[DAFx 2022] Tutorial

Musical Instrument Sound Analysis, Transformation, and Resynthesis in the Sound Morphing Toolbox

Marcelo Caetano

[DAFx 2022] Tutorial

- The [MORPH] Project
- Motivation
  - Timbre Spaces & Musical Instruments
  - Musical Instrument Timbre & Sound Morphing
- Sinusoidal Modeling
- The Source-Filter Model
- The Sound Morphing Toolbox

Overview of the [MORPH] Project

From Timbre Perception to the Creative Exploration of Musical Instrument Sound Morphing

Overview of the [MORPH] Project

- H2020 MSCA-IF-GF 2018
- Horizon 2020
- Marie-Skłodowska Curie Actions
- Individual Fellowships
- Global Fellowship

Overview of the [MORPH] Project

- Researcher: Dr. Marcelo Caetano
- Investigation: use sound morphing to study musical instrument timbre perception
- Outgoing phase: CIRMMT-McGill (Montreal, Canada)
  - Prof. Philippe Depalle (1st year)
  - Prof. Stephen McAdams (2nd year)
- Returning phase: CNRS-PRISM (Marseille, France)
  - Dr. Richard Kronland-Martinet (3rd year)

Overview of the [MORPH] Project

- CIRMMT-McGill
  - 1st year (Oct 2020-Sep 2021): modeling and transformation
  - 2nd year (Oct 2021-Sep 2022): musical instrument sound morphing and timbre perception
- CNRS-PRISM
  - 3rd year (Oct 2022-Sep 2023): creative exploration of musical instrument sound morphing

[MORPH] Proposed Research

- Use sound morphing to investigate timbre perception
- Focus on musical instruments
- Create continuous timbre spaces
- Break categorical perception of musical instrument timbre

Timbre Spaces & Musical Instruments

Timbre Perception of Musical Instrument Sounds

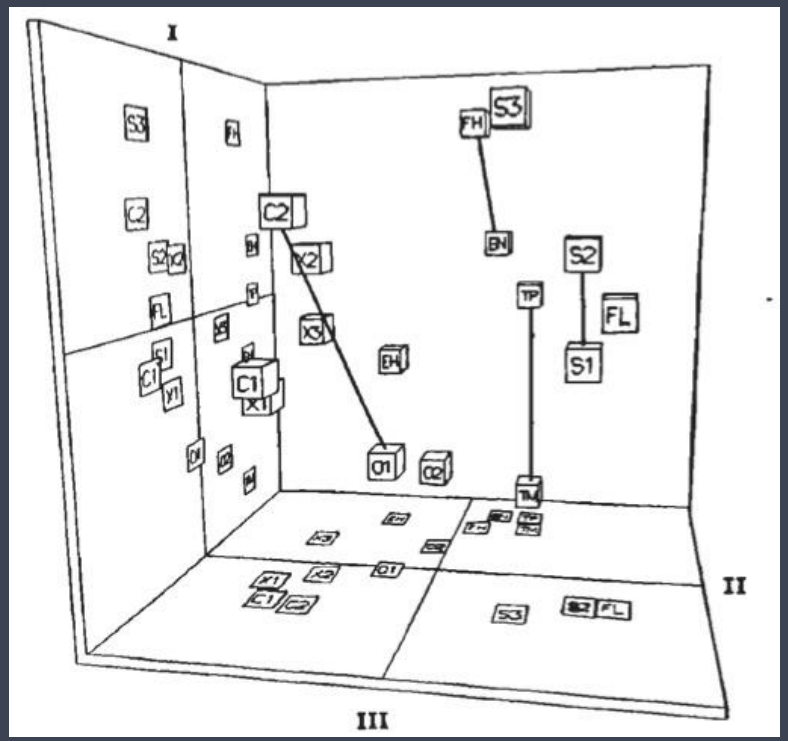
Musical Instrument Timbre [1]

- The ways in which sounds are perceived to differ
- Perceptual difference not accounted for by *pitch*, *loudness*, *spatial position*, *duration*, and environmental characteristics
- **Categorical view:** sound source recognition, identification, tracking in time
- **Sensory view:** multidimensional set of attributes associated with timbre spaces

[1] S McAdams & M Goodchild (2017). Musical Structure: Sound and Timbre. *The Routledge Companion to Music Cognition*, 129–139, New York, NY.

Musical Instrument Timbre

MDS Timbre Space [2]



Features of Timbre Spaces

- Timbre spaces are sparse
- Each musical instrument is represented by one note
- Each musical instrument sound occupies a point in space
- Void space between musical instruments
- Limitations of representation:
  - Dimensions are highly dependent on stimuli
  - No information about temporal variations

[MORPH] Research Objectives

- **Central question:** can we fill the gaps between musical instruments and create a *sonic continuum* [6]?
- Use **sound morphing** to investigate the relationship between the *sensory* and the *categorical* facets of timbre perception
- **Sound morphing** alters both the *sensory* and the *categorical* perception of sounds

[6] T Wishart (1996). On Sonic Art. Simon Emmerson Ed., Routledge, New York, NY.

Sound Morphing & Timbre

- **Sound morphing** is a transformation that gradually blurs the *categorical* distinction between sounds by blending their *sensory* attributes
- **Sound morphing** can potentially bridge the gaps between acoustic musical instruments in timbre spaces
- **Sound morphing** can potentially break the categorical perception of musical instrument timbre and create hybrid musical instruments
- **Sound morphing** potentially allows musicians and composers to explore continuous timbre spaces

Musical Instrument Timbre & Sound Morphing

Timbre Perception of Morphed Musical Instrument Sounds

What is Sound Morphing?

α=0 α=0.25 α=0.5 α=0.75 α=1

Musical Instrument Sound Morphing

- Musical instrument sound perception: categorical or continuous [7]?
- What sound features should be manipulated [8]?
- Is there a “musical instrument continuum”?

[7] S Carral (2011) Determining the Just Noticeable Difference in Timbre Through Spectral Morphing: A Trombone Example. Acta Acoust. United Ac., 97:466–476.

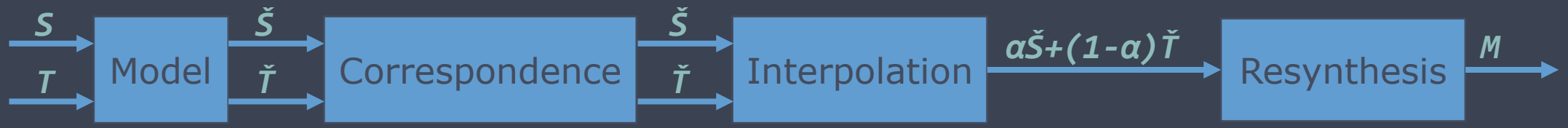
[8] K Siedenburg, K J-Mollerup, S McAdams (2015) Acoustic and Categorical Dissimilarity of Musical Timbre: Evidence from Asymmetries Between Acoustic and Chimeric Sounds. Front Psychol, 6.

What is (not) Sound Morphing?

- Image morphing is an imperfect analogy
- Sounds have a temporal dimension
- Morphed sound should be perceptually intermediate
- Result should fuse into a single percept 
  - Usually not enough to mix or crossfade
- Morph is a single entity with intermediate features 
  - Cross-synthesis applies the spectral envelope of one sound onto the other

Morphing and the Interpolation Principle

- Model source S and target T as $\check{S}$ and $\check{T}$
- Establish correspondence between elements
- Interpolate model parameters: $M = \alpha\check{S} + (1-\alpha)\check{T}$
- Recreate morph M



Morphing Sounds

- Multiple possible transformations along the temporal dimension
  - Dynamic morph 
  - Warped morph 
  - Static morph 
  - Cyclostationary morph [9] 

[9] M Caetano & X Rodet (2013) Musical Instrument Sound Morphing Guided by Perceptually Motivated Features. IEEE Trans Audio, Speech, Lang Proc, 21(8), 1666-1675.

[MORPH]

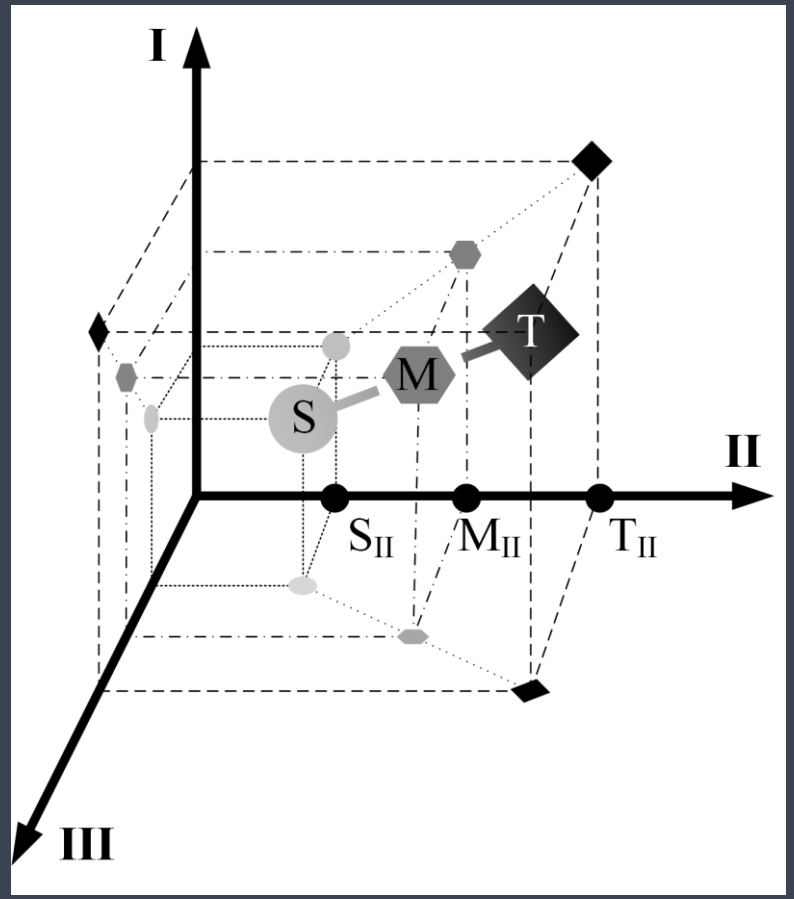
From Timbre Perception to the Creative Exploration of Musical Instrument Sound Morphing

Morphing Musical Instrument Sounds

- In the [MORPH] project
  - Focus on musical instrument sounds
  - Focus on timbre
- What kind of morph?
  - Morphed musical instrument sounds that are perceptually intermediate
  - Morph spectral, temporal, and spectro-temporal features
  - Static morph

Morphing & the Sonic Continuum

Sound Morphing



Sonic Continuum

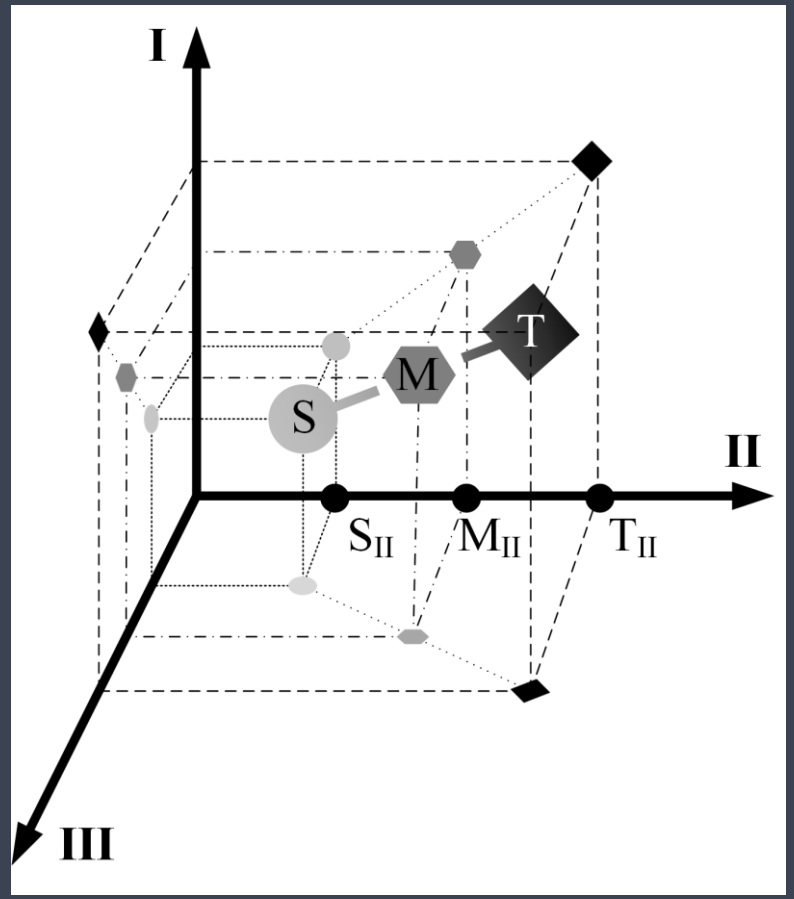
- Metric timbre space
- Model and transformation with perceptual impact
- M is perceptually intermediate between S and T
- Continuous timbre space

Research objectives

- Modeling and transformation of musical instrument sounds
  - High quality model
  - Perceptually natural transformations
- Musical instrument sound morphing and timbre perception
  - Continuous timbre spaces
  - Categorization of musical instrument timbre
- Creative exploration of musical instrument sound morphing
  - Interface for interactive morphing
  - Real-time morphing transformations

Morphing & the Sonic Continuum

Sound Morphing



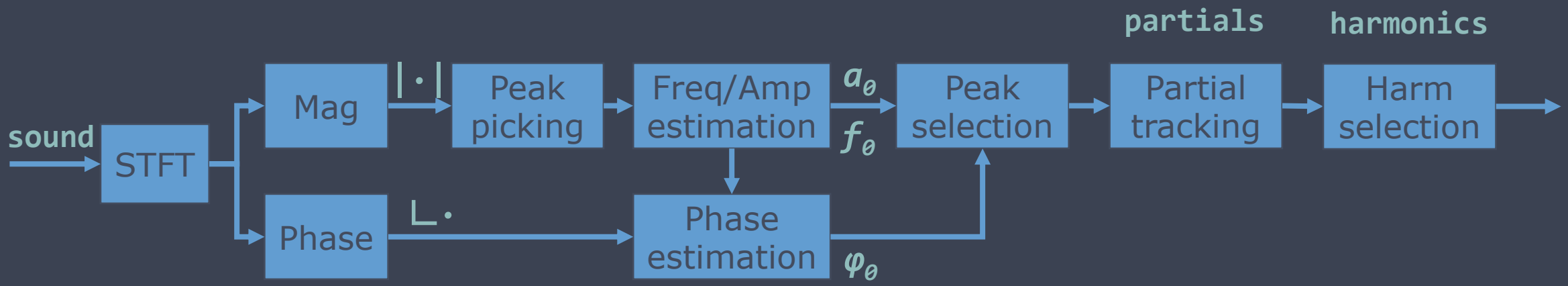
Timbre Perception

- Smooth transitions between instruments
- Will $\alpha = 0.5$ break categorical perception of musical instrument timbre?
- Will sound morphing produce continuous timbre spaces?
- Will sound morphing generate hybrid musical instruments?

Sinusoidal Modeling

Modeling the oscillatory modes of musical instrument sounds

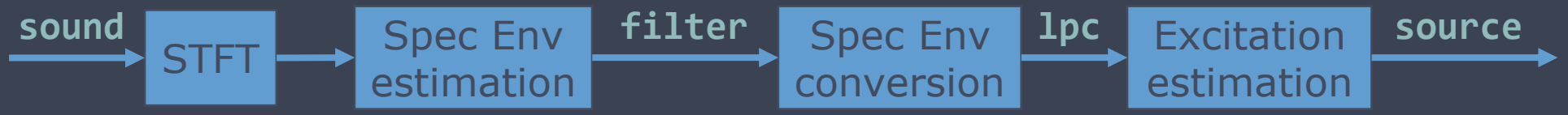
Sinusoidal Model Algorithmic Steps



Source-Filter Model

Modeling the resonating properties of musical instrument sounds

Source-Filter Model Algorithmic Steps



The Sound Morphing Toolbox (SMT)

Morphing musical instrument sounds

What's in the Sound Morphing Toolbox

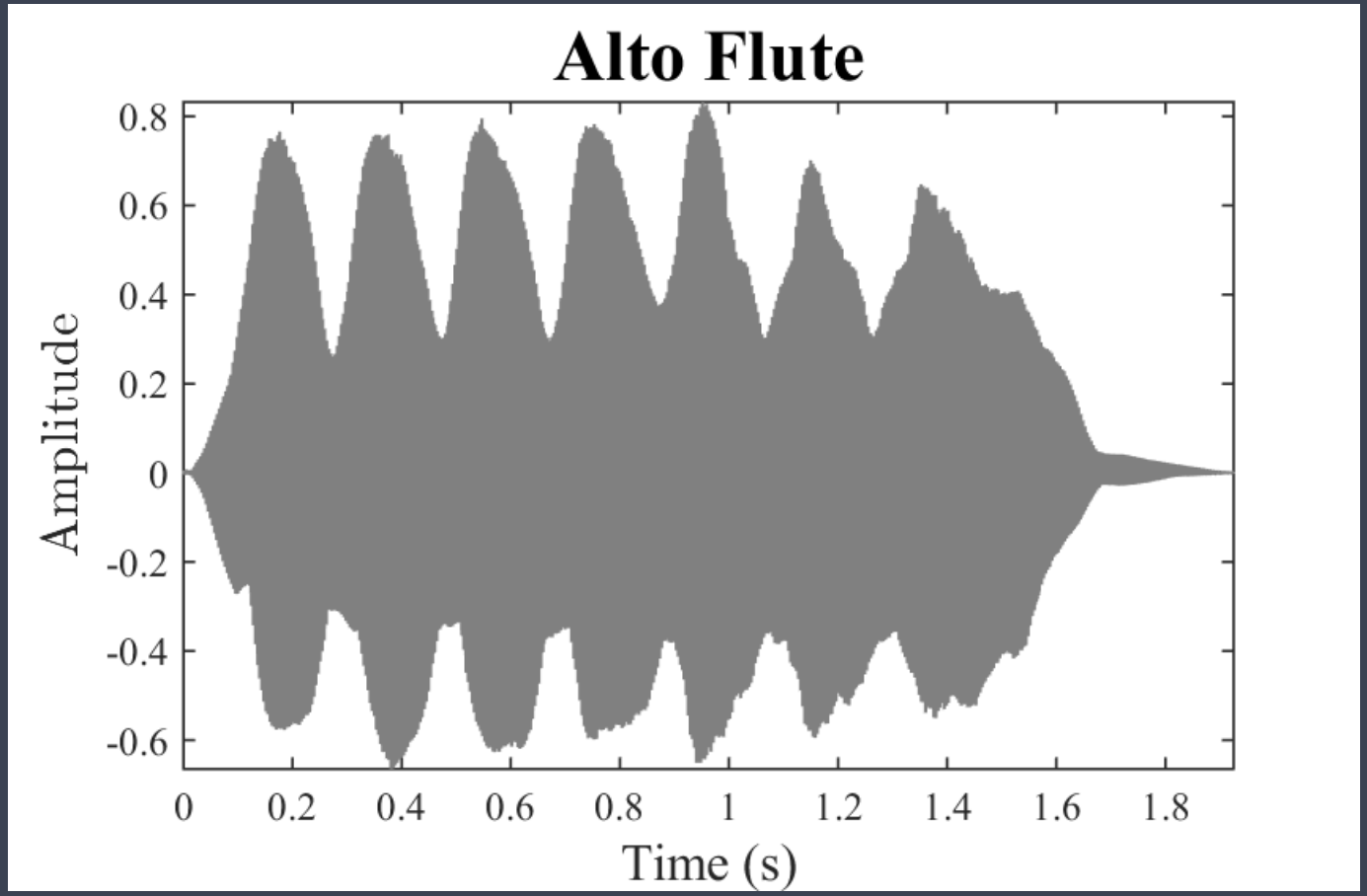
- Time Scaling Algorithm
  - Synchronized Overlap-Add: Fixed Synthesis (SOLA-FS)
- Musical Instrument Sound Model
  - Sinusoidal Model (SM)
  - Source-Filter Model (SFM)
- Sound Morphing Algorithm
  - Sinusoidal morphing (SM)

https://github.com/marcelo-caetano/sound-morphing/

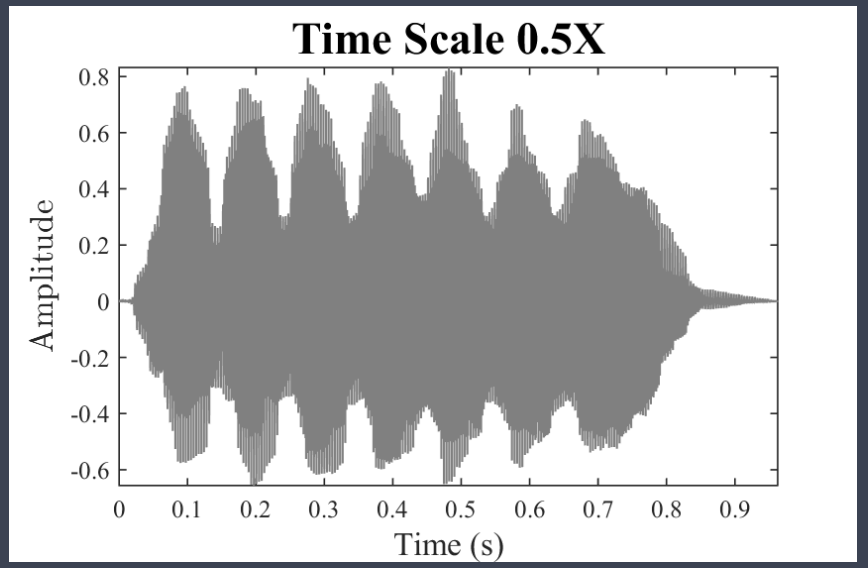
Morphing Musical Instrument Sounds with the SMT

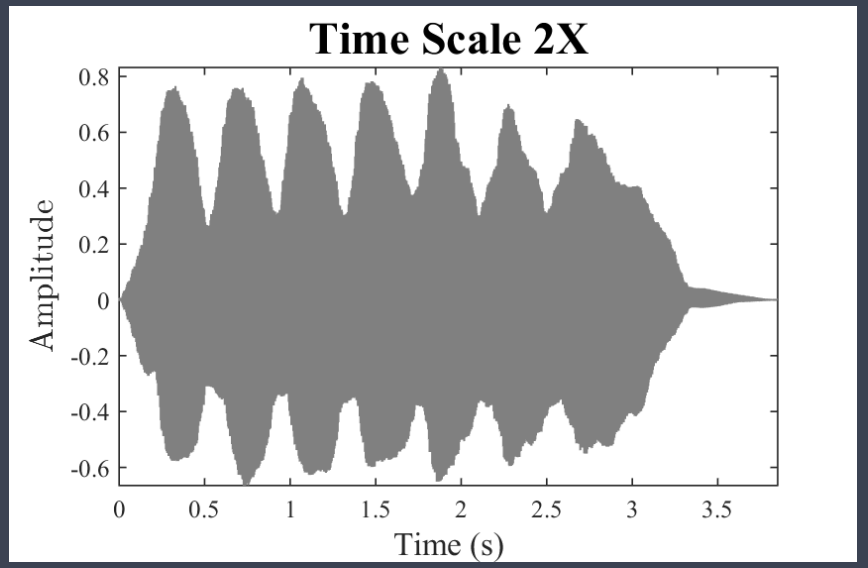
- Time scale *S* and *T* to the same duration
- Model both sounds with sinusoids
- Connect corresponding partials
- Interpolate the amplitudes and frequencies of the partials
- Resynthesize *M* from the interpolated partials

Time Scaling with SOLA-FS



Extreme time scales with SOLA-FS

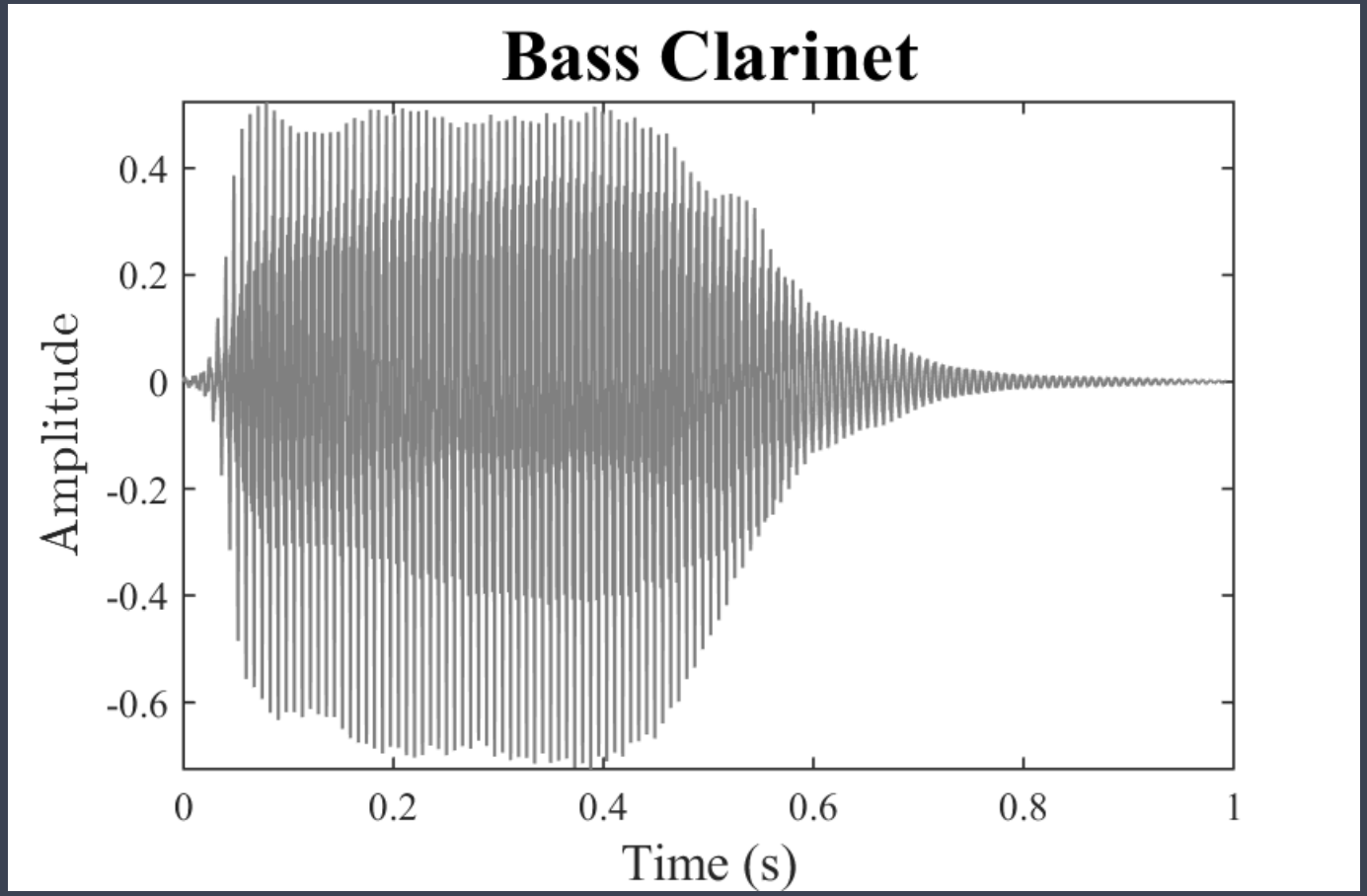




Sinusoidal Modeling in the SMT

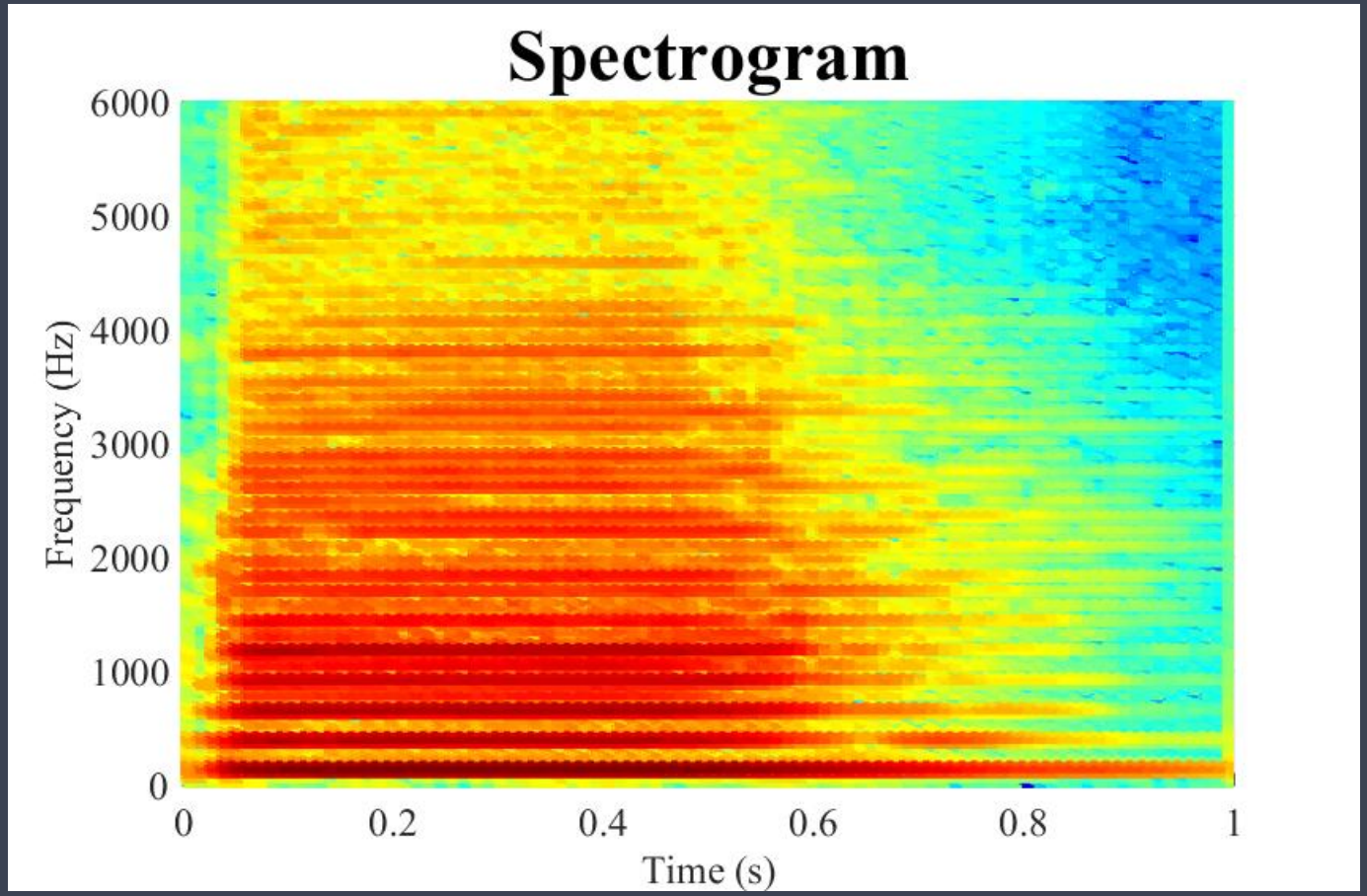
- Represent oscillatory modes of musical instruments with time-varying sinusoids
- Oscillatory modes of musical instrument sounds are also called partials
- Each partial appears as a peak in the Fourier spectrum
- Model each peak as a sinusoid
- Connect peaks across frames to get time-varying sinusoids
- Powerful representation of musical instrument sounds

Waveform

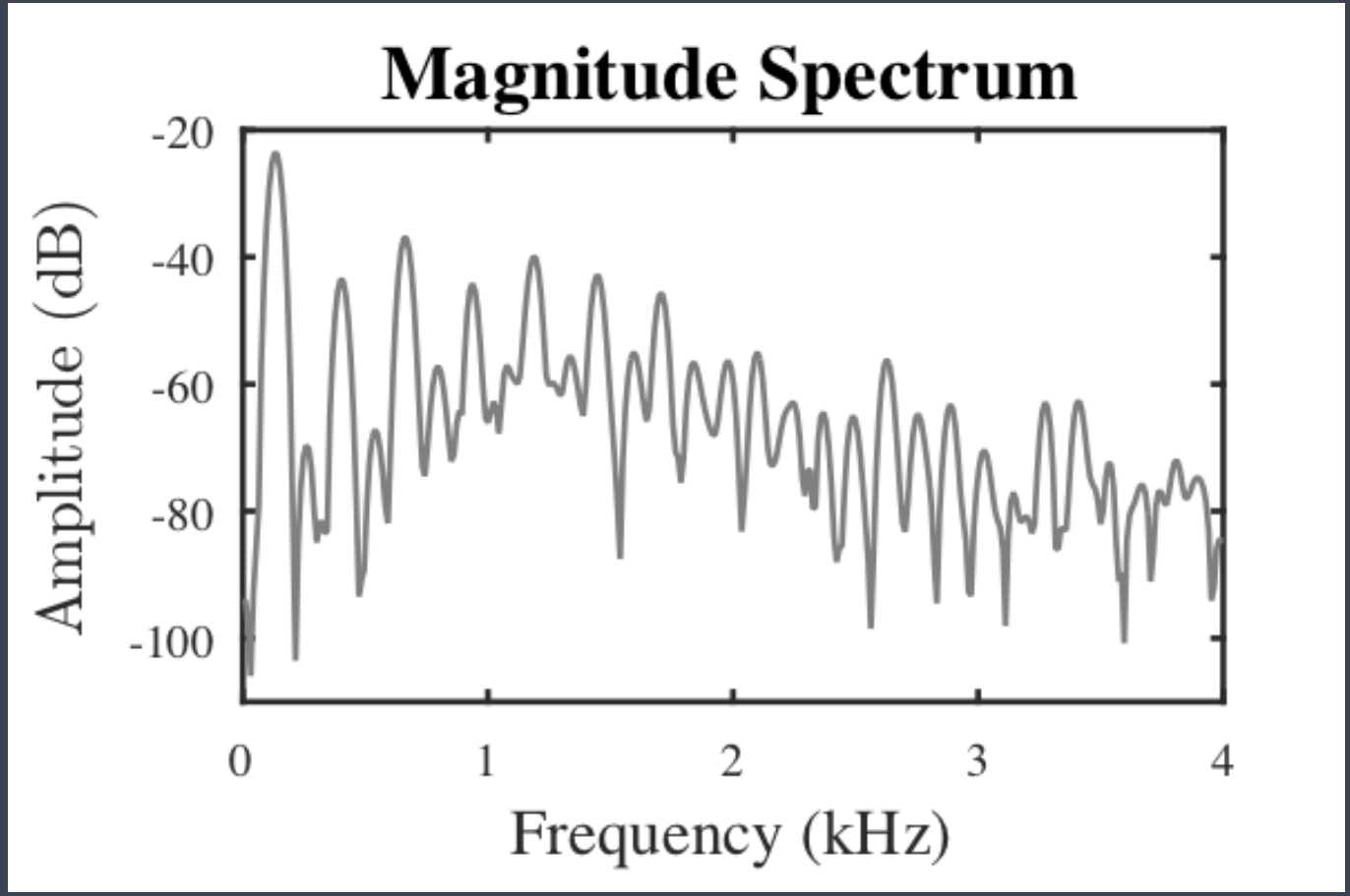


original

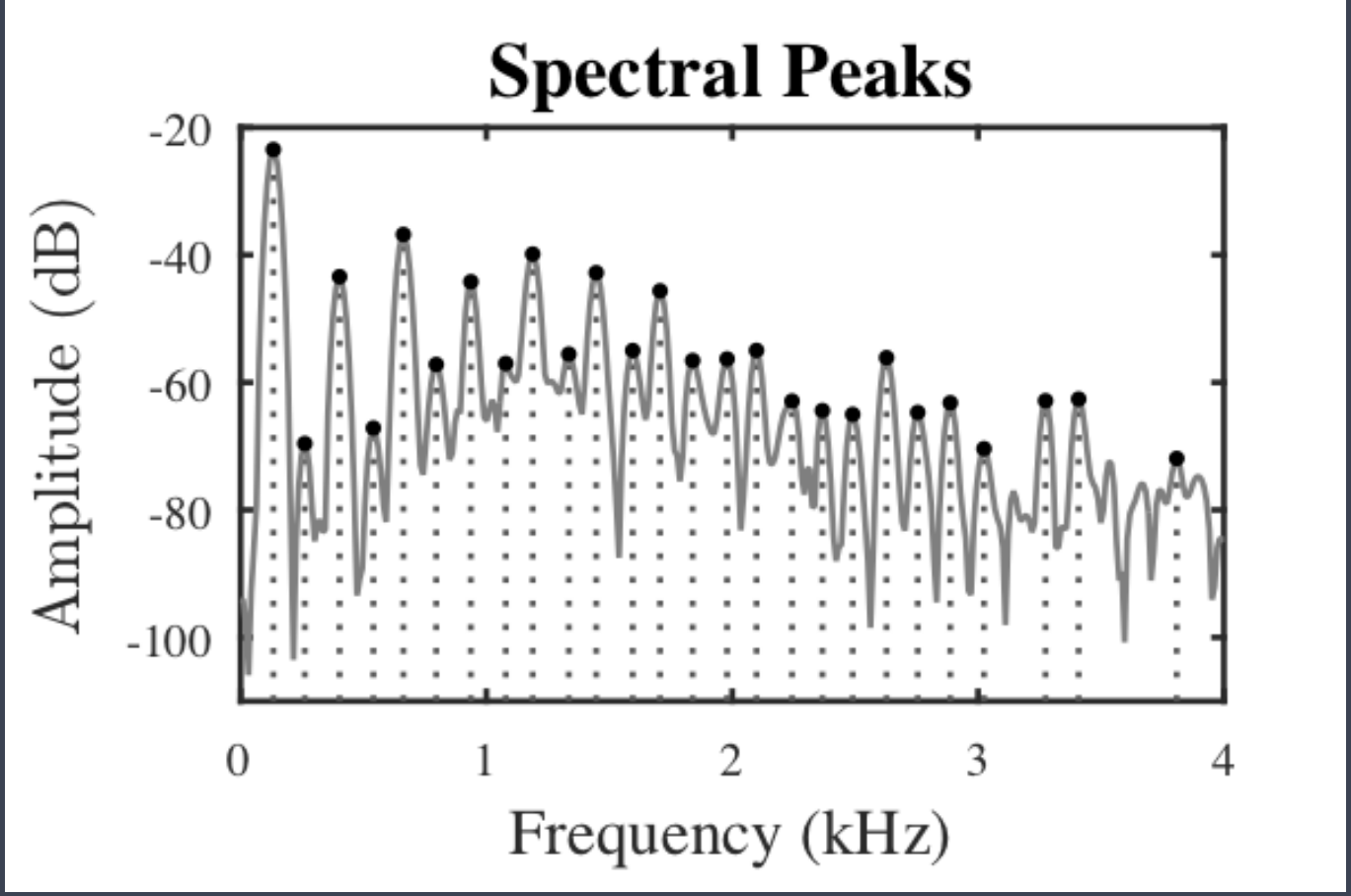
Spectrogram



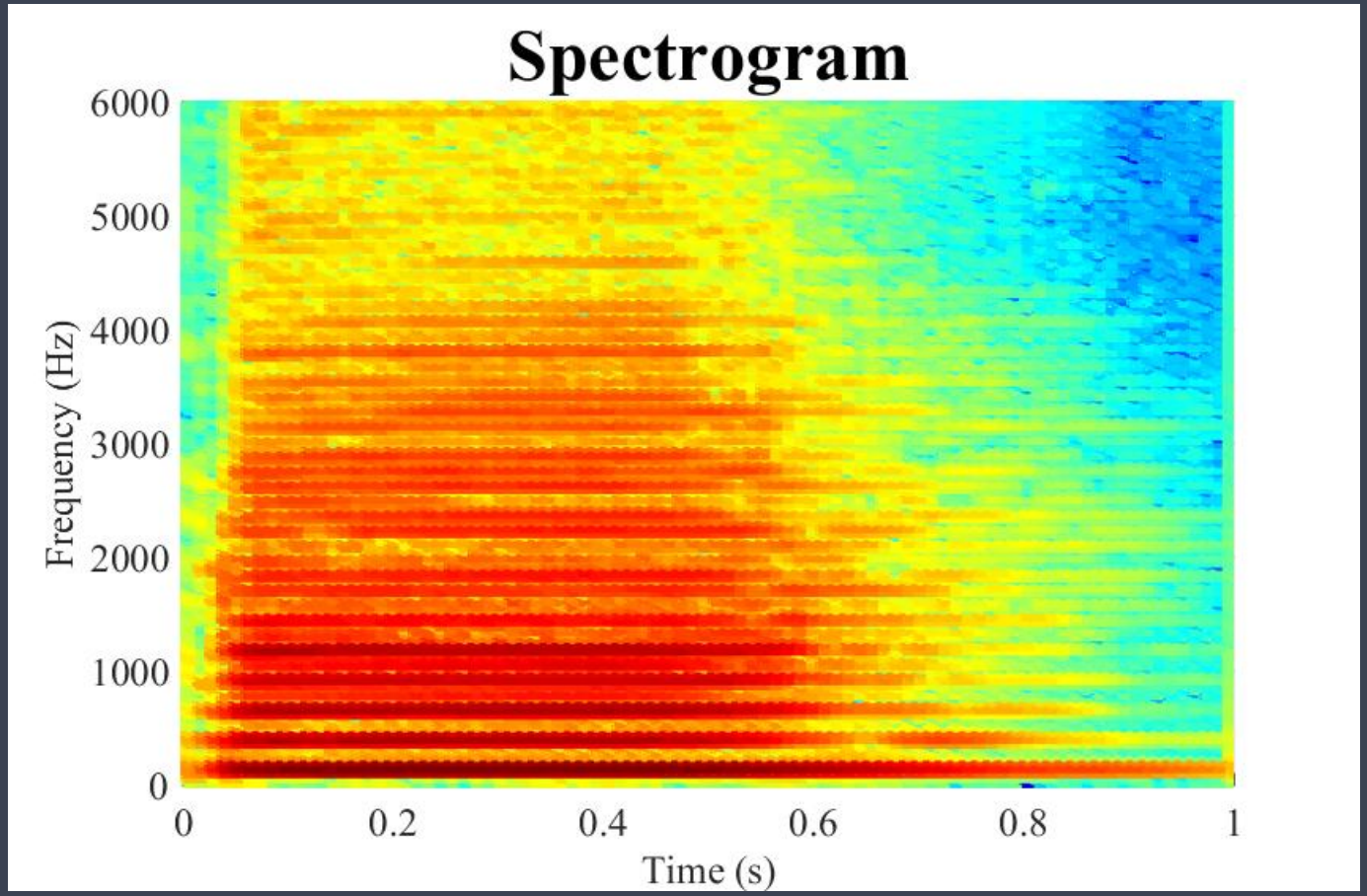
Fourier Spectrum



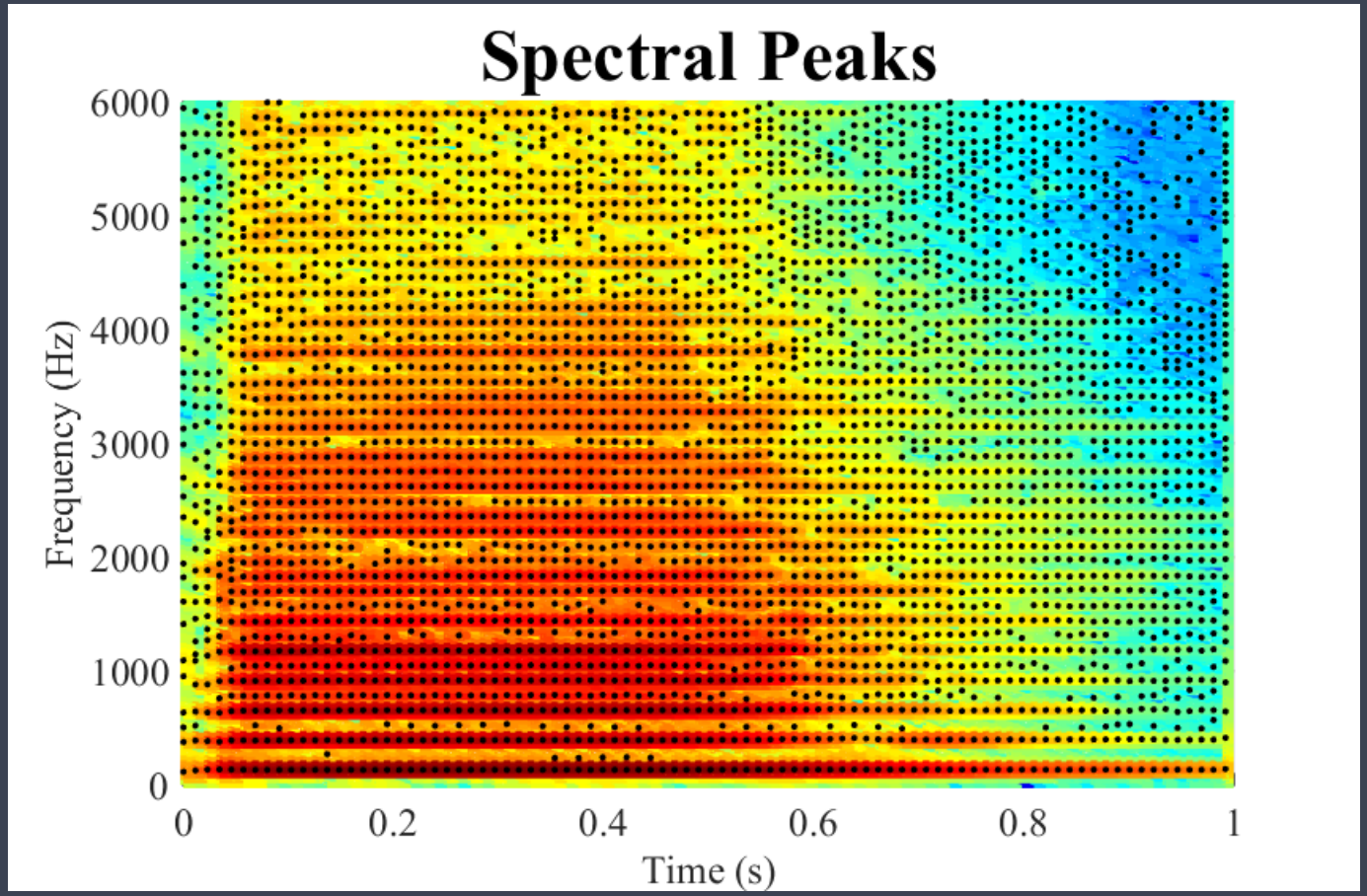
Peak Picking



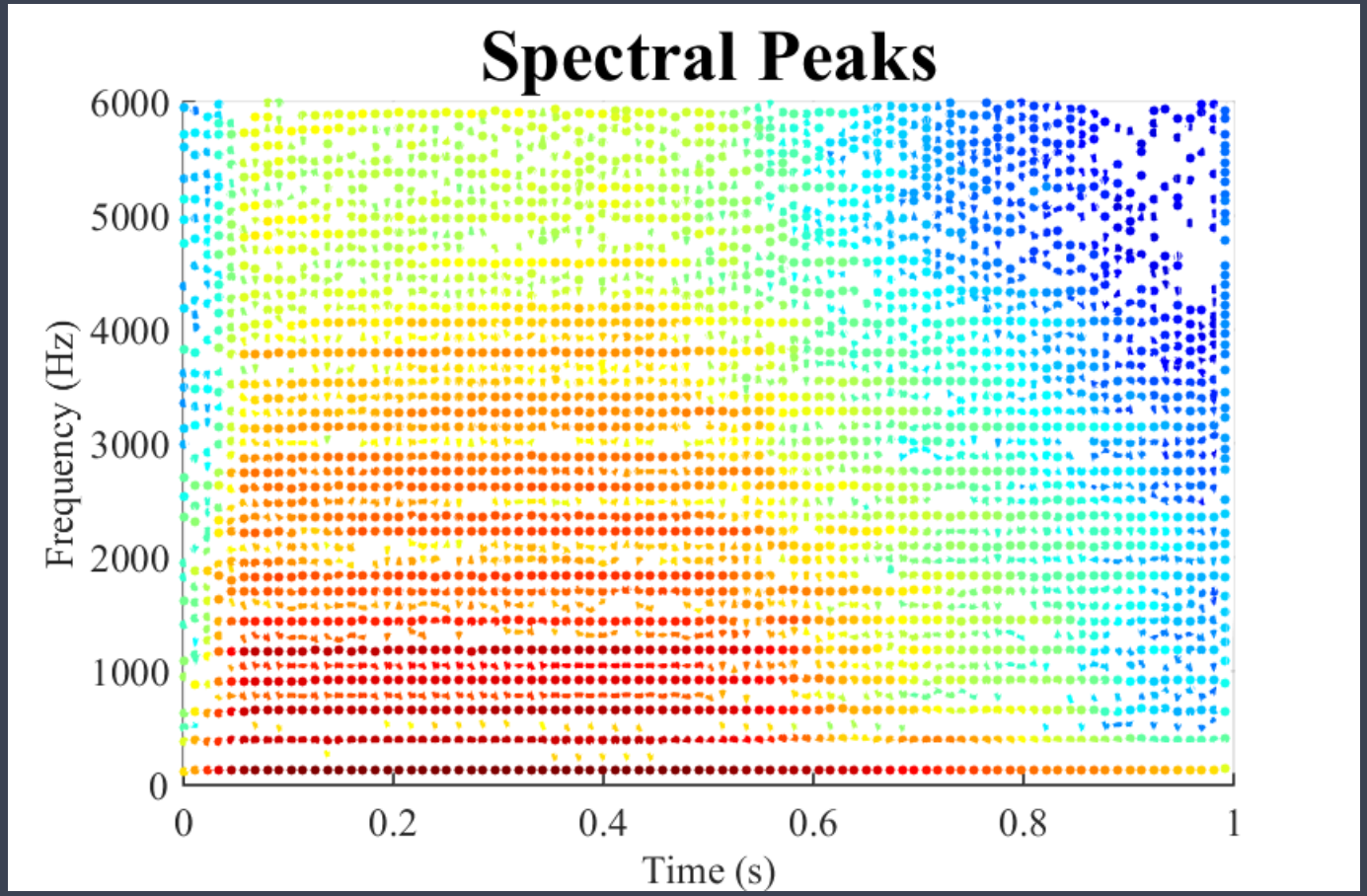
Spectrogram



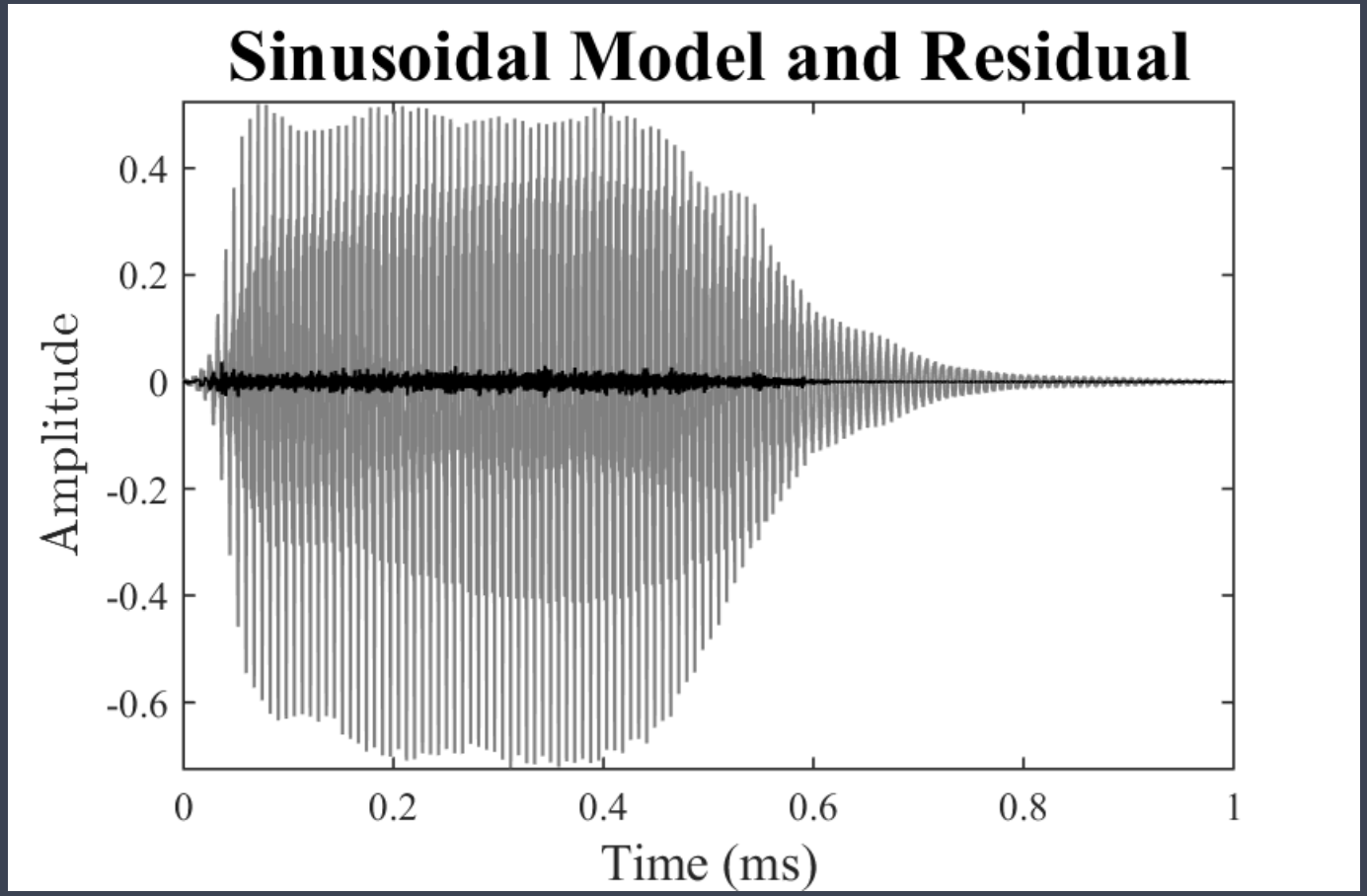
Spectral Peaks over the Spectrogram



Spectro-Temporal Representation with Spectral Peaks



Resynthesis with all peaks and phases

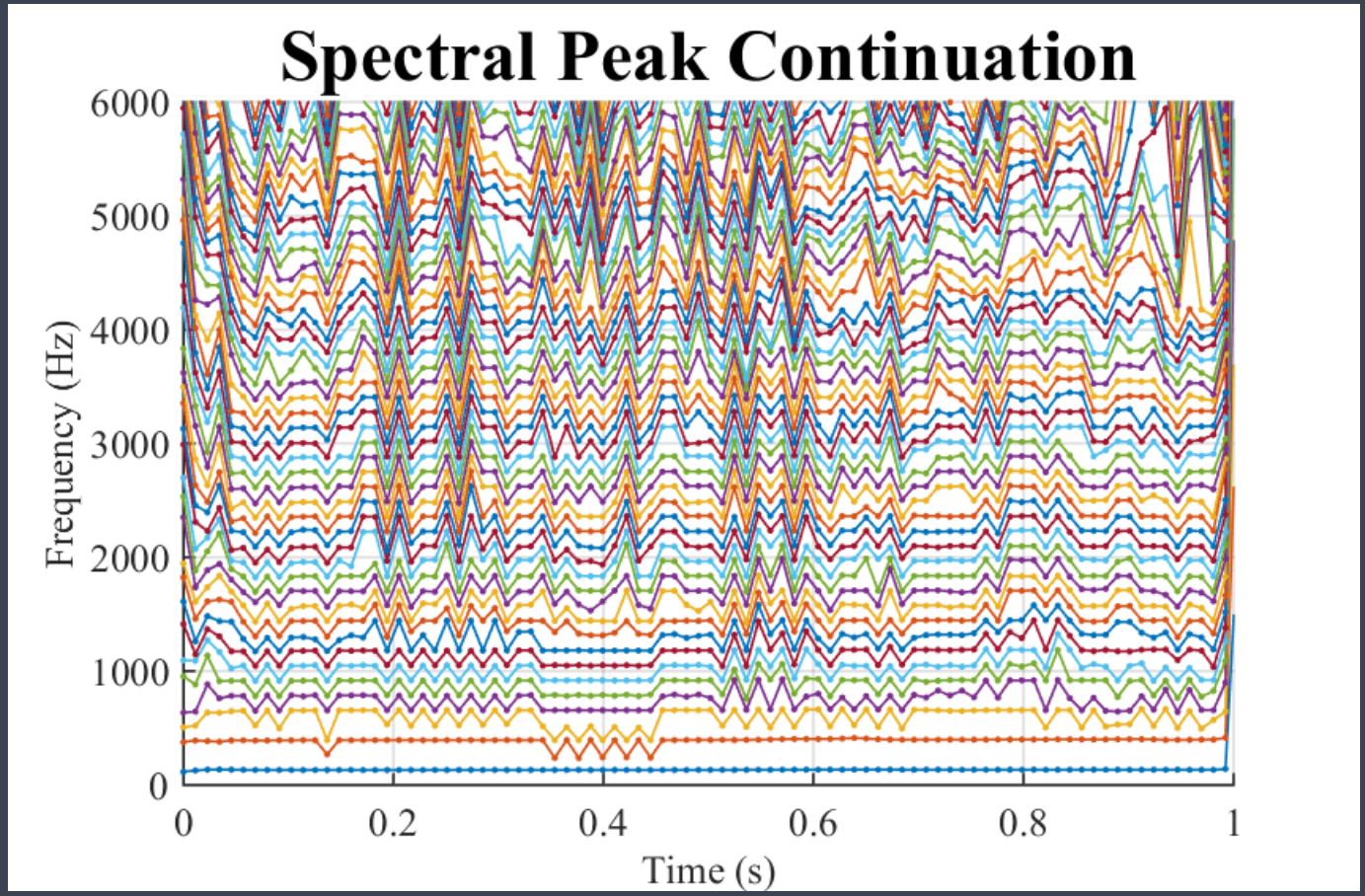


sinusoidal

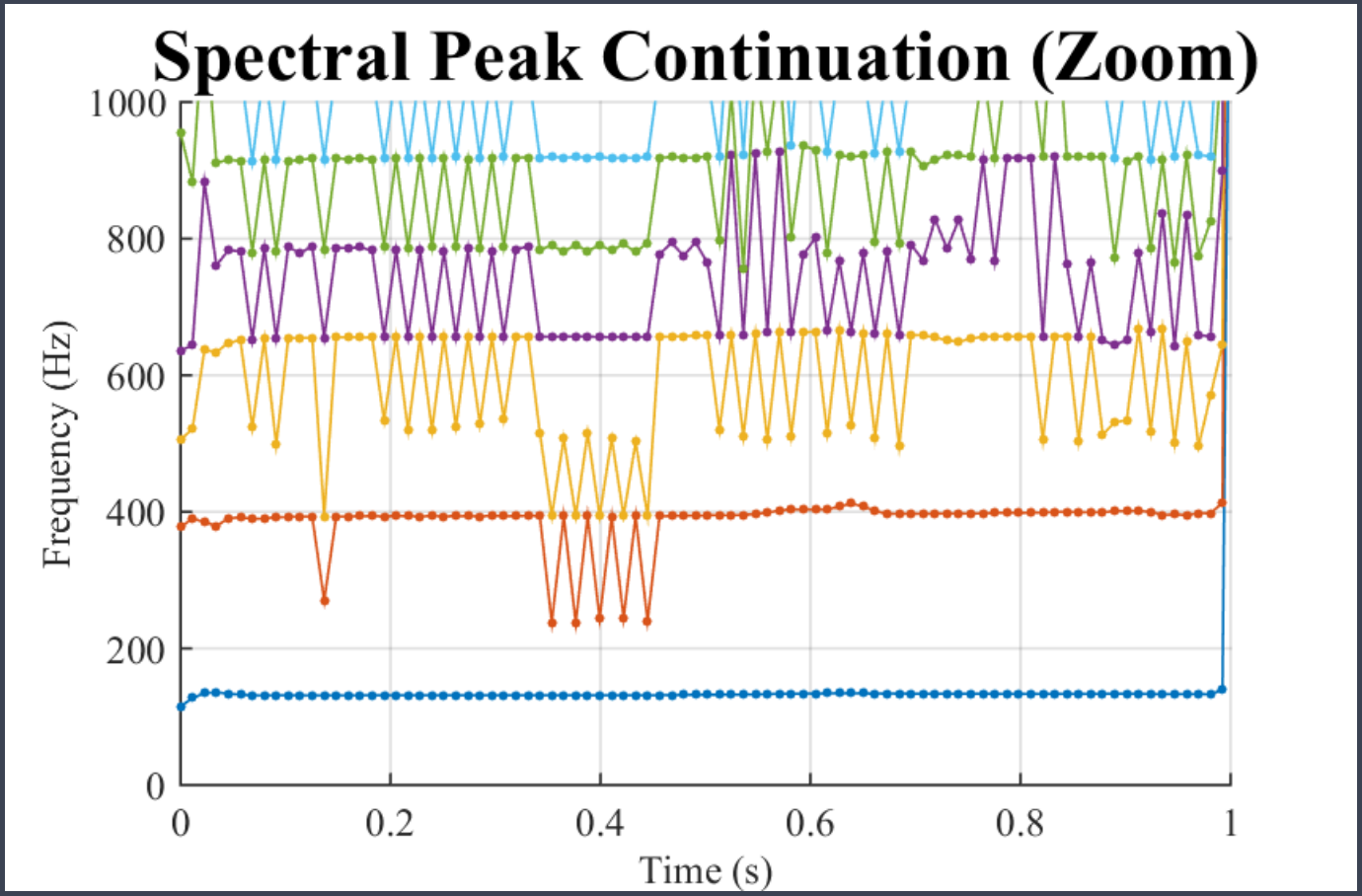
residual

Normalized residual

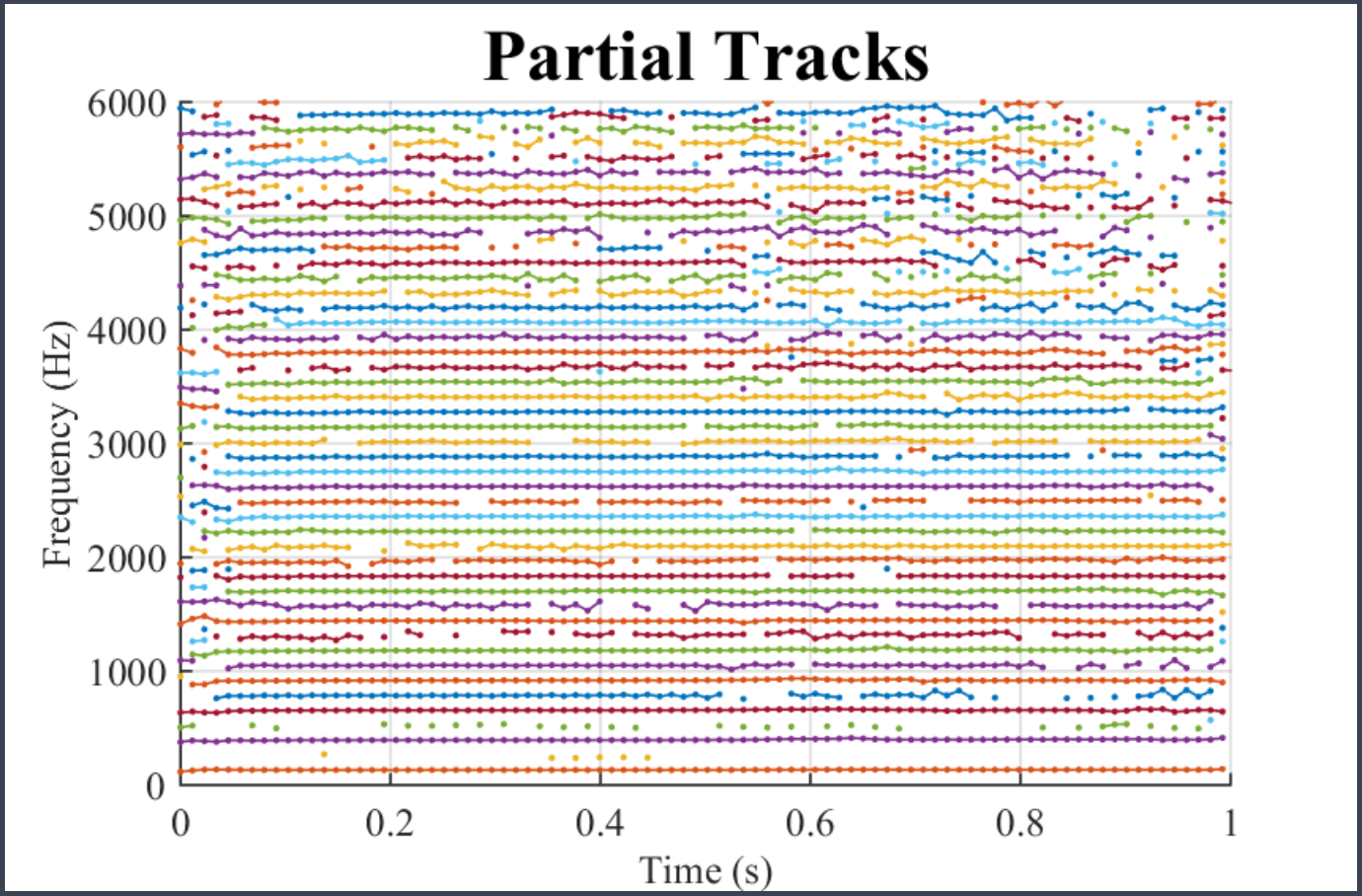
Spectral peak continuation



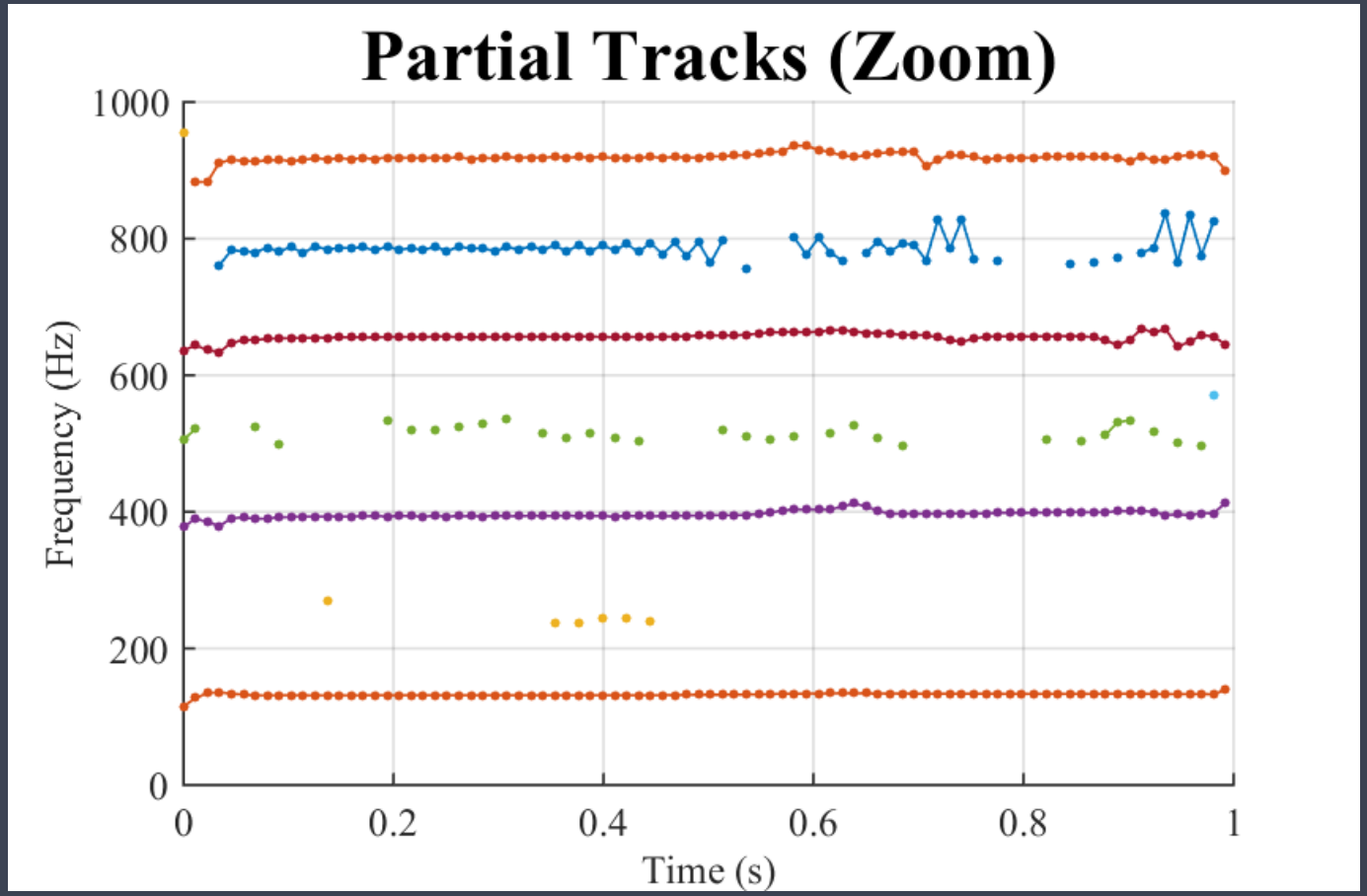
Spectral peak continuation



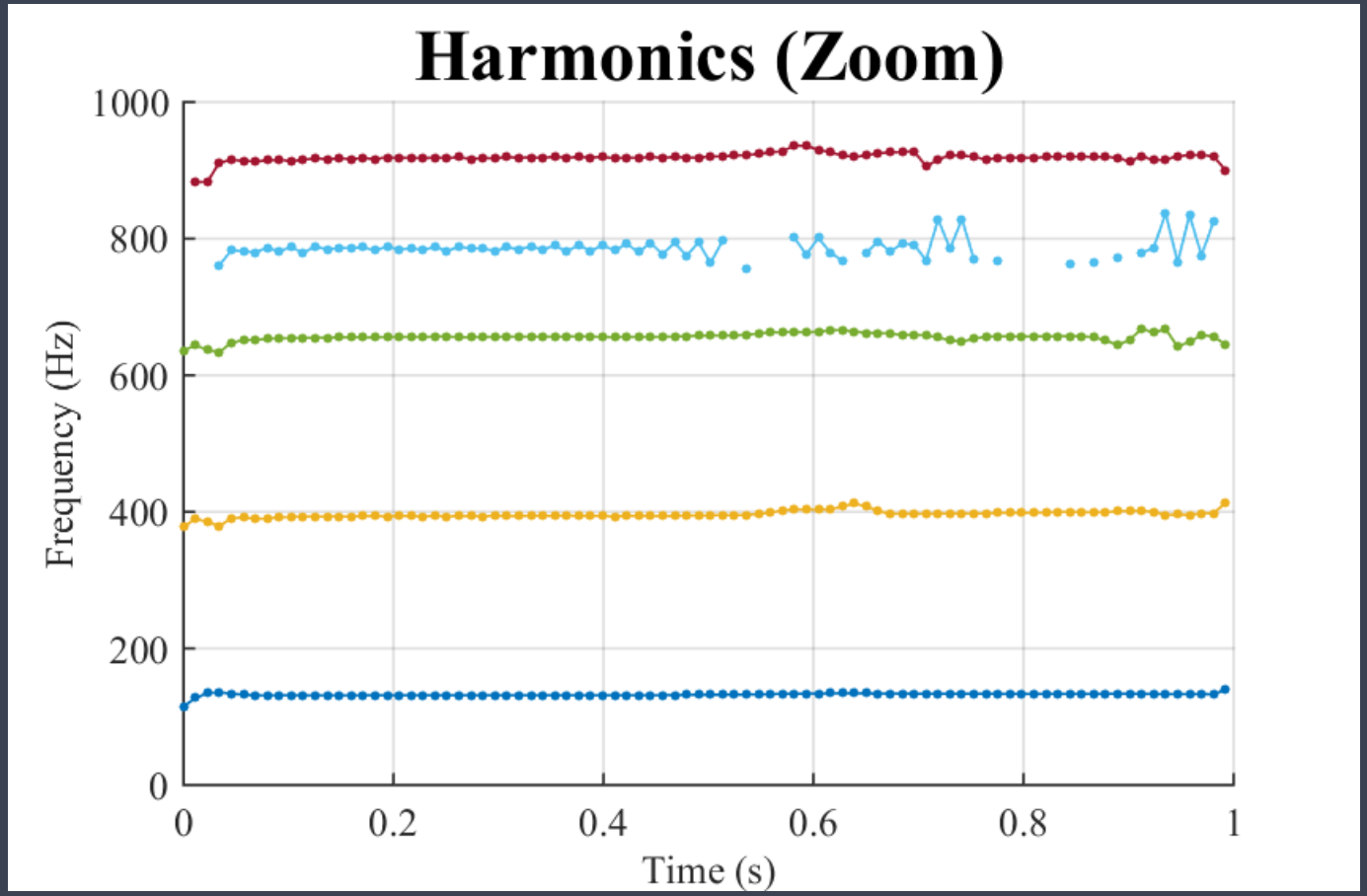
Make partial tracks



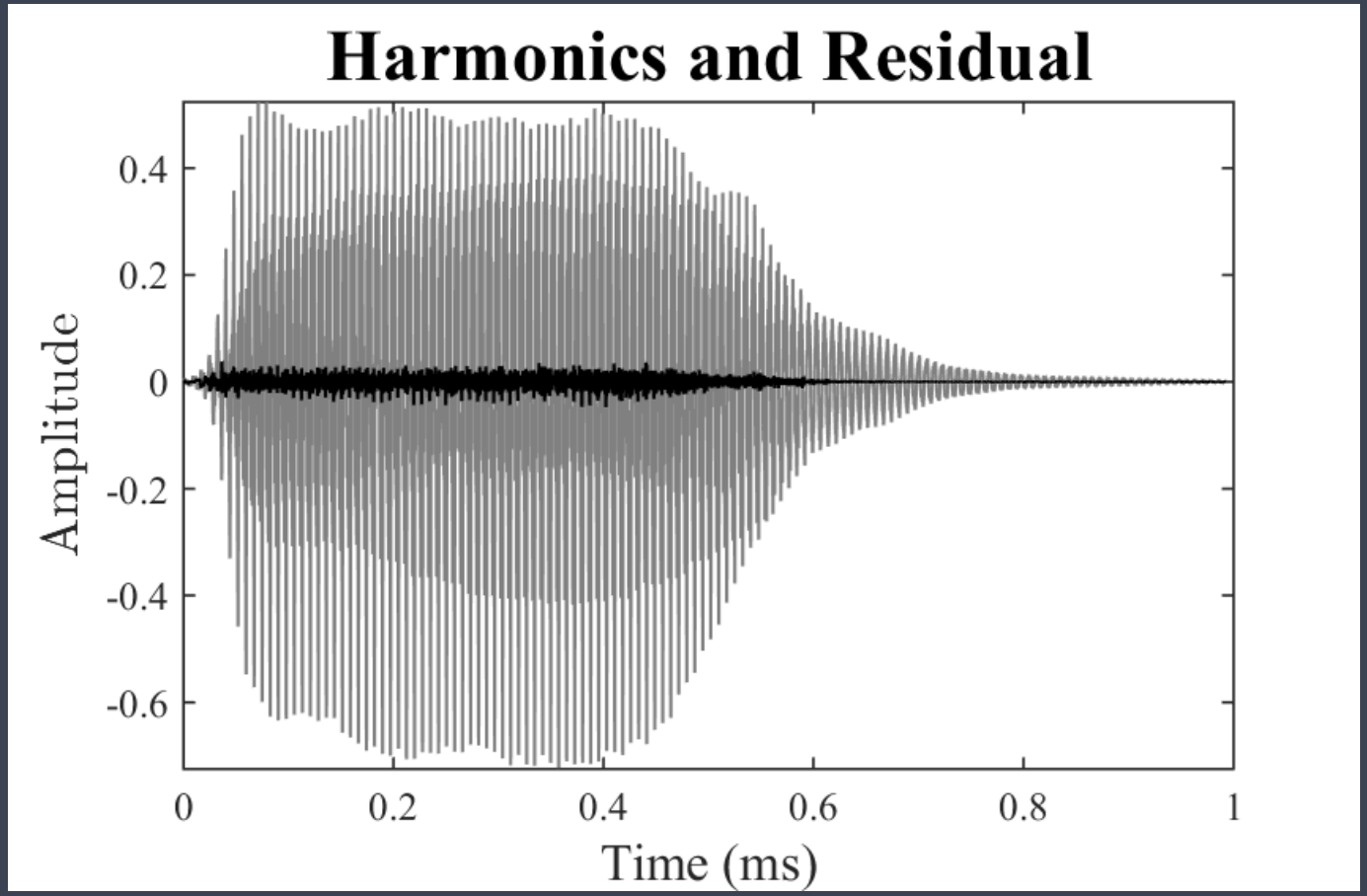
Make partial tracks



Only harmonics of the fundamental



Resynthesis: Only Harmonics



harmonics

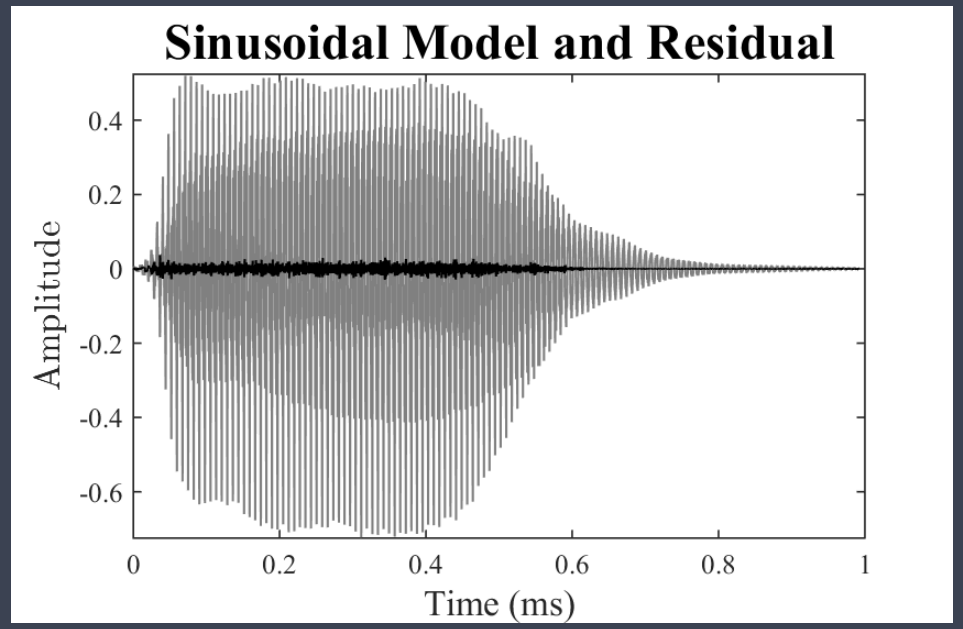
residual

Resynthesis without Phase

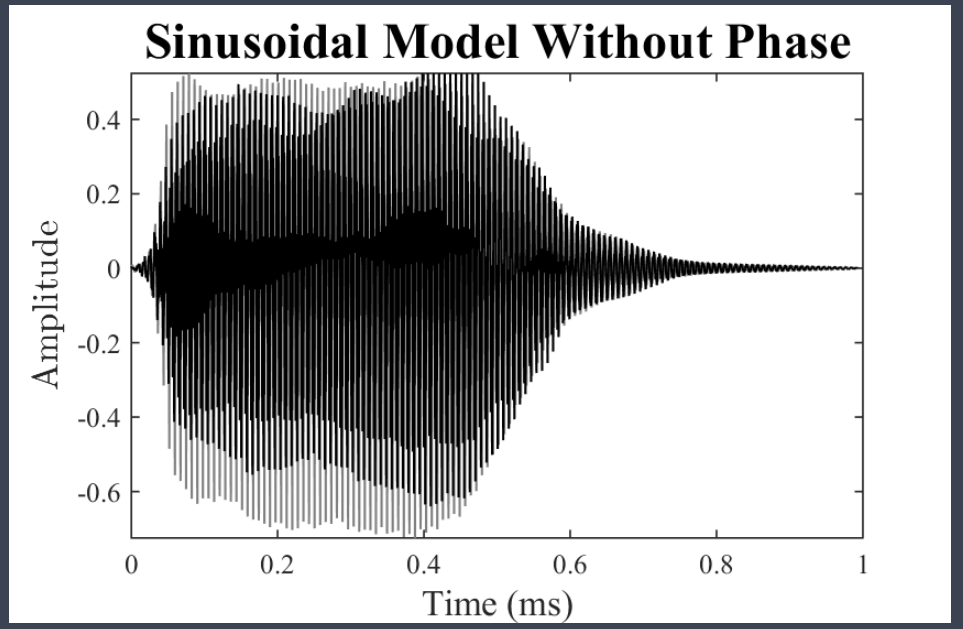
- Phase is used to preserve original waveform
- Morphed sound has interpolated amplitudes and frequencies
- What to do with the phase?
  - Discard original phase
  - Integrate frequency track
  - Different waveform
  - Perceptually different?

Comparison: All Peaks with Phase vs without Phase

All peaks with phase

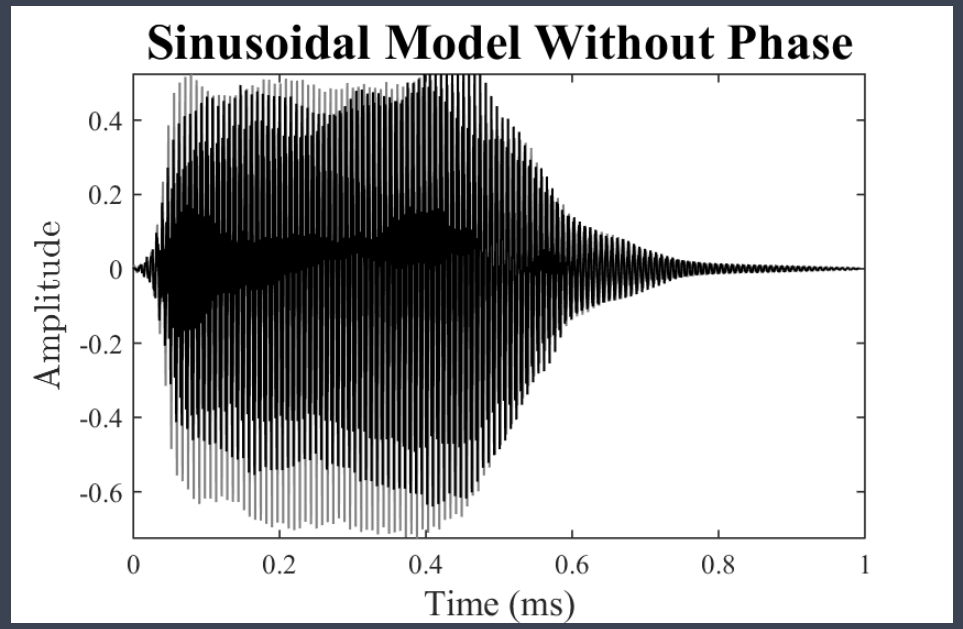


All peaks without phase

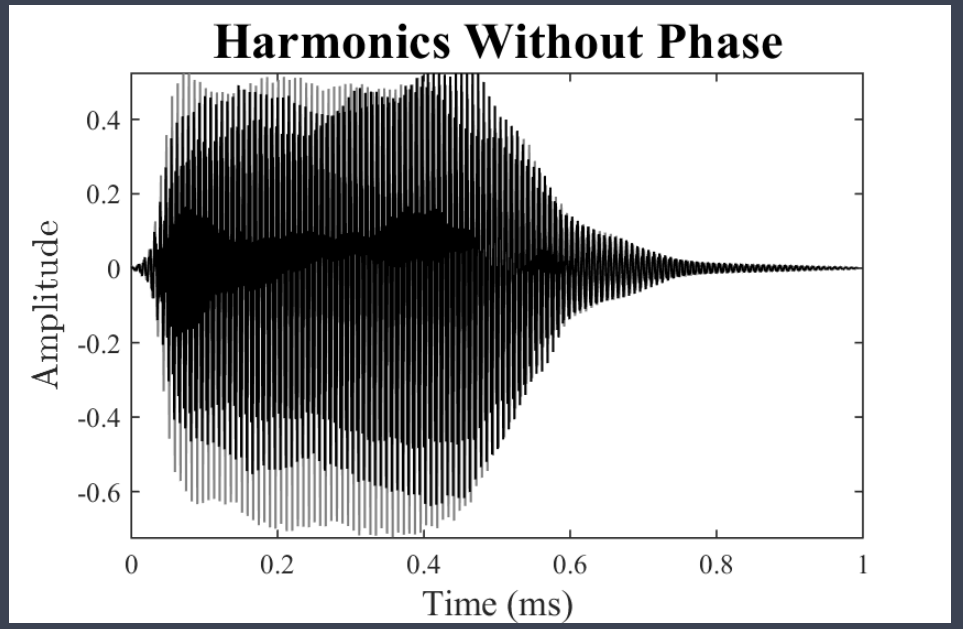


Comparison: No phase all peaks vs no phase harmonics

All peaks without phase

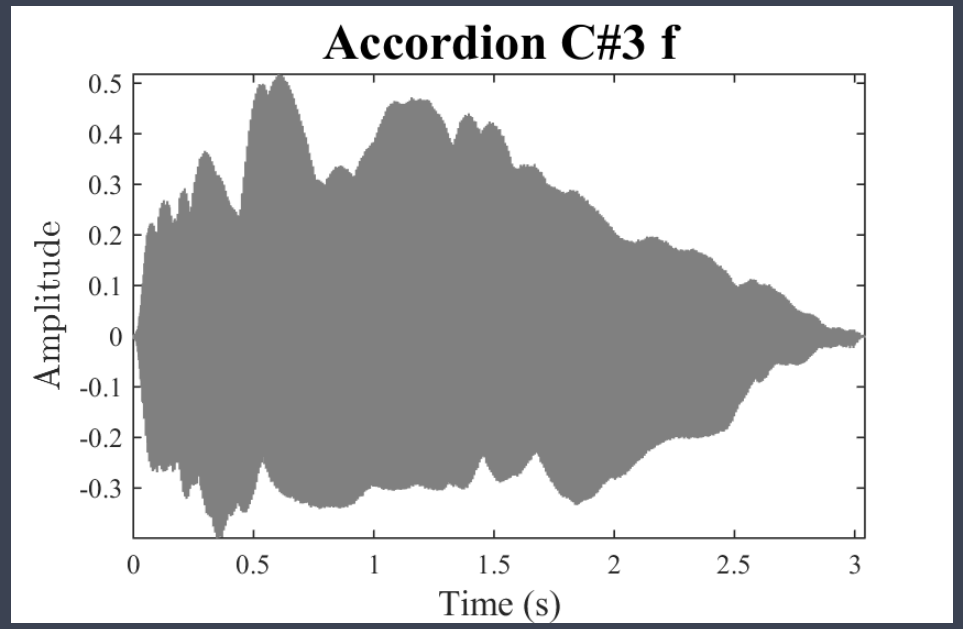


Harmonics without phase

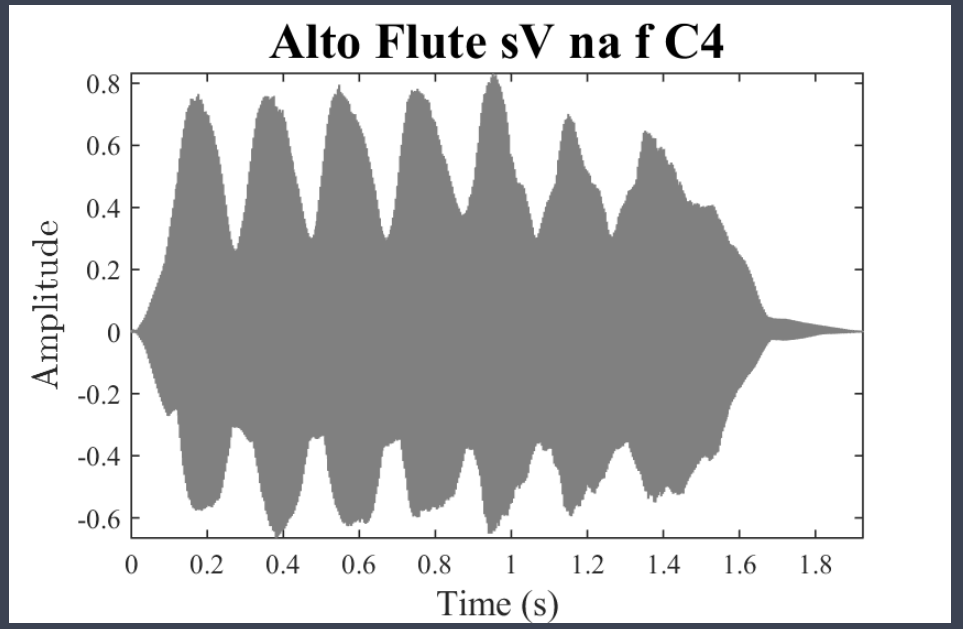


Morphing with the SMT

Accordion

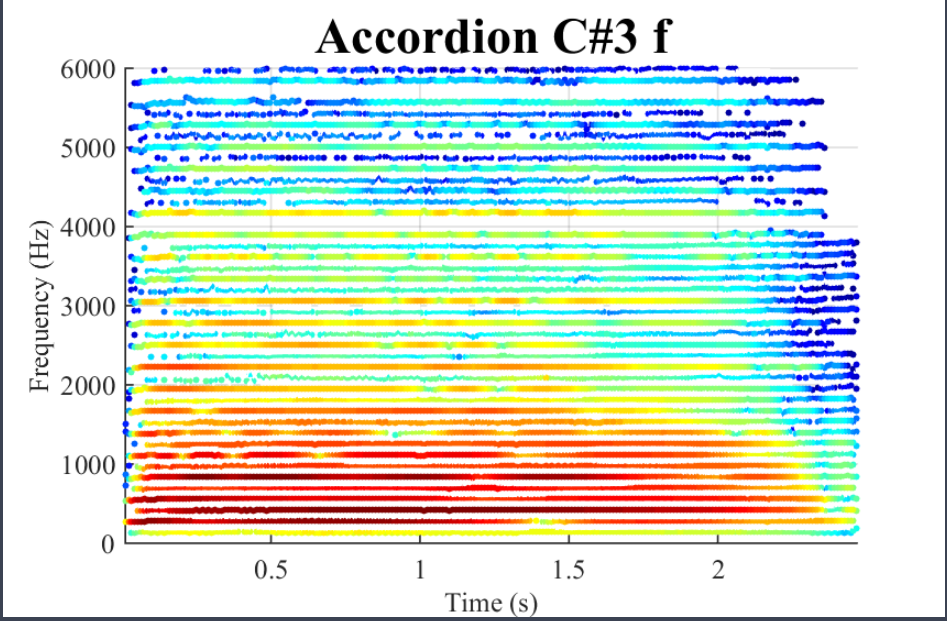


Alto Flute

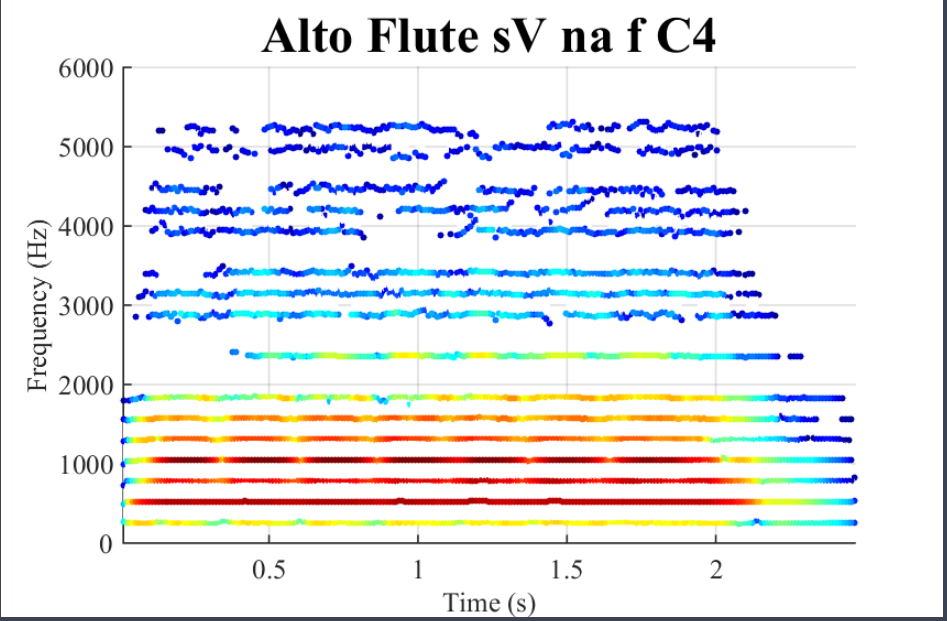


Morphing with the SMT

Accordion

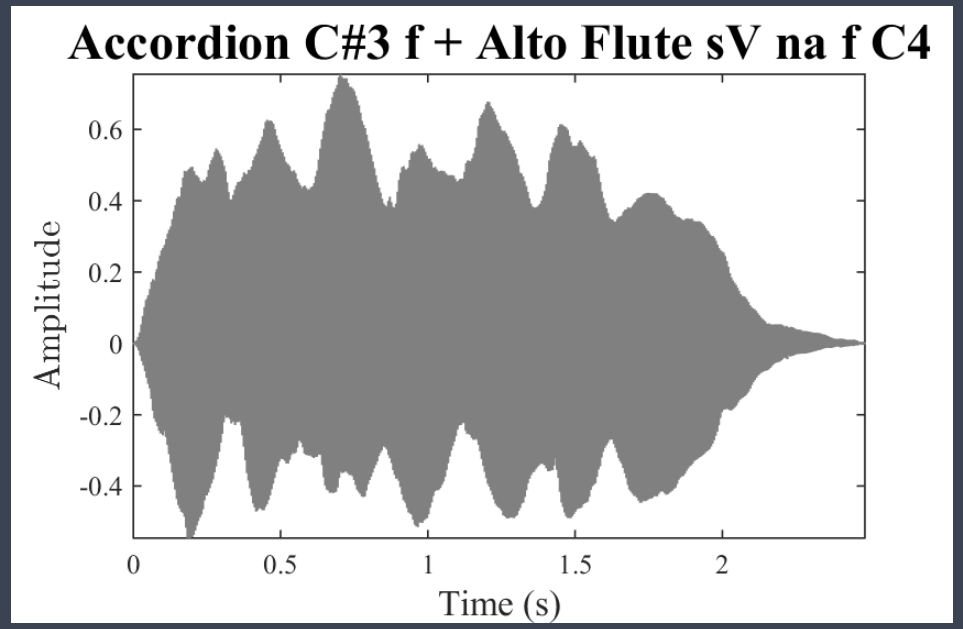


Alto Flute

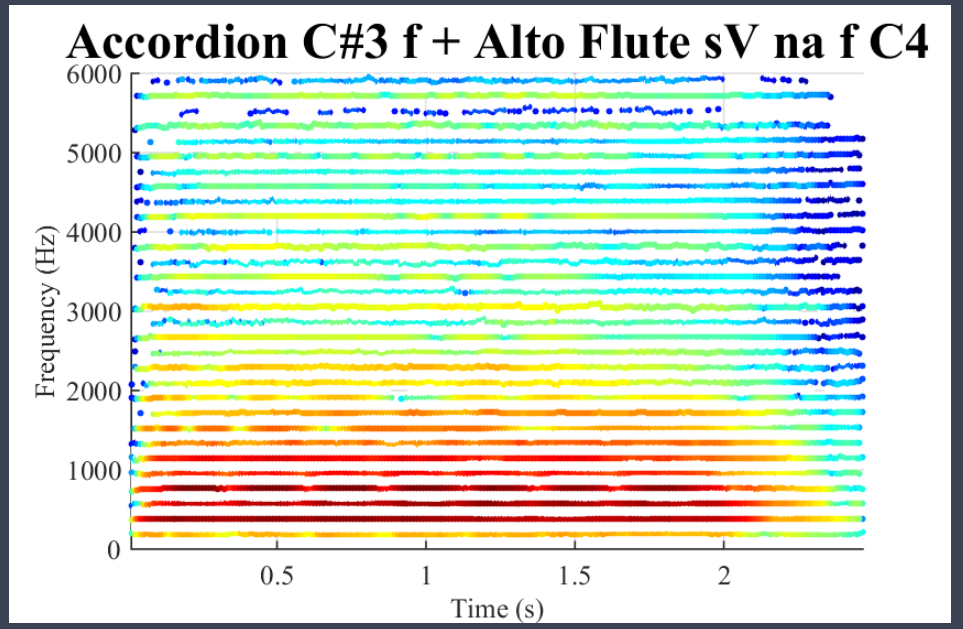


Morphing with the SMT

Morph waveform



Morph spectrogram



α=0.5

Morphing example: Different pitches and specificities

Accordion

- Original
- Time-scaled
- Harmonics

Alto Flute

- Original
- Time-scaled
- Harmonics

α=0.5

[MORPH]

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 831852 (MORPH)