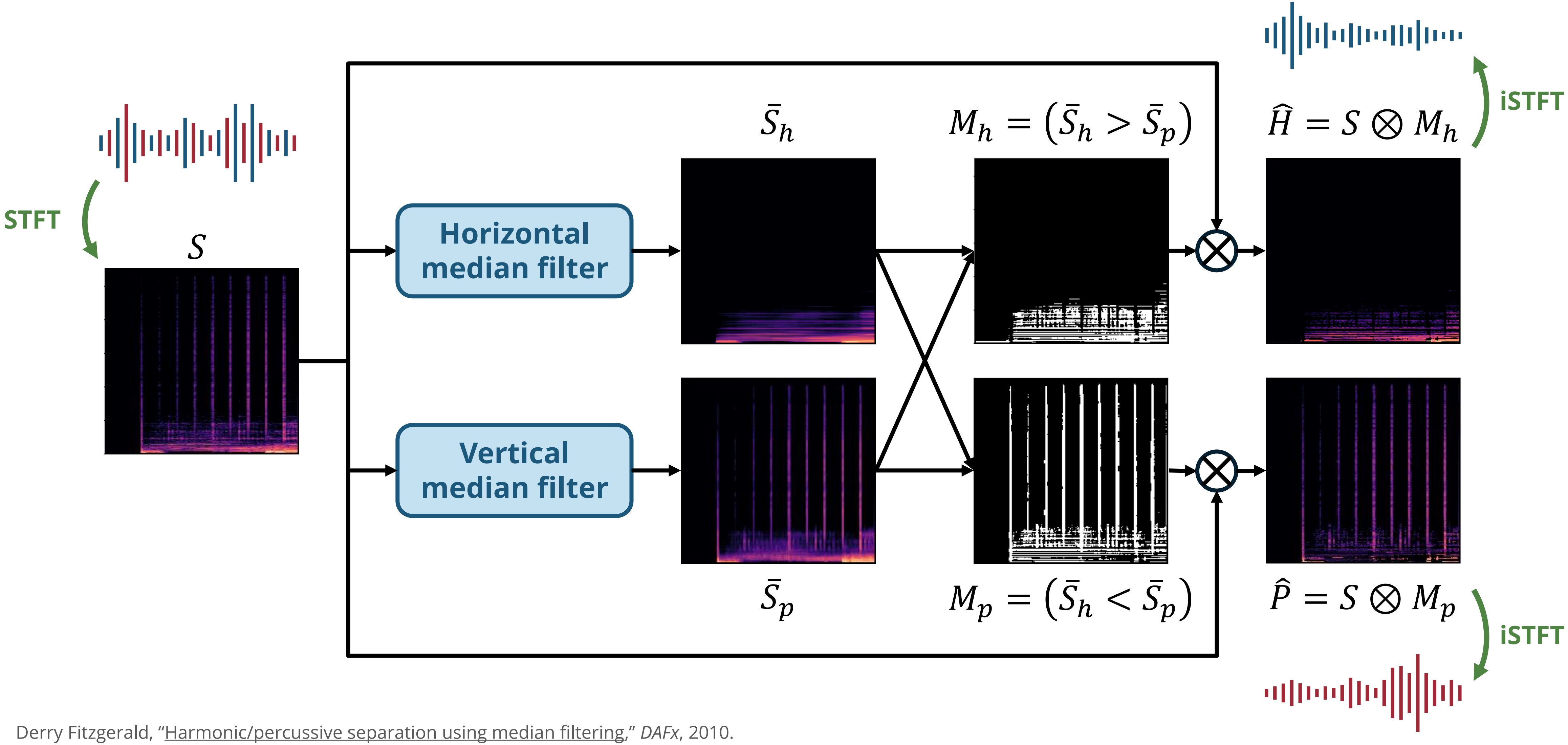
- There are **more than one solution** to $x = s_1 + s_2 + \cdots + s_N$
  - In fact, there are infinite possibilities
- However, we do know **what's more likely than another**!



(Recap) Time-Frequency Masking

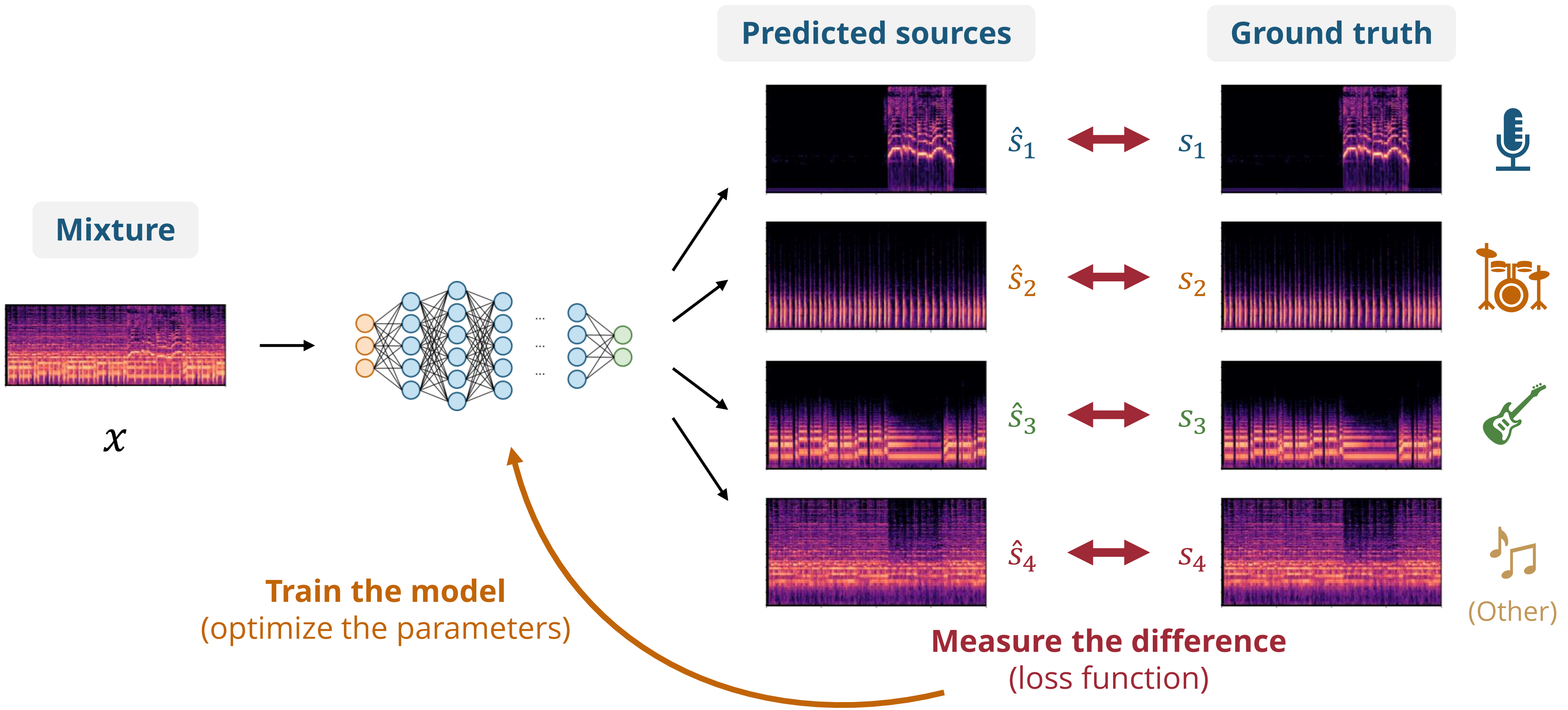


(Recap) Harmonic-Percussive Separation (Fitzgerald et al., 2010)

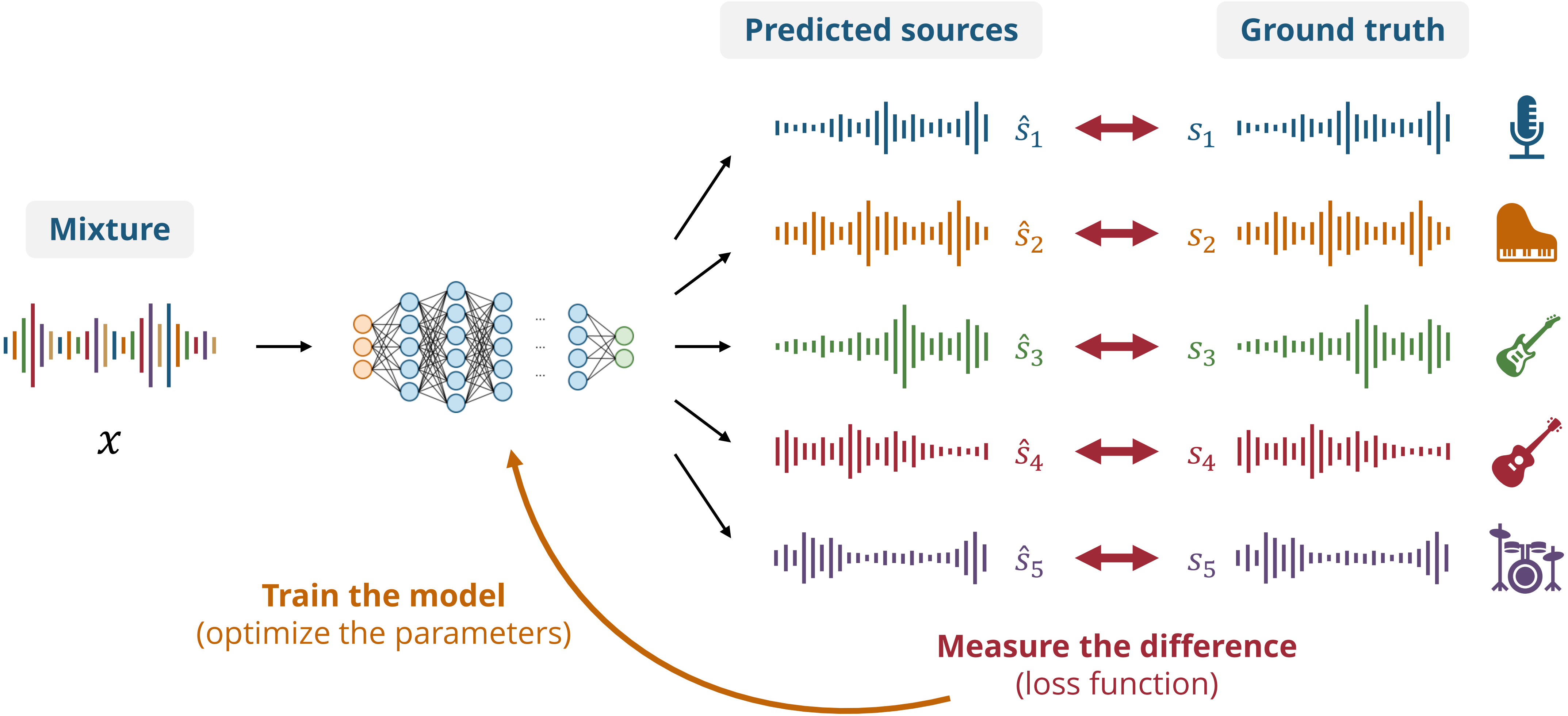


Derry Fitzgerald, "Harmonic/percussive separation using median filtering," *DAFx*, 2010.
Jonathan Driedger, Meinard Müller, Sascha Disch, "Extending Harmonic-Percussive Separation of Audio Signals," *ISMIR*, 2014.

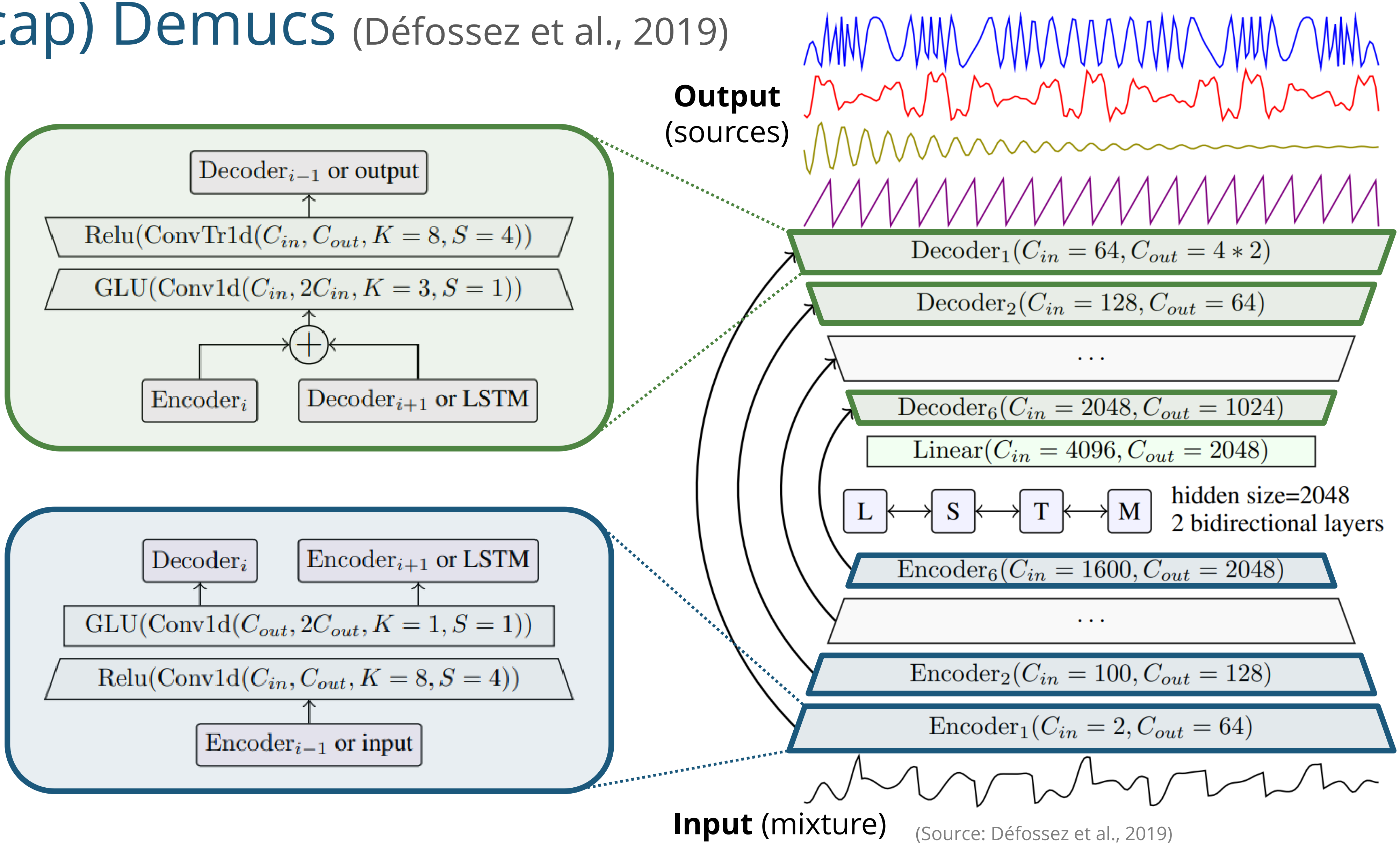
(Recap) Deep Learning Based Source Separation



(Recap) Deep Learning Based Source Separation



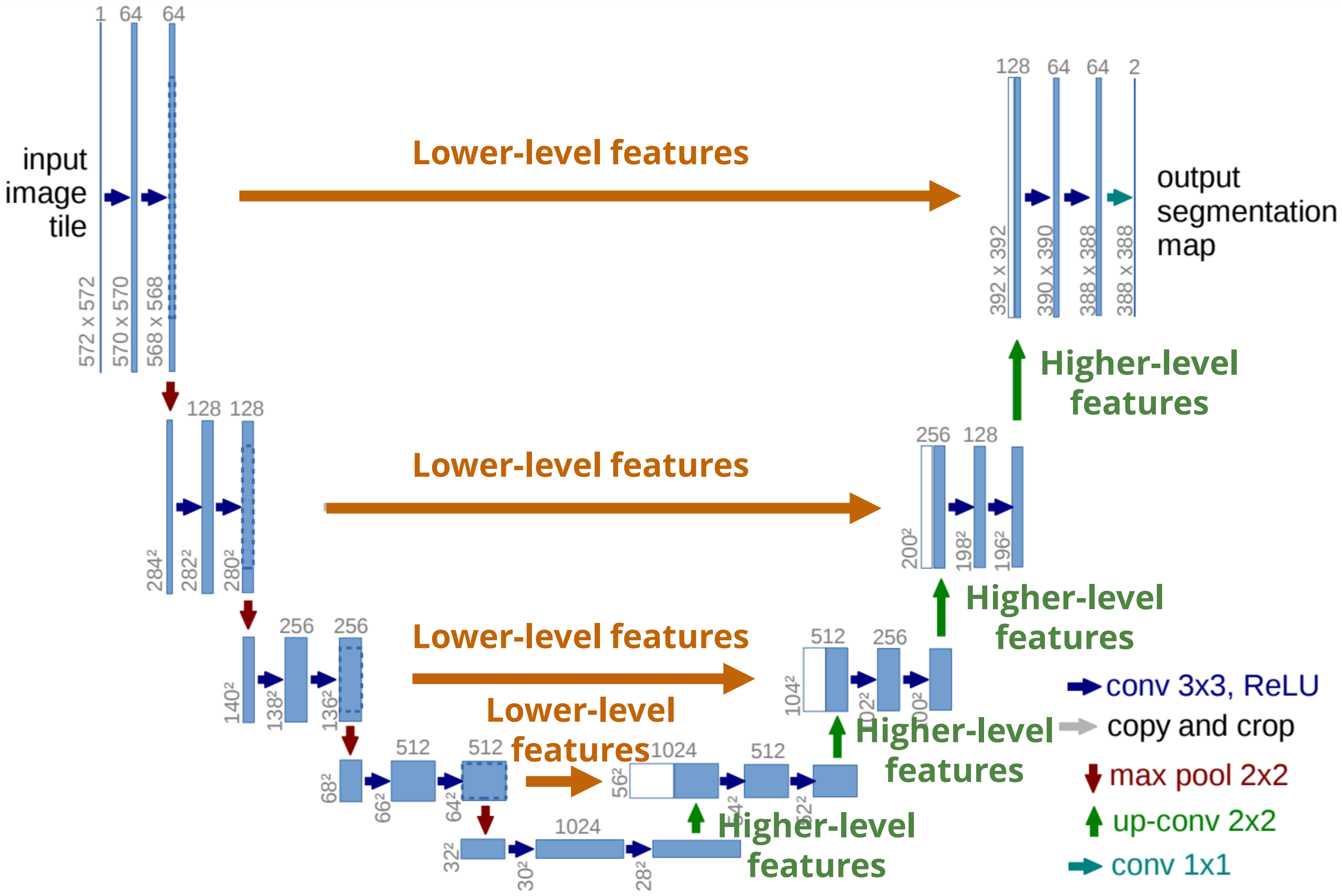
(Recap) Demucs (Défossez et al., 2019)



(Source: Défossez et al., 2019)

Alexandre Défossez, Nicolas Usunier, Léon Bottou, and Francis Bach, "Music Source Separation in the Waveform Domain," *arXiv preprint arXiv:1911.13254*, 2019.

(Recap) U-Net (Ronneberger et al., 2015)



(Source: Ronneberger et al., 2015)

Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *MICCAI*, 2015.

Datasets

- MIR-1K
- MedleyDB
- iKala
- DSD100
- MUSDB18 & MUSDB18-HQ
- MoisesDB
- Synthetic: Slakh2100, SynthSOD

Ethan Manilow, Prem Seetharman, and Justin Salamon, "Open Source Tools & Data for Music Source Separation," *Tutorials of ISMIR*, 2020.

(Recap) Choral Separation (Chen et al., 2022)

Demo

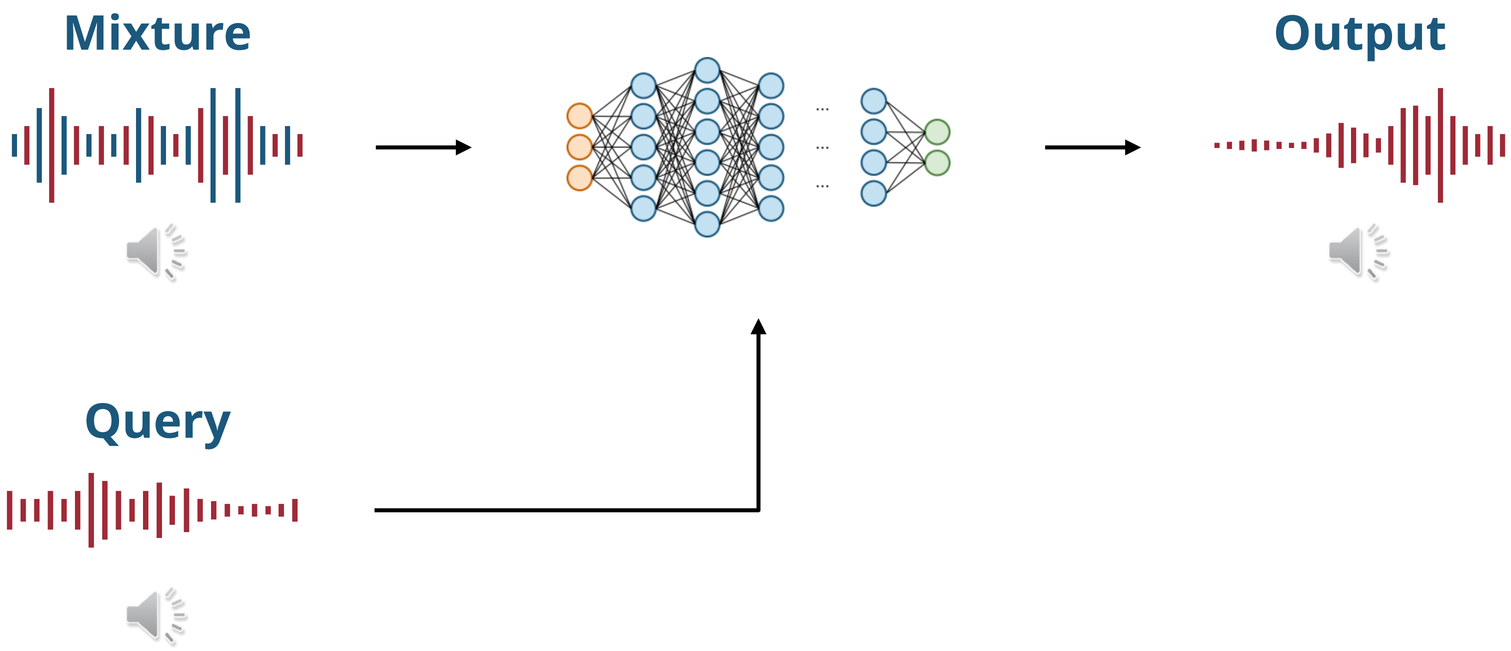
Mixture		**Soprano**	**Alto**	**Tenor**	**Bass**
	→				
	→				

Data Augmentation

SoundFont	**Standard**	**Expressive** (vowels only)	**Expressive** (words)

Ke Chen, Hao-Wen Dong, Yi Luo, Julian McAuley, Taylor Berg-Kirkpatrick, Miller Puckette, and Shlomo Dubnov, "Improving Choral Music Separation through Expressive Synthesized Data from Sampled Instruments," *ISMIR*, 2022.
retrocirce.github.io/cms_demo/

(Recap) Query-by-Audio Source Separation (Chen et al., 2022)

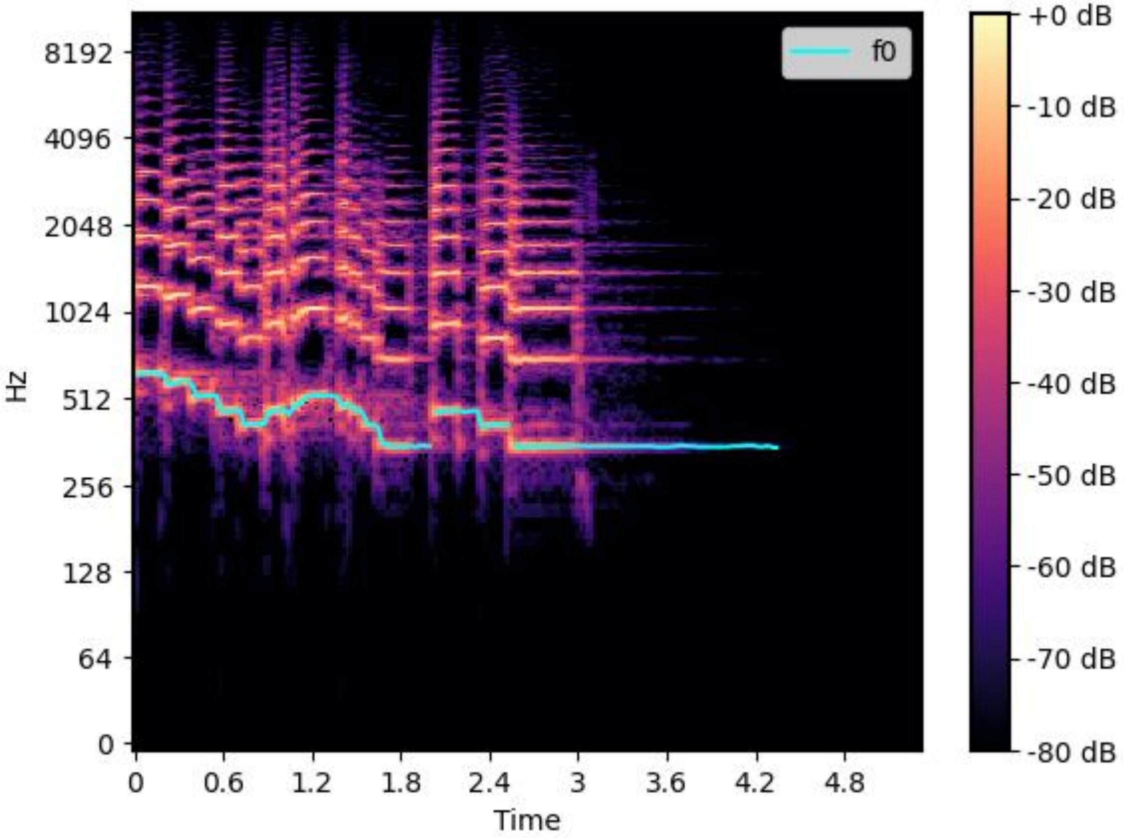


Ke Chen, Xingjian Du, Bilei Zhu, Zejun Ma, Taylor Berg-Kirkpatrick, and Shlomo Dubnov, "Zero-shot Audio Source Separation through Query-based Learning from Weakly-labeled Data," *AAAI*, 2022.

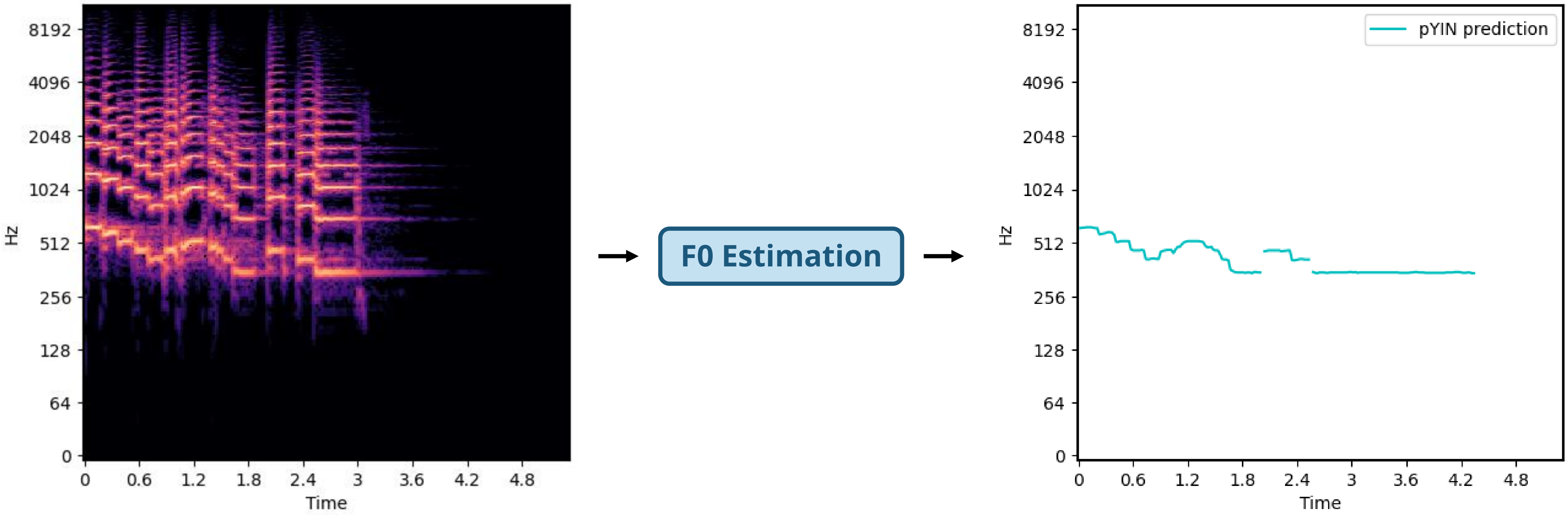
Music Transcription

Fundamental Frequency (F0)

- **Definition**: The **lowest frequency component** of a waveform



Fundamental Frequency (F0) Estimation



Auto-tune Artist

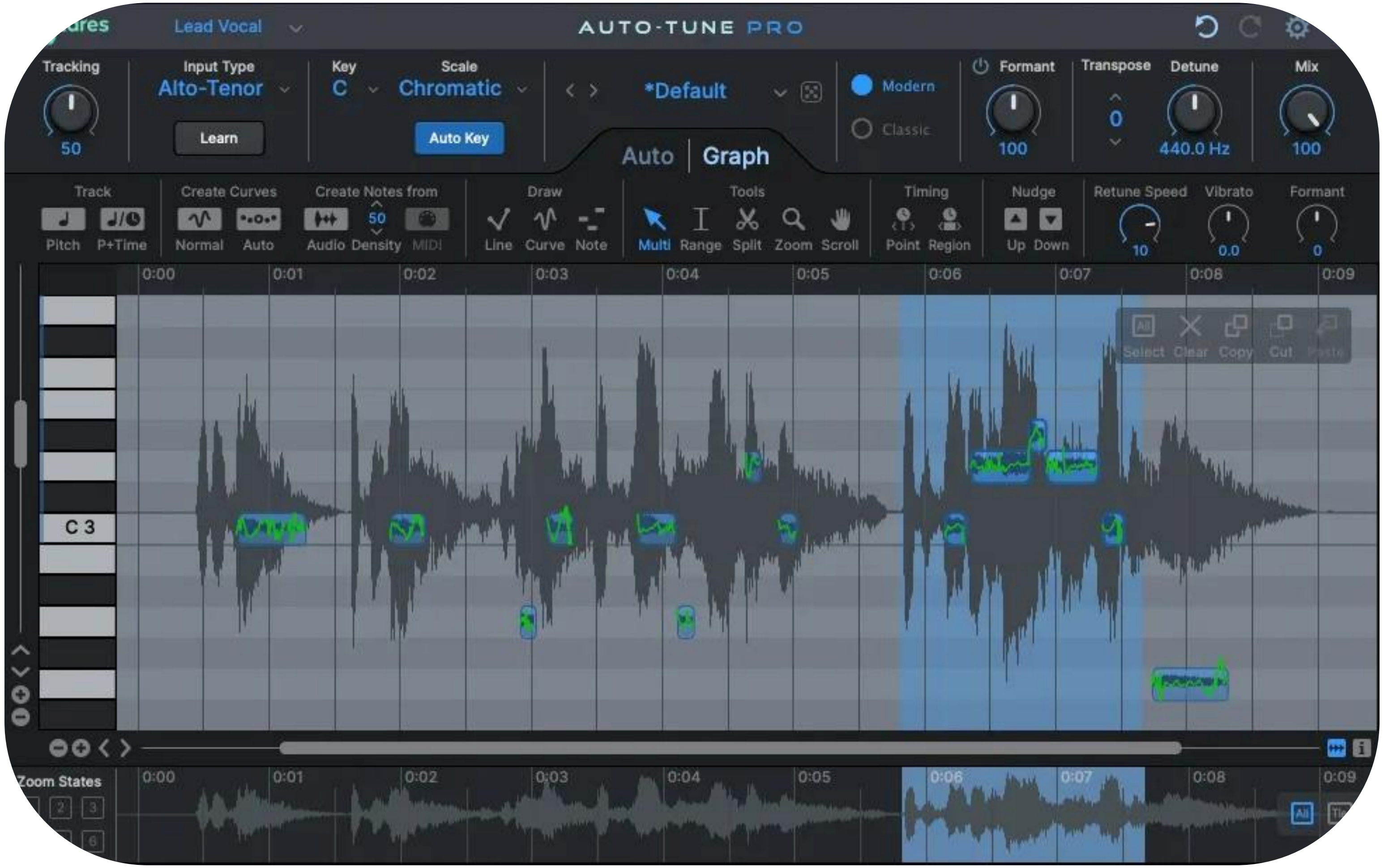
(Source: Antares Audio Technologies)

antarestech.com/documentation/auto-tune-artist/basic-view-controls

Auto-tune

youtube.com/shorts/Kg8OSbKRETA

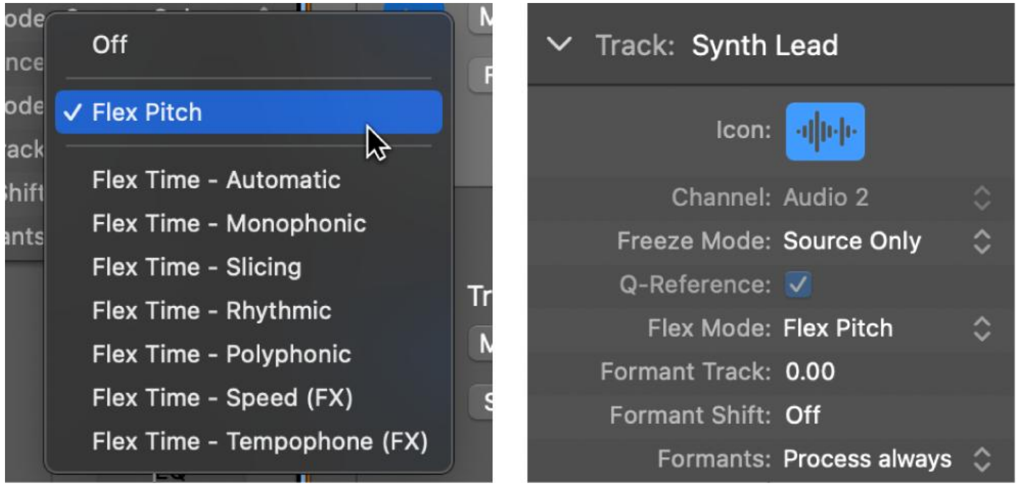
Auto-tune Pro



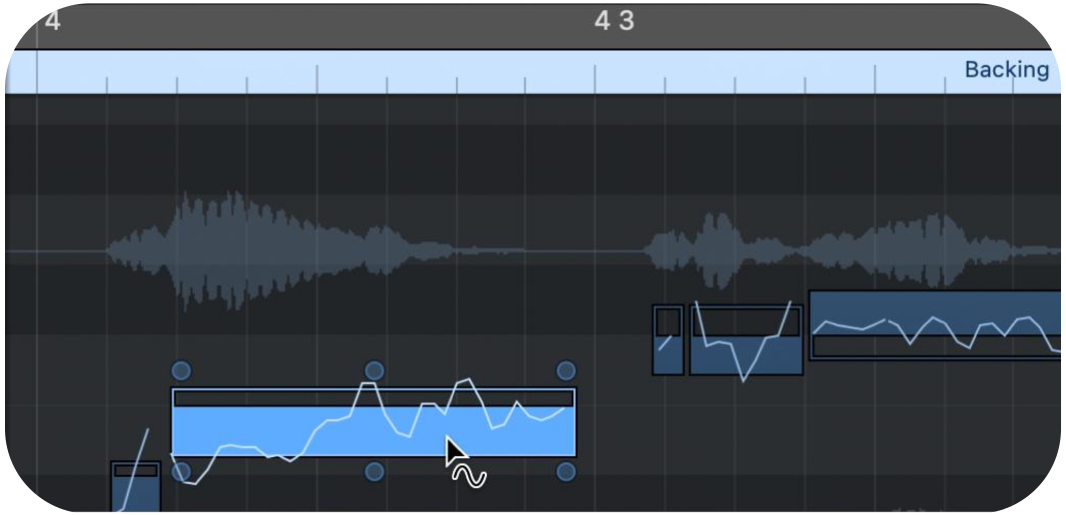
(Source: Antares Audio Technologies)

antarestech.com/documentation/auto-tune-pro/chapter-2

Pitch Correction in Logic Pro



(Source: Logic Pro User Guide)

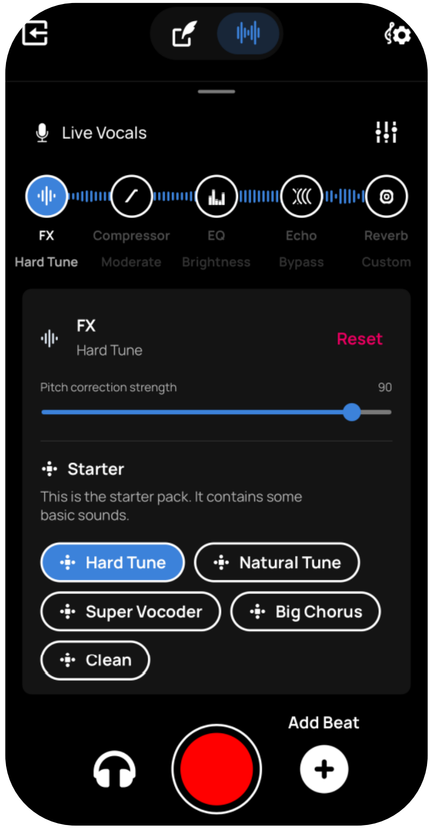


(Source: Logic Pro User Guide)

support.apple.com/guide/logicpro/edit-pitch-and-timing-with-flex-pitch-lgcpc53e6bef/mac
support.apple.com/guide/logicpro/flex-pitch-algorithm-and-parameters-lgcpba8e3301/mac

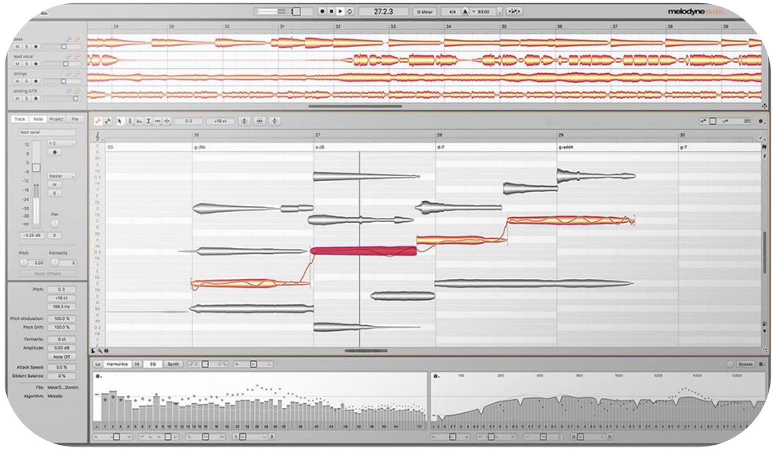
Other Auto-tune & Pitch Correction Tools

Voloco



(Source: Voloco User Manual)

Melodyne



(Source: Celemony)

iZotope Nectar

(Source: iZotope Team)

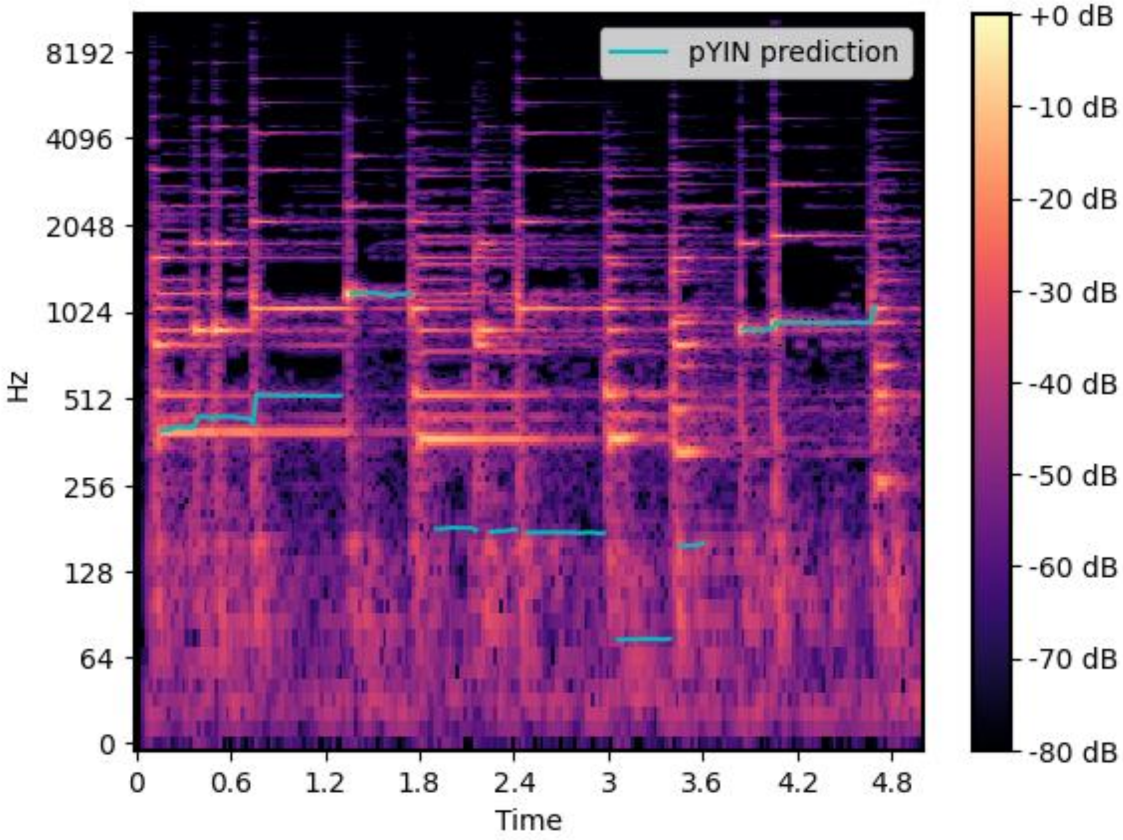
celemony.com/en/melodyne/what-is-melodyne
izotope.com/en/learn/why-upgrade-to-nectar-4.html
resonantcavity.com/assets/docs/voloco-user-manual.pdf

F0 Estimation Models

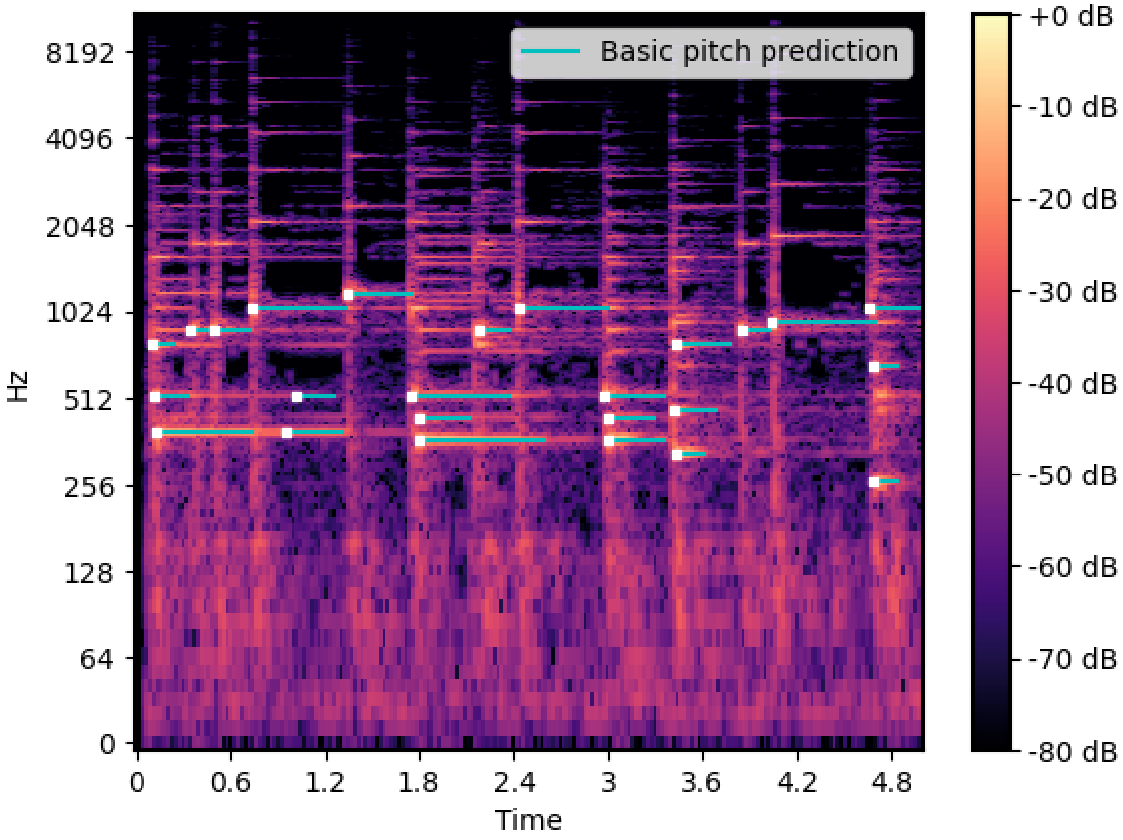
- **pYIN** (de Cheveigné et al., 2002)
  - `librosa.pyin`
- **CREPE** (Kim et al., 2018)
  - github.com/marl/crepe
- **PESTO** (Riou et al., 2023)
  - github.com/SonyCSLParis/pesto

Alain de Cheveigné and Hideki Kawahara, "YIN, a fundamental frequency estimator for speech and music," *Journal of the Acoustical Society of America,* 111(4):1917-1930, 2002.
Matthias Mauch and Simon Dixon. "pYIN: A fundamental frequency estimator using probabilistic threshold distributions." *ICASSP*, 2014.
Jong Wook Kim, Justin Salamon, Peter Li, and Juan Pablo Bello, "CREPE: A Convolutional Representation for Pitch Estimation," *ICASSP*, 2018.
Alain Riou, Stefan Lattner, Gaëtan Hadjeres, and Geoffroy Peeters, "PESTO: Pitch Estimation with Self-supervised Transposition-equivariant Objective," *ISMIR*, 2023.

Polyphonic F0 Estimation



Polyphonic F0 Estimation



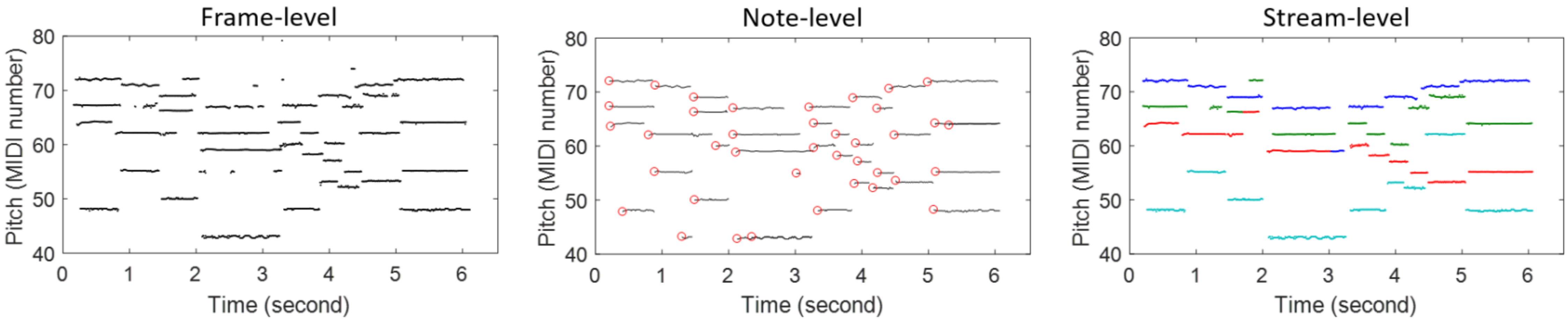
basicpitch.spotify.com

Polyphonic F0 Estimation Models

- **Deep Salience** (Bittner et al., 2017)
  - github.com/rabitt/ismir2017-deepsalience
- **Onset and Frames** (Hawthorne et al., 2018)
  - Piano only
  - github.com/jongwook/onsets-and-frames
- **Basic Pitch** (Bittner et al., 2022)
  - github.com/spotify/basic-pitch
  - basicpitch.spotify.com

Rachel M. Bittner, Brian McFee, Justin Salamon, Peter Li, and Juan P. Bello, "Deep Salience Representations for F0 Estimation in Polyphonic Music," *ISMIR*, 2017.
Curtis Hawthorne, Erich Elsen, Jialin Song, Adam Roberts, Ian Simon, Colin Raffel, Jesse Engel, Sageev Oore, and Douglas Eck, "Onsets and Frames: Dual-Objective Piano Transcription," *ISMIR*, 2018.
Rachel M. Bittner, Juan José Bosch, David Rubinstein, Gabriel Meseguer-Brocal, and Sebastian Ewert, "A Lightweight Instrument-Agnostic Model for Polyphonic Note Transcription and Multipitch Estimation," *ICASSP*, 2022.

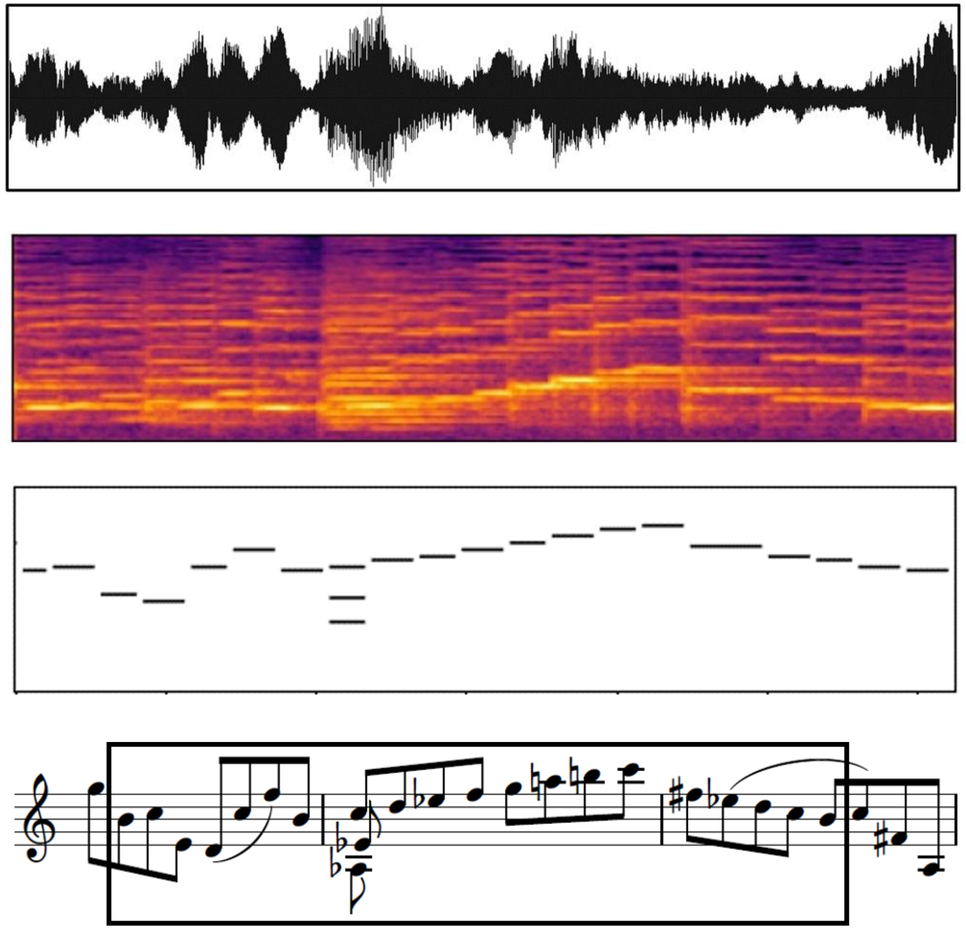
F0 Estimation vs Music Transcription



(Source: Benetos et al., 2019)

Emmanouil Benetos, Simon Dixon, Zhiyao Duan, and Sebastian Ewert, "Automatic Music Transcription: An Overview," *IEEE Signal Processing Magazine*, 36(1):20-30, 2019.

Music Transcription

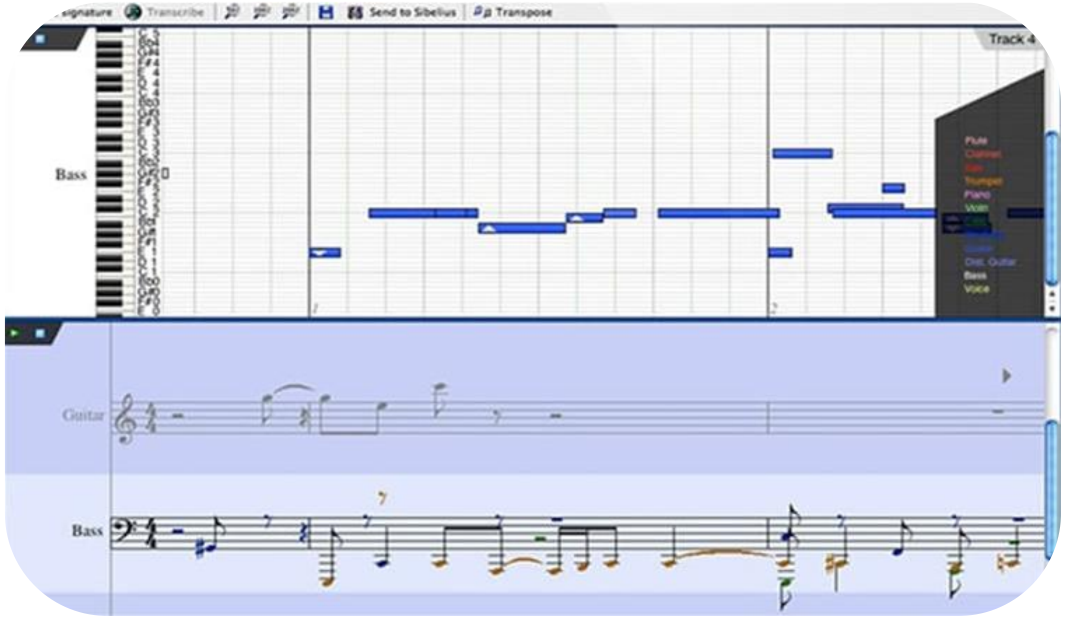


(Source: Dong et al., 2022)

Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, "Deep Performer: Score-to-Audio Music Performance Synthesis," *ICASSP*, 2022.

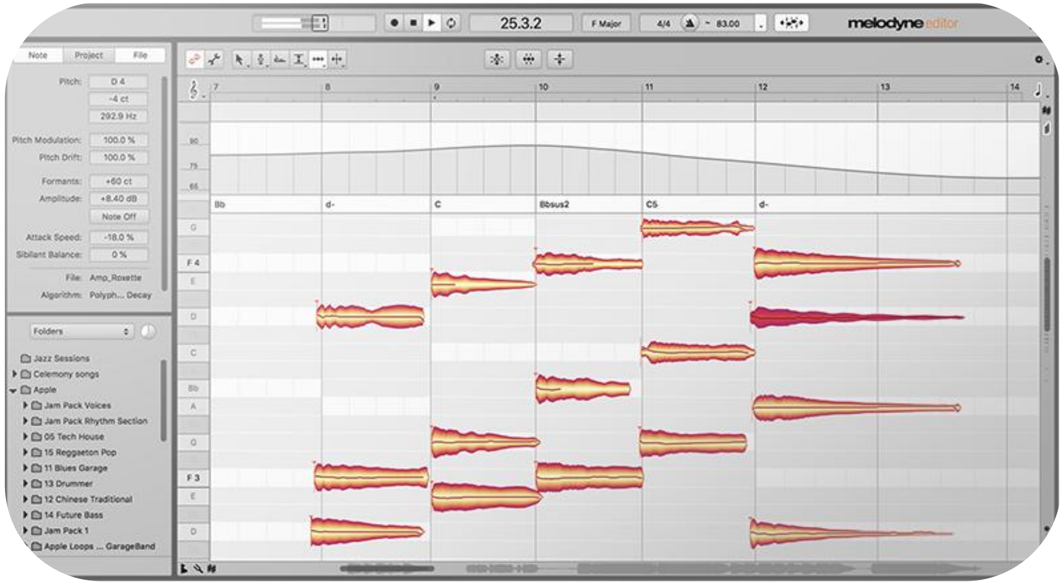
Commercial Music Transcription Software

AudioScore in Sibelius



(Source: Avid)

Melodyne Editor

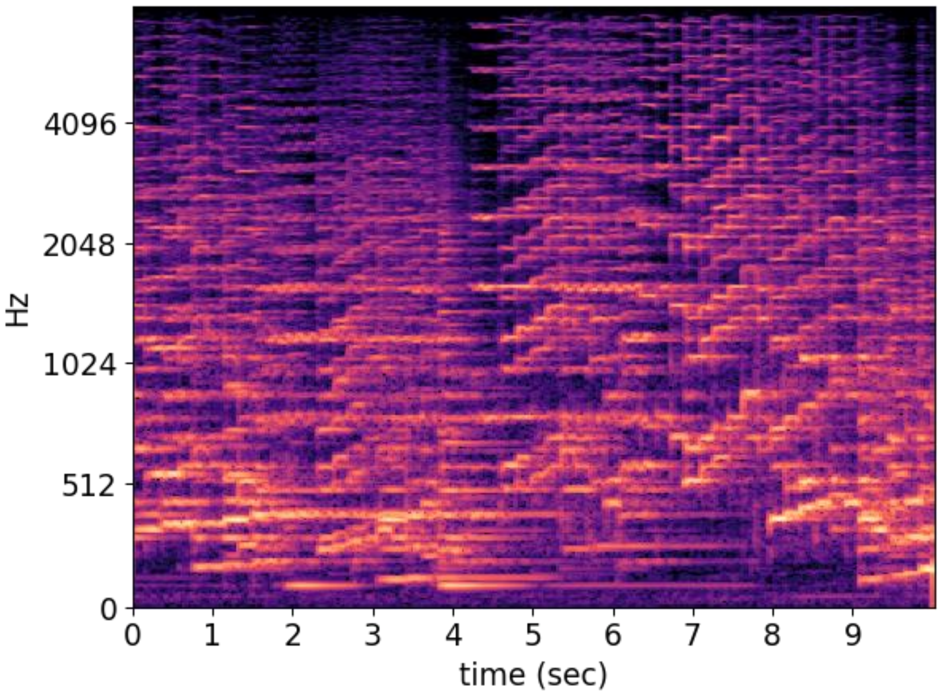


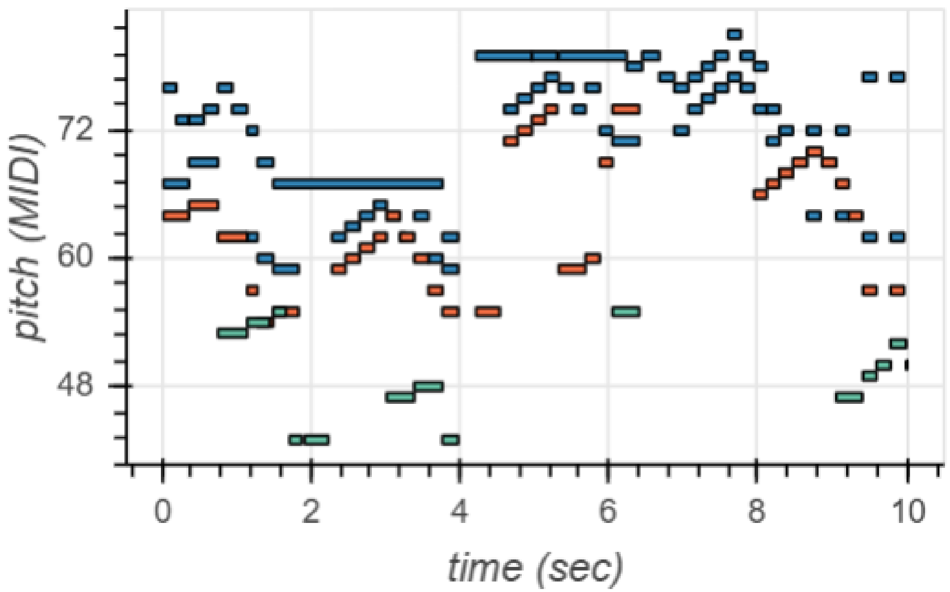
(Source: Celemony)

avid.com/products/audioscore-ultimate
shop.celemony.com/cgi-bin/WebObjects/CelemonyShop.woa/wo/kT0RWfafDLue8eUCipTXJw/0.0.31.23.5.21.2.3

Multitrack Transcription Models

- **MT3** (Gardner et al., 2022)
  - github.com/magenta/mt3





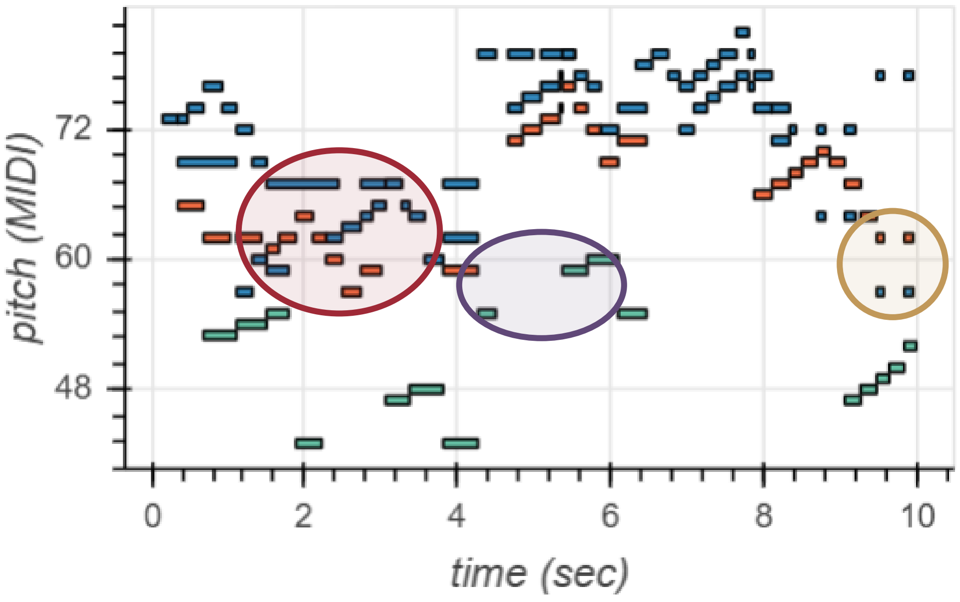
(Source: Gardner et al., 2022)

Josh Gardner, Ian Simon, Ethan Manilow, Curtis Hawthorne, and Jesse Engel, "MT3: Multi-Task Multitrack Music Transcription," *ICLR*, 2022.

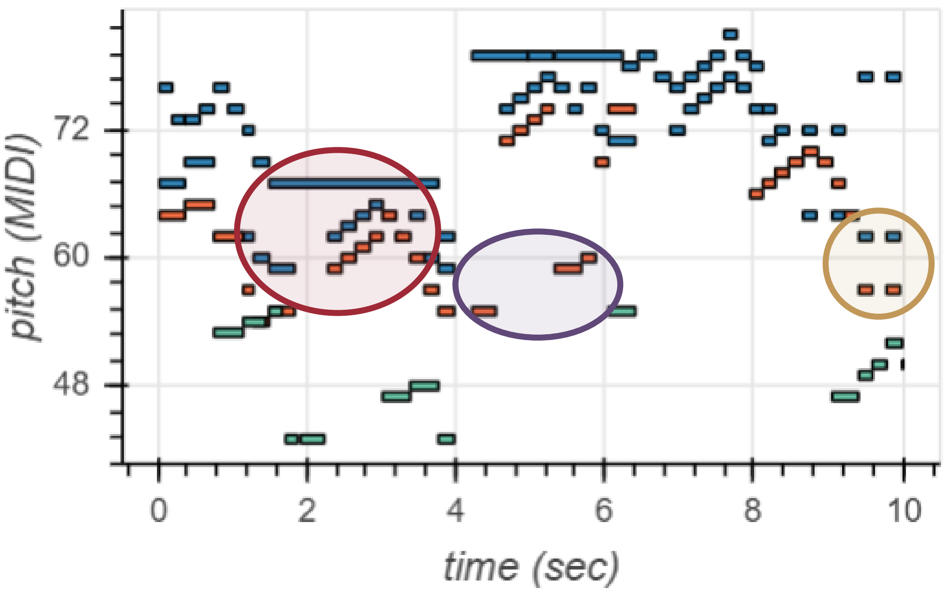
Multitrack Transcription Models

- **MT3** (Gardner et al., 2022)
  - github.com/magenta/mt3

Ground truth



MT3 prediction



(Source: Gardner et al., 2022)

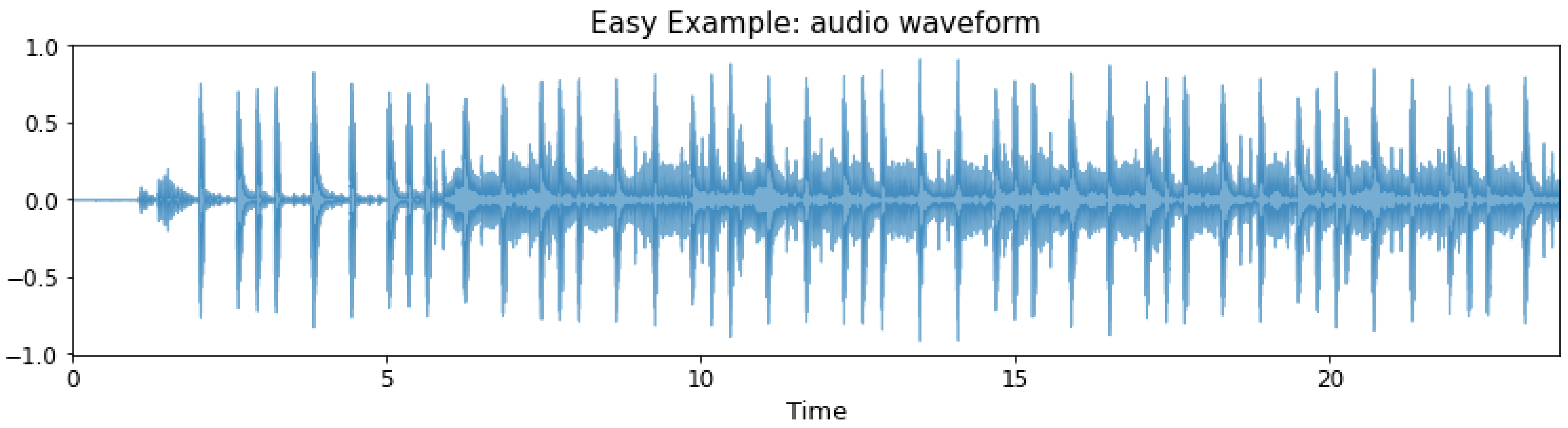
Josh Gardner, Ian Simon, Ethan Manilow, Curtis Hawthorne, and Jesse Engel, "MT3: Multi-Task Multitrack Music Transcription," *ICLR*, 2022.

Resources

- Rachel Bittner, Mark Cartwright, and Ethan Manilow, "Programming MIR Baselines from Scratch: Three Case Studies," *Tutorials of ISMIR*, 2021.
  - Part 1: Transcription with NMF (Ethan Manilow)
  - Part 2: Pitch Tracking with pytorch (Rachel Bittner)
  - Part 3: Instrument Classification with OpenL3 & Tensorflow (Mark Cartwright)
- Rachel Bittner, Alain de Cheveigné, and Johana Devaney, "Fundamental Frequency Estimation in Music," *Tutorials of ISMIR*, 2018.
  - Part 1: Pitch (Alain de Cheveigné)
  - Part 2: Polyphonic fundamental frequency estimation (Rachel Bittner)
  - Part 3: Applications (Johana Devaney)

Rhythm Analysis

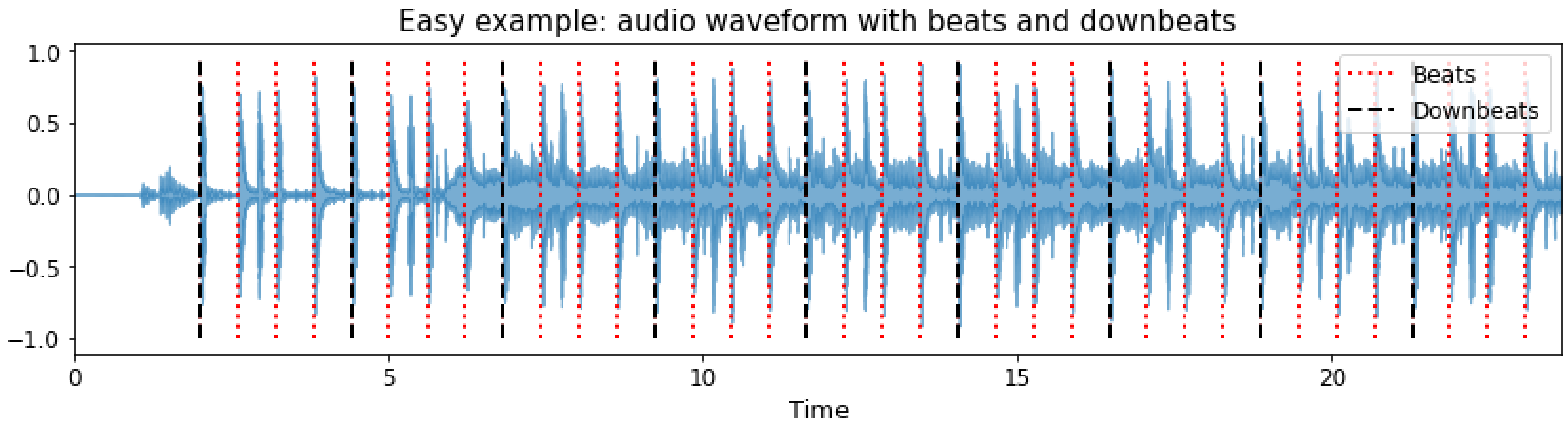
Beat & Downbeat Estimation

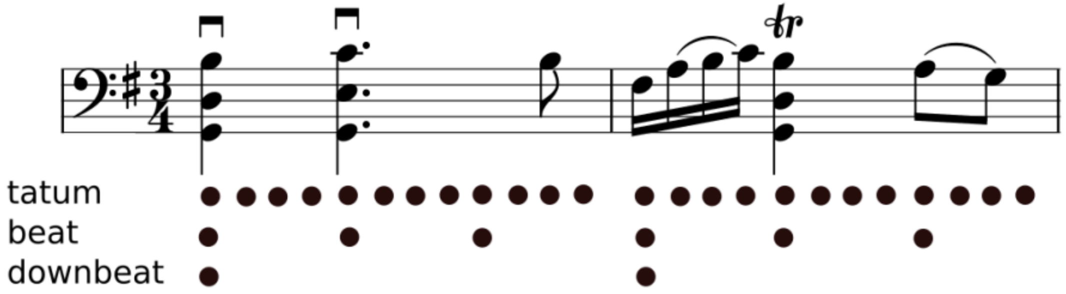


(Source: Davies et al., 2021)

Matthew E. P. Davies, Sebastian Böck, and Magdalena Fuentes, "Tempo, Beat and Downbeat Estimation," *Tutorials of ISMIR*, 2021.

Beat & Downbeat Estimation

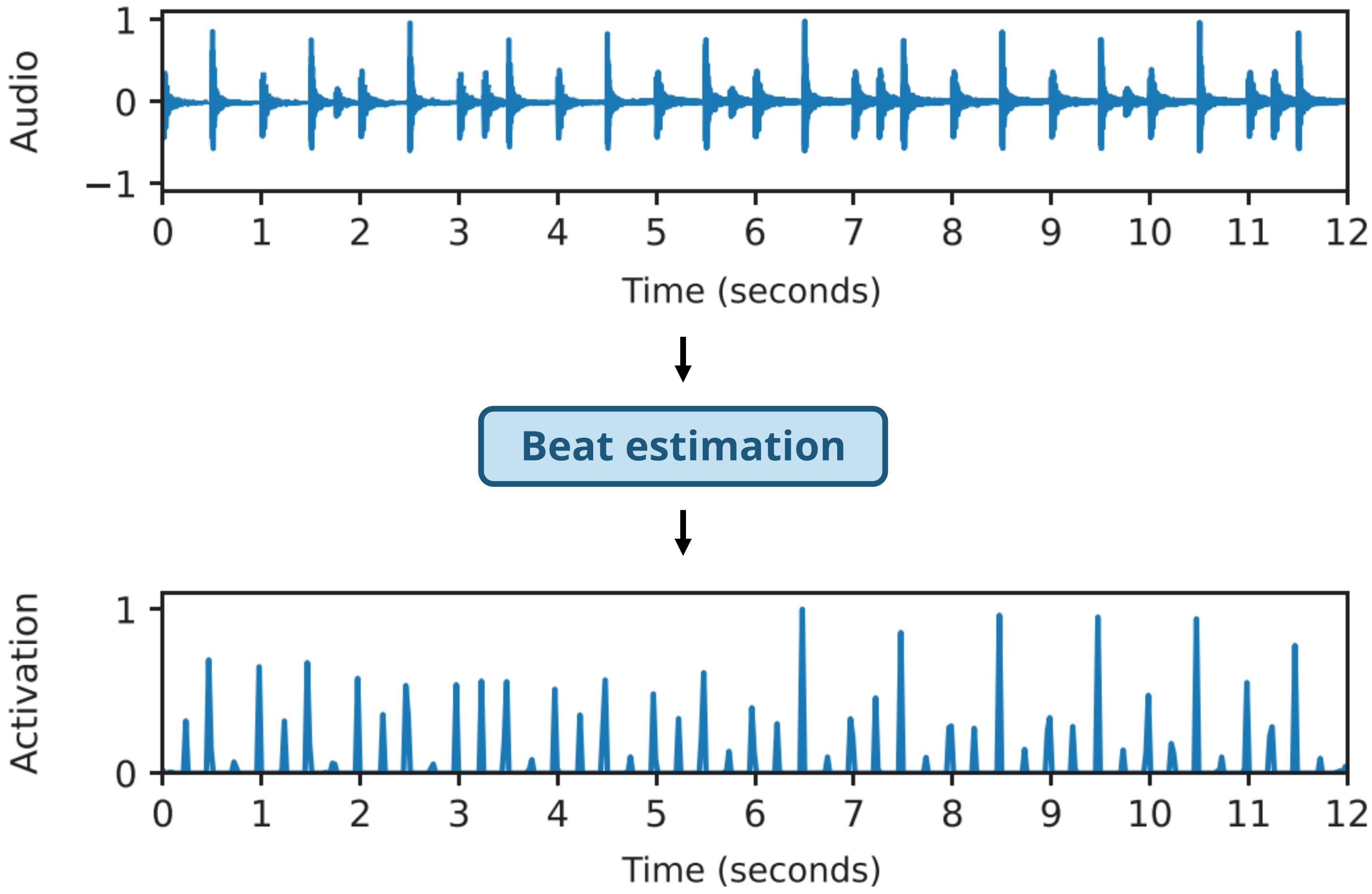




(Source: Davies et al., 2021)

Matthew E. P. Davies, Sebastian Böck, and Magdalena Fuentes, "Tempo, Beat and Downbeat Estimation," *Tutorials of ISMIR*, 2021.

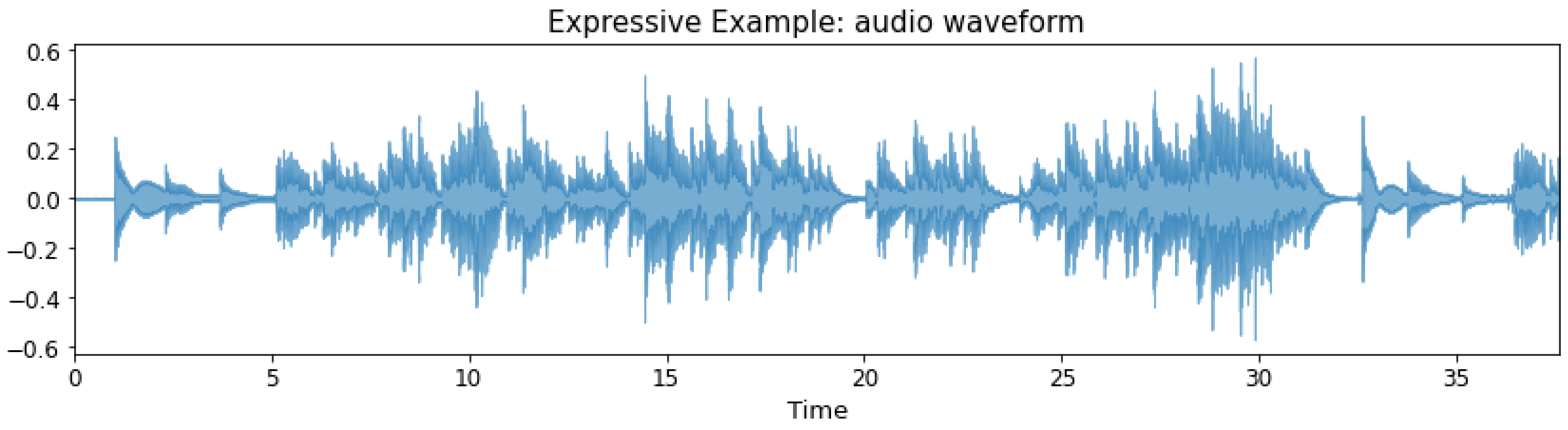
Beat & Downbeat Estimation



(Source: Meier et al., 2024)

Peter Meier, Ching-Yu Chiu, and Meinard Müller, "A Real-Time Beat Tracking System with Zero Latency and Enhanced Controllability," *TISMIR*, 7(1):213-227, 2024.

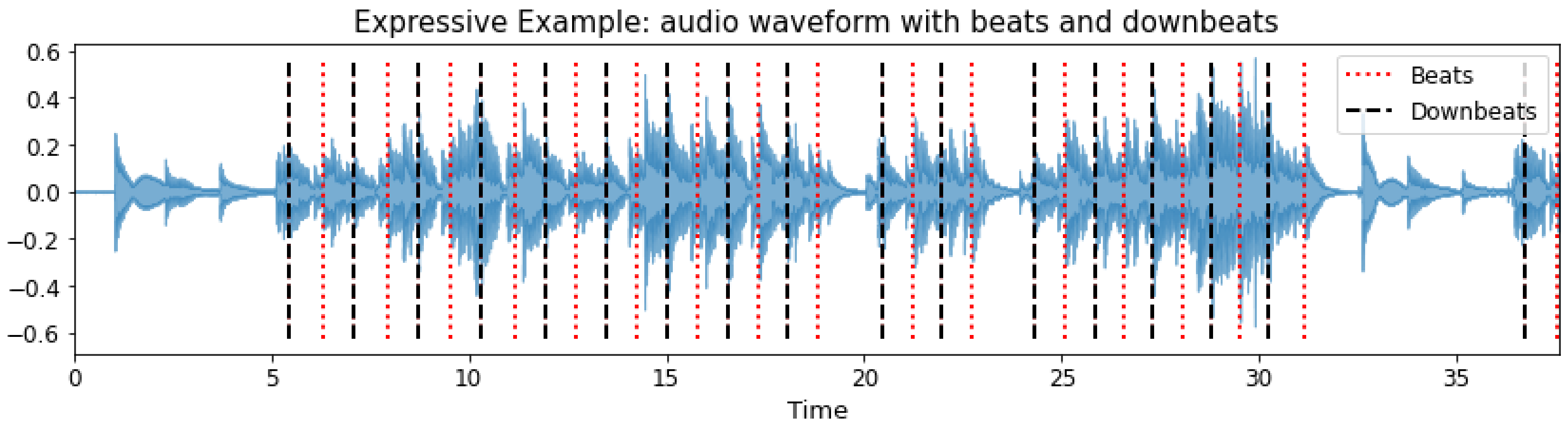
Beat & Downbeat Estimation

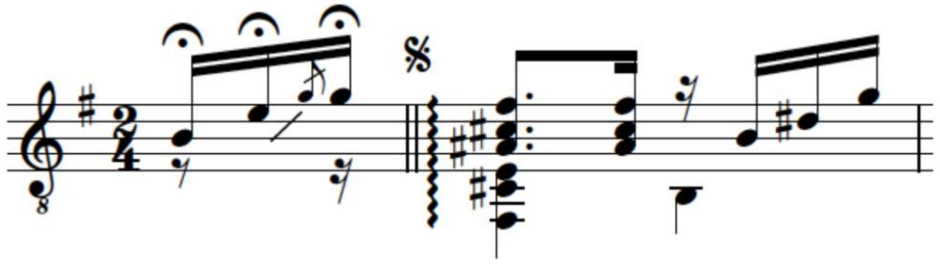


(Source: Davies et al., 2021)

Matthew E. P. Davies, Sebastian Böck, and Magdalena Fuentes, "Tempo, Beat and Downbeat Estimation," *Tutorials of ISMIR*, 2021.

Beat & Downbeat Estimation





(Source: Davies et al., 2021)

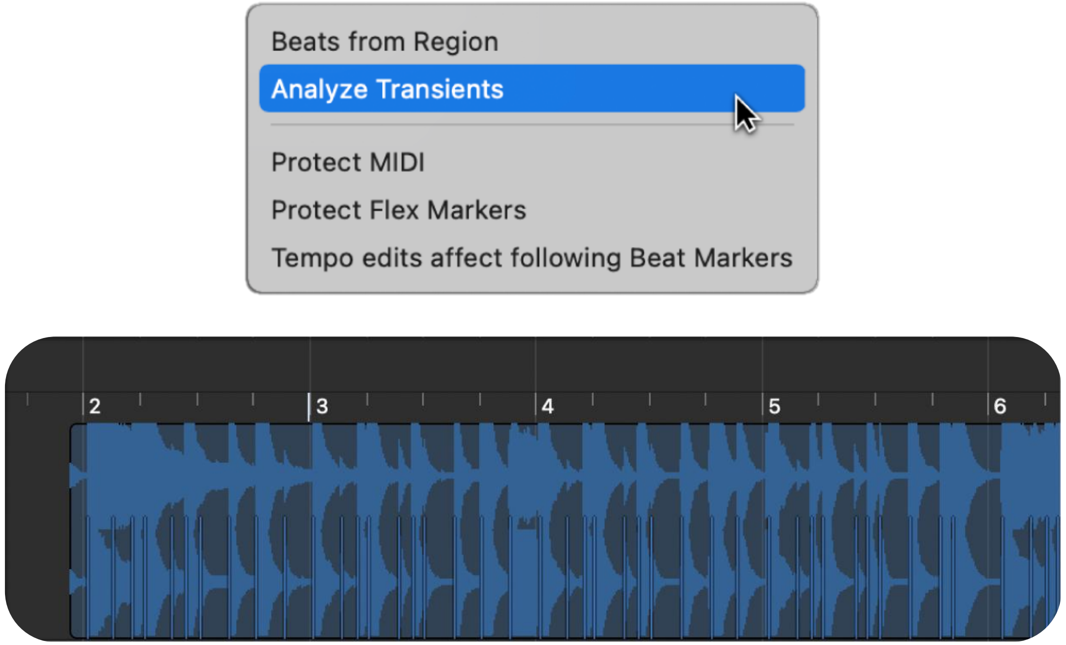
Matthew E. P. Davies, Sebastian Böck, and Magdalena Fuentes, "Tempo, Beat and Downbeat Estimation," *Tutorials of ISMIR*, 2021.

Beat Tracking in Pro Tools & Logic Pro

Beat Detective in Pro Tools

(Source: Logic Pro User Guide)

Beat Mapping in Logic Pro

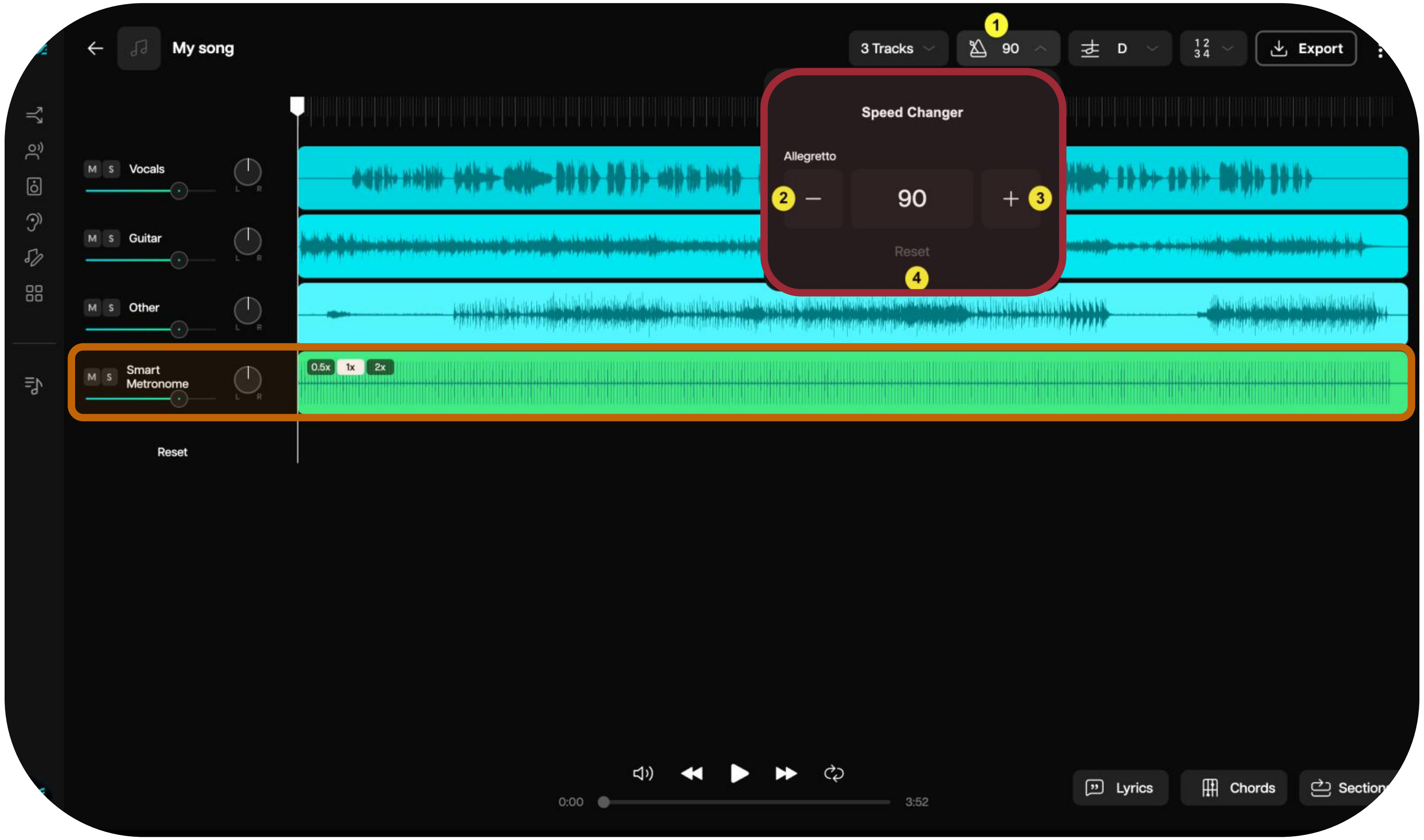


(Source: Logic Pro User Guide)

macprovideo.com/article/logic-pro/beat-detection-in-pro-tools-vs-logic-pro-part-1---creating-tempo-maps
support.apple.com/en-in/guide/logicpro/lgcp6338b8d1/mac

Tempo Estimation & Beat Tracking in Moises

Tempo estimation



Beat tracking

(Source: Moises)

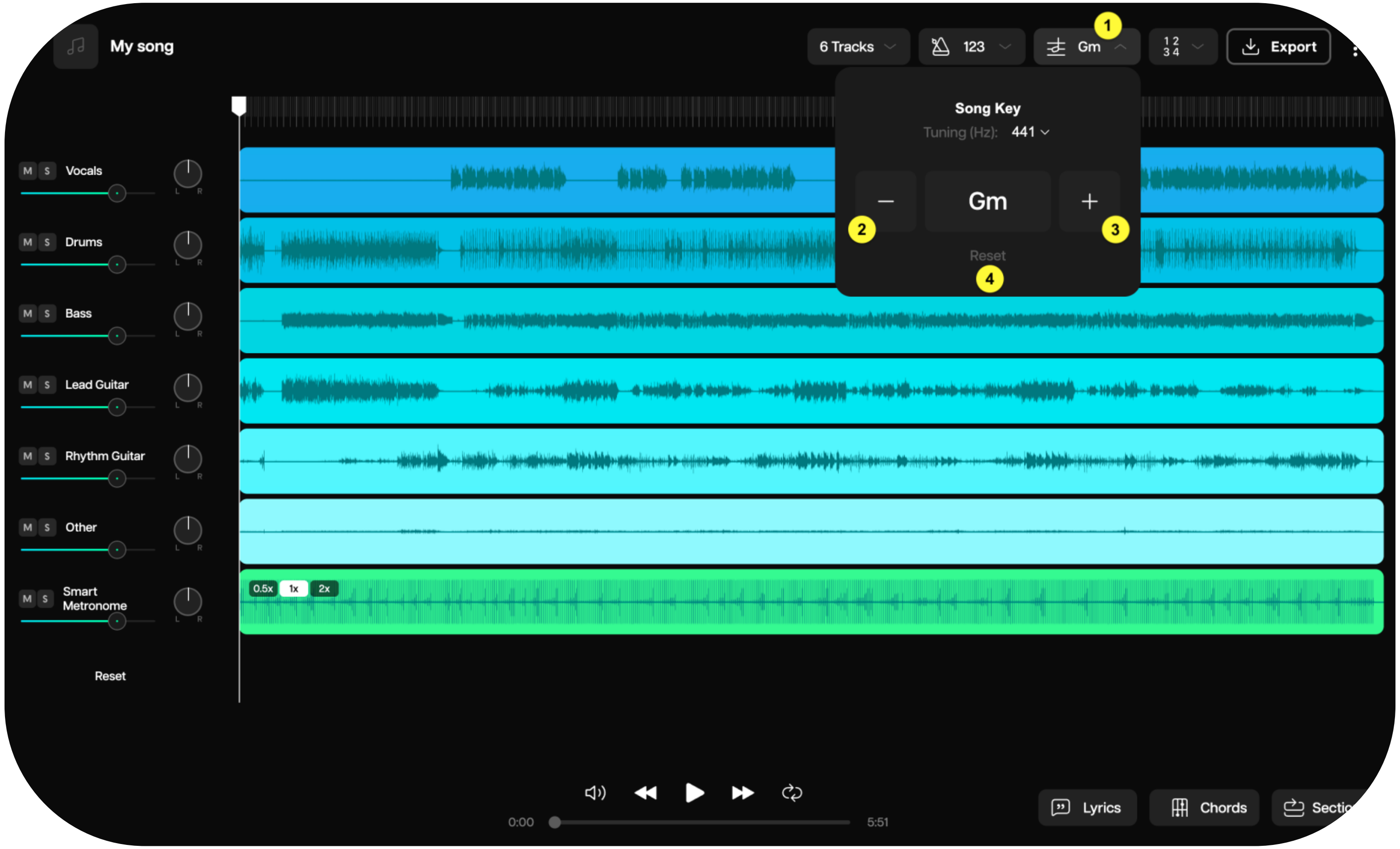
help.moises.ai/hc/en-us/articles/6582847813148-How-do-I-change-the-song-s-tempo

Resources on Rhythm Analysis

- Matthew E. P. Davies, Sebastian Böck, and Magdalena Fuentes, "Tempo, Beat and Downbeat Estimation," *Tutorials of ISMIR*, 2021.
- Hendrik Schreiber, Julián Urbano, and Meinard Müller, "Music Tempo Estimation: Are We Done Yet?," *TISMIR*, 3(1):111-125, 2020.

Other Music Analysis

Key Detection in Moises



(Source: Moises)

help.moises.ai/hc/en-us/articles/6582517238044-How-do-I-use-the-Key-Detection-with-the-Pitch-Changer

Chord Detection

youtube.com/shorts/_N3b_GARMfA

Structure Analysis

Music segmentation

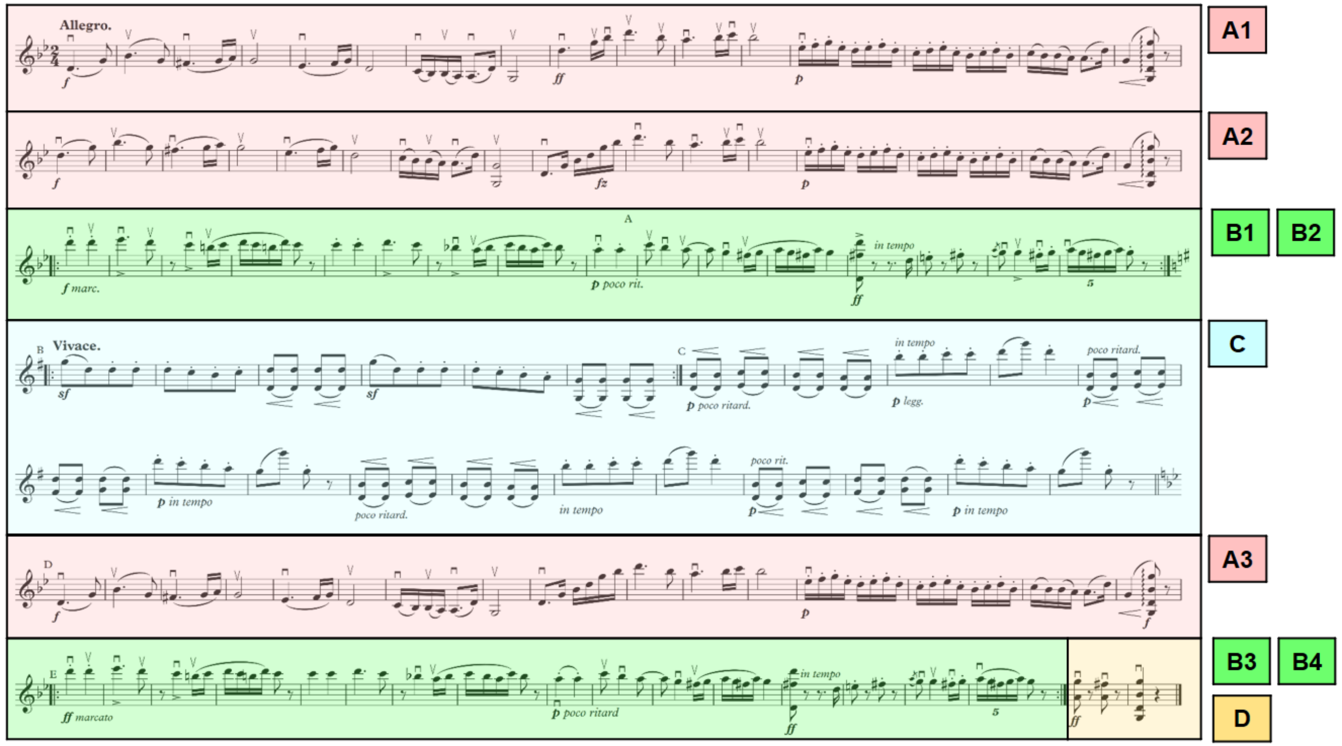


Figure 4.5 following [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Hierarchical music segmentation

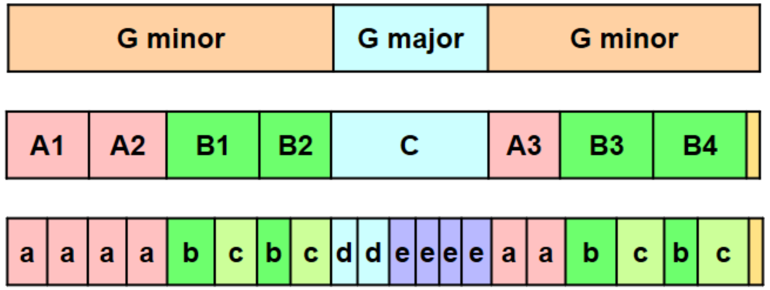
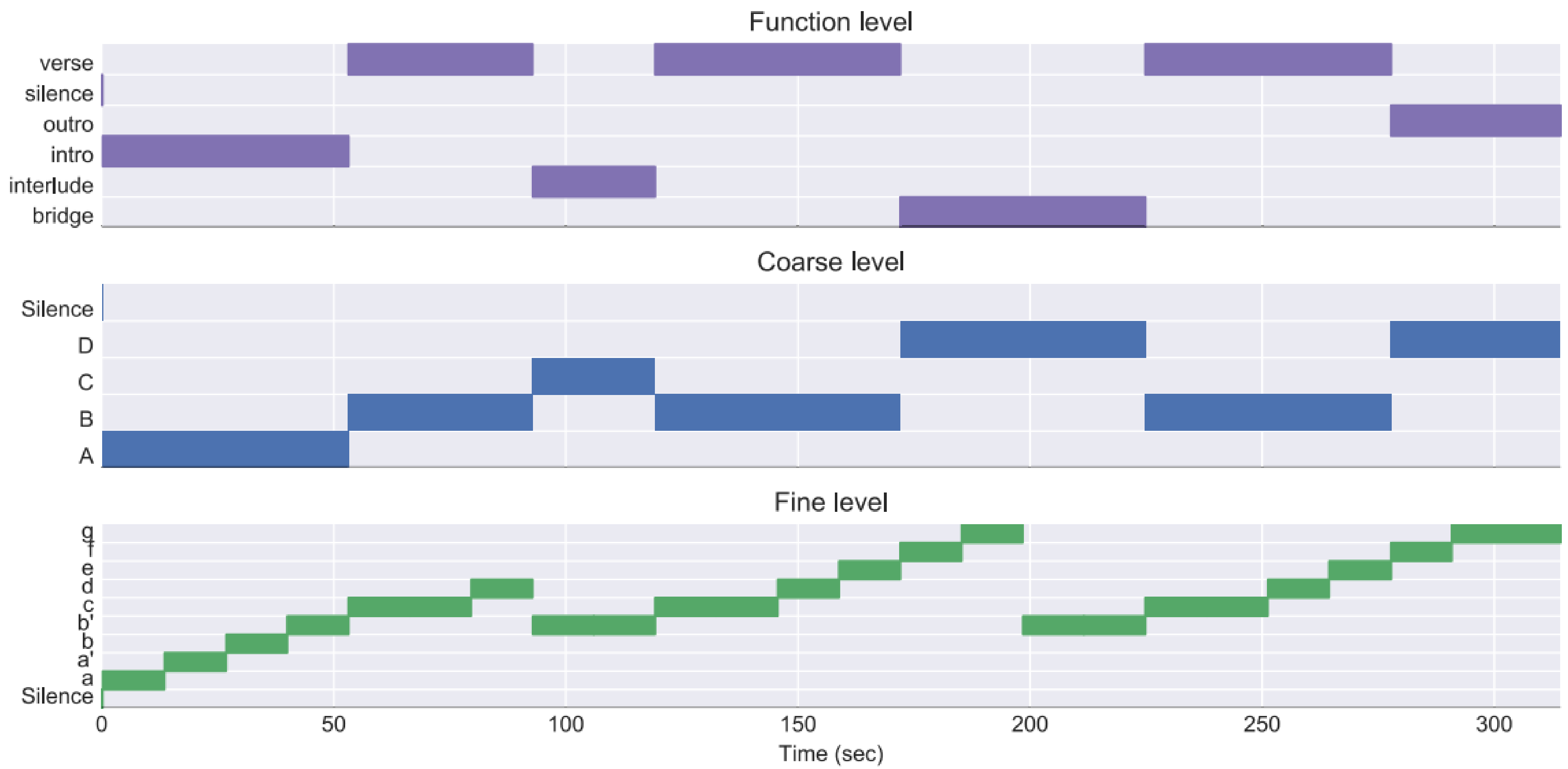


Figure 4.28 from [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

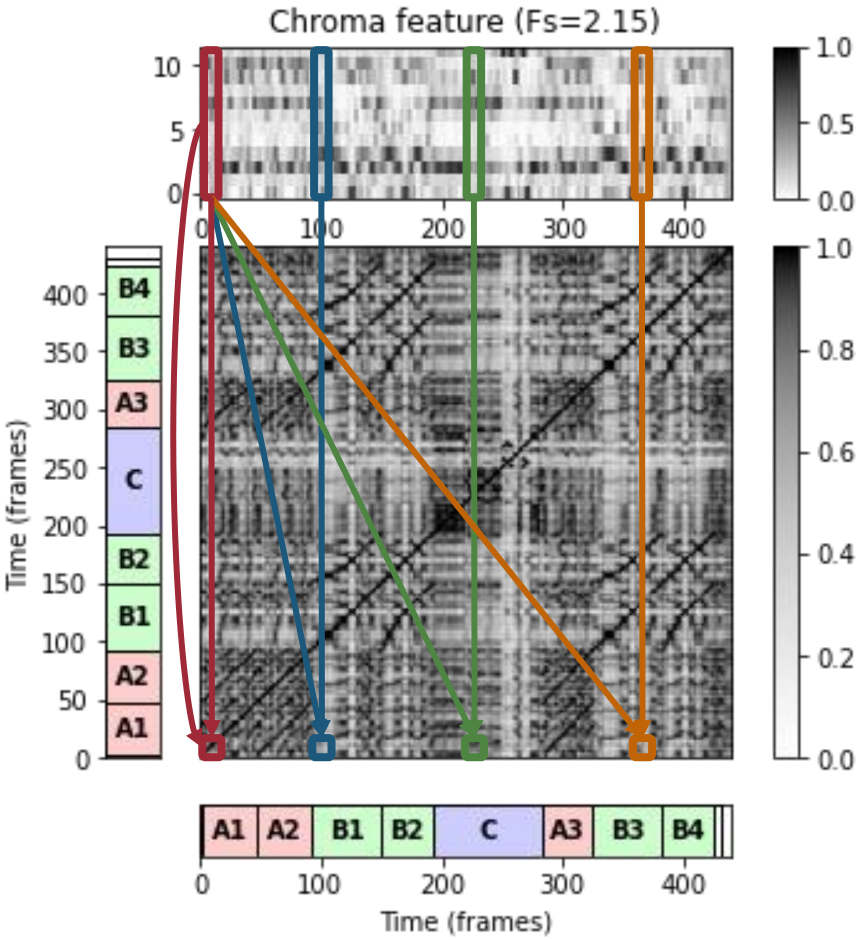
Hierarchical Music Segmentation



(Source: Müller & Zalkow, 2019)

Oriol Nieto, Gautham J. Mysore, Cheng-i Wang, Jordan B. L. Smith, Jan Schlüter, Thomas Grill, and Brian McFee, "Audio-Based Music Structure Analysis: Current Trends, Open Challenges, and Applications," *TISMIR*, 3(1):246-263, 2020.

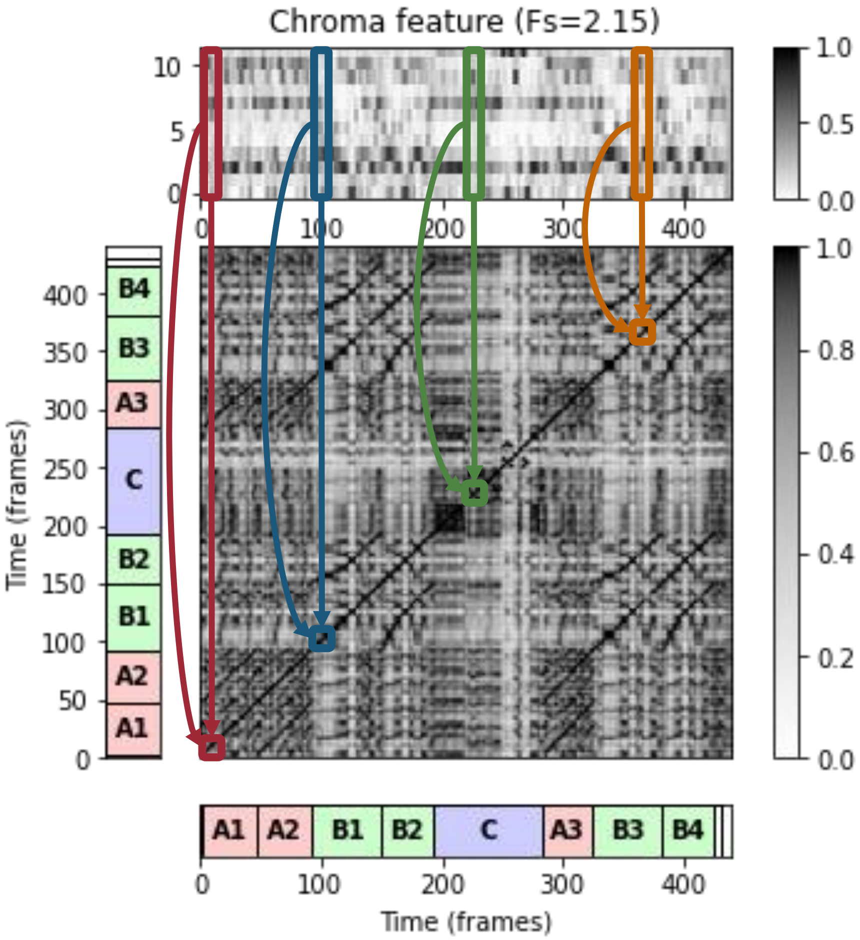
Self-Similarity Matrices (SSMs)



(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

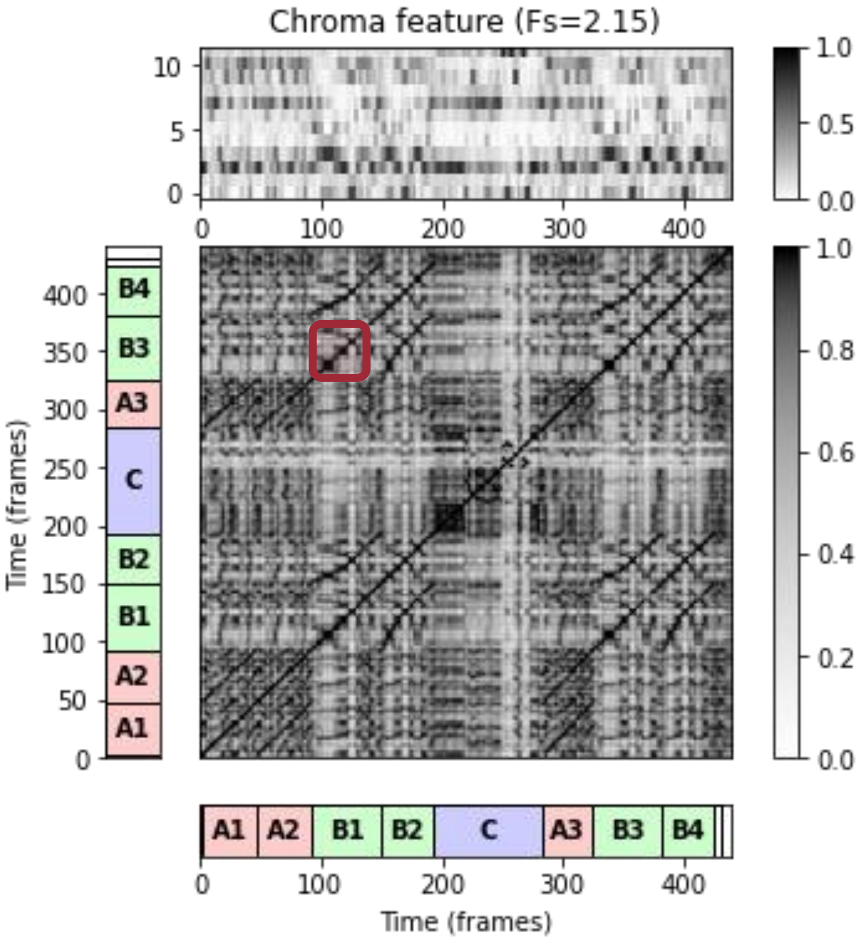
Self-Similarity Matrices (SSMs)



(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

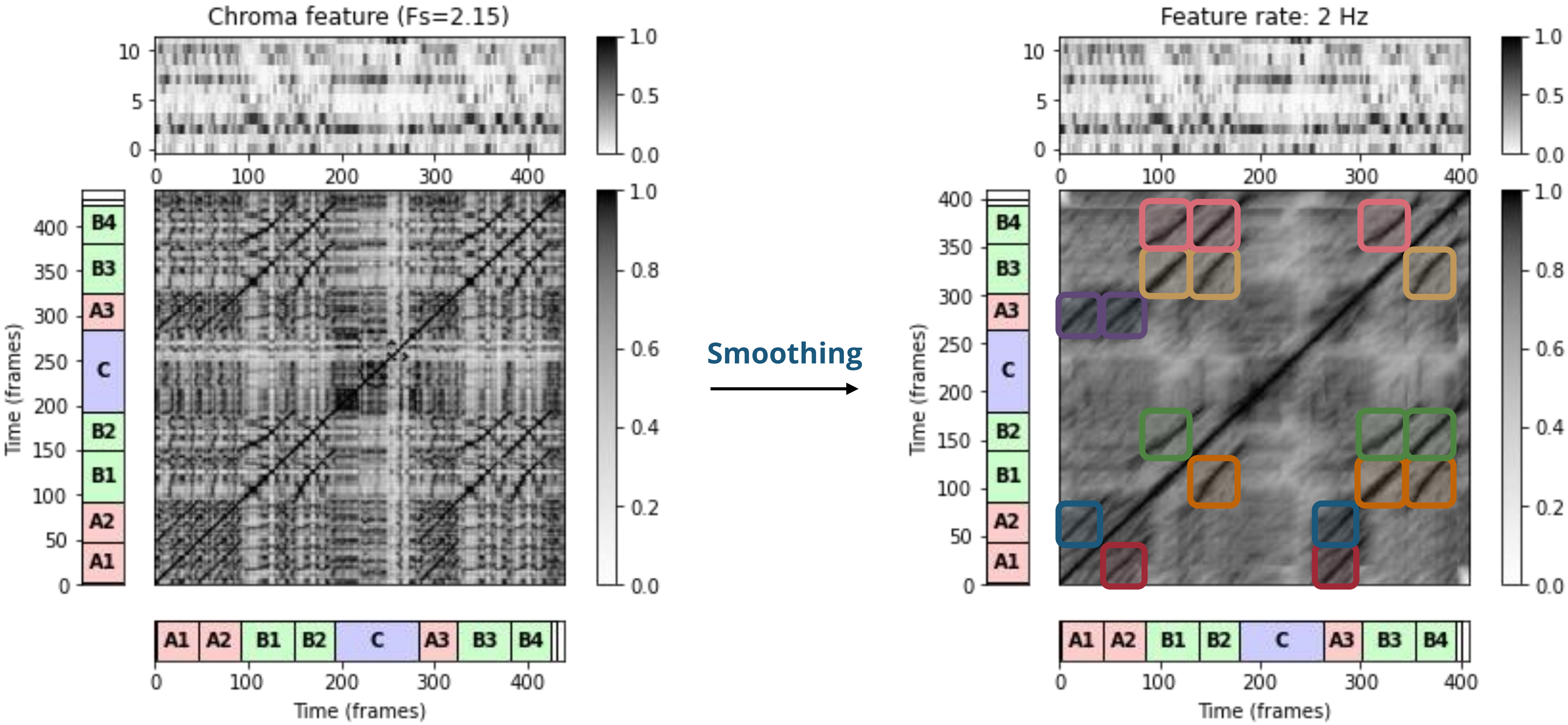
Self-Similarity Matrices (SSMs)



(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

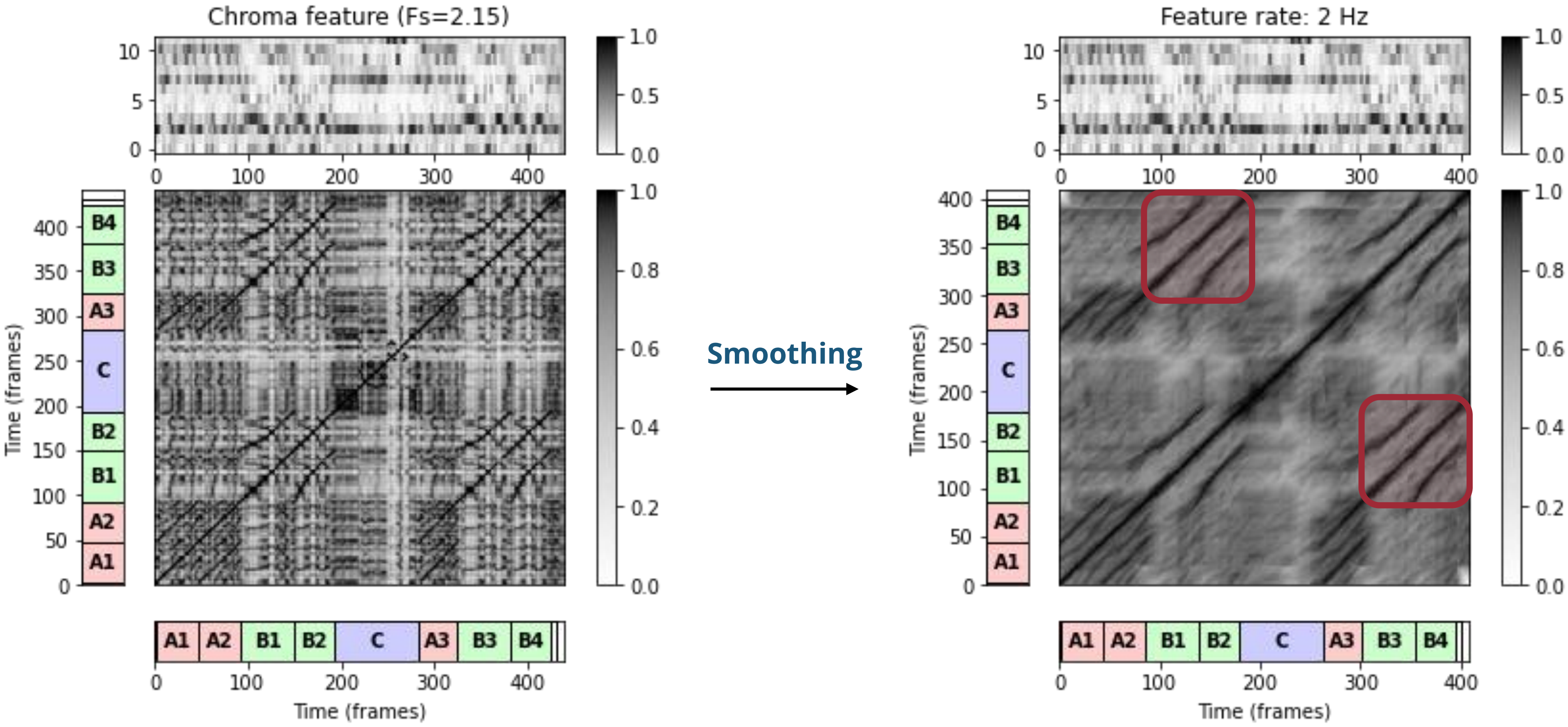
Self-Similarity Matrices (SSMs)



(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

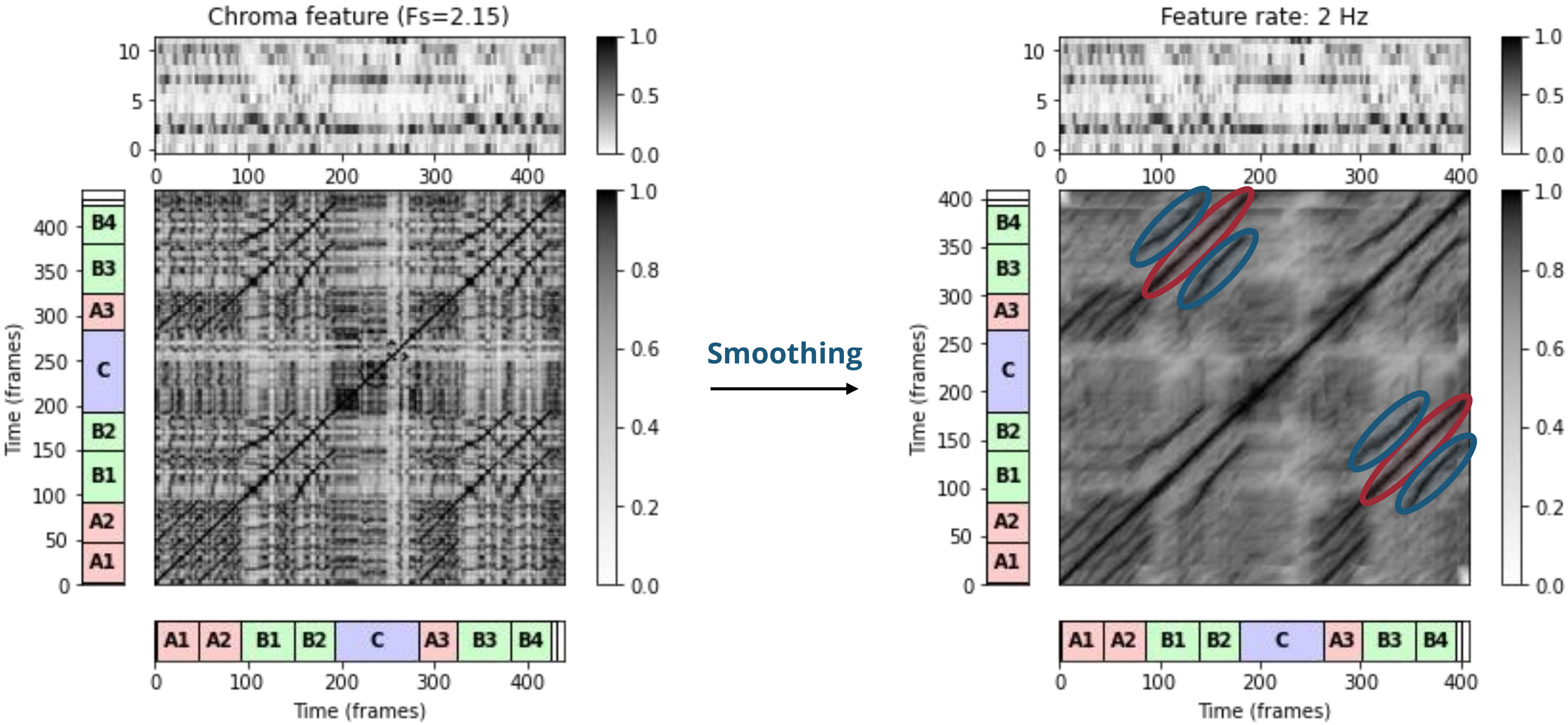
Self-Similarity Matrices (SSMs)



(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

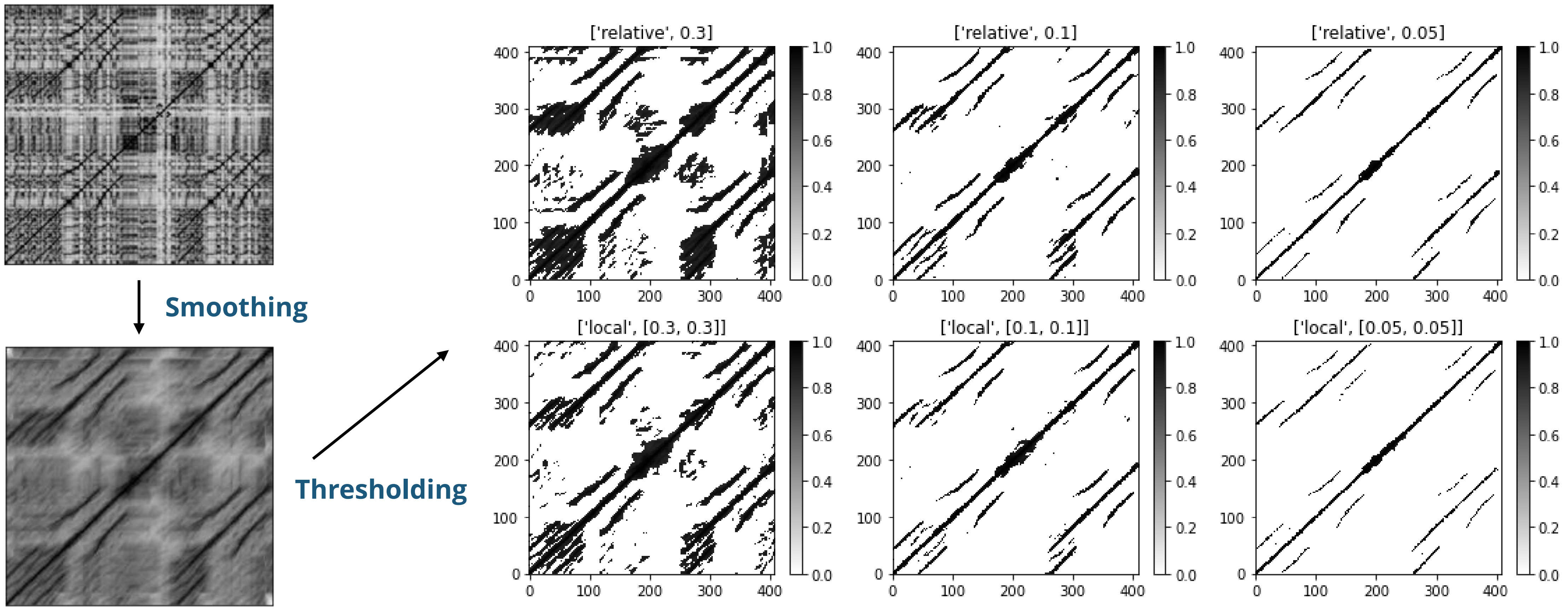
Self-Similarity Matrices (SSMs)



(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

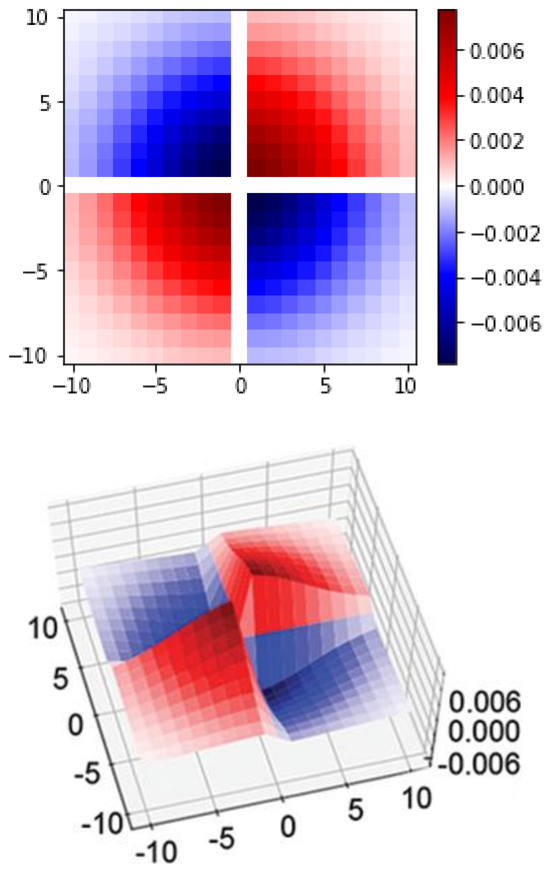
Self-Similarity Matrices (SSMs)



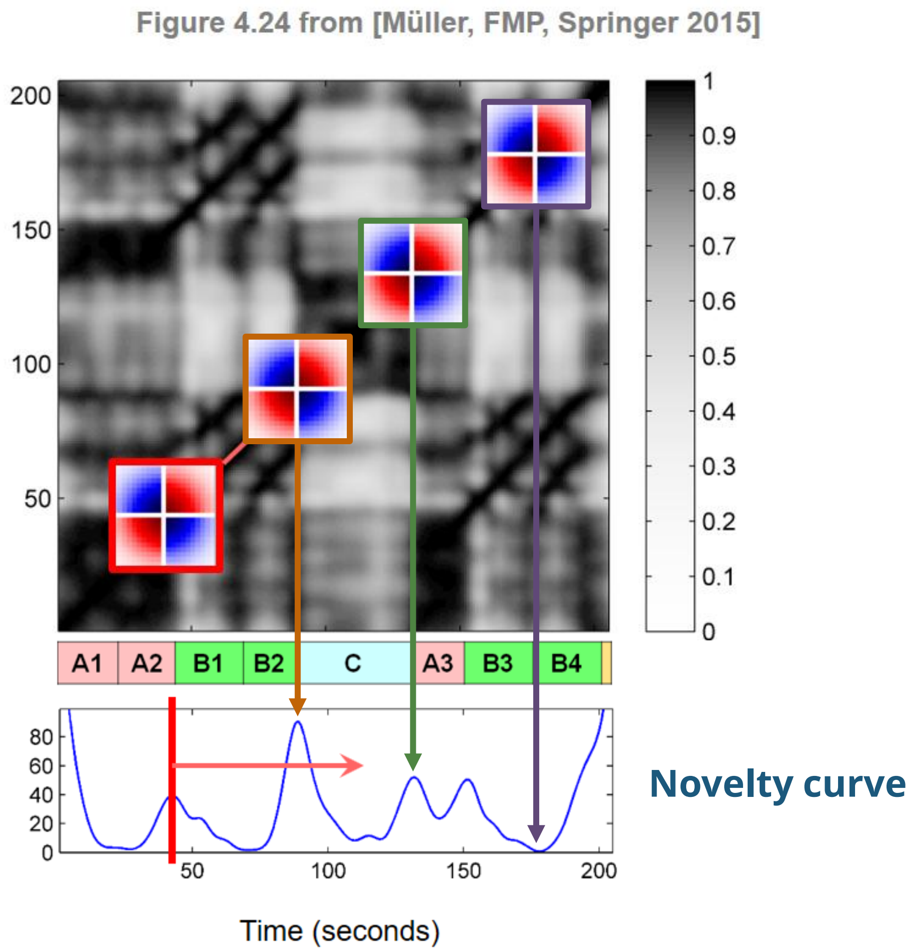
(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019

Self-Similarity Matrices (SSMs)



(Source: Müller & Chiu, 2024)



(Source: Müller & Zalkow, 2019)

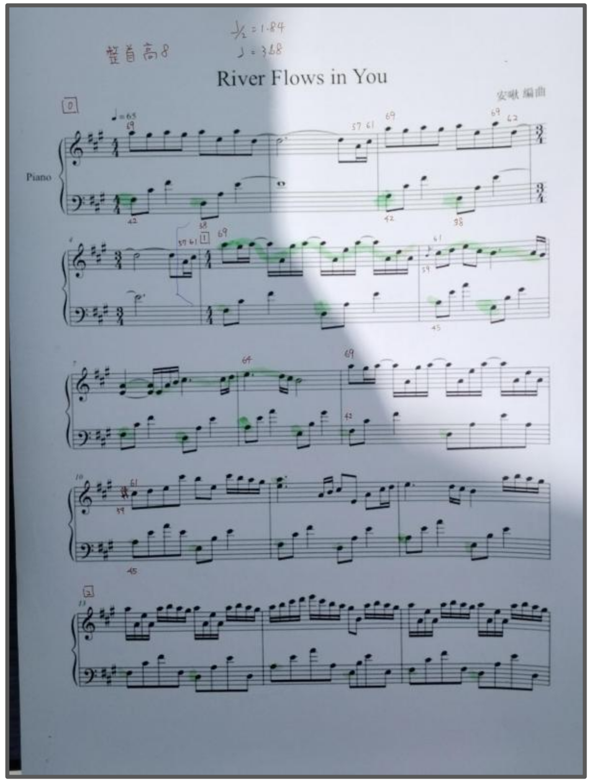
Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.
Meinard Müller and Ching-Yu Chiu, "A Basic Tutorial on Novelty and Activation Functions for Music Signal Processing," *TISMIR*, 7(1):179-194, 2024.

Resources on Music Structure Analysis

- Meinard Müller and Ching-Yu Chiu, "A Basic Tutorial on Novelty and Activation Functions for Music Signal Processing," *TISMIR*, 7(1):179-194, 2024.
- Oriol Nieto, Gautham J. Mysore, Cheng-i Wang, Jordan B. L. Smith, Jan Schlüter, Thomas Grill, and Brian McFee, "Audio-Based Music Structure Analysis: Current Trends, Open Challenges, and Applications," *TISMIR*, 3(1):246-263, 2020.
- Meinard Müller & Jordan B. L. Smith, "Music Structure Analysis," *Tutorials of ISMIR*, 2014. (part 1, part 2, part 3)

Optical Music Recognition (OMR)

- **Goal**: Convert **scanned sheet music** into **digital musical notation**



→ **Optical Music Recognition** →

breezewhite.github.io/oemer/

Challenges

Violin Sonata No. 1 in G minor
(BWV 1001)

Challenges

Weihnachtsoratorium
(Christmas Oratorio; BWV 248)

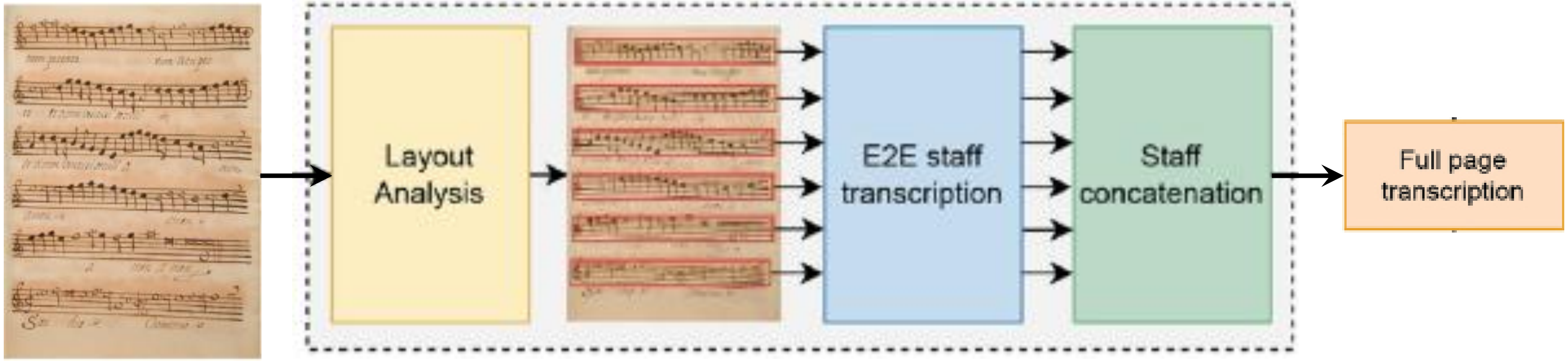
Challenges of OMR

(Source: Calvo-Zaragoza et al., 2018)

(Source: Novotný & Pokorný, 2015)

Jorge Calvo-Zaragoza, Juan C. Martinez-Sevilla, Carlos Penarrubia, and Antonio Rios-Vila, "Optical Music Recognition: Recent Advances, Current Challenges, and Future Directions," *ICDAR*, 2023.
Jiří Novotný and Jaroslav Pokorný, "Introduction to Optical Music Recognition: Overview and Practical Challenges," *DATESO*, 2015.

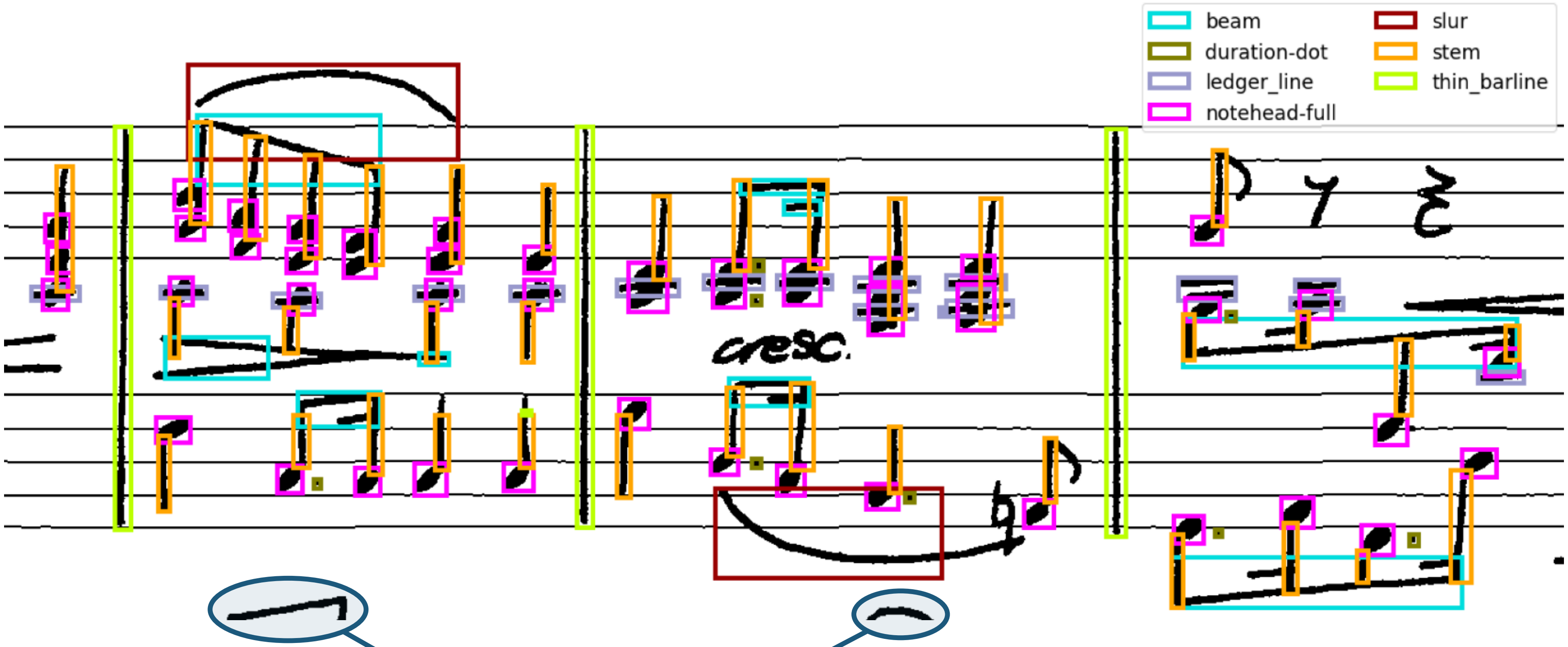
Common Pipeline of OMR Systems



(Source: Calvo-Zaragoza et al., 2018)

Jorge Calvo-Zaragoza, Juan C. Martinez-Sevilla, Carlos Penarrubia, and Antonio Rios-Vila, "Optical Music Recognition: Recent Advances, Current Challenges, and Future Directions," *ICDAR*, 2023.

Musical Object Recognition



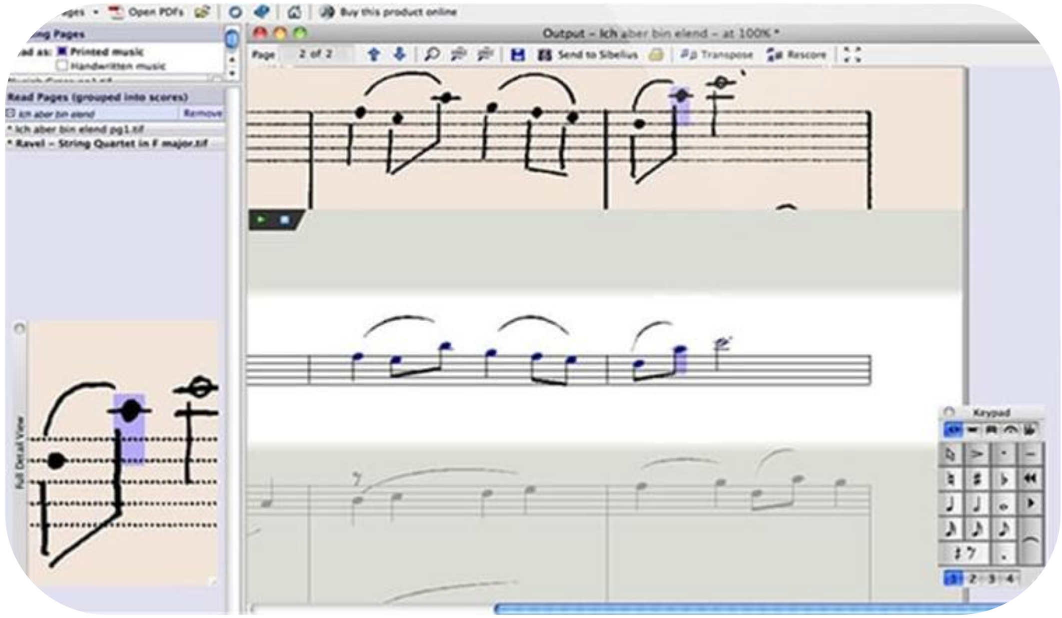
From the staff below!

(Source: Pacha et al., 2018)

Alexander Pacha, Jan Hajič, and Jorge Calvo-Zaragoza, "A Baseline for General Music Object Detection with Deep Learning," *Applied Science*, 8(9):1488, 2018.

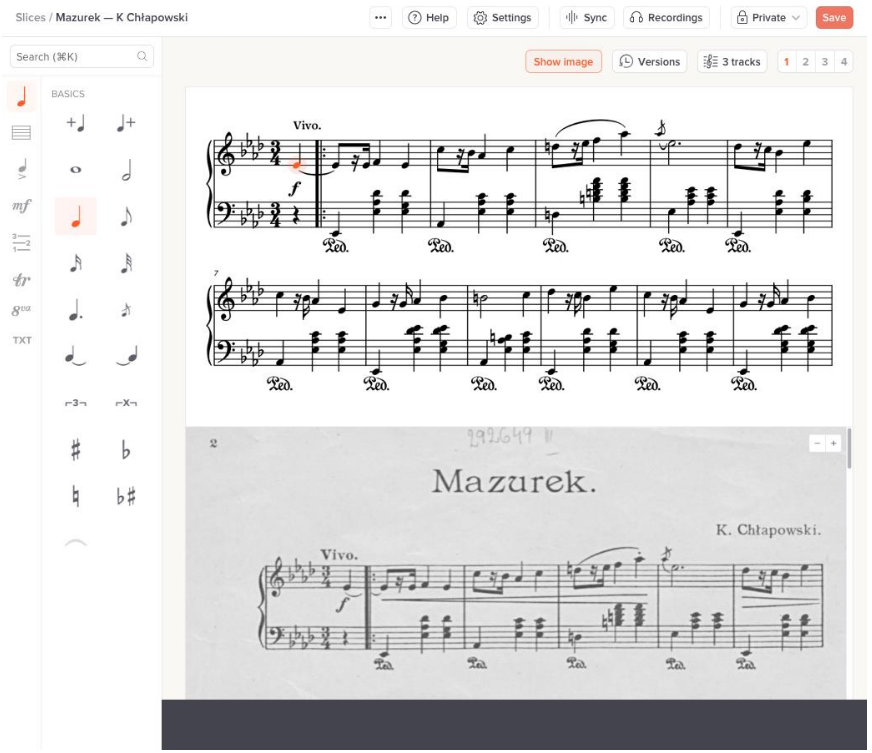
Commercial OMR Software

PhotoScore & NotateMe in Sibelius



(Source: Avid)

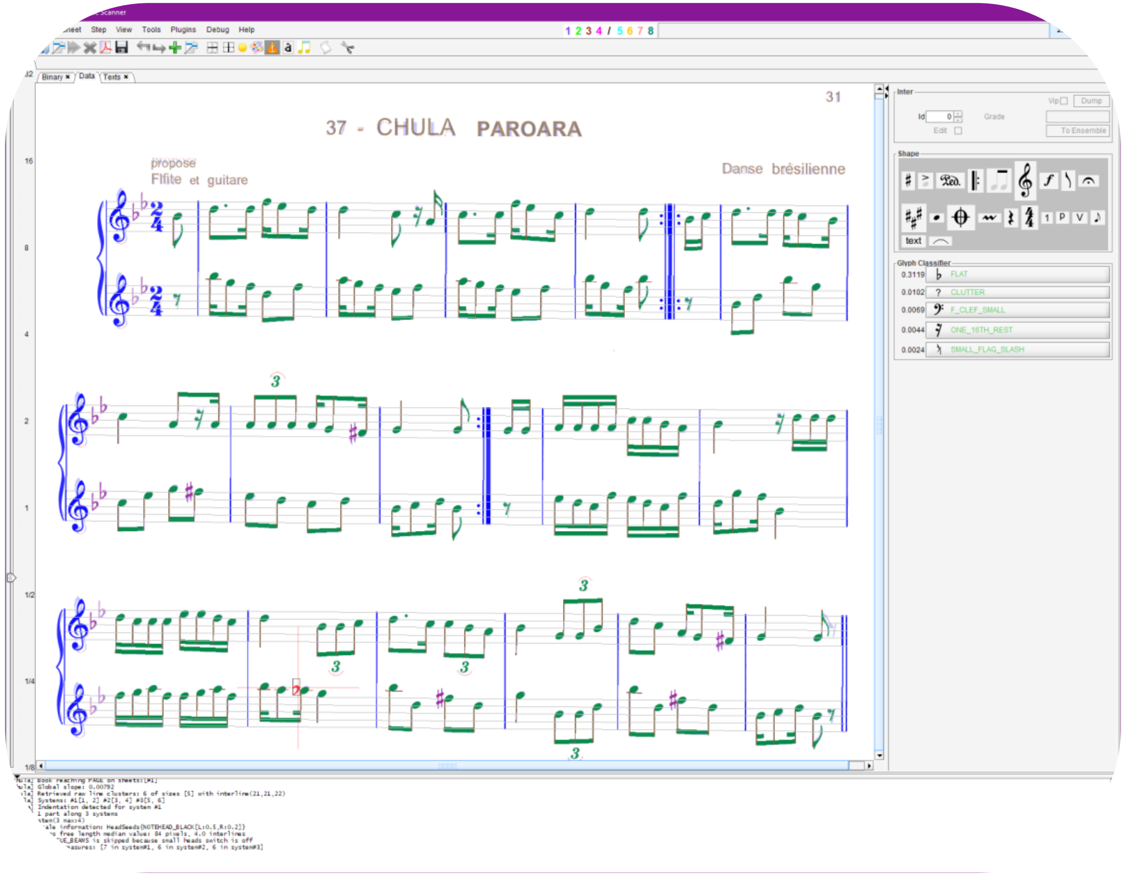
Soundslice



(Source: Soundslice)

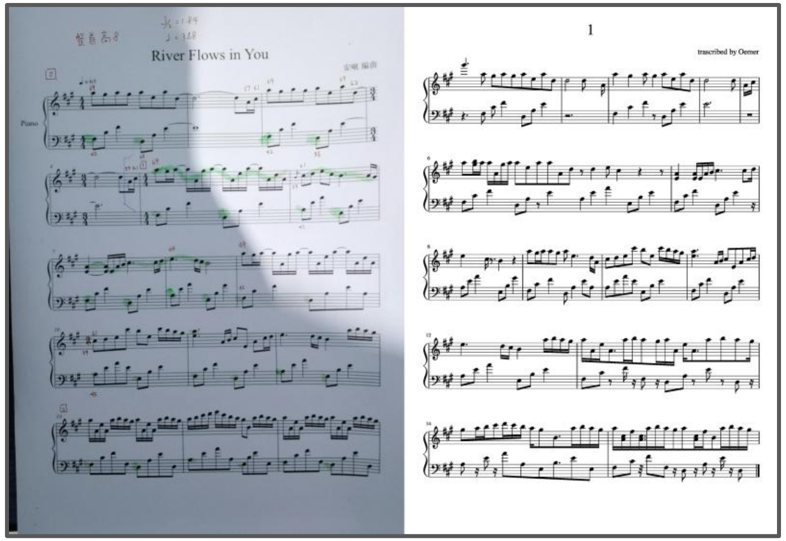
avid.com/products/photoscore-and-notateme-ultimate
soundslice.com/sheet-music-scanner/

Open-source OMR Software: Audiveris



github.com/Audiveris/audiveris

Open-source OMR Software: Oemer



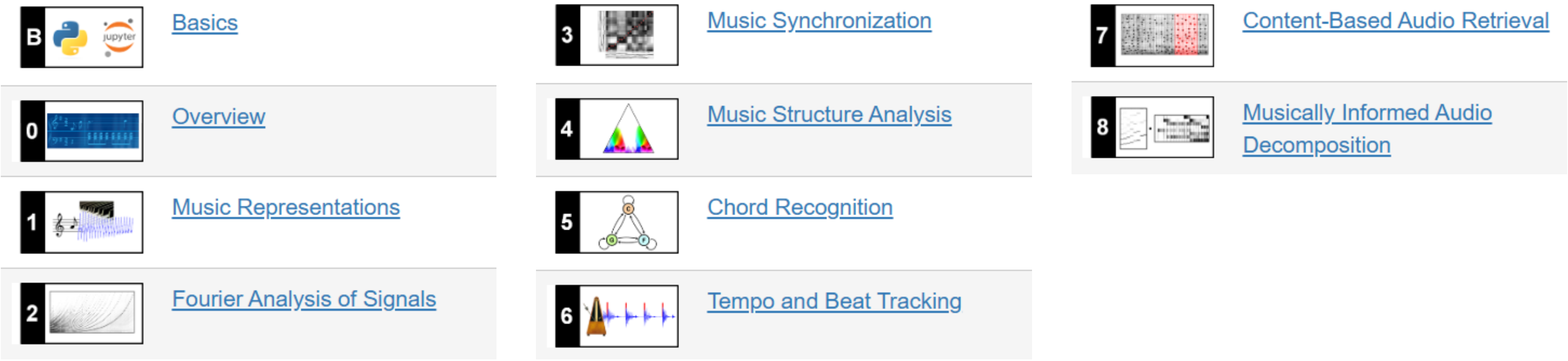
github.com/BreezeWhite/oemer

Resources on Optical Music Recognition (OMR)

- Jorge Calvo-Zaragoza, Jan Hajič jr., Alexander Pacha, and Ichiro Fujinaga, "Optical Music Recognition for Dummies," *Tutorials of ISMIR*, 2021. (slides)
- OMR Datasets: apacha.github.io/OMR-Datasets/

Resources on Music Information Research (MIR)

- Meinard Müller, "Fundamentals of Music Processing – Using Python and Jupyter Notebooks," *Springer*, 2021.
- Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.
  - Jupyter notebooks available at audiolabs-erlangen.de/FMP



(Source: Müller & Zalkow, 2019)

Resources on Music Information Research (MIR)

- Masataka Goto, Jin Ha Lee, and Meinard Müller, "Exploring 25 Years of Music Information Retrieval: Perspectives and Insights," *Tutorials of ISMIR*, 2024.
- Geoffroy Peeters, Gabriel Meseguer-Brocal, Alain Riou, and Stefan Lattner, "Deep Learning 101 for Audio-based MIR," *Tutorials of ISMIR*, 2024. (book)
- Keunwoo Choi, György Fazekas, Kyunghyun Cho, and Mark Sandler, "A Tutorial on Deep Learning for Music Information Retrieval," *arXiv preprint arXiv:1709:04396*, 2017. (code)