

PAT 463/563 (Fall 2026)

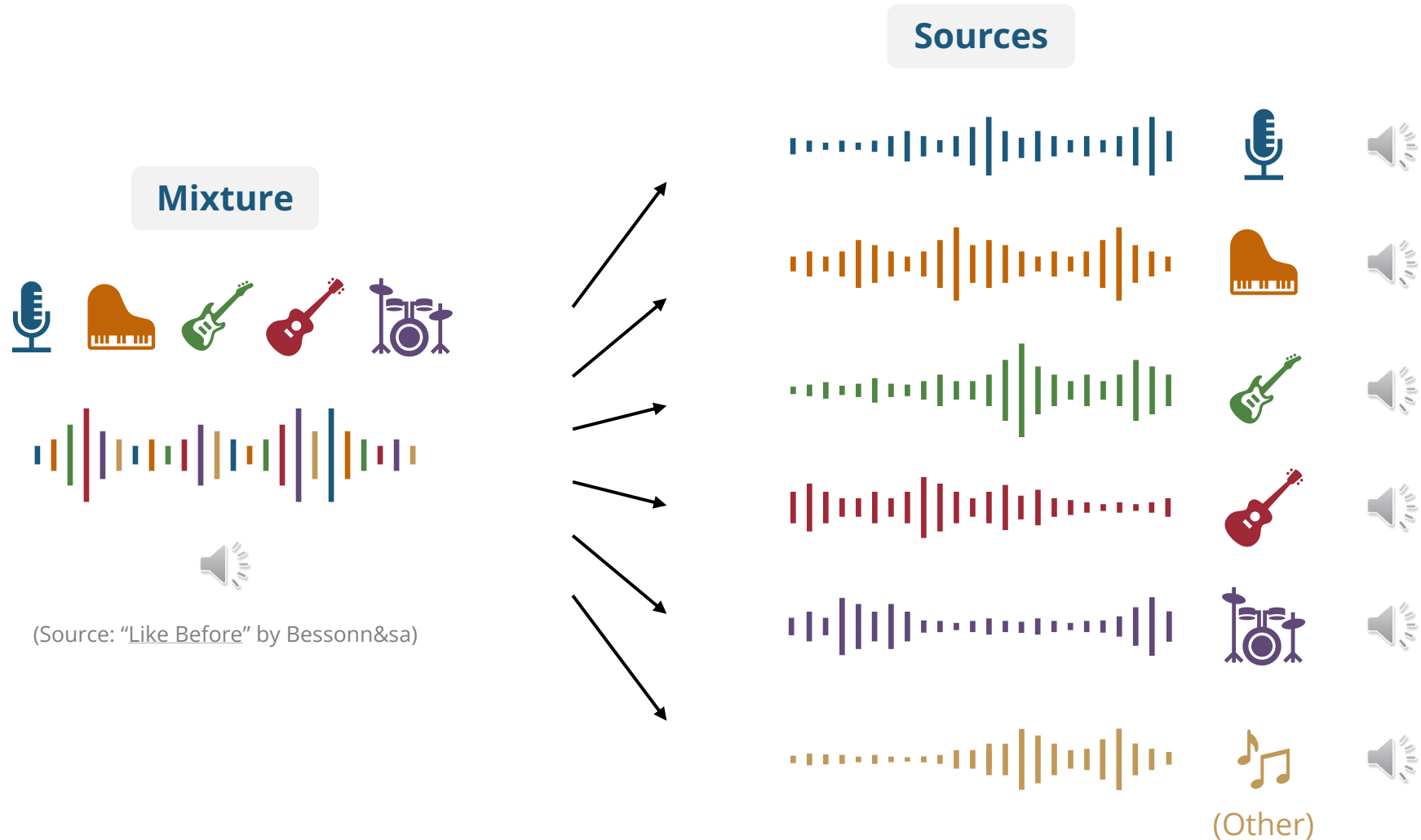
Music & AI

Lecture 9: Source Separation

Instructor: Hao-Wen (Herman) Dong

What is Source Separation?

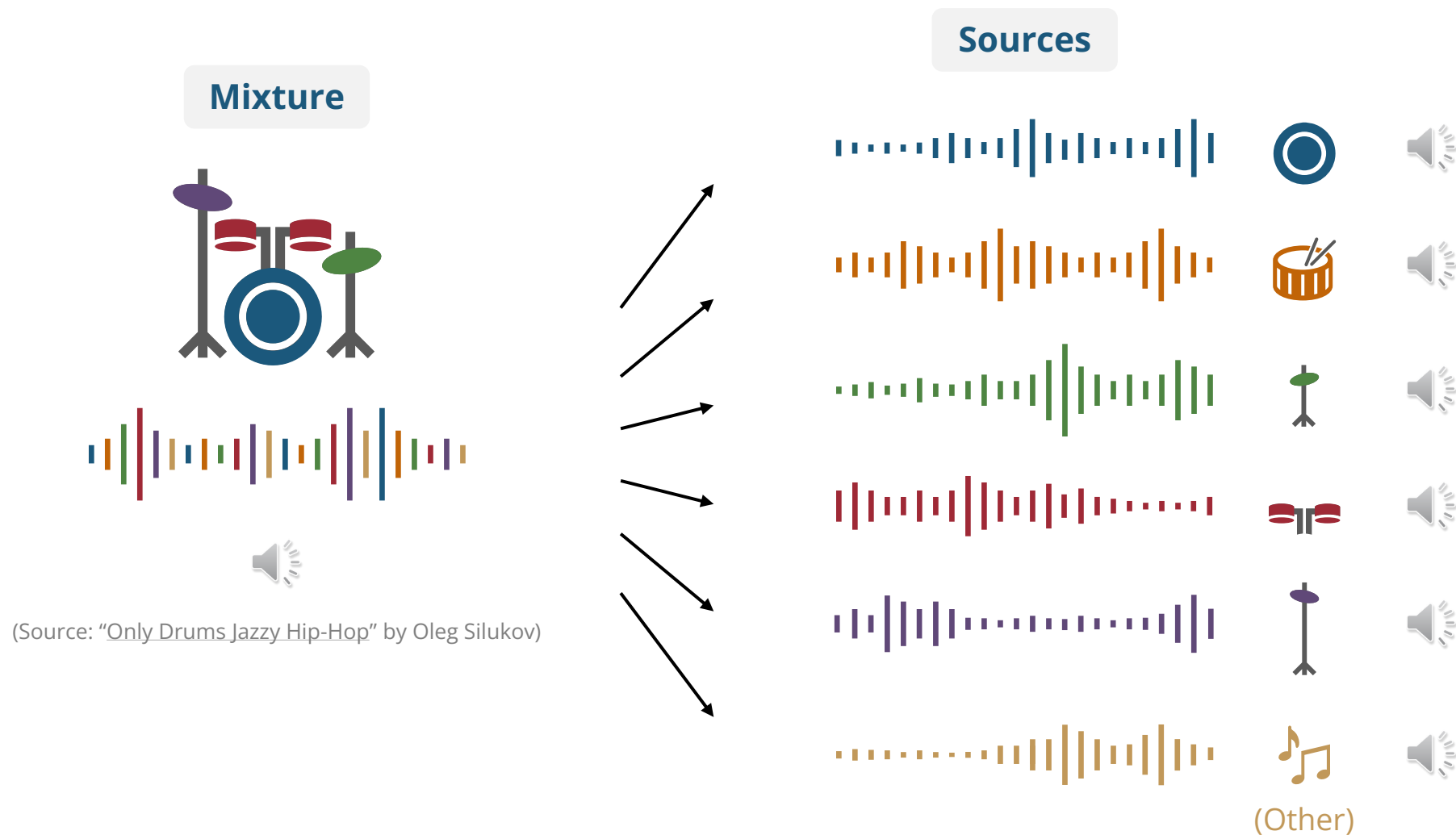
Source Separation



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.

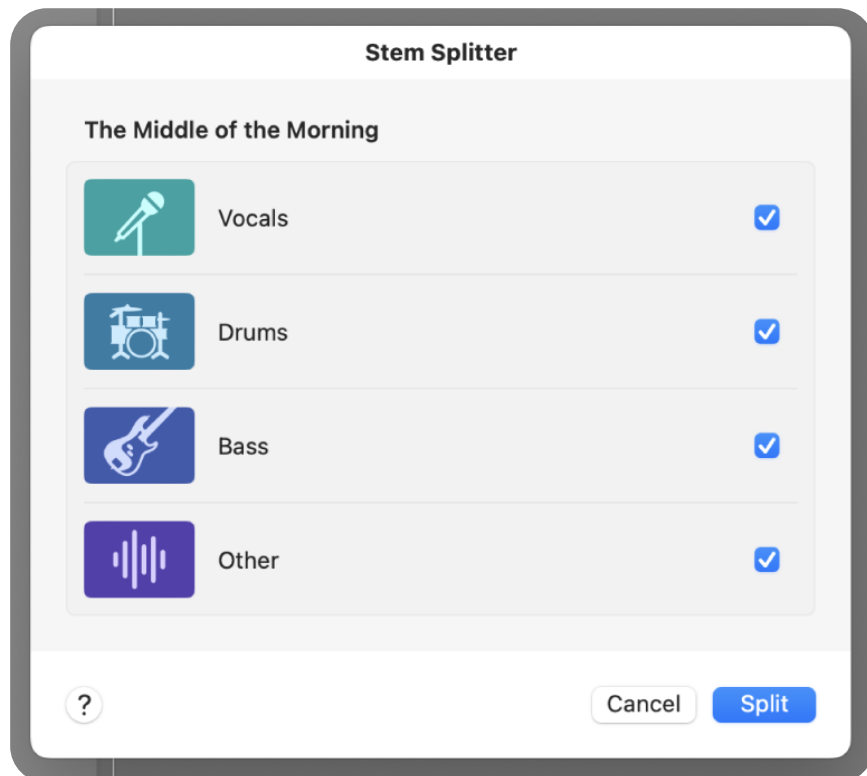
| Source Separation



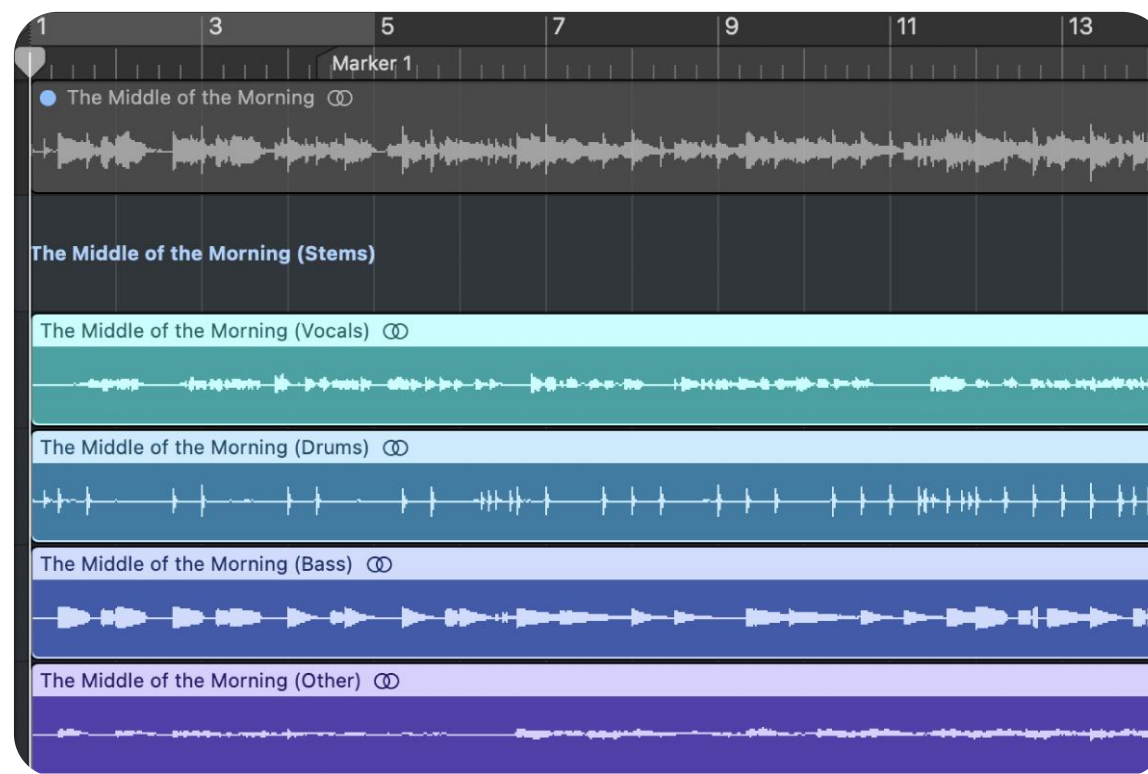
Oleg Silukov, [CC BY-NC-SA](#), via [Jamendo](#)

Illustration inspired by Ethan Manilow, Prem Seetharman, and Justin Salamon's Tutorial on "Open Source Tools & Data for Music Source Separation" at ISMIR 2020.

Stem Splitter in Logic Pro

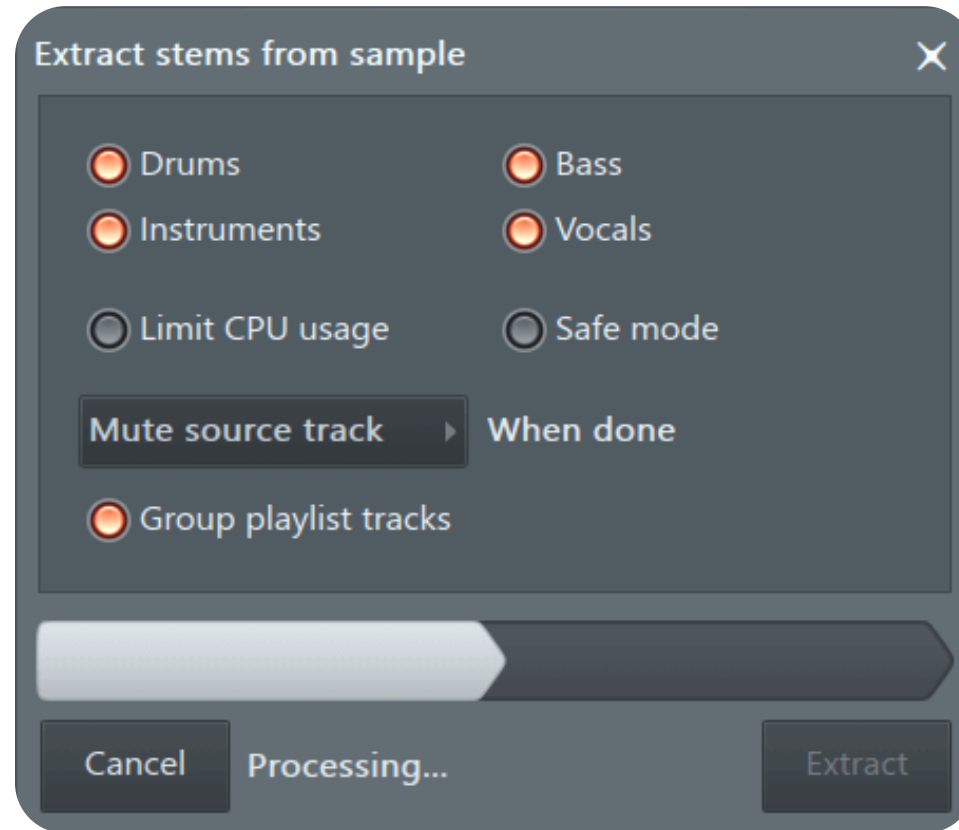


(Source: Logic Pro User Guide)



(Source: Logic Pro User Guide)

Extracting Stems from Sample in FL Studio

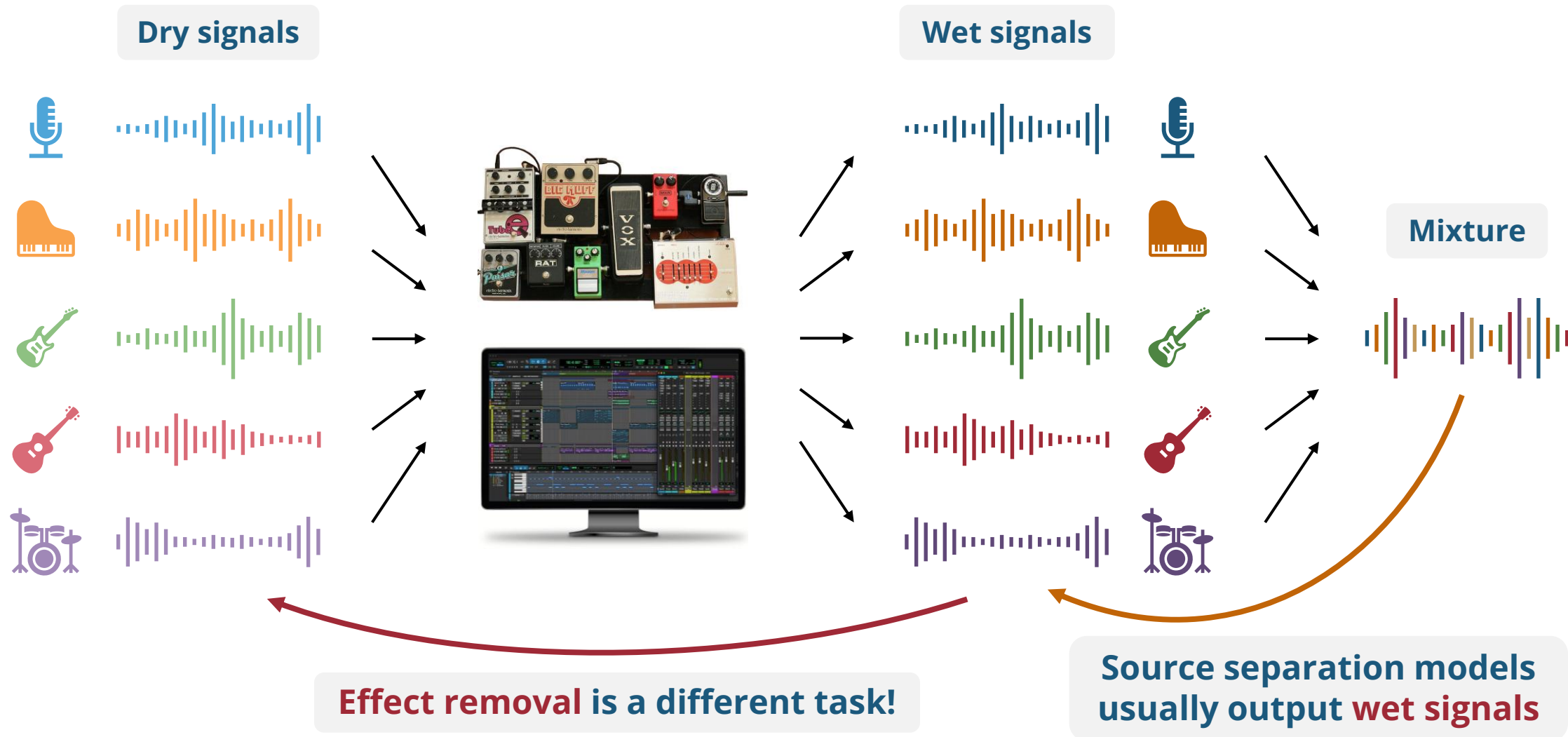


(Source: FL Studio Reference Manual)

Source Separation: The Many Names

- Also known as
 - **Stem separation**
 - **Stem splitter**
 - **Music demixing** → slightly different meaning
 - **Stem extraction** → slightly different meaning

Source Separation does NOT Remove Effects

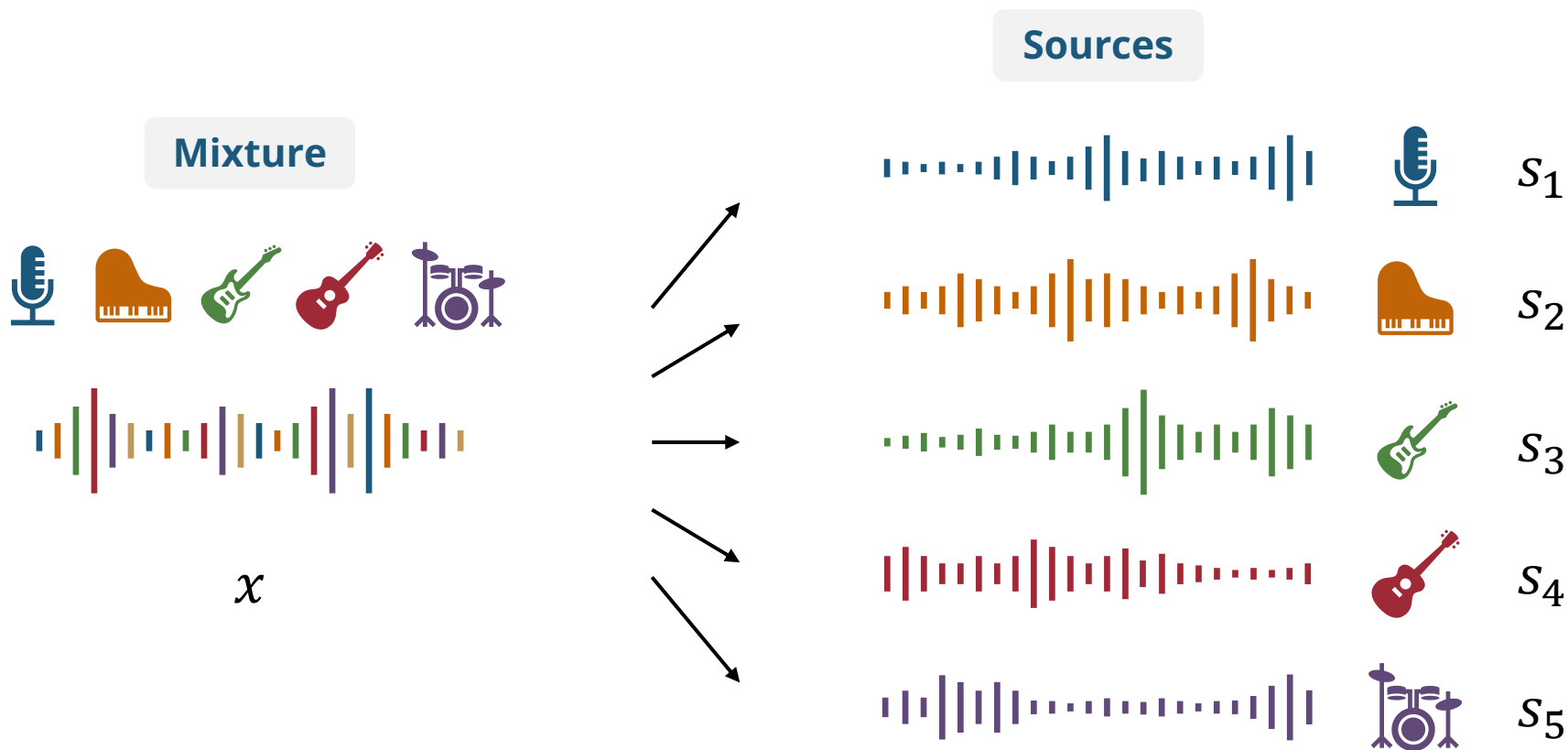


| Applications of Source Separation

- **Musical applications**
 - Remixing & sampling
 - Music practicing & education
 - Karaoke accompaniment generation
- **MIR tasks** (Oftentimes source separation is the first step)
 - Music transcription
 - Musical instrument & vocal detection
 - Singer identification
 - Lyric recognition
 - Lyric-to-music alignment

How does it work?

Mathematical Formulation



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

Source Separation is an Ill-posed Problem

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

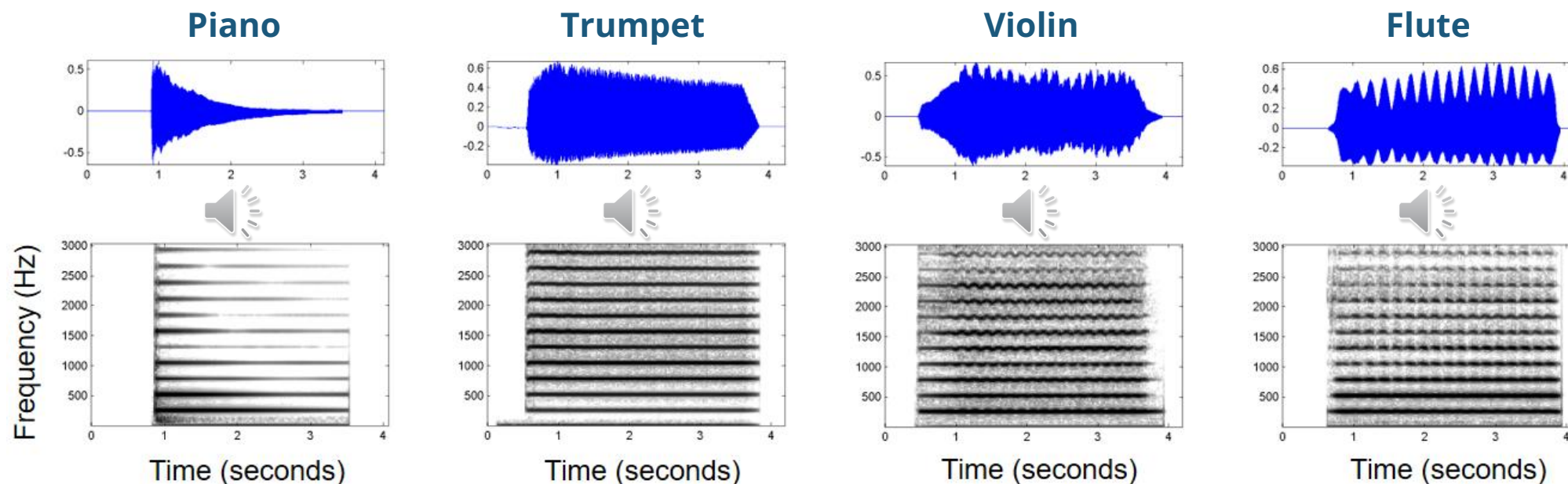
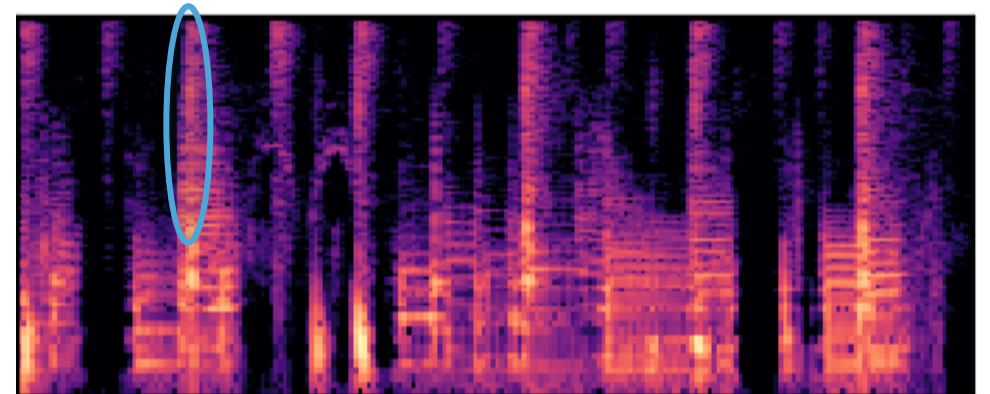
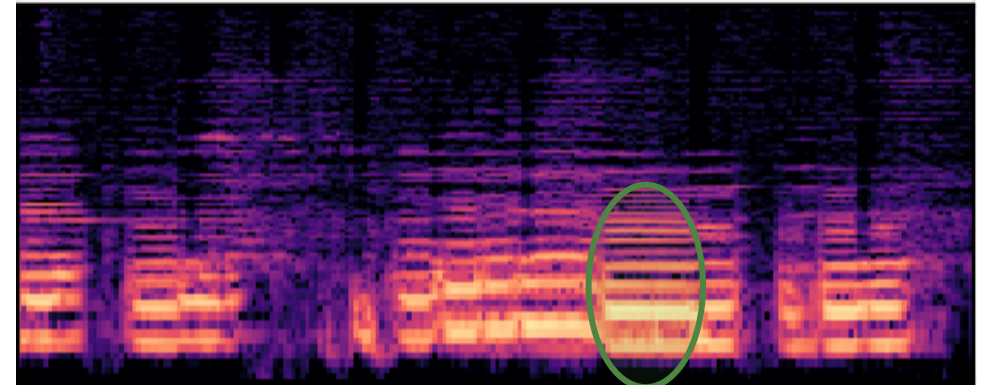
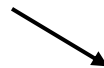
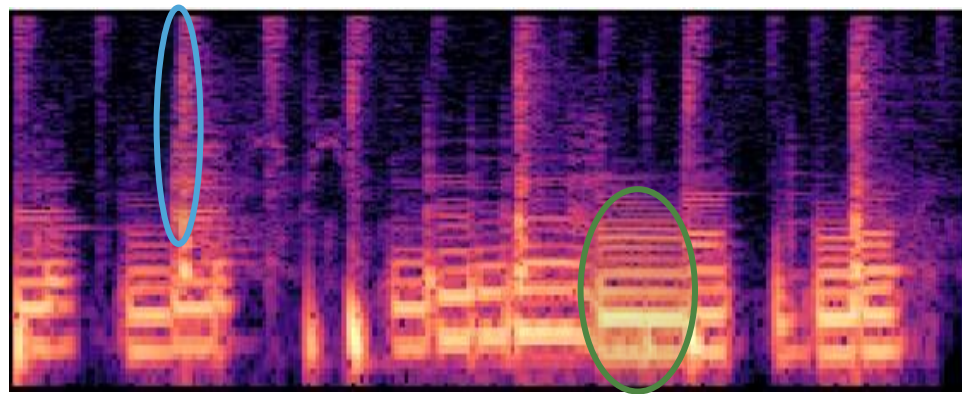


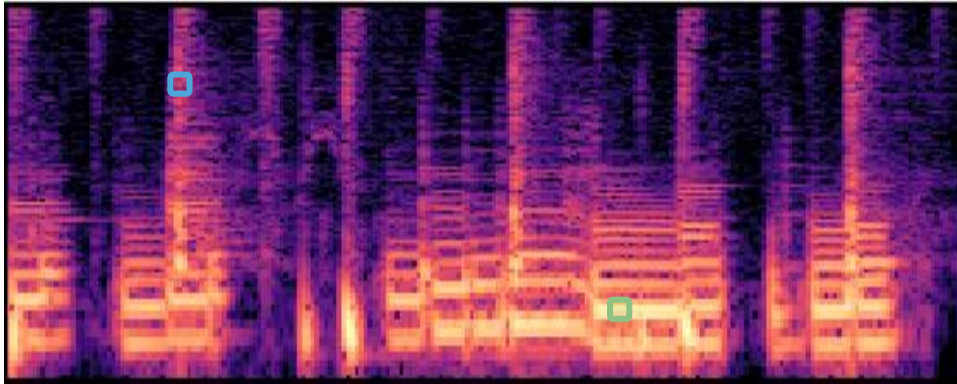
Figure 1.23 from [Müller, FMP, Springer 2015]

(Source: Müller et al., 2021)

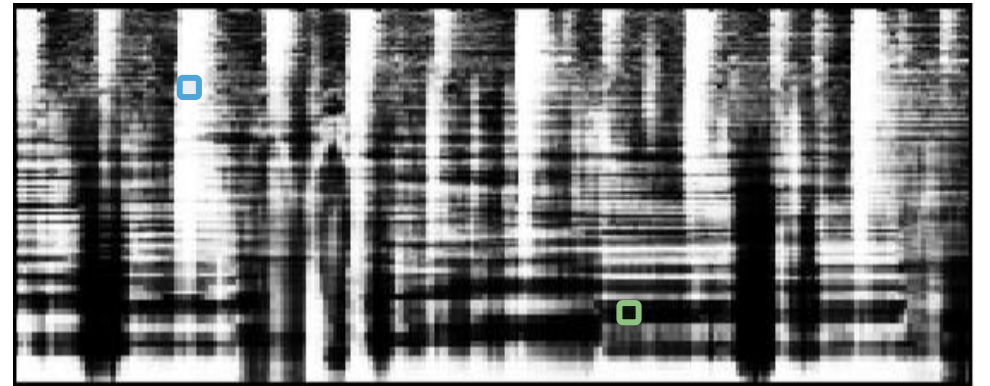
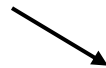
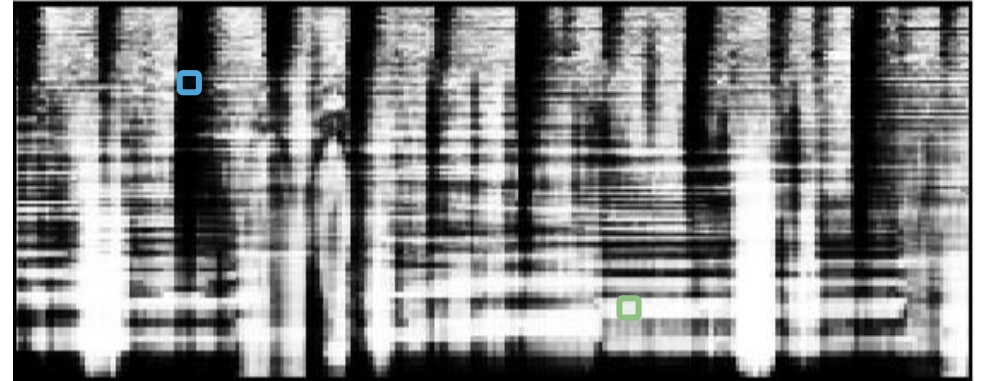
| Time-Frequency Masking



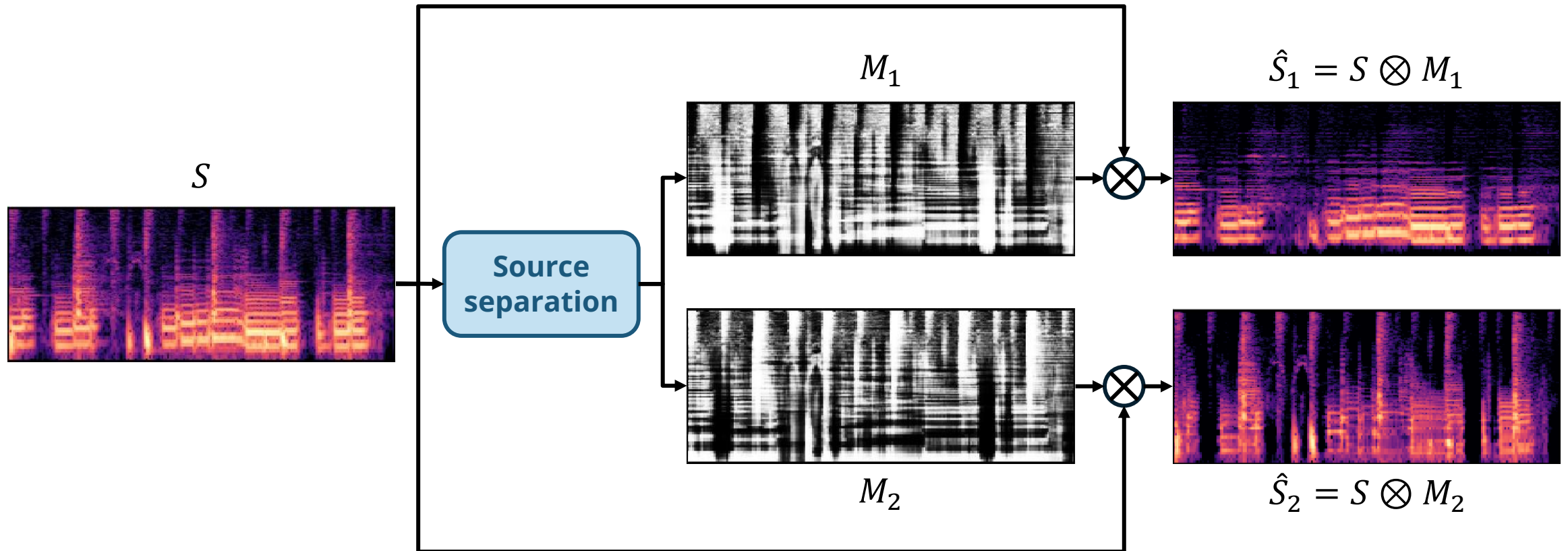
| Time-Frequency Masking



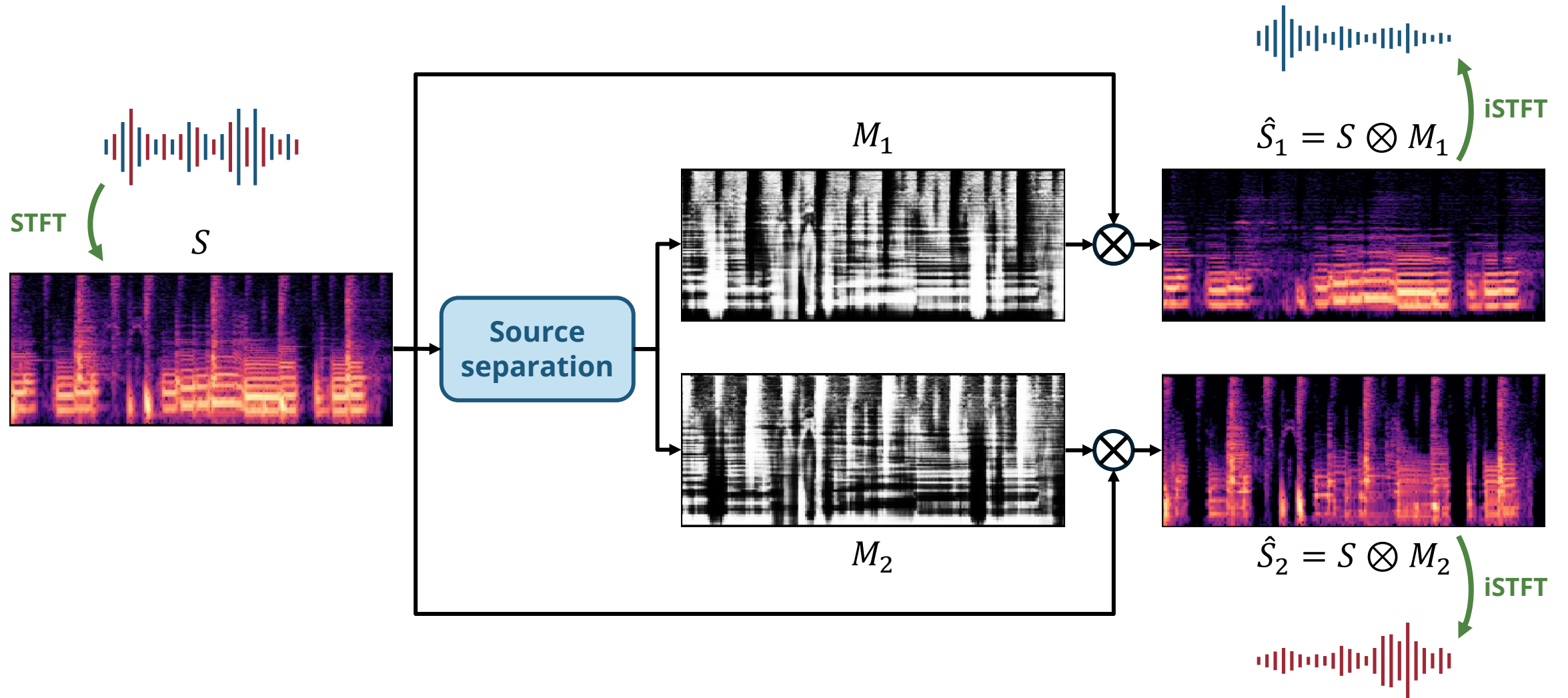
Classification on each time-frequency bin



Time-Frequency Masking

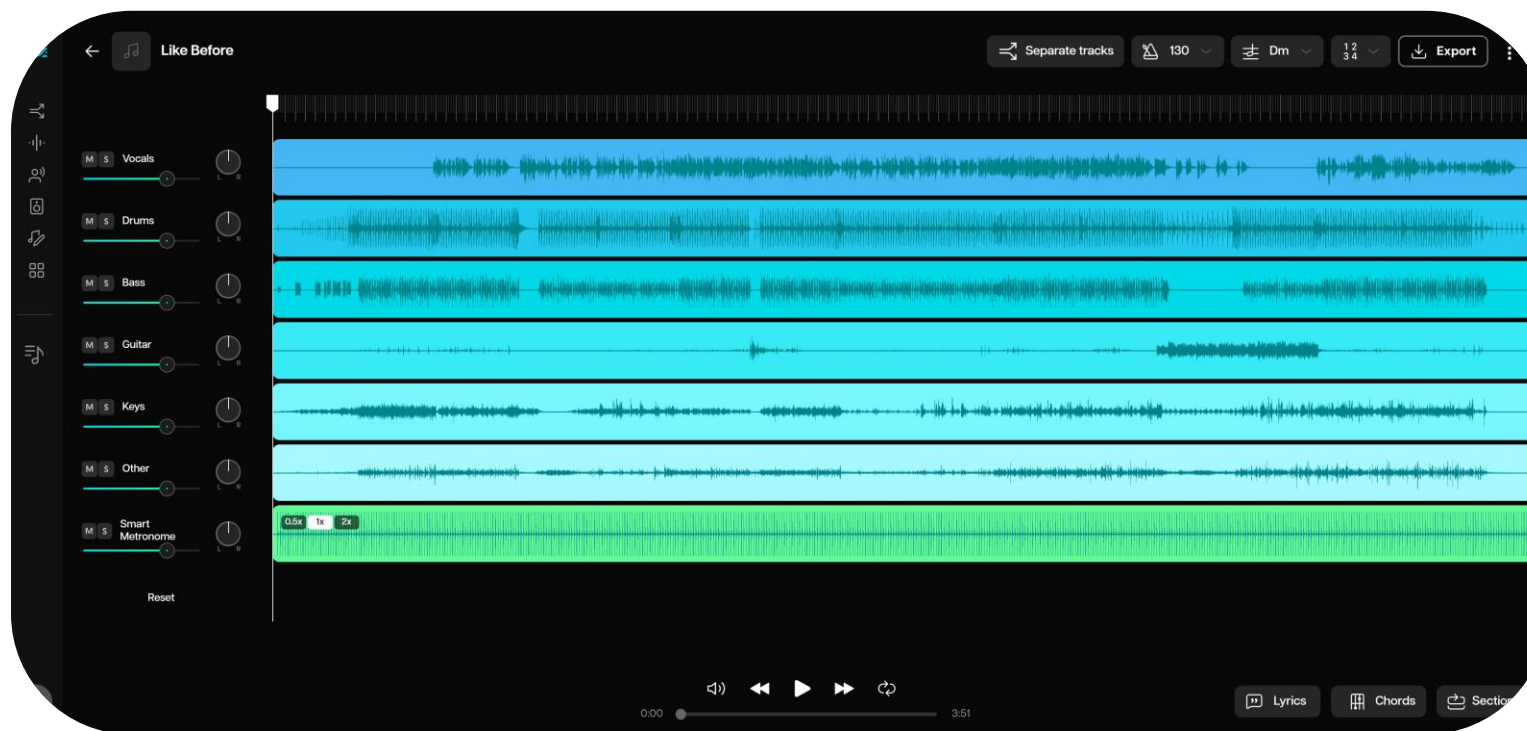


Time-Frequency Masking



🔥 HW 2: Source Separation

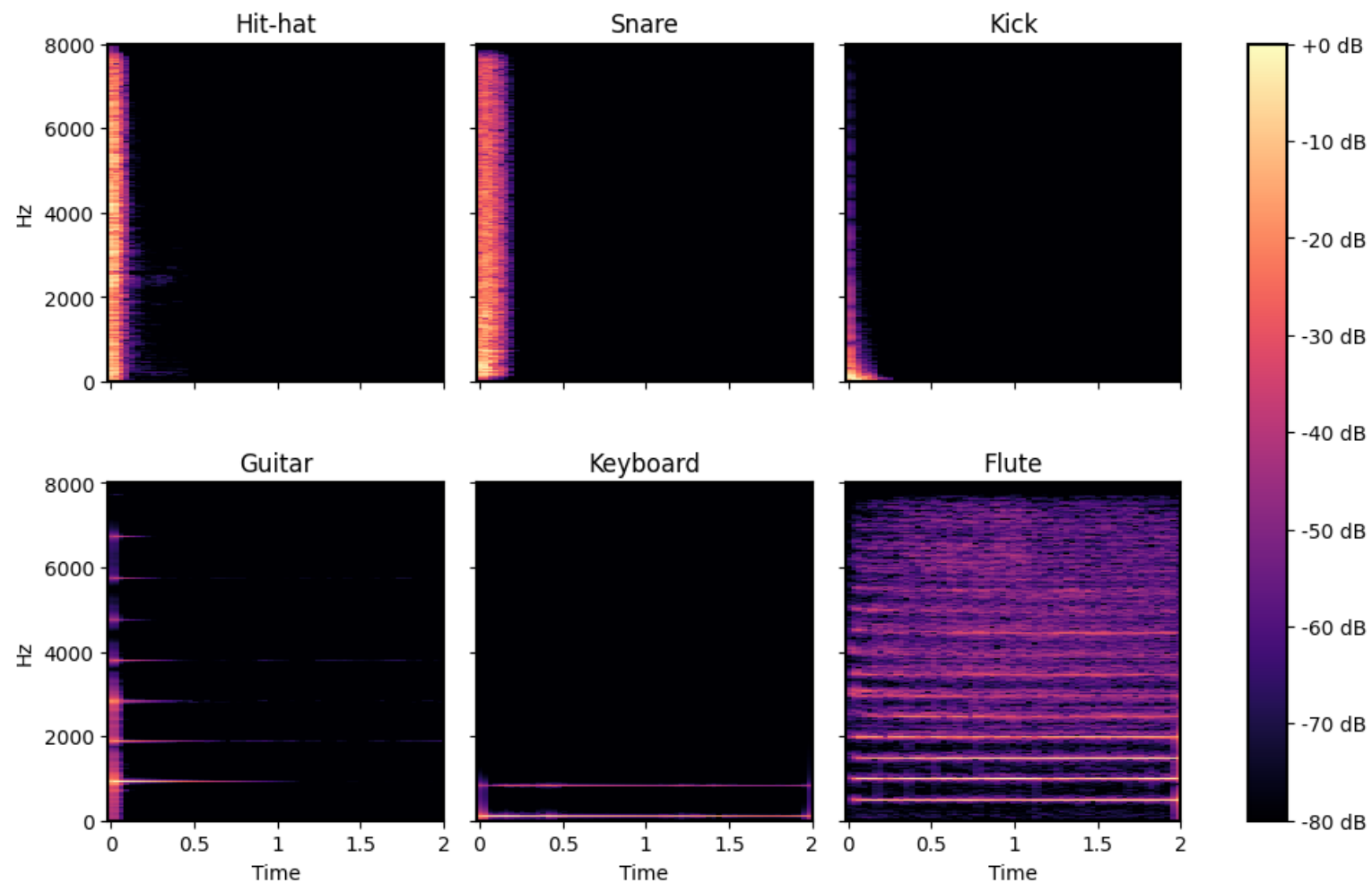
- Instructions will be sent by **emails** and released on the **course website**
- **Part 1:** Music Source Separation



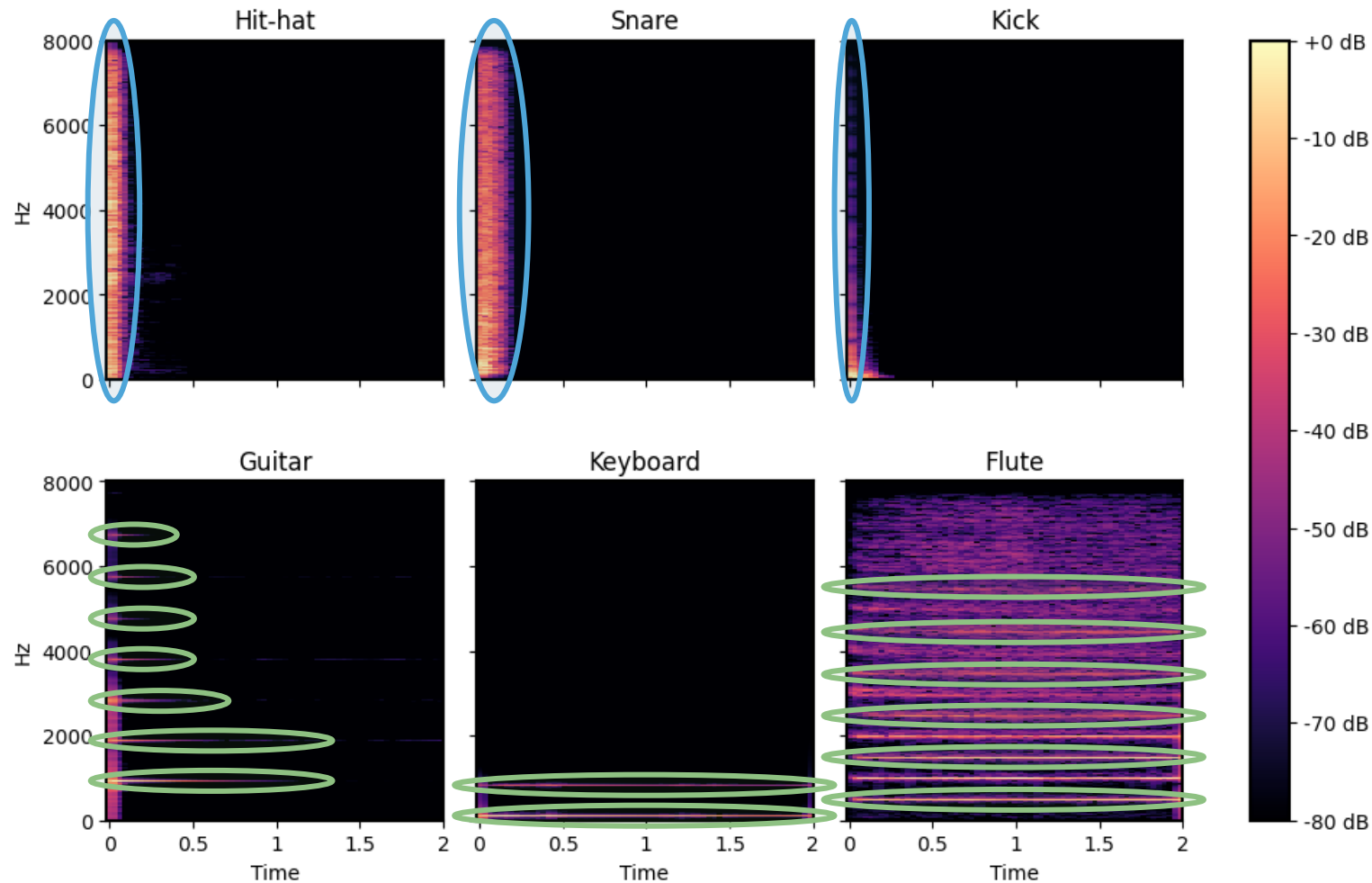
(Source: Moises)

Traditional Approaches

Percussive vs. Harmonic Components



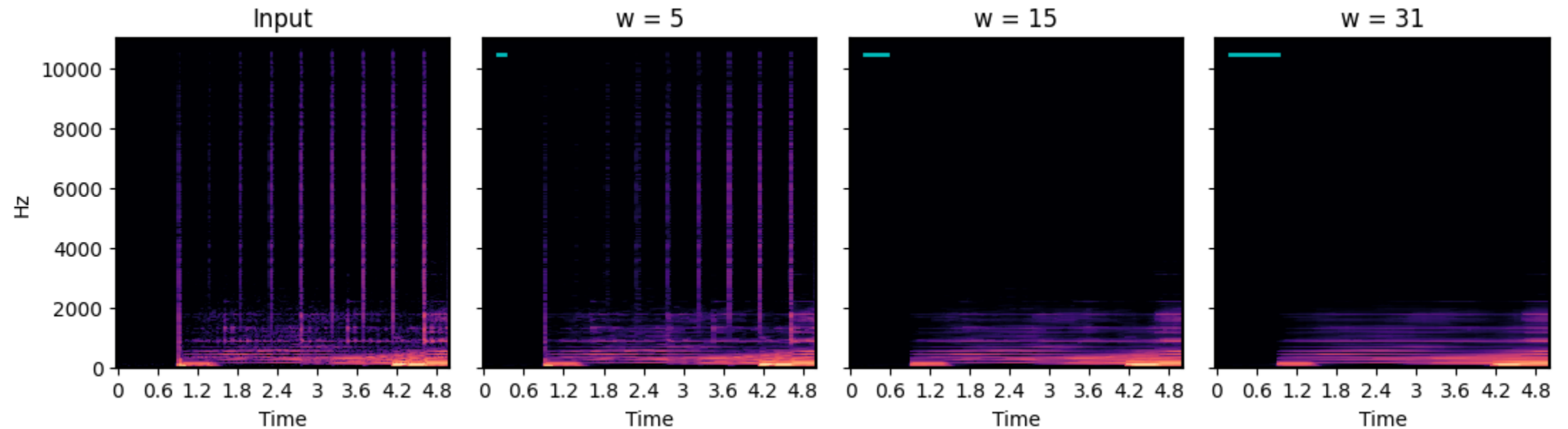
Percussive vs. Harmonic Components



Percussive sounds
tend to look **vertical**
in a spectrogram

Harmonic sounds
tend to look **horizontal**
in a spectrogram

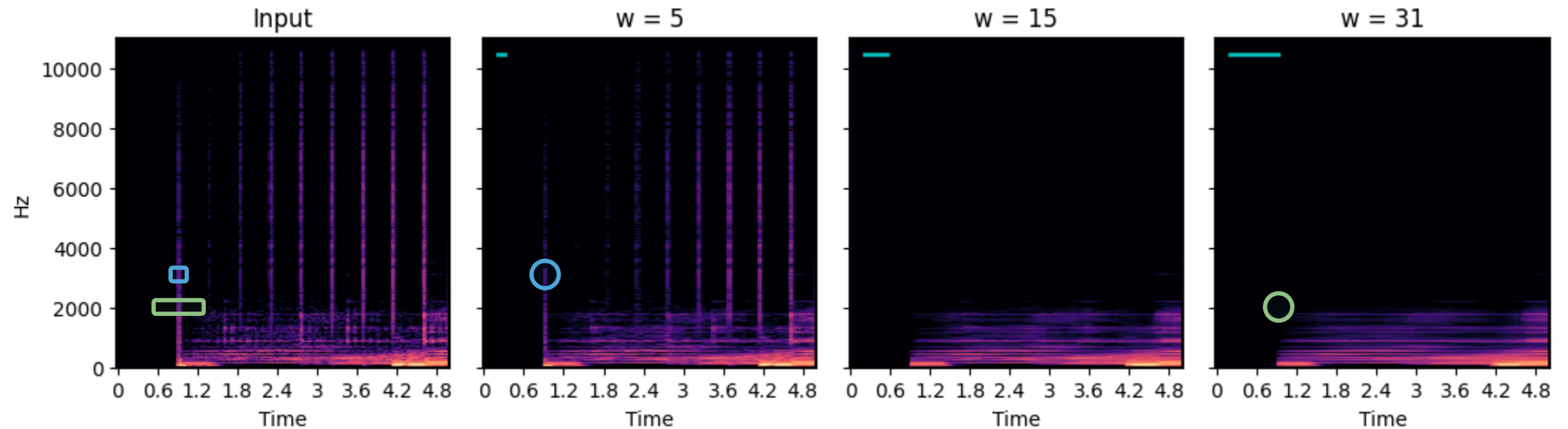
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



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.

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

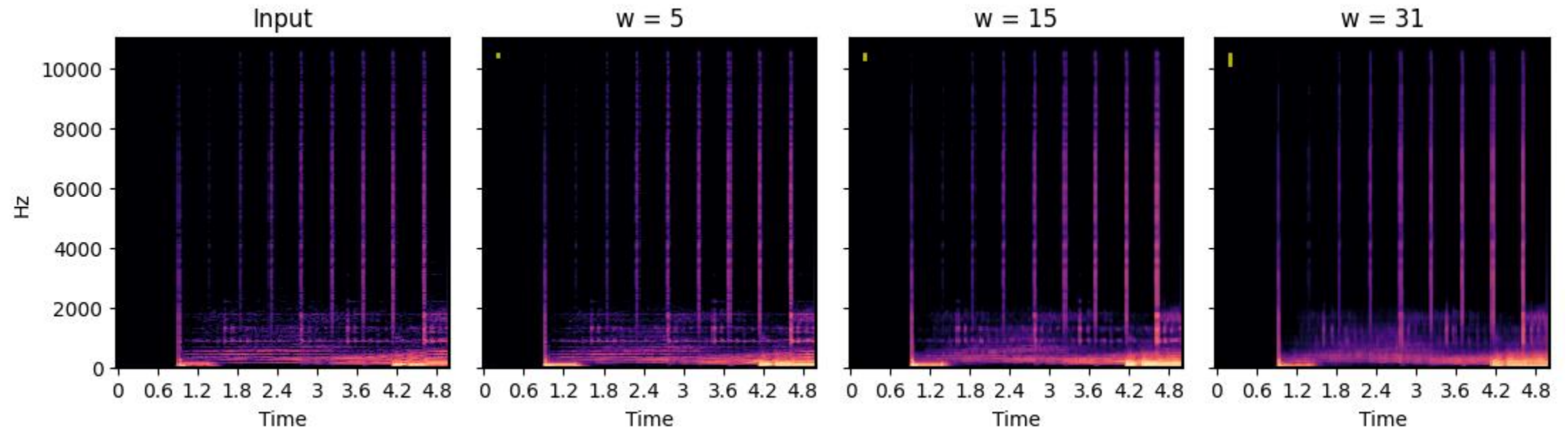


Applying a median filter over the time axis makes percussive patterns less prominent!

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.

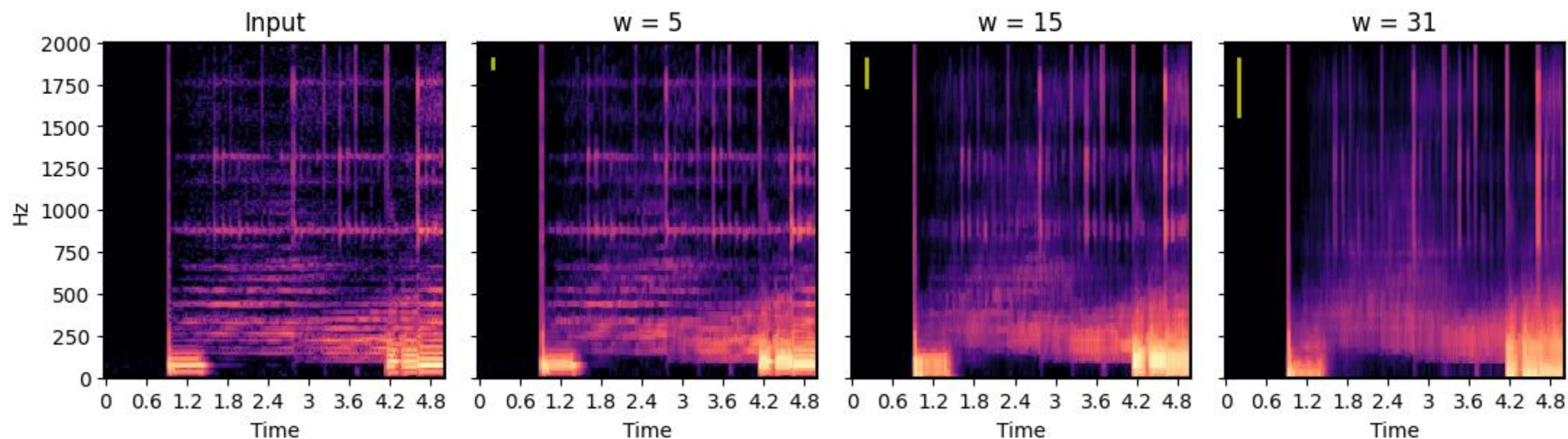
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



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.

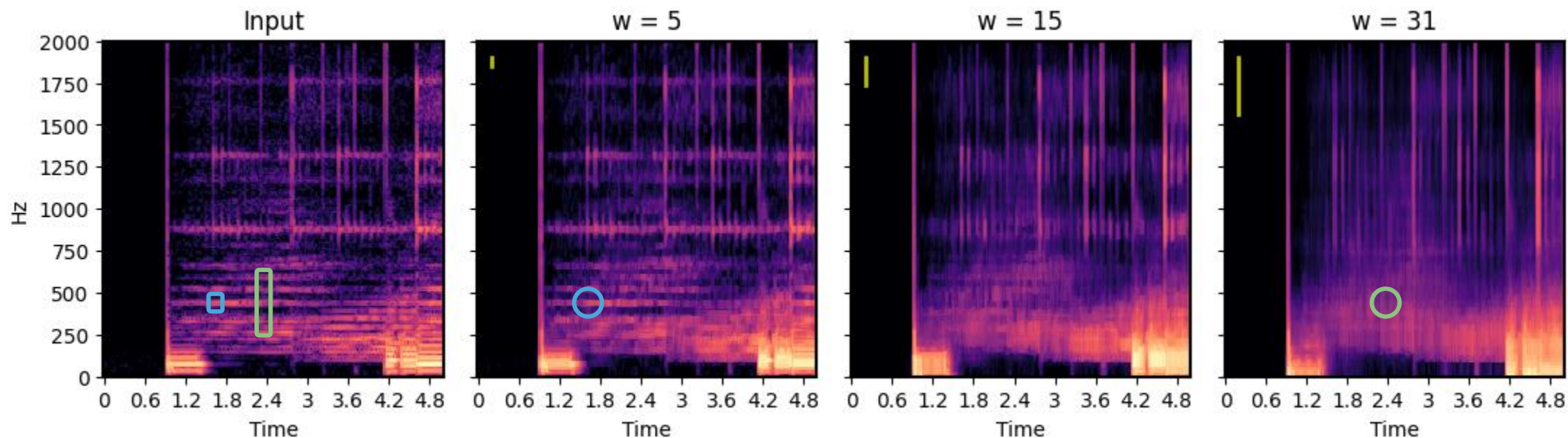
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



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.

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

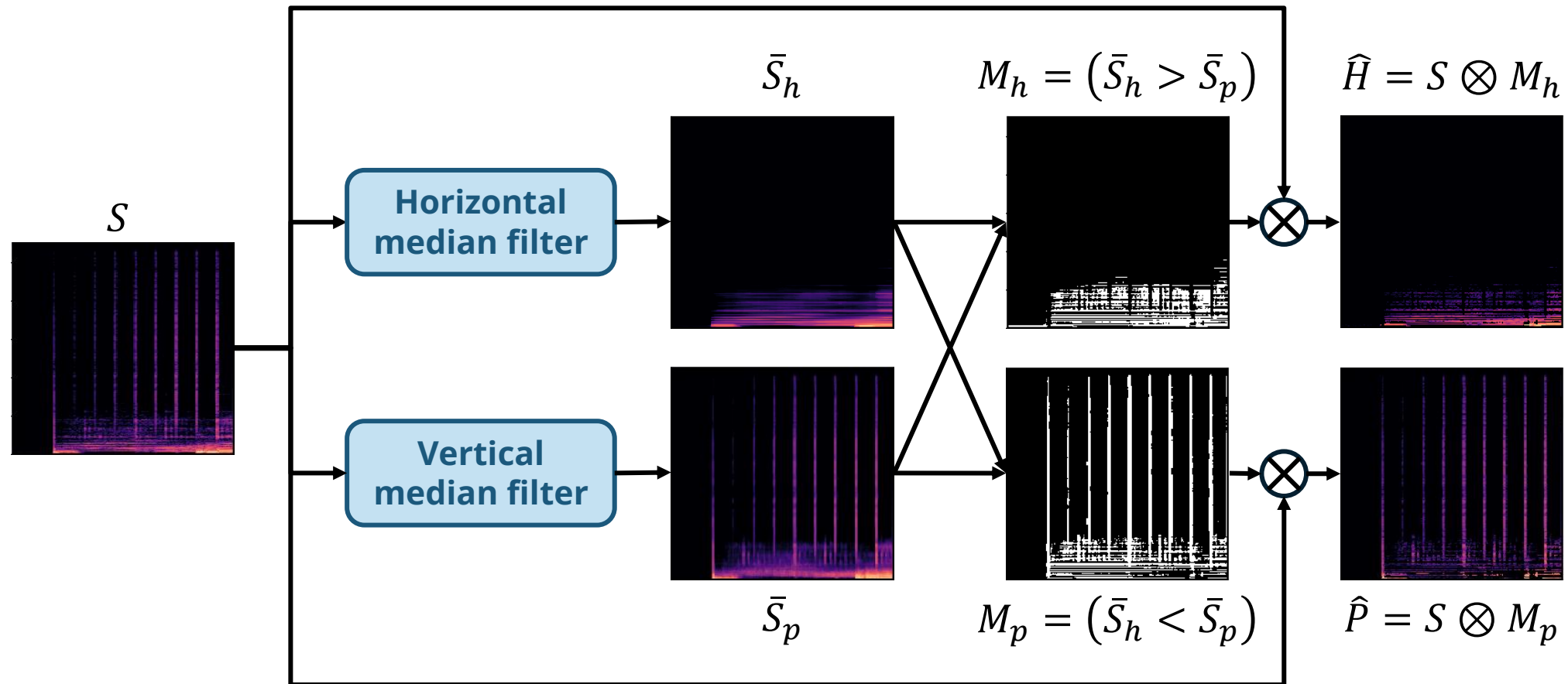


Applying a median filter over the frequency axis makes harmonic patterns less prominent!

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.

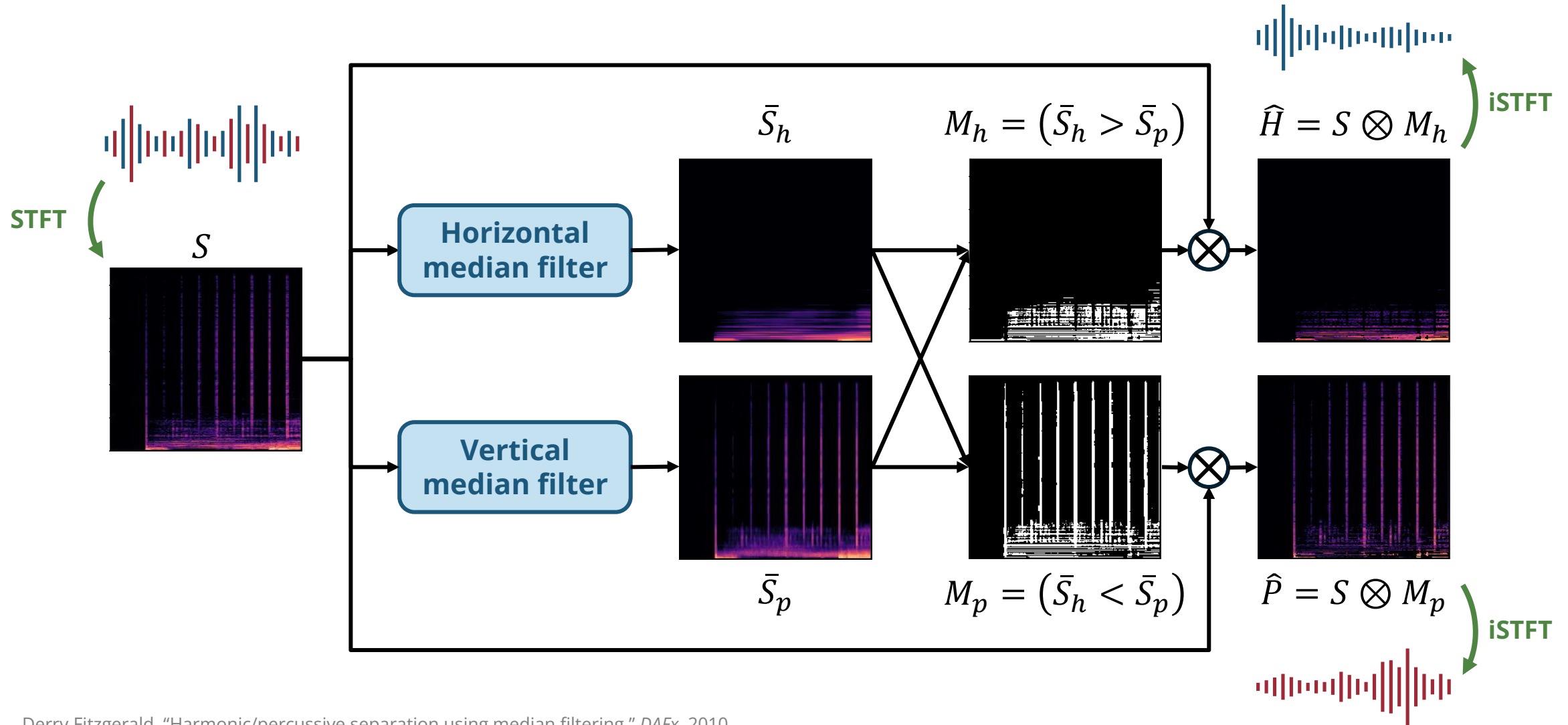
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



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.

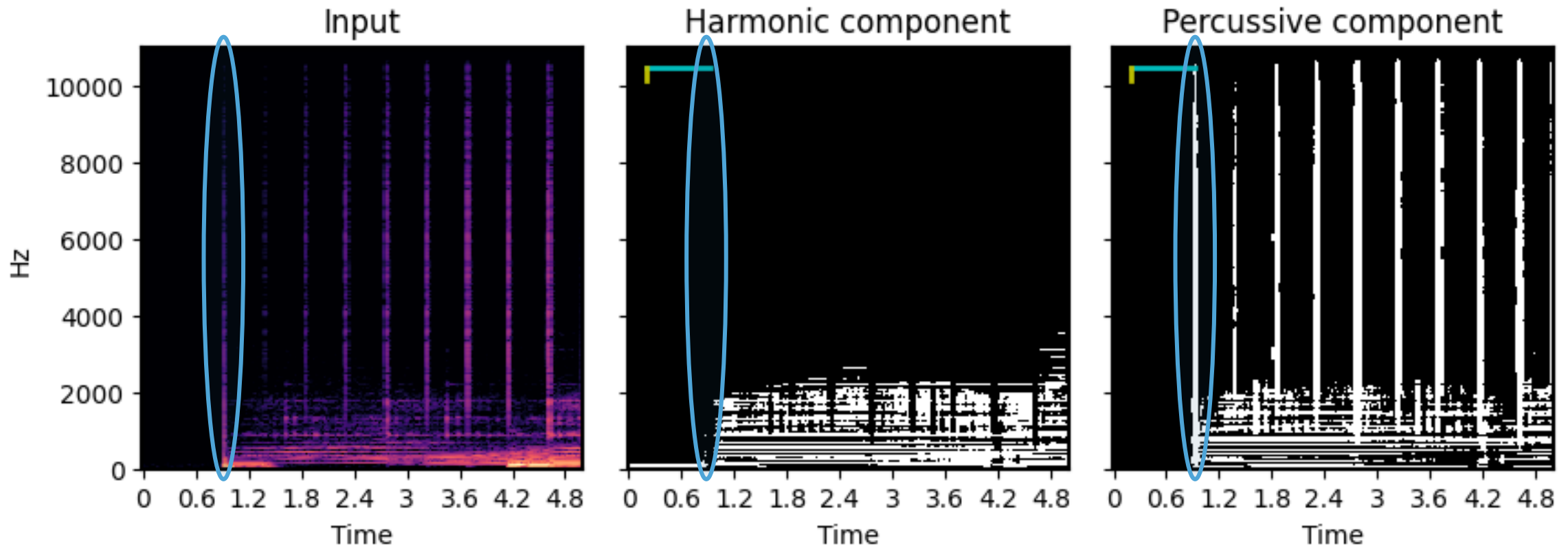
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



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.

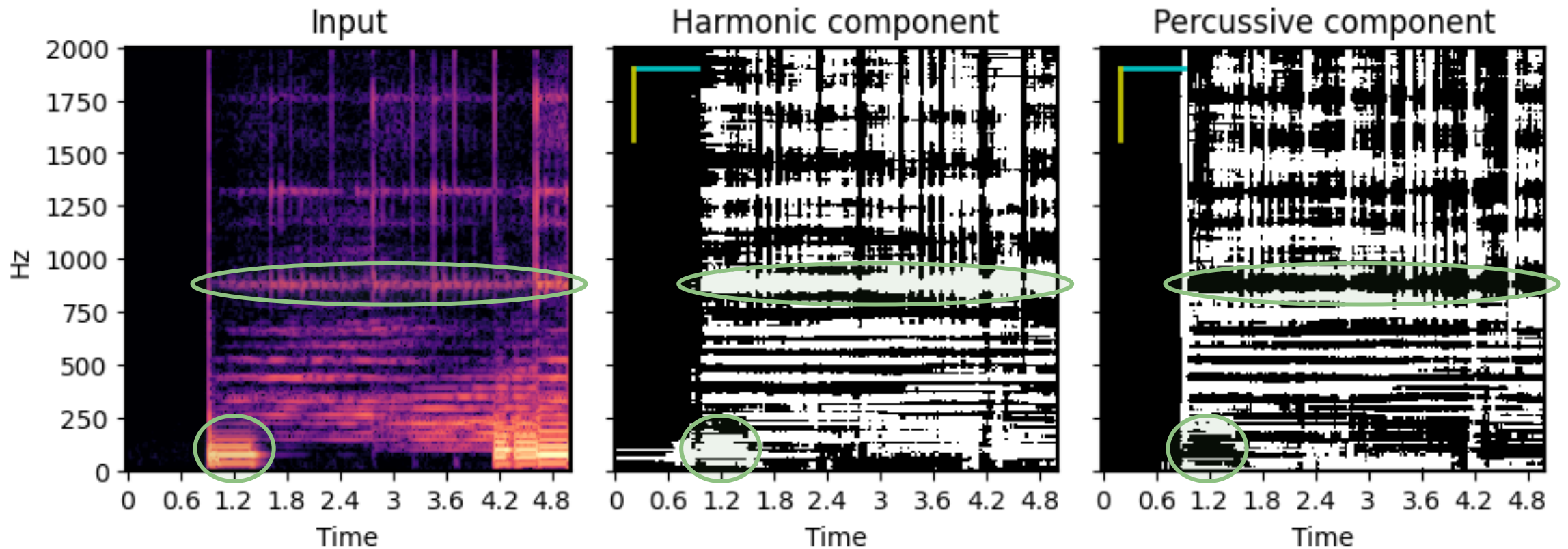
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



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.

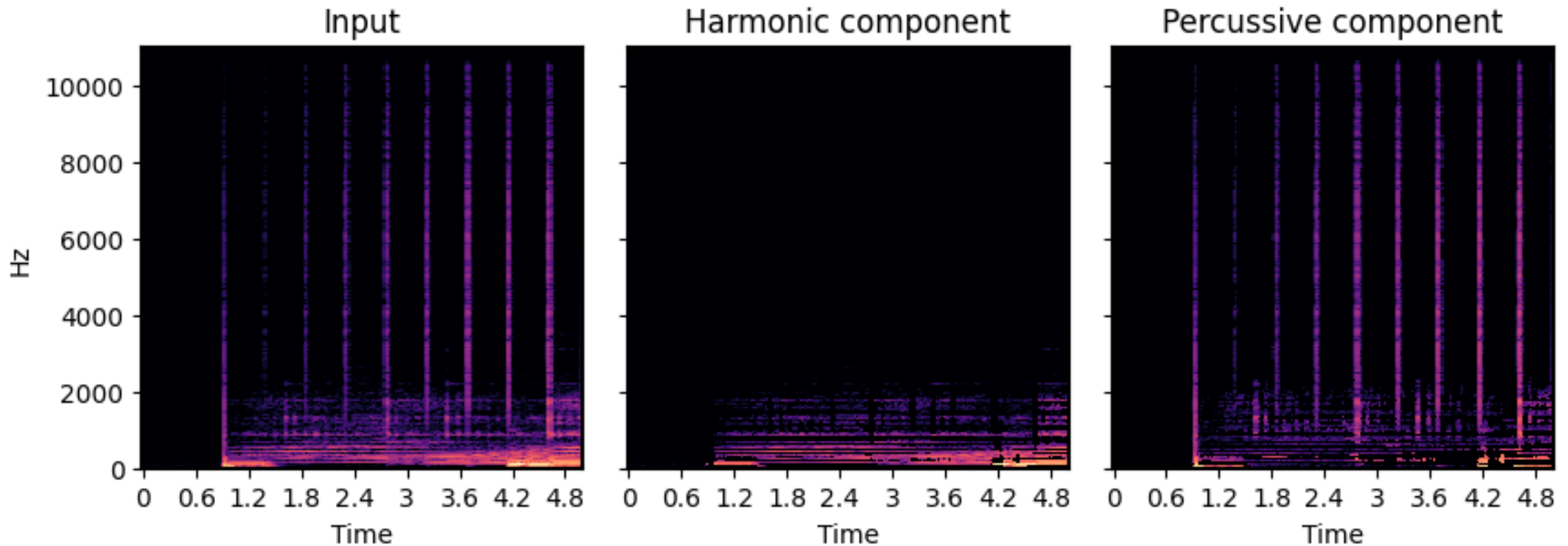
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



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.

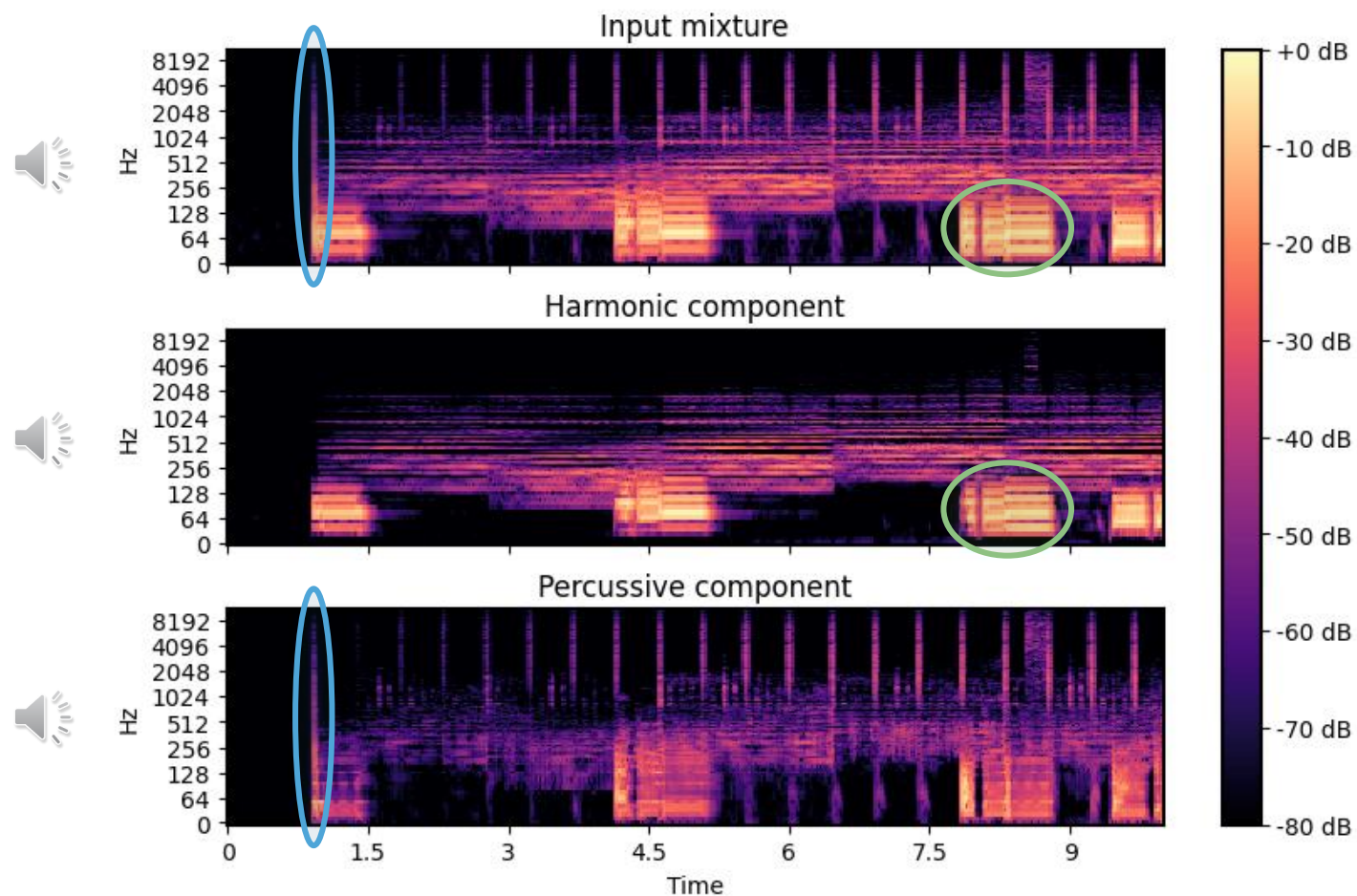
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



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.

HPSS: Example Result



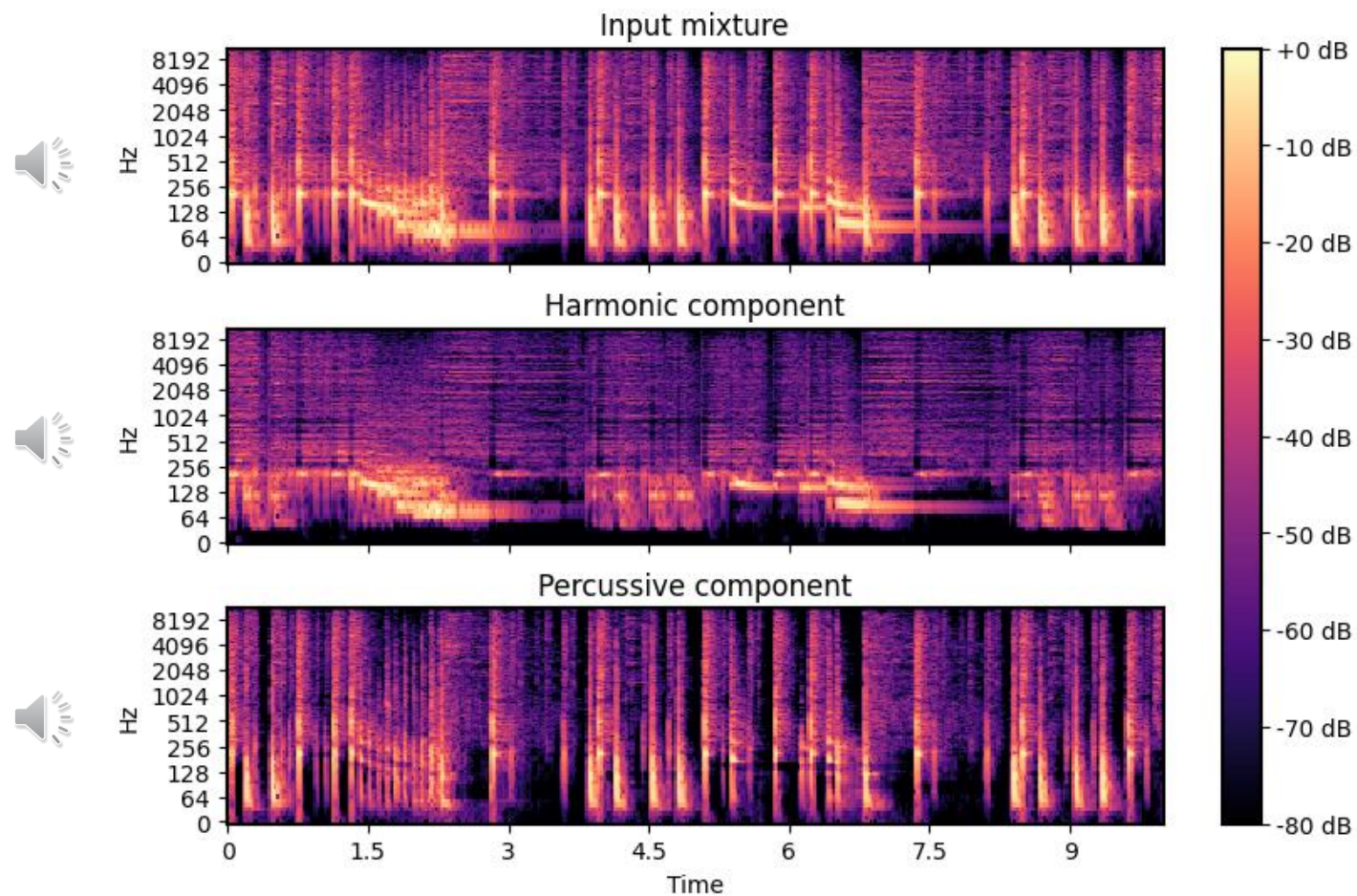
(Source: "Like Before" by Bessonn&sa)

Bessonn&sa, CC BY-NC-SA, via Jamendo

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.

HPSS: Example Result



(Source: "Only Drums Jazzy Hip-Hop" by Oleg Silukov)

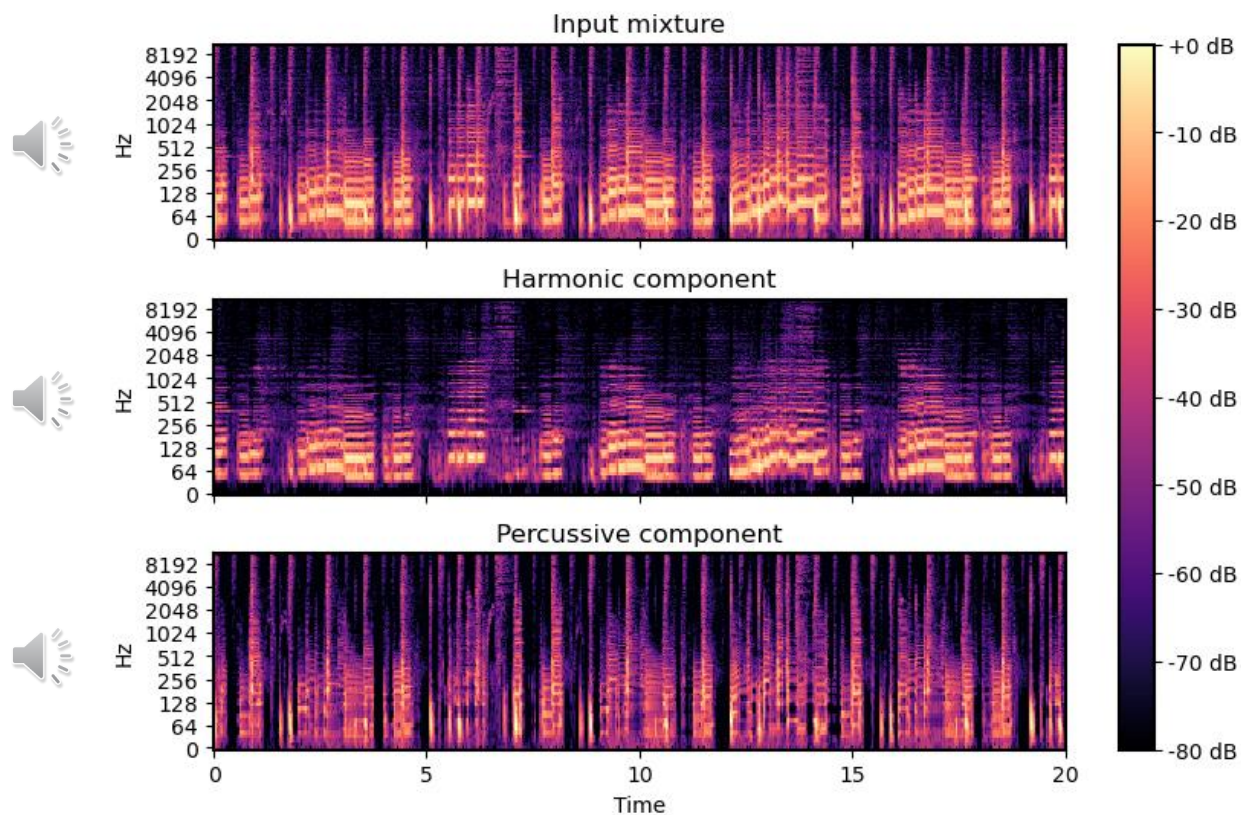
Oleg Silukov, [CC BY-NC-SA](#), via [Jamendo](#)

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.

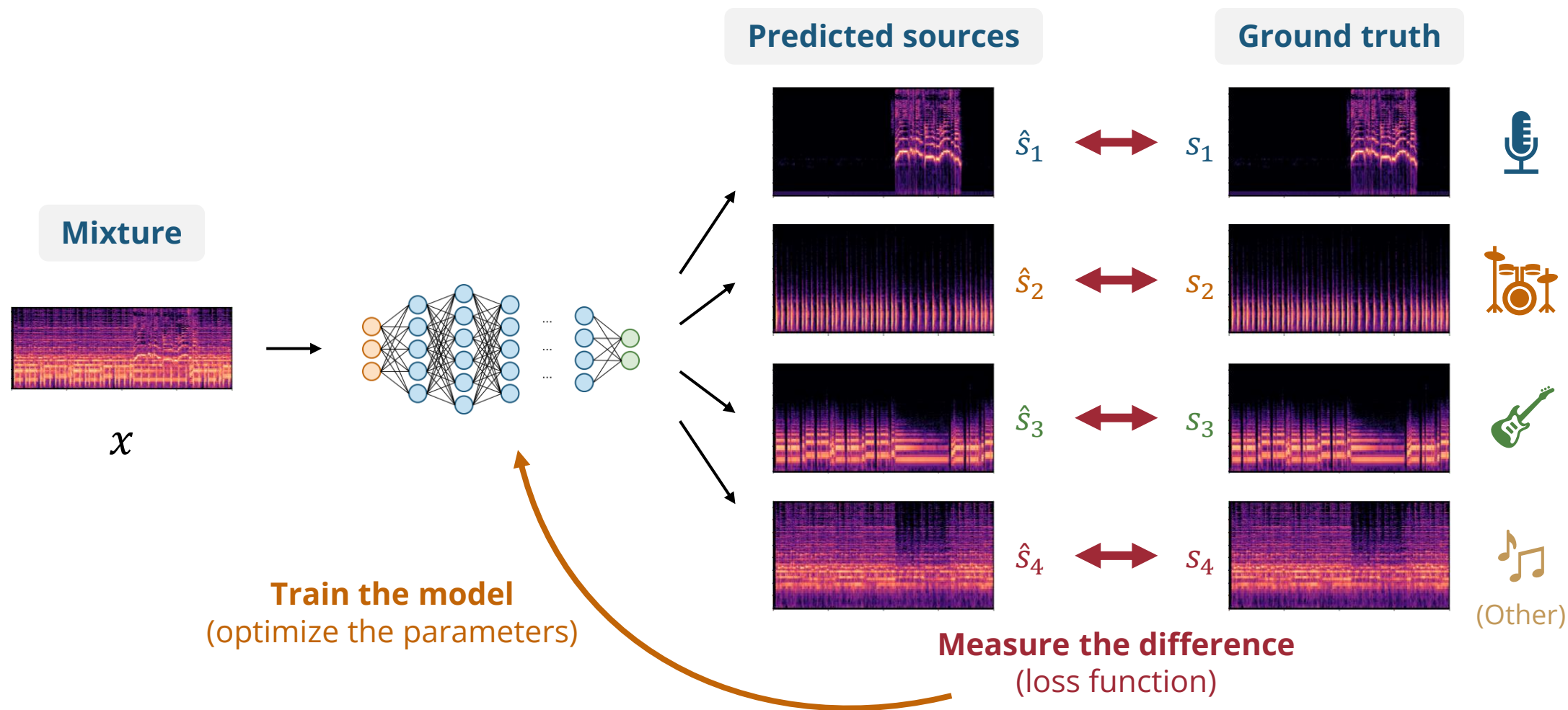
🔥 PA 3: Source Separation

- Instructions will be sent by **emails** and released on the **course website**
- **Part 1:** Harmonic-Percussive Source Separation (HPSS) using **librosa**

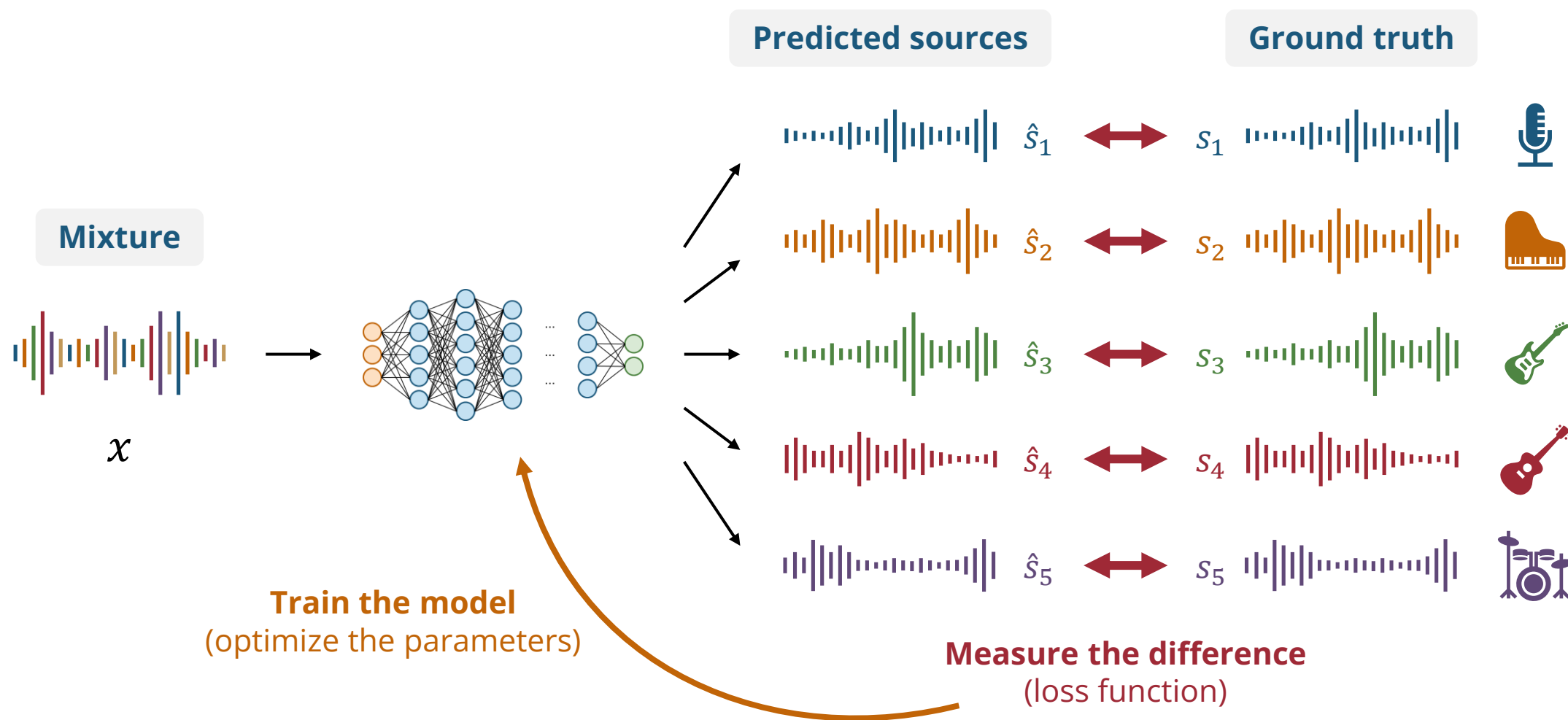


Deep Learning Approaches

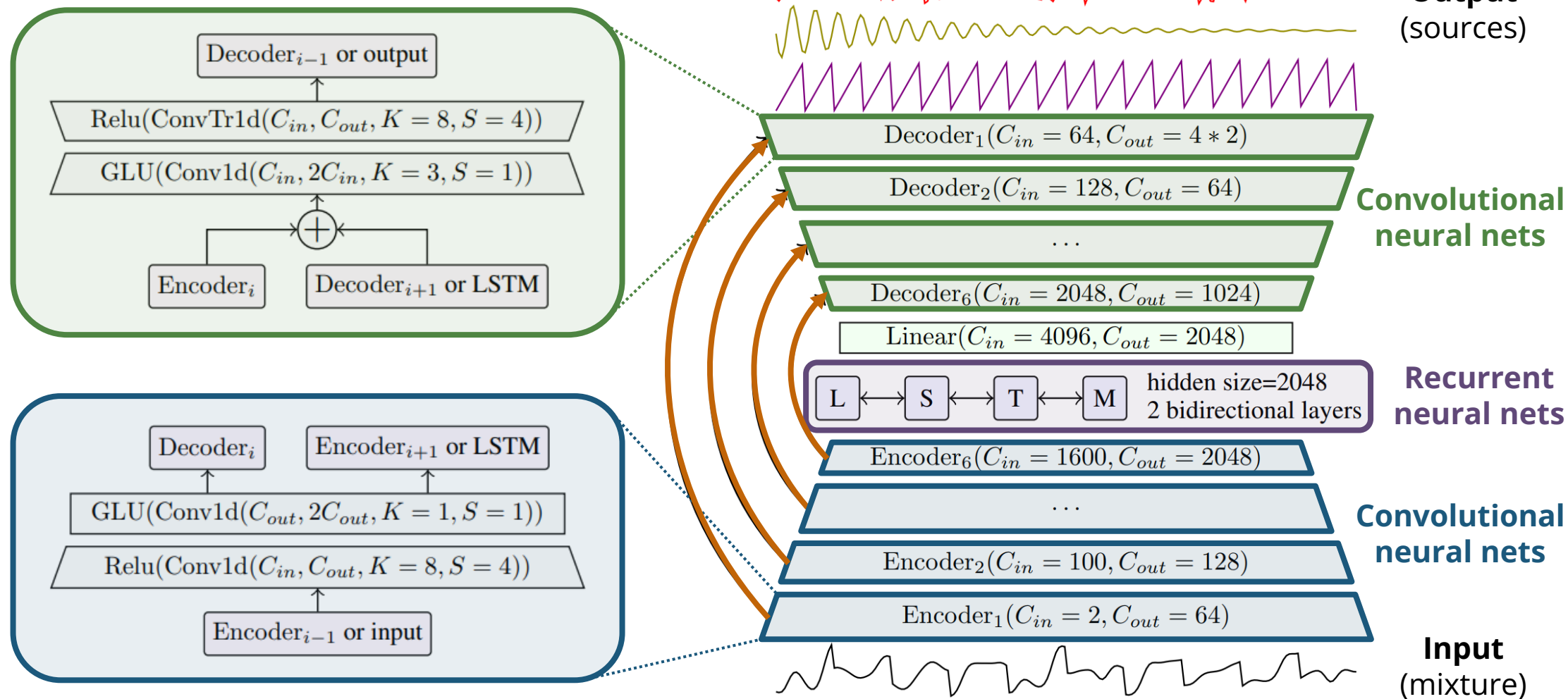
Deep Learning Based Source Separation



Deep Learning Based Source Separation



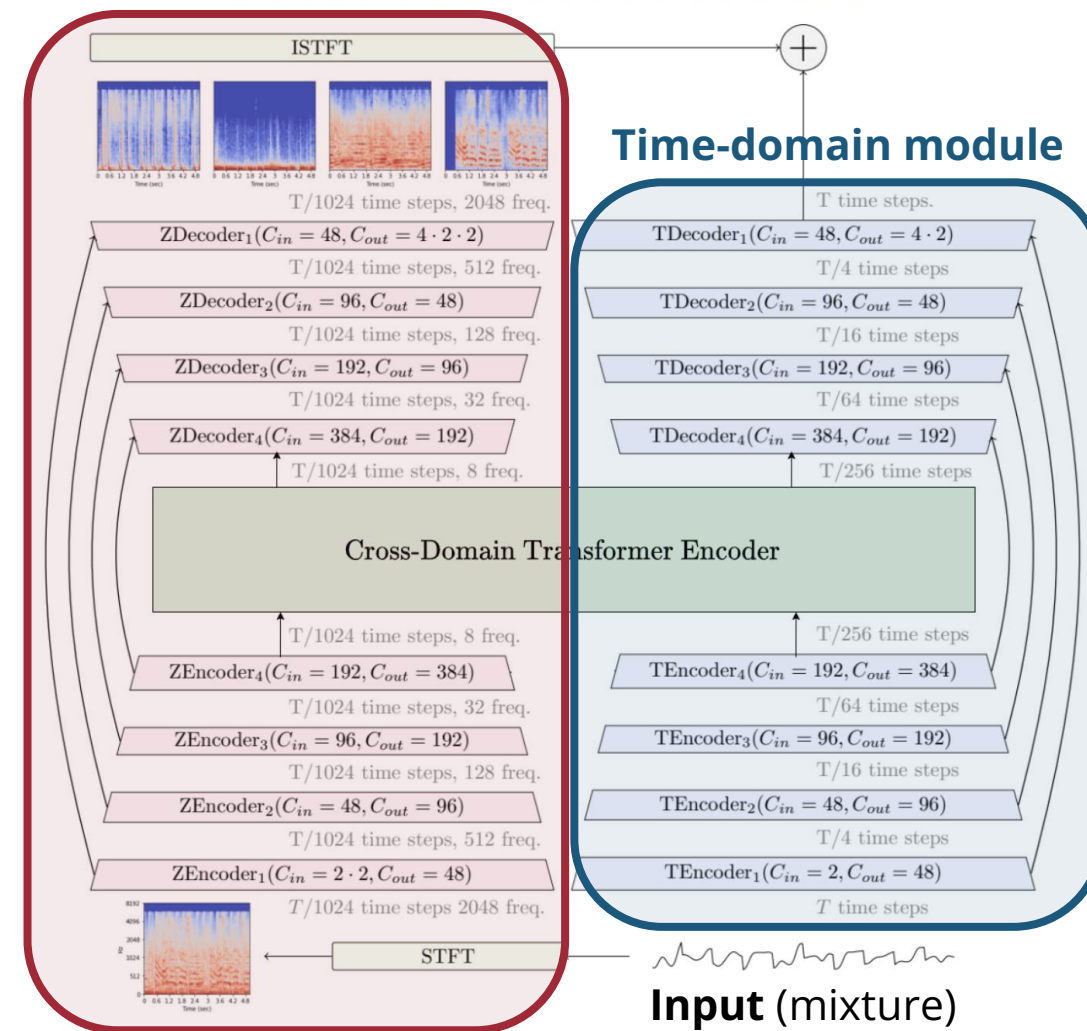
Demucs (Défossez et al., 2019)



(Source: Défossez et al., 2019)

Demucs-Hybrid (Rouard et al., 2023)

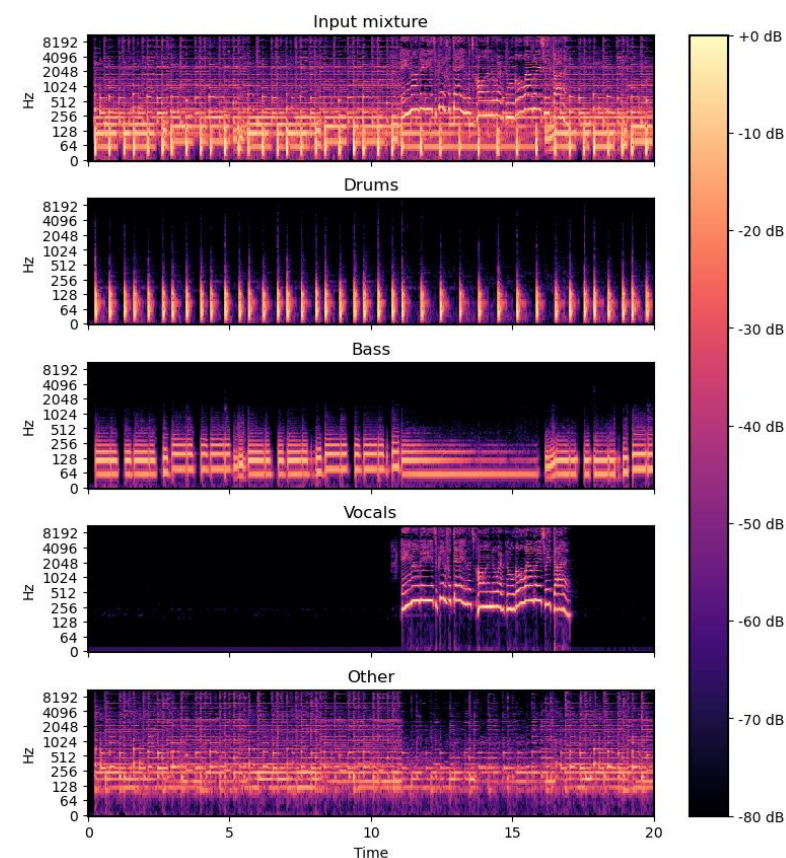
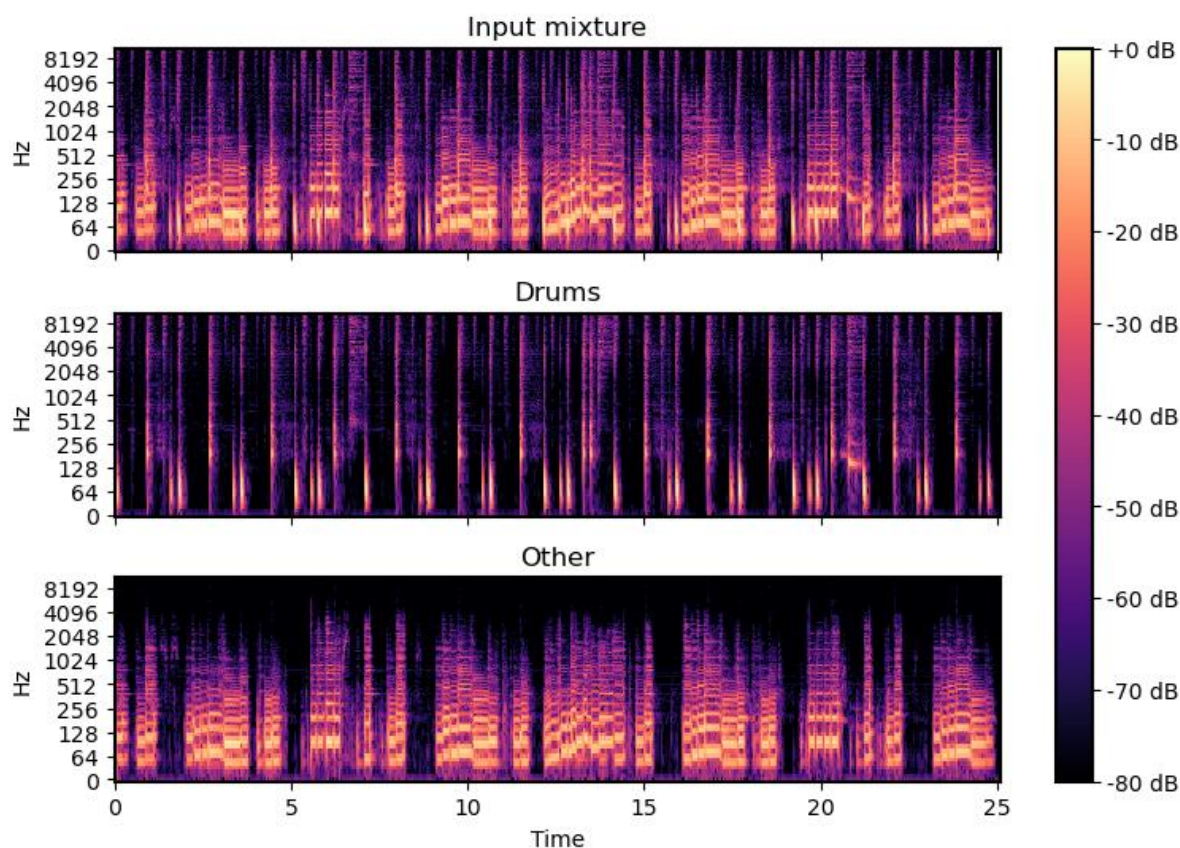
Frequency-domain module



(Source: Rouard et al., 2023)

PA 3: Source Separation

- Part 2: Music Source Separation using Demucs



More Example Systems

Choral Music 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/

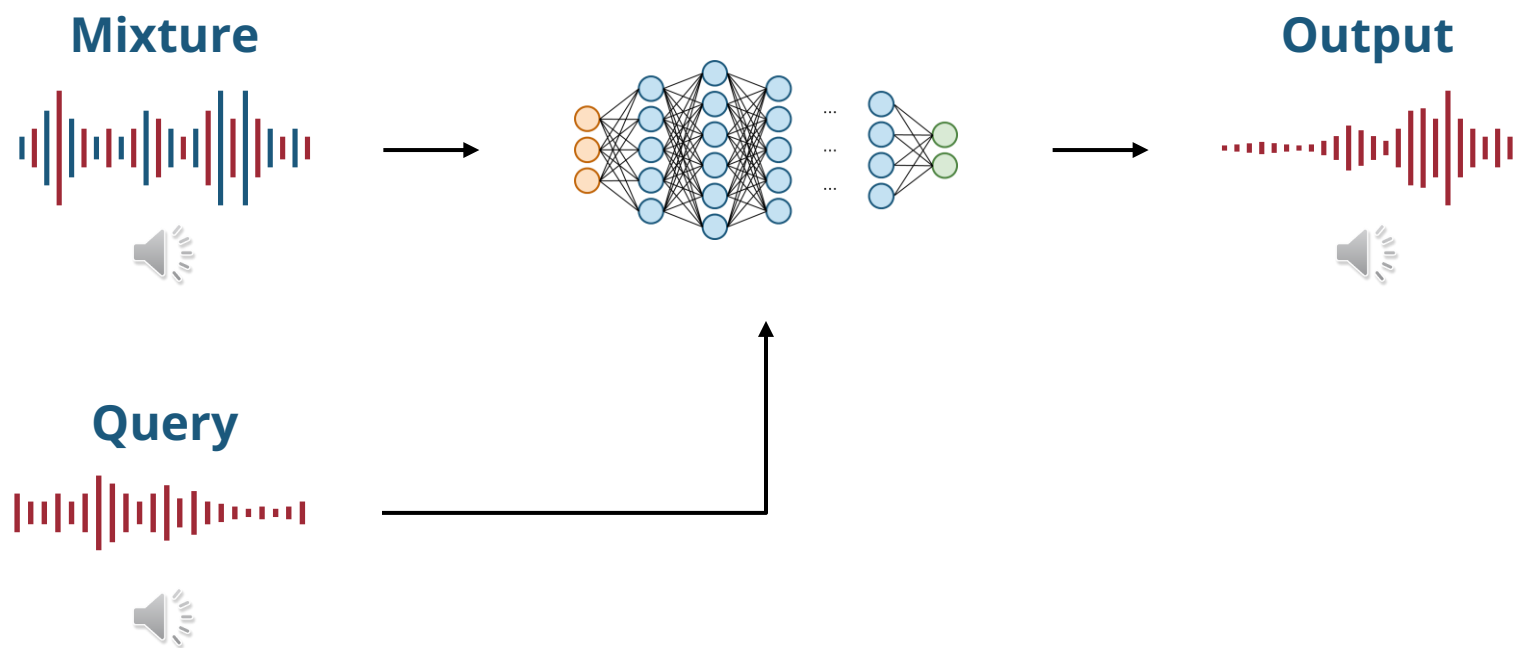
Blob Opera

- This is **NOT** based on source separation
- Sharing this simply **because it's cool!** 😎
- It's based on a **ML-based music harmonization model!**



artsandculture.google.com/experiment/blob-opera/AAHWrq360NcGbw

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



Cinematic Sound Separation

Speech



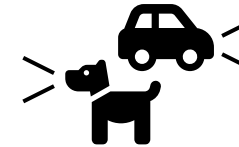
(Source: Wikimedia Commons)

Music



(Source: Wikimedia Commons)

Sound effects



(Source: Wikimedia Commons)

BPJ Media Inc, [CC BY-SA 3.0](#), via Wikimedia Commons.
Vancouver Film School Retouched version by User:Quenhitrn., [CC BY 2.0](#), via Wikimedia Commons.
The Blackbird Academy, [CC BY-SA 2.0](#), via Wikimedia Commons.
One Man Films, "One Shot - WAR ACTION SHORT FILM," *YouTube*, September 11, 2022.

| Cinematic Sound Separation



youtu.be/tY9n2XMF5IQ

Cinematic Sound Separation



youtu.be/1BR4SAKDhMk

🔥 HW 2: Source Separation

- Instructions will be sent by **emails** and released on the **course website**
- **Part 2:** Cinematic Sound Separation



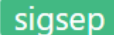
(Source: Moises)

Resources

Datasets

- [MIR-1K](#)
- [MedleyDB](#)
- [iKala](#)
- [DSD100](#)
- [MUSDB18](#) & [MUSDB18-HQ](#)
- [MoisesDB](#)
- Synthetic: [Slakh2100](#), [SynthSOD](#)

Datasets

Dataset	Year	Tracks	Track duration (s)	Full/stereo?
MASS	2008	9	16 ± 7	✗ / ✓
MIR-1K	2010	1,000	(8 ± 8)	✗ / ✗
QUASI	2011	5	(206 ± 21)	✓ / ✓
ccMixer	2014	50	(231 ± 77)	✓ / ✓
MedleyDB	2014	63	(206 ± 121)	✓ / ✓
iKala	2015	206	30	✗ / ✗
 DSD100	2015	100	(251 ± 60)	✓ / ✓
 MUSDB18	2017	150	(236 ± 95)	✓ / ✓
 MUSDB18-HQ	2019	150	(236 ± 95)	✓ / ✓

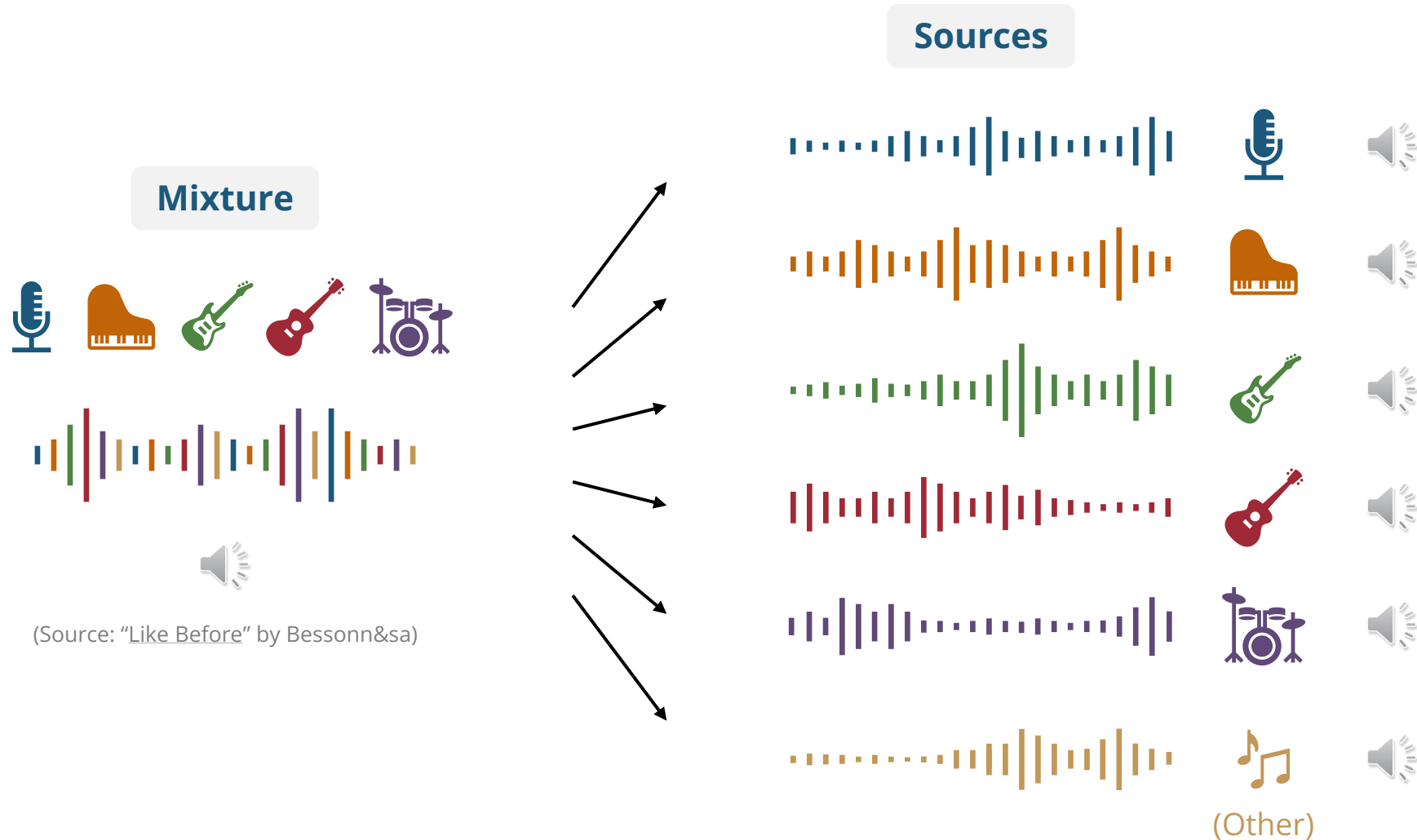
(Source: SigSep)

| Optional Reading

- Ethan Manilow, Prem Seetharman, and Justin Salamon, "[Open Source Tools & Data for Music Source Separation](#)," *Tutorials of ISMIR*, 2020.

Recap

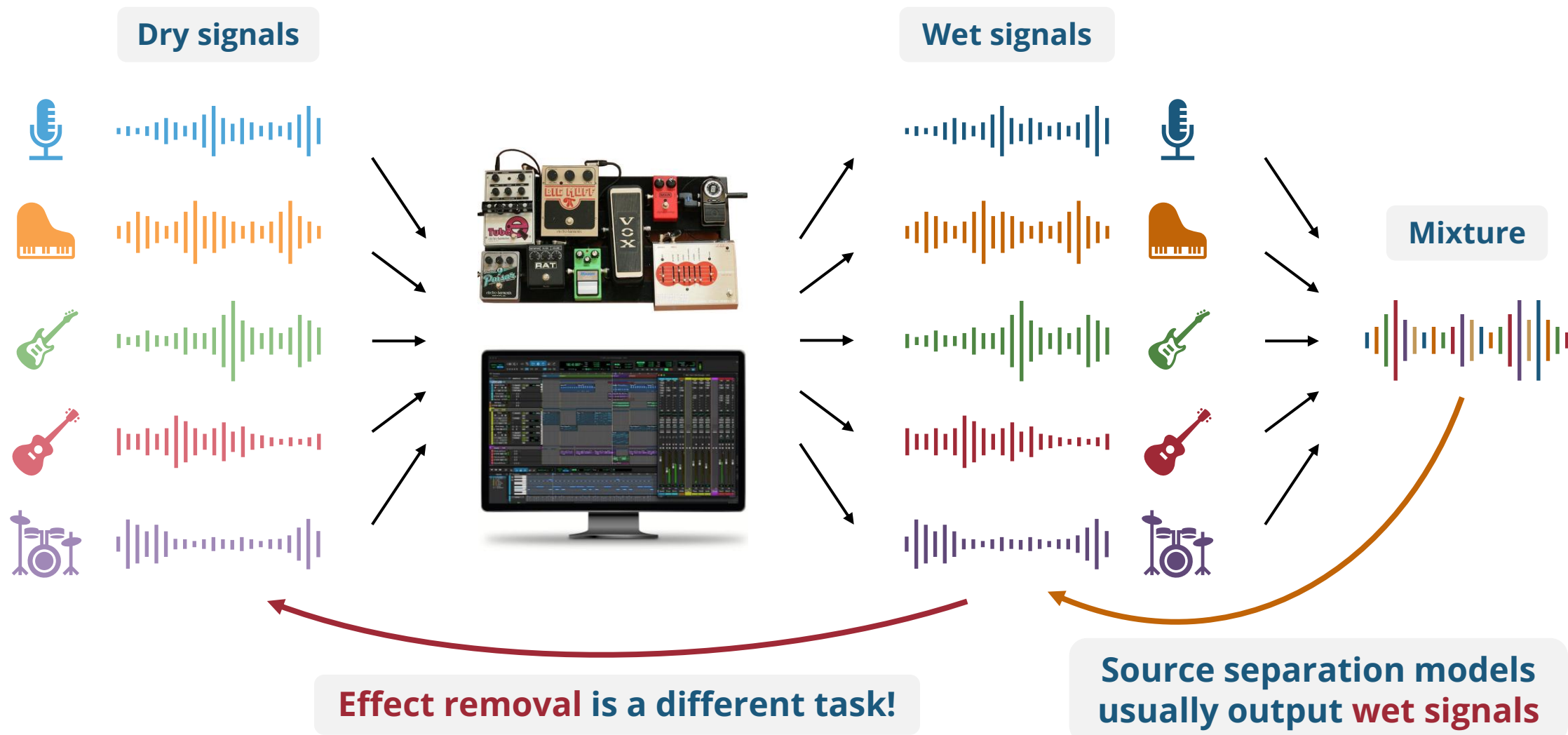
Source Separation



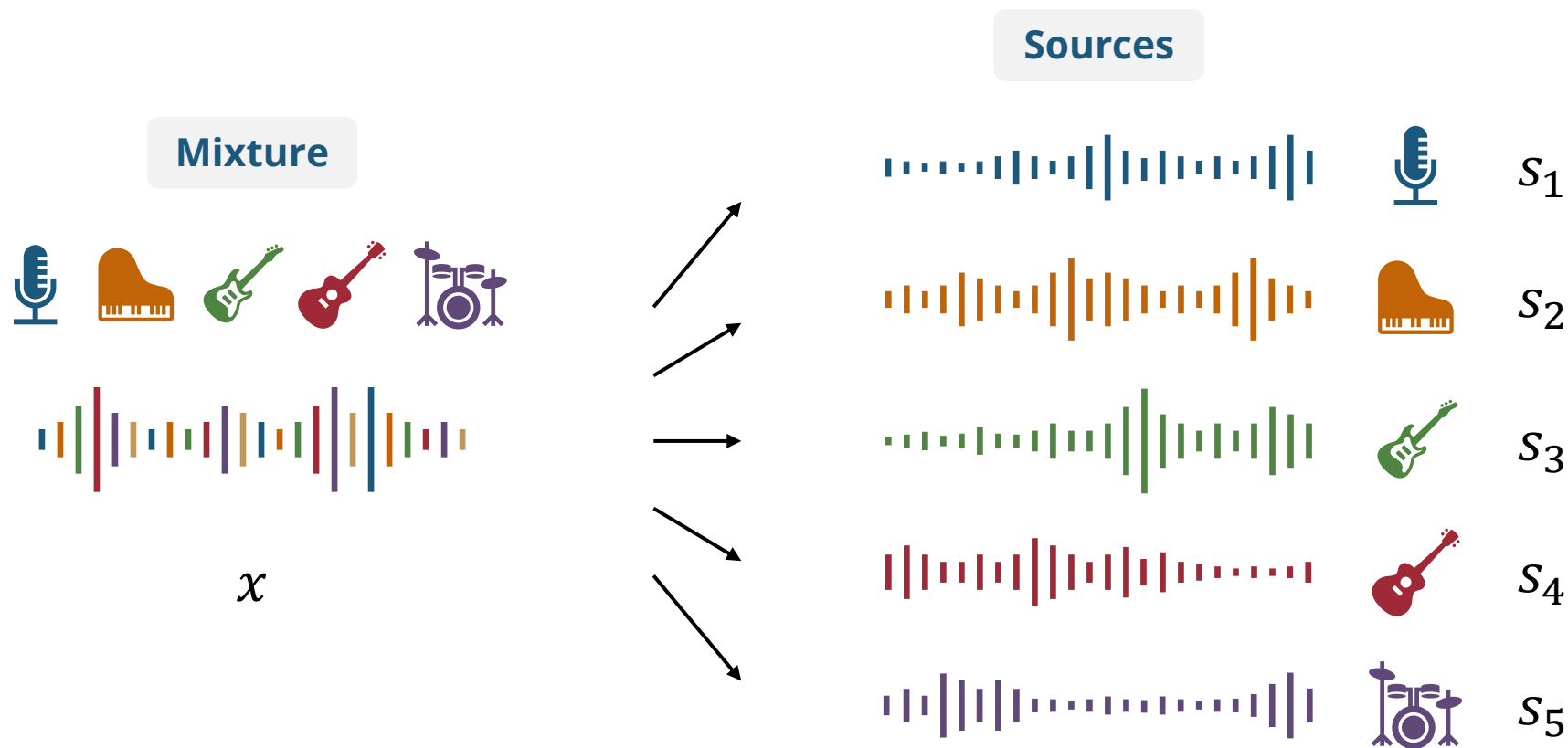
Bessonn&sa, CC BY-NC-SA, via Jamendo

Illustration inspired by Ethan Manilow, Prem Seetharman, and Justin Salamon's Tutorial on "Open Source Tools & Data for Music Source Separation" at ISMIR 2020.

Source Separation does NOT Remove Effects



Mathematical Formulation



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

Source Separation is an Ill-posed Problem

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

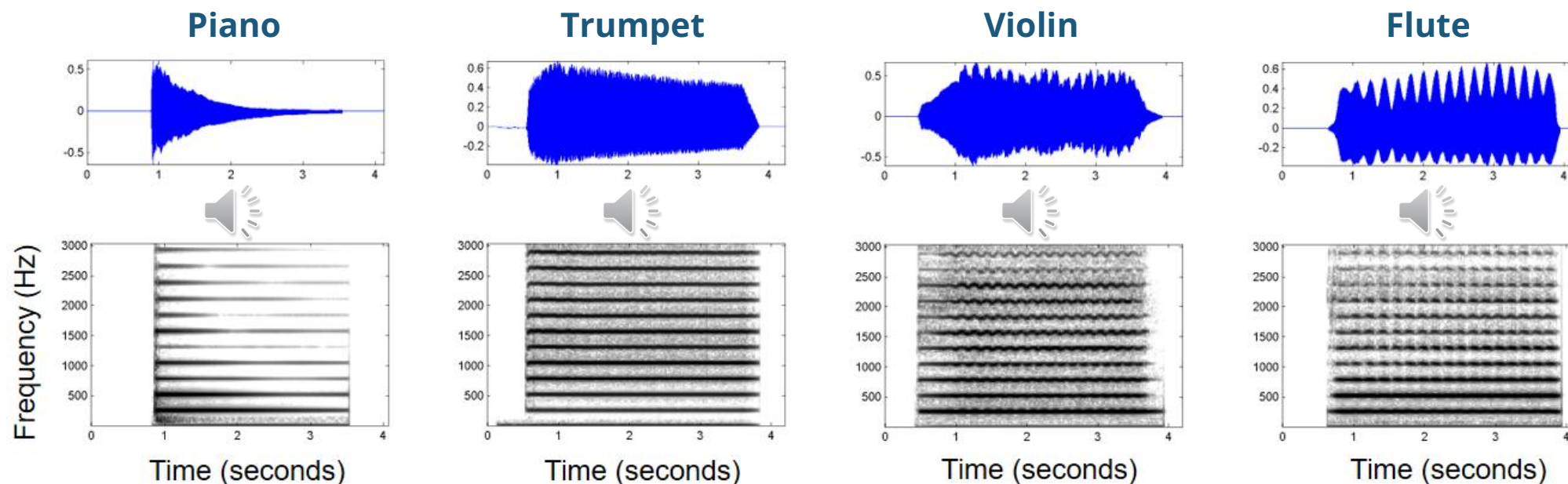
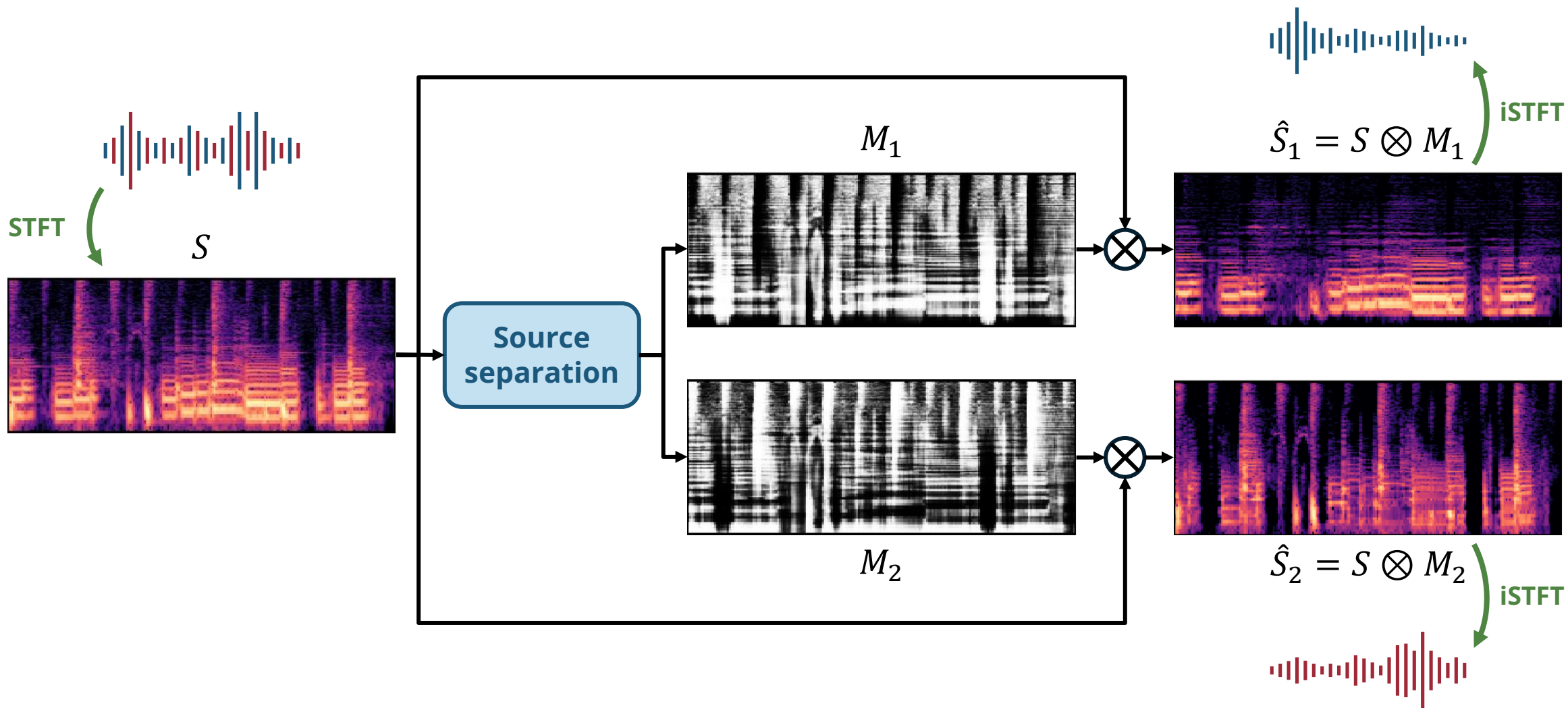


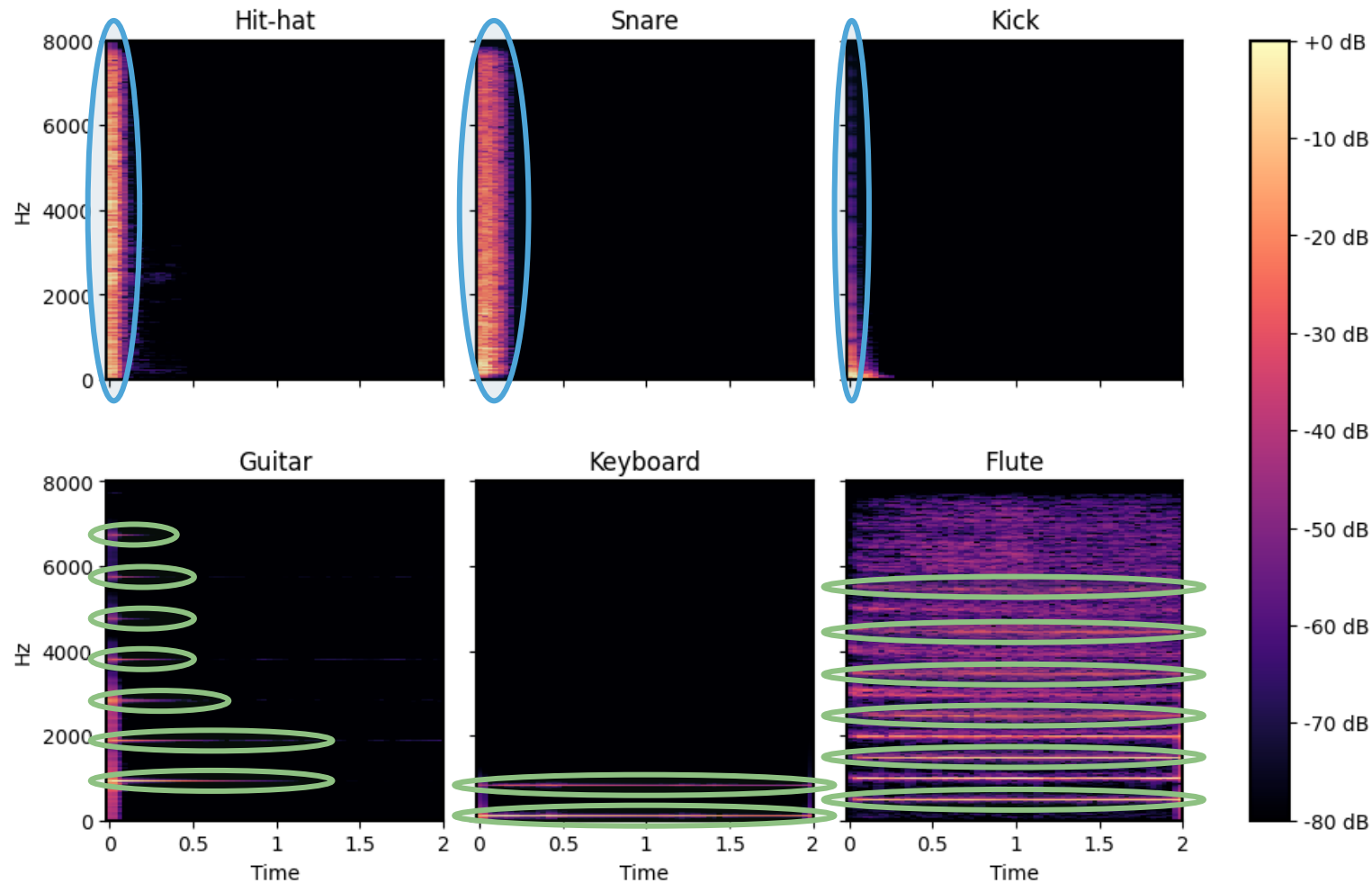
Figure 1.23 from [Müller, FMP, Springer 2015]

(Source: Müller et al., 2021)

Time-Frequency Masking



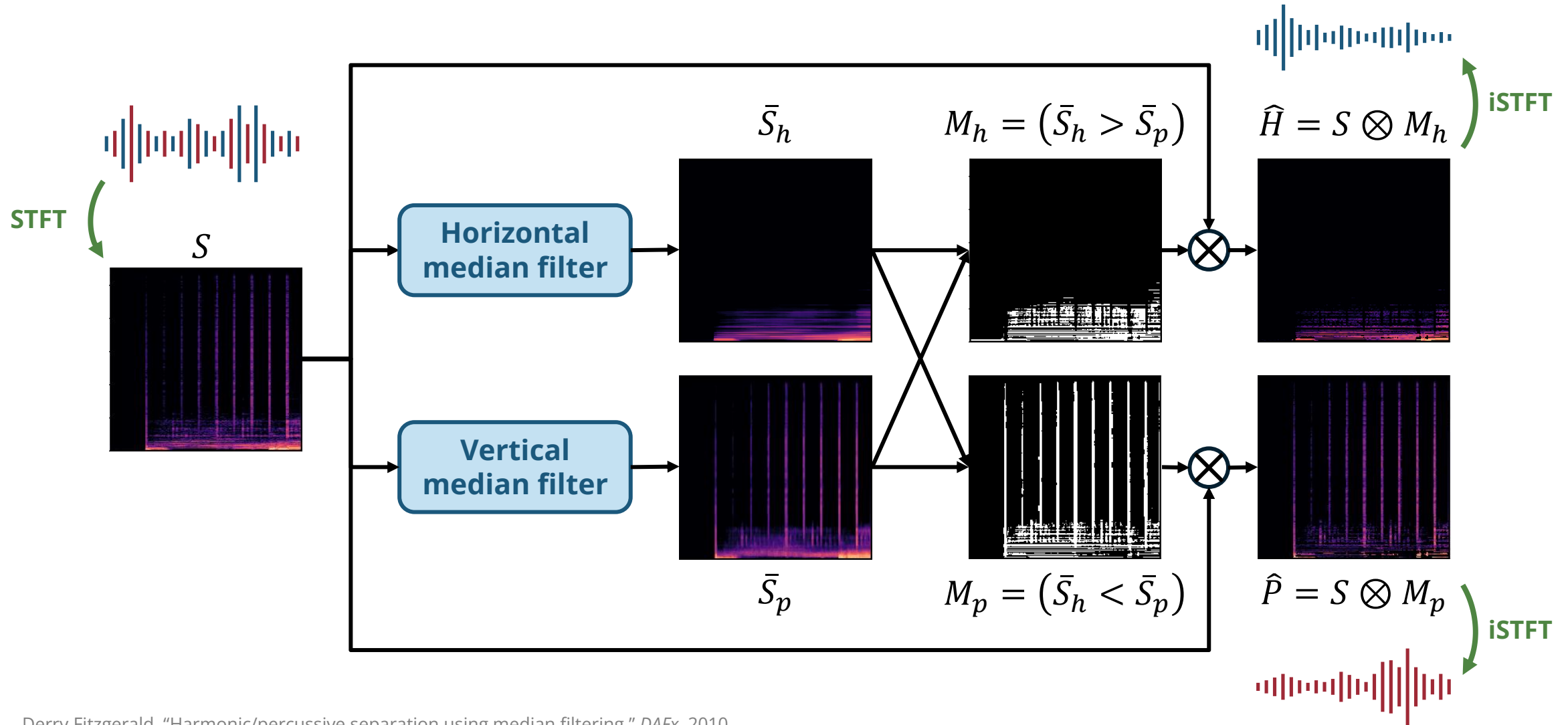
Percussive vs Harmonic Components



Percussive sounds
tend to look **vertical**
in a spectrogram

Harmonic sounds
tend to look **horizontal**
in a spectrogram

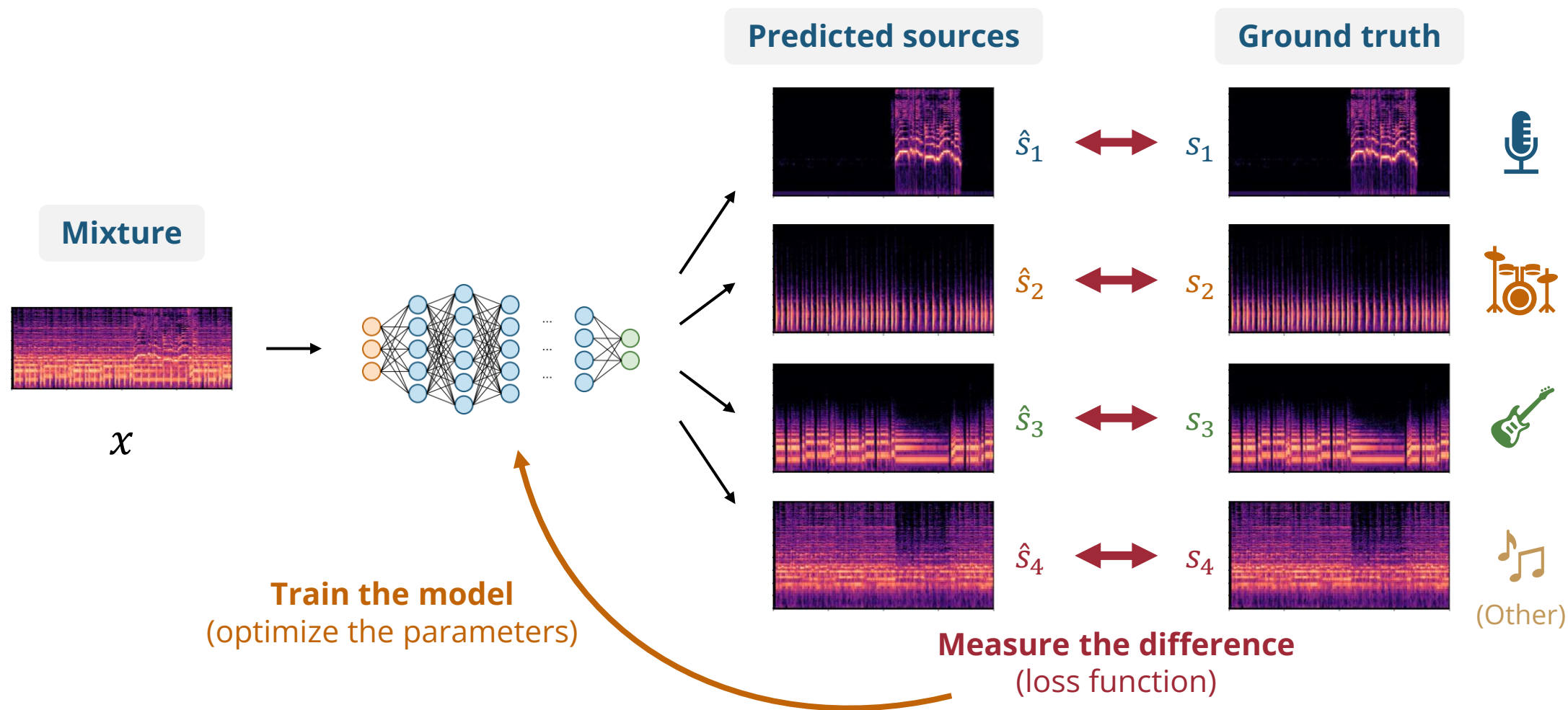
Harmonic-Percussive Separation (Fitzgerald et al., 2010)



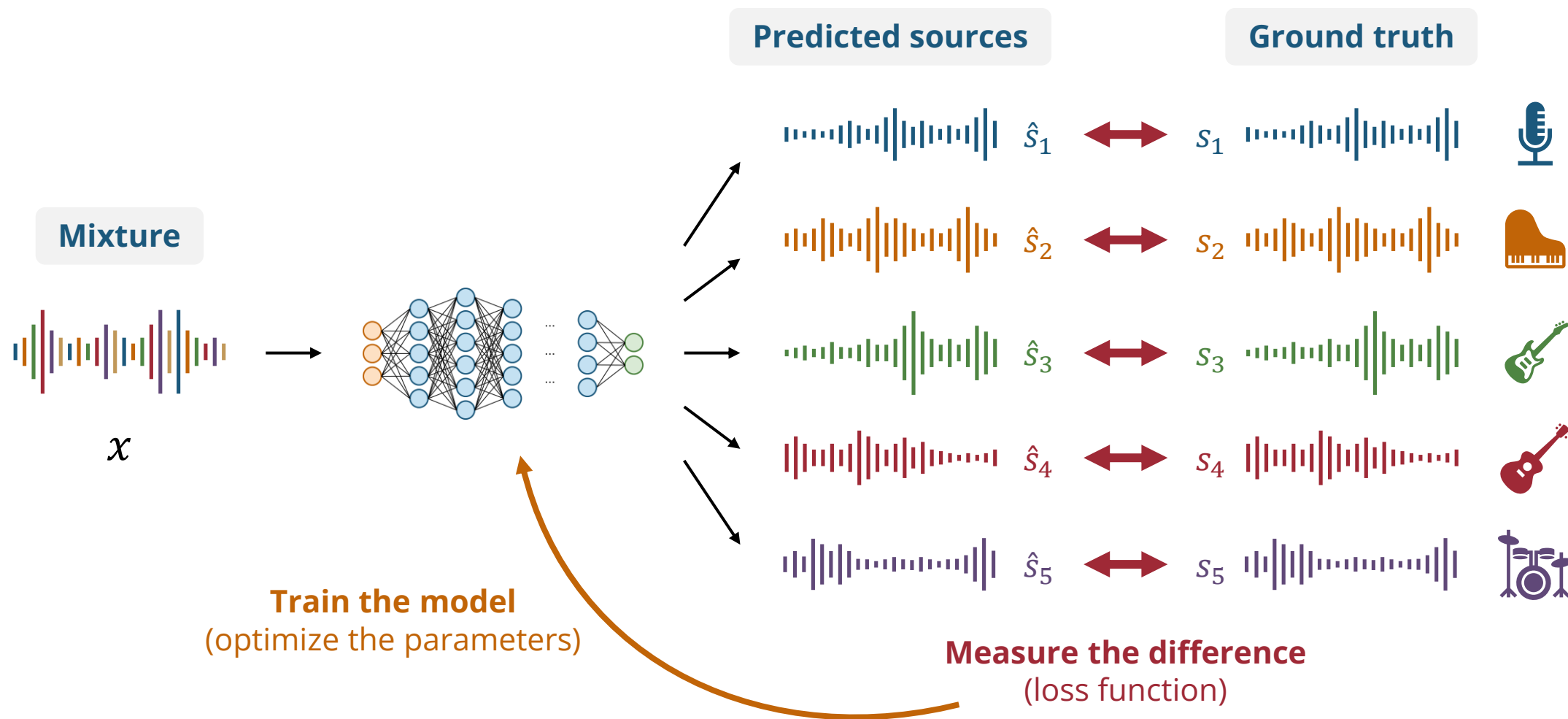
Derry Fitzgerald, "Harmonic/percussive separation using median filtering," *DAFx*, 2010.

Jonathan Driedger, Meinard Müller, Sascha Disch, "Extending Harmonic-Percussive Separation of Audio Signals," *ISMIR*, 2014.

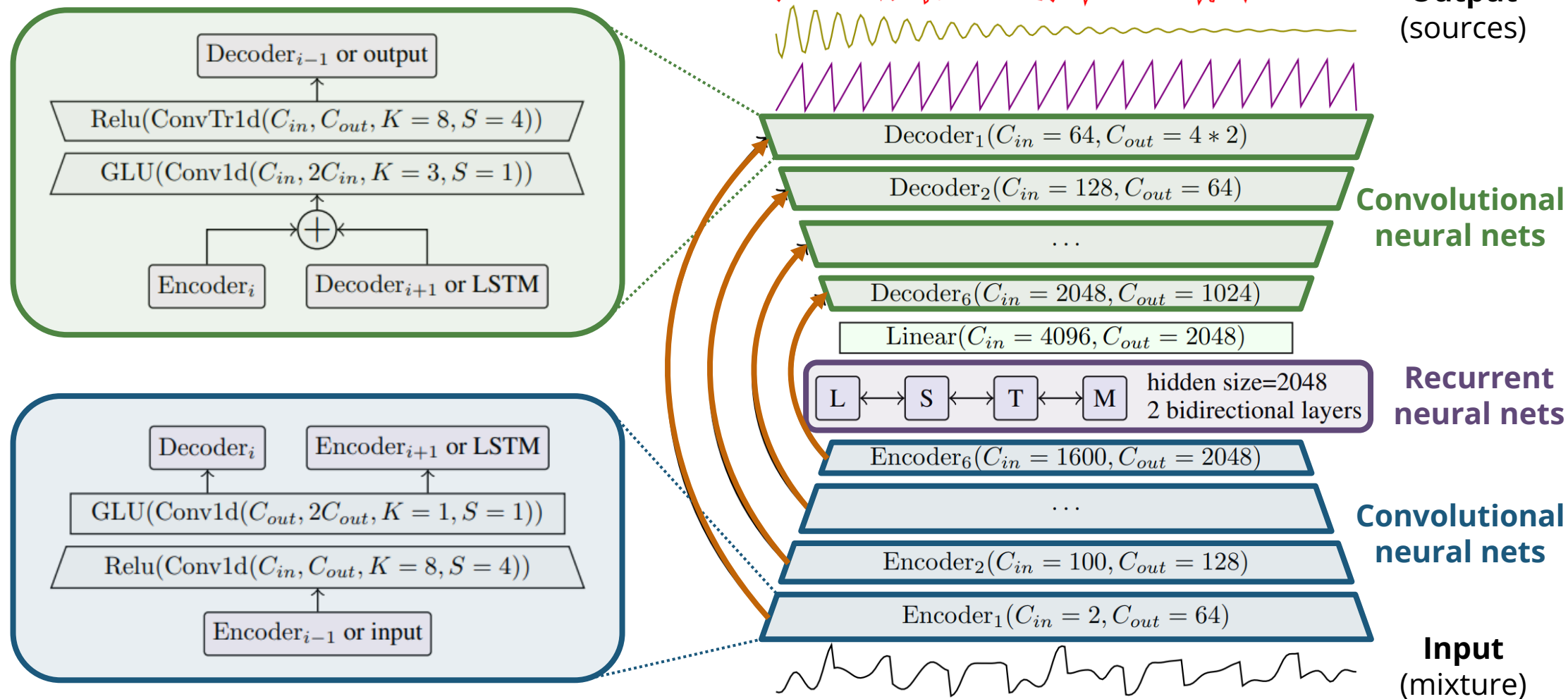
Deep Learning Based Source Separation



Deep Learning Based Source Separation



Demucs (Défossez et al., 2019)



(Source: Défossez et al., 2019)

Cinematic Sound Separation

Speech



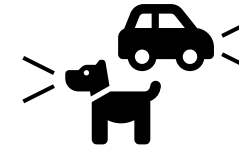
(Source: Wikimedia Commons)

Music



(Source: Wikimedia Commons)

Sound effects



(Source: Wikimedia Commons)

BPJ Media Inc, [CC BY-SA 3.0](#), via Wikimedia Commons.
Vancouver Film School Retouched version by User:Quenhitrn., [CC BY 2.0](#), via Wikimedia Commons.
The Blackbird Academy, [CC BY-SA 2.0](#), via Wikimedia Commons.
One Man Films, "[One Shot - WAR ACTION SHORT FILM](#)," *YouTube*, September 11, 2022.

Next Lecture

Convolutional Neural Networks

