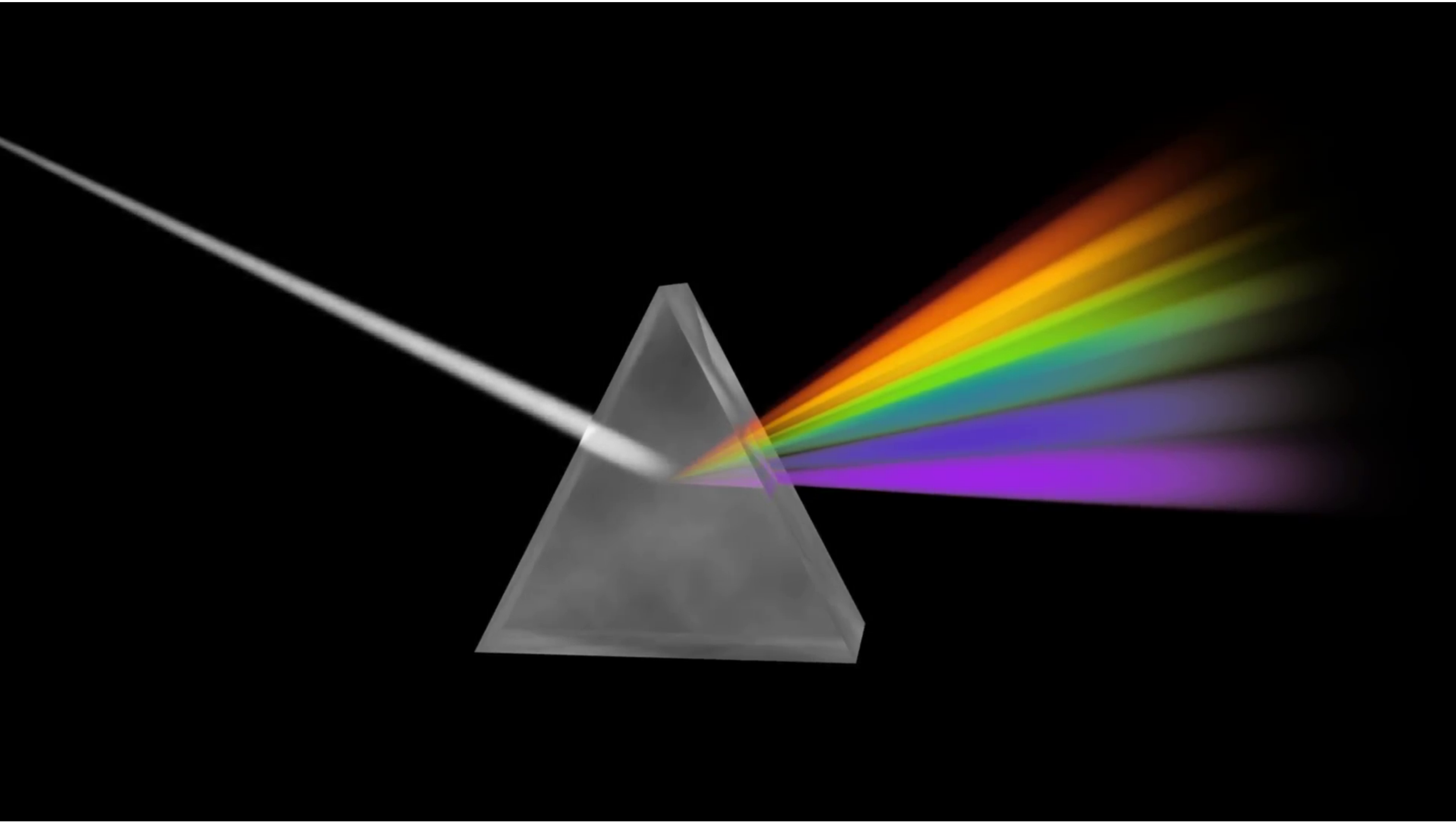
Periodicity analysis

Olivier
Sound, Audio & Music Analysis (SAMA)

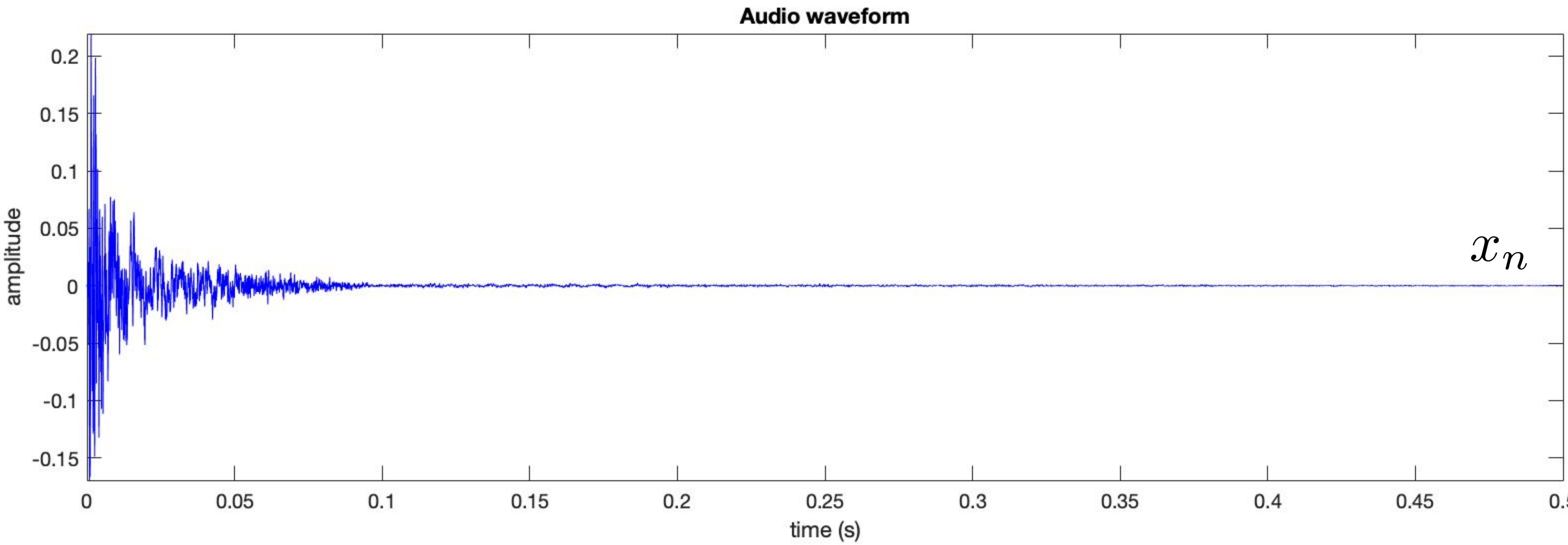
Overview

- Fourier Transform
- Autocorrelation function
- Wavelet
- Comb filter resonator
- Deep learning

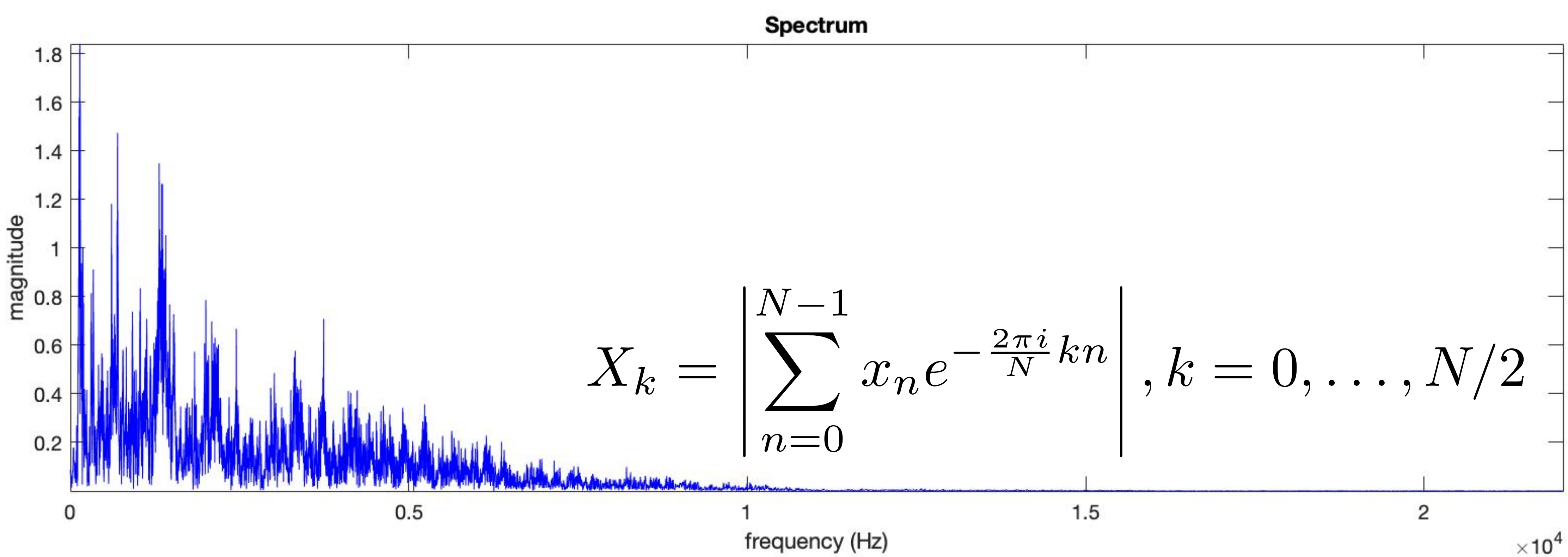
Spectrum



Fourier Transform



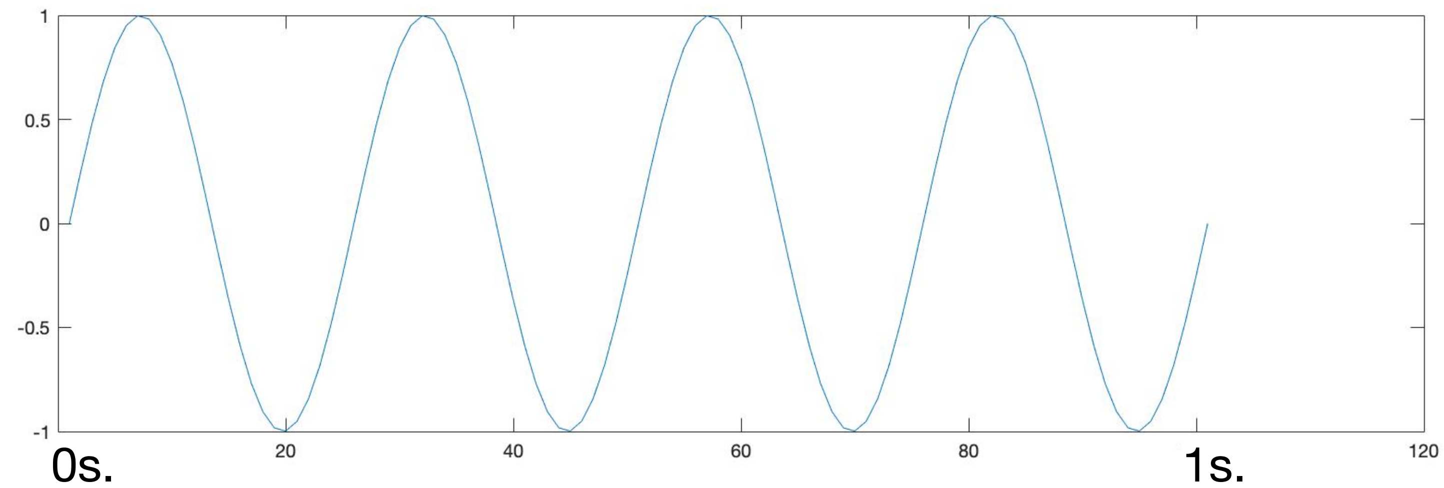
$$x_n$$



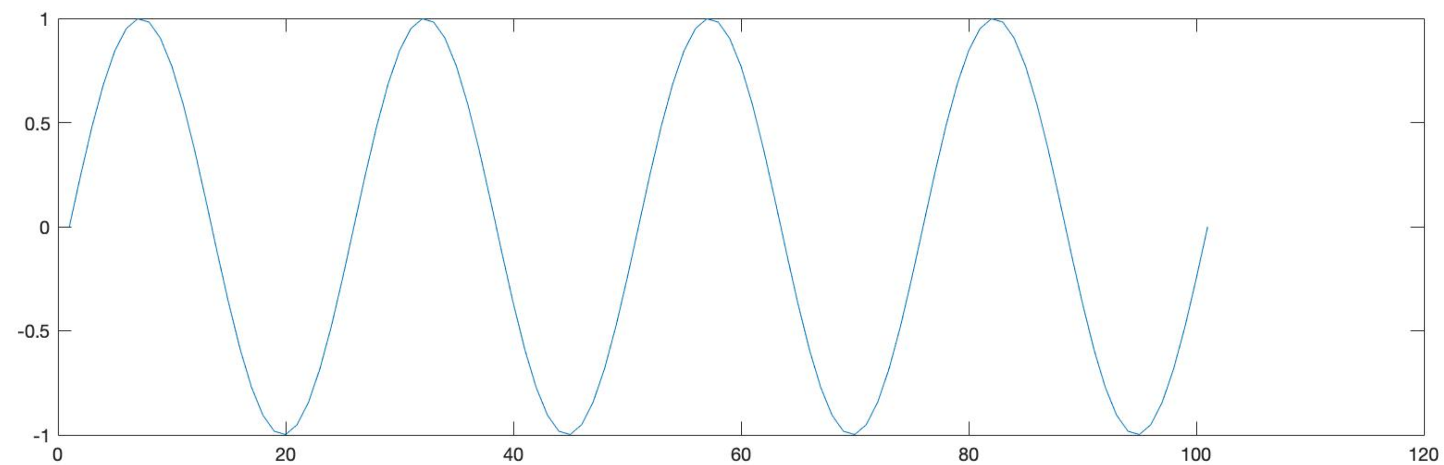
$$X_k = \left| \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \right|, k = 0, \ldots, N/2$$

Fourier Transform

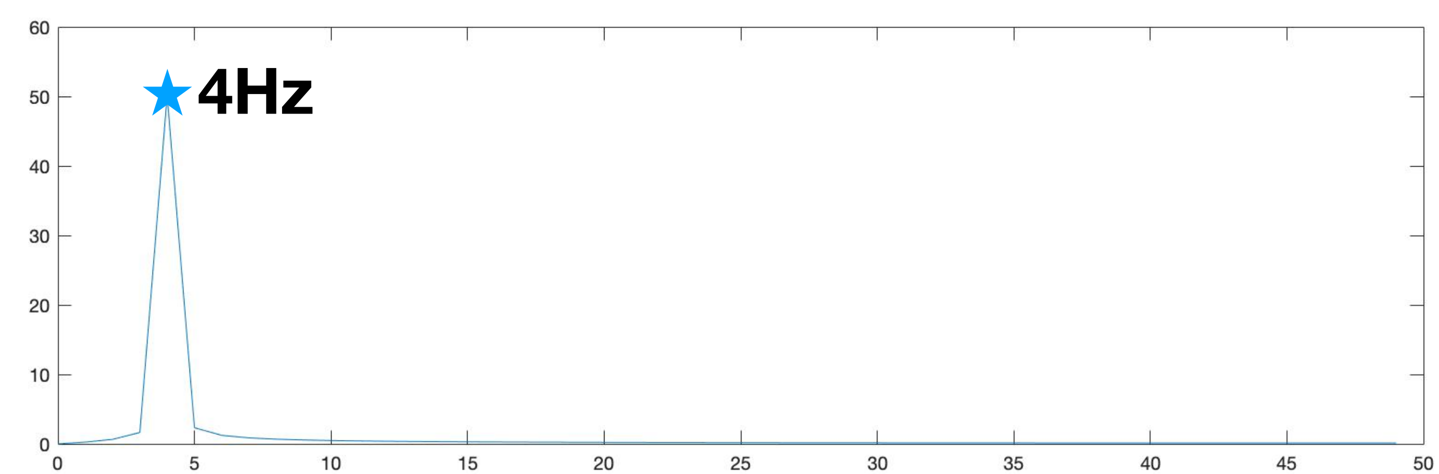
Signal to analyze:



4Hz sinusoid:

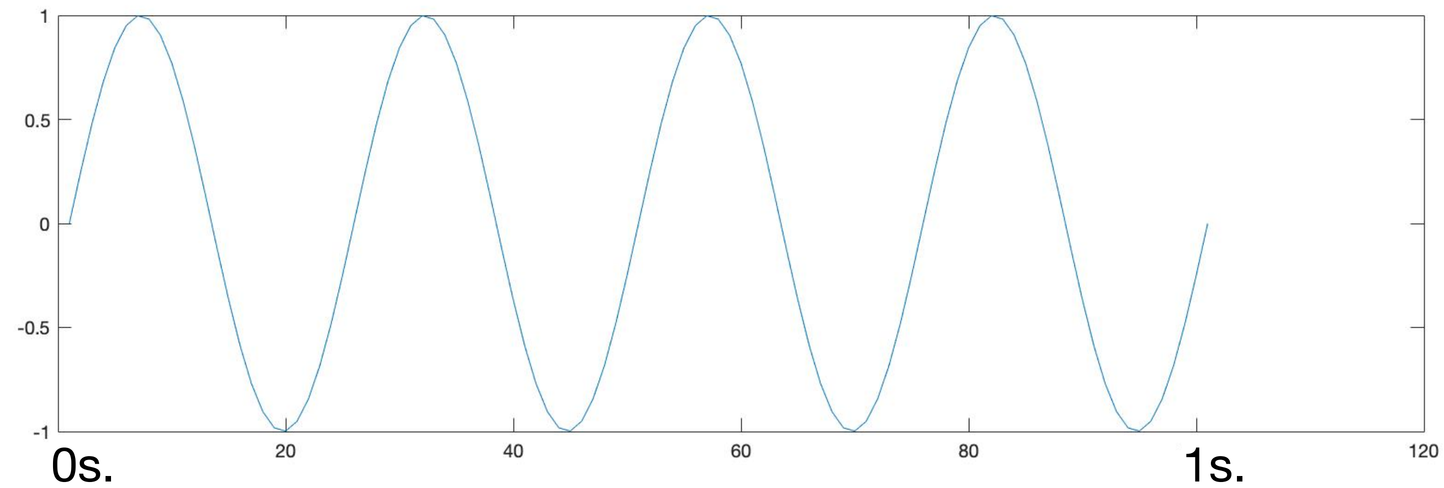


Fourier Transform:

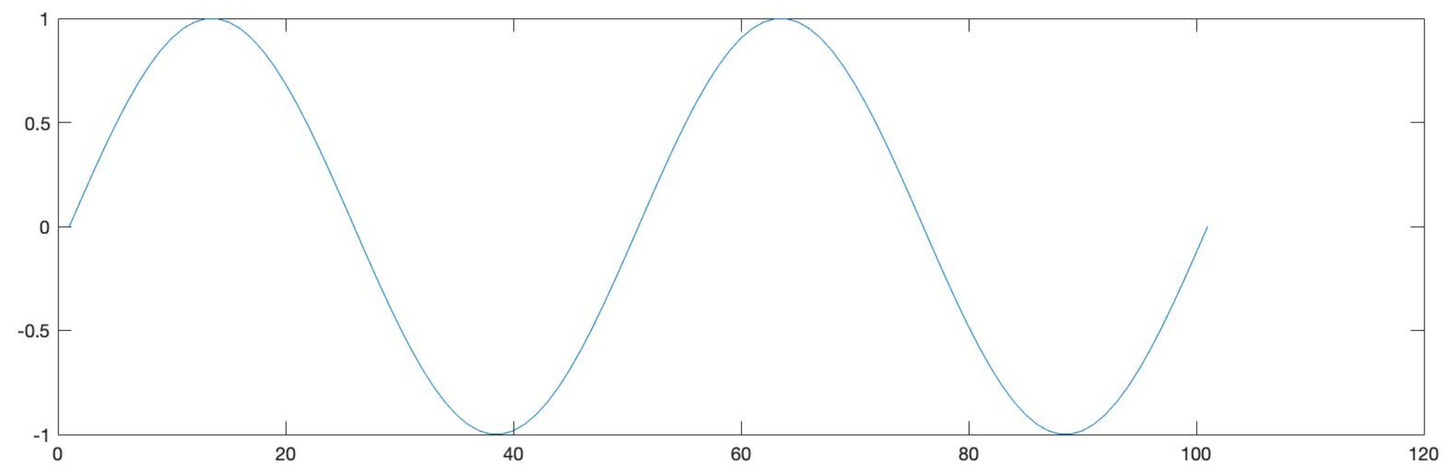


Fourier Transform

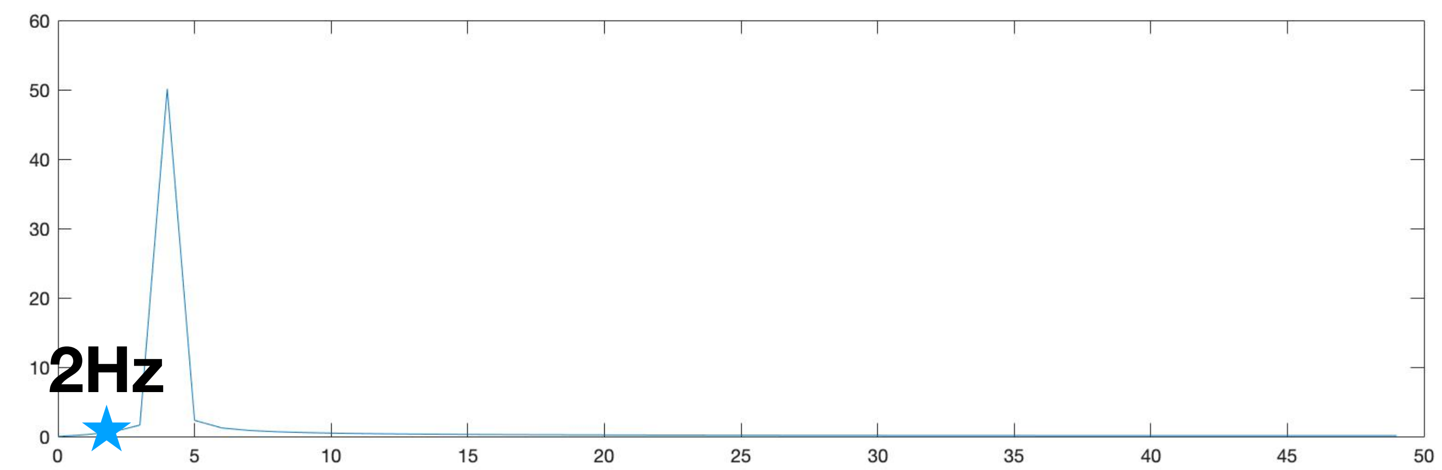
Signal to analyze:



2Hz sinusoid:

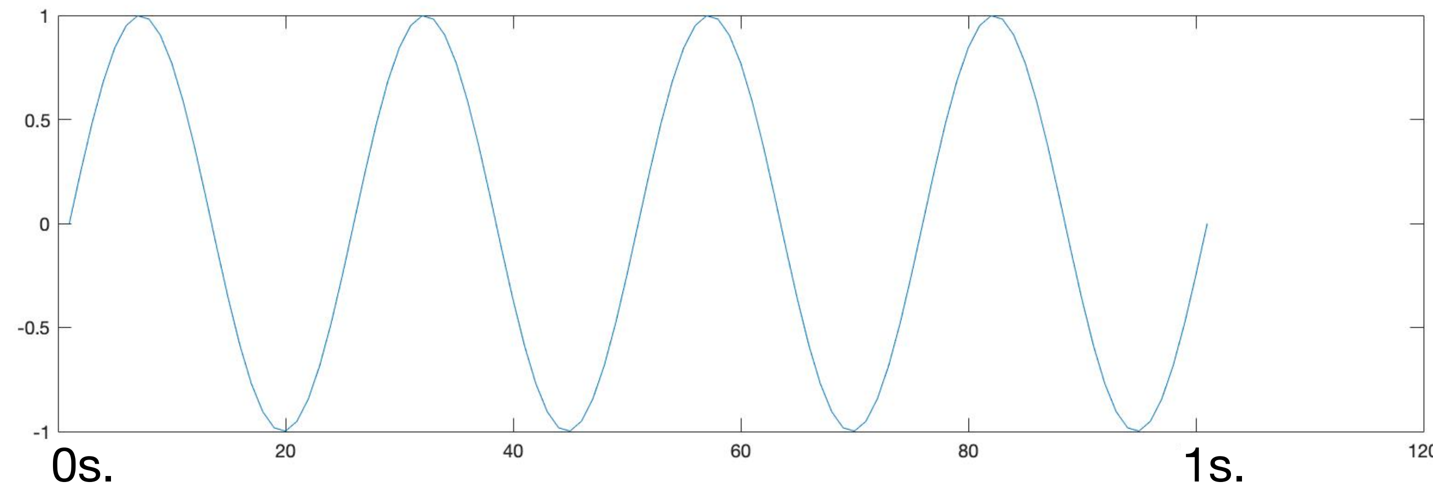


Fourier Transform:

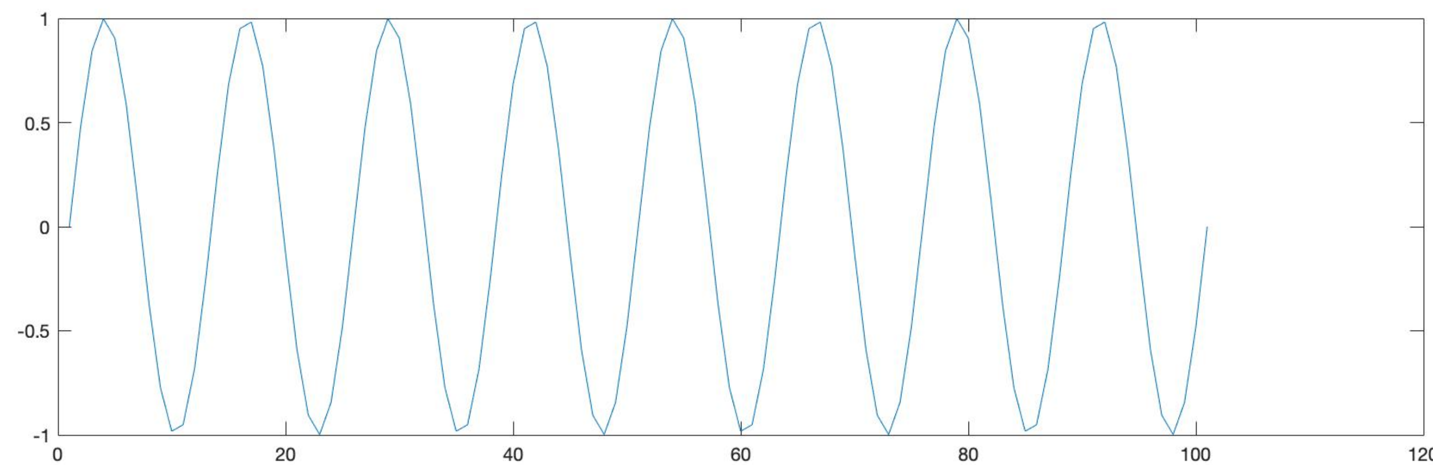


Fourier Transform

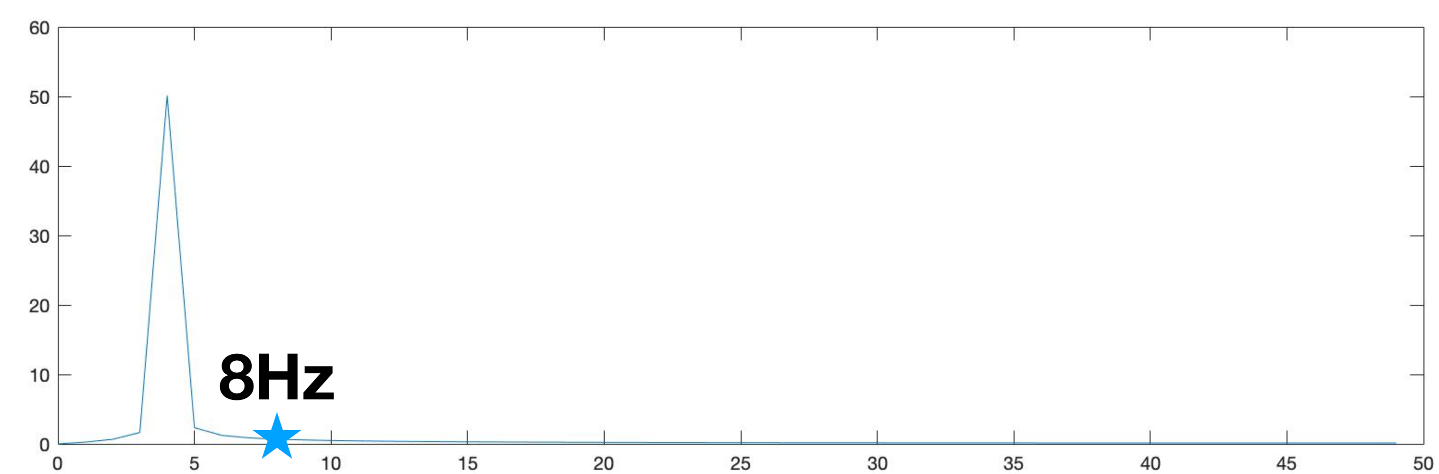
Signal to analyze:



8Hz sinusoid:

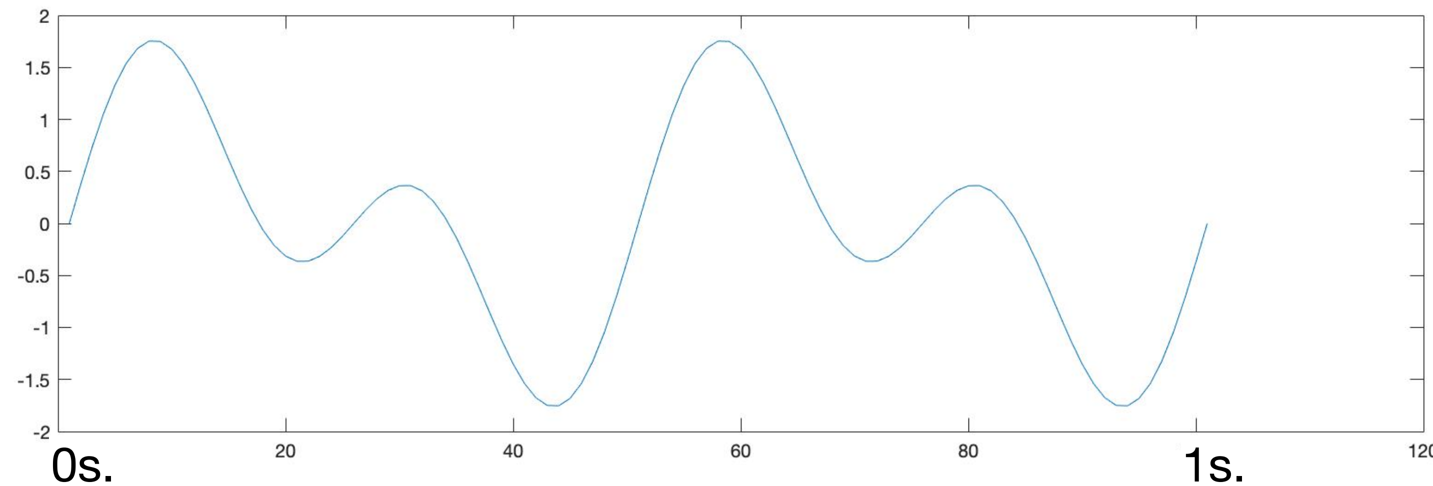


Fourier Transform:

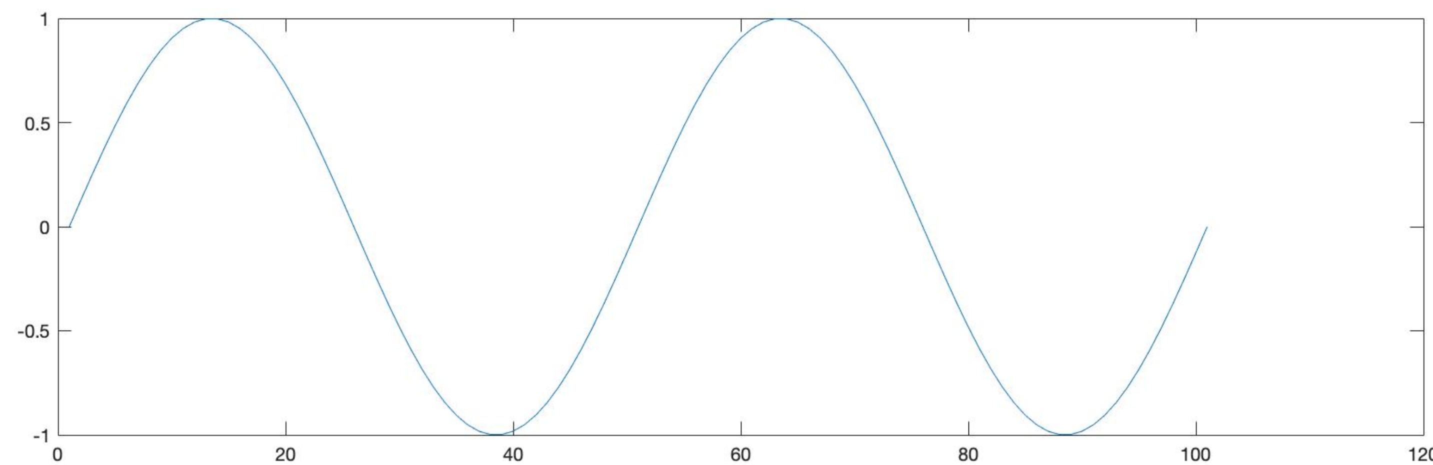


Fourier Transform

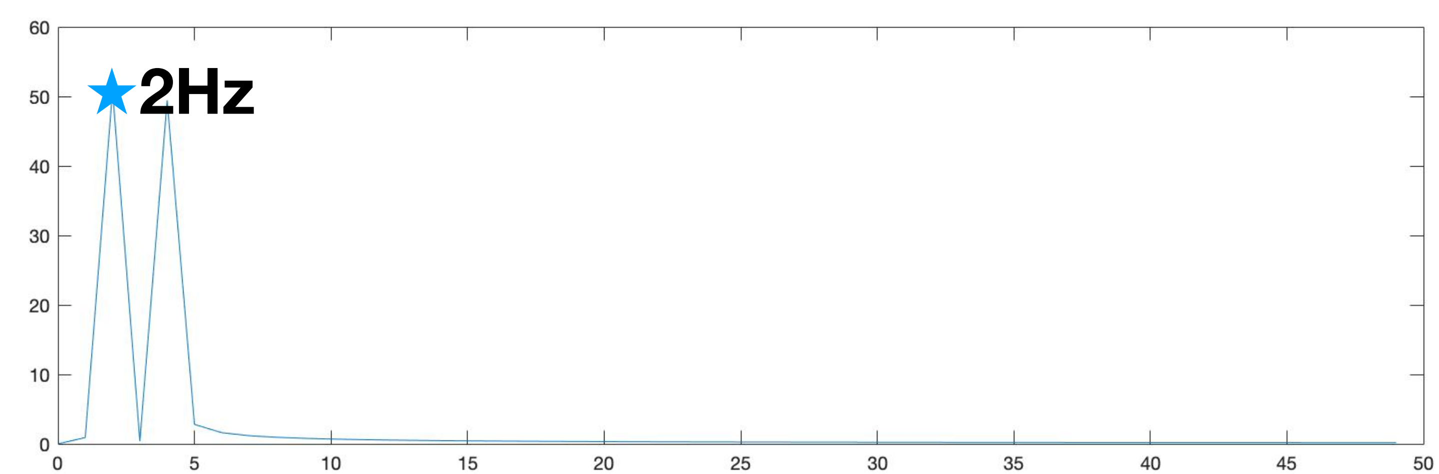
Signal to analyze:



2Hz sinusoid:

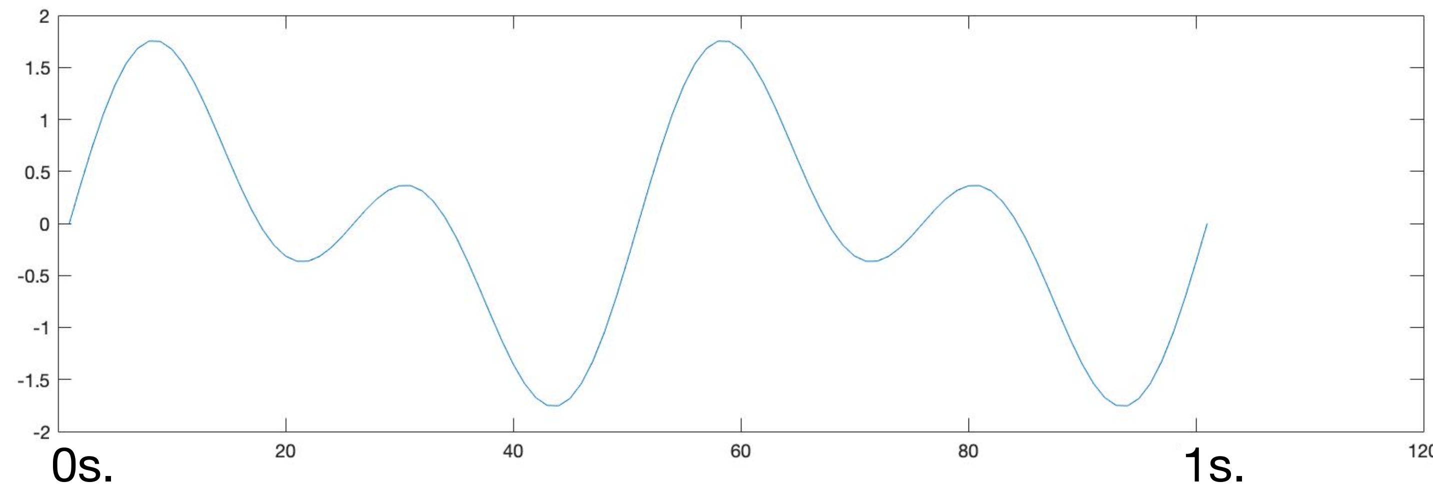


Fourier Transform:

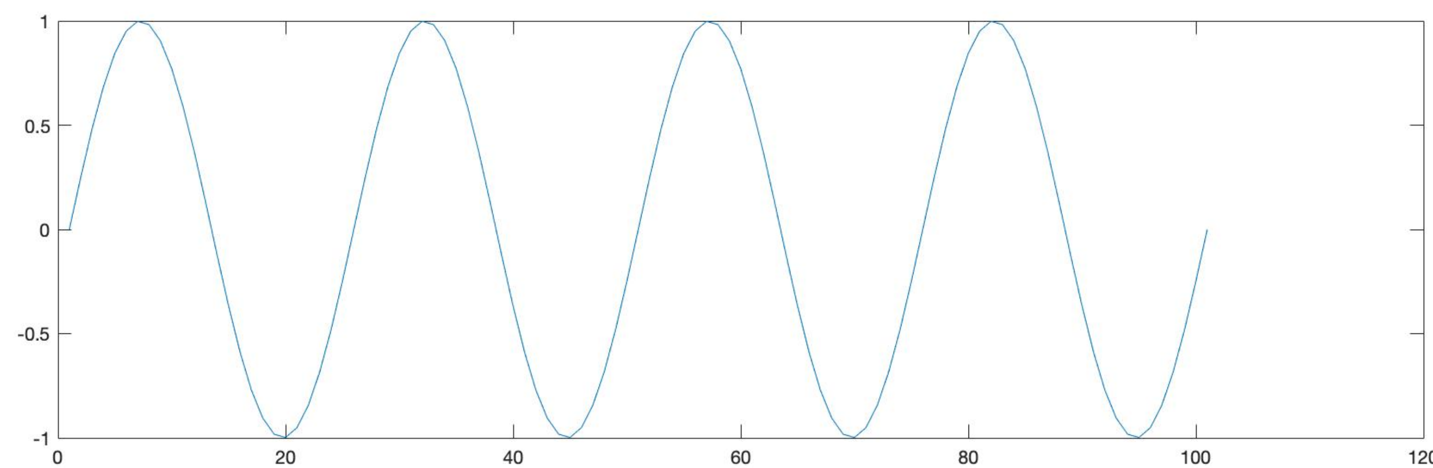


Fourier Transform

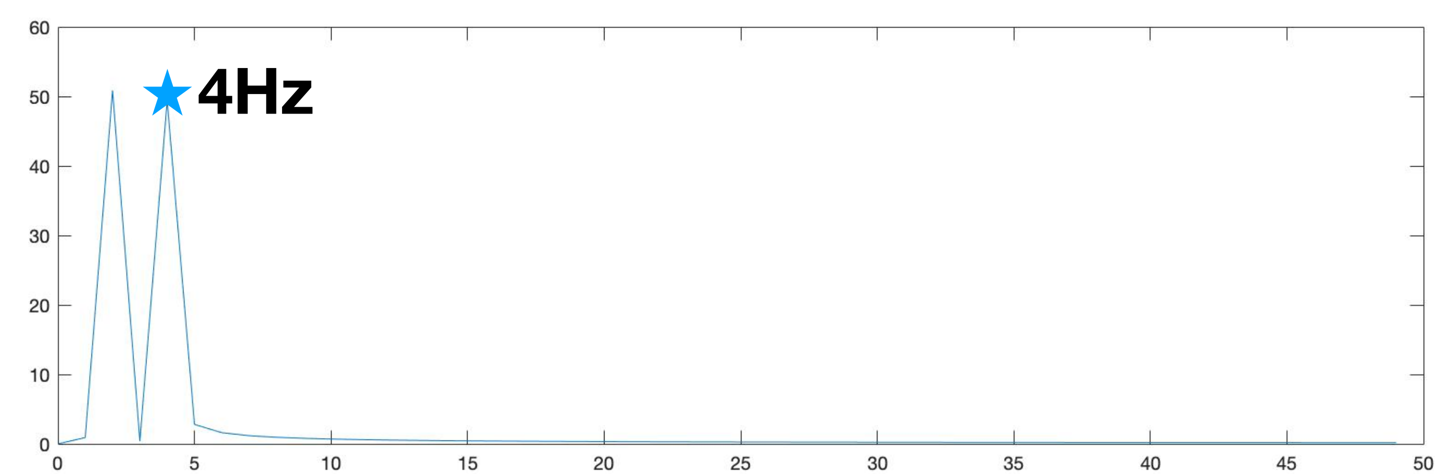
Signal to analyze:



4Hz sinusoid:



Fourier Transform:



Behavioral/Cognitive

Neural Entrainment to the Beat: The “Missing-Pulse” Phenomenon

Idan Tal,[1] Edward W. Large,[2] Eshed Rabinovitch,[1] Yi Wei,[2] Charles E. Schroeder,[3,4] David Poeppel,[5,6] and Elana Zion Golumbic[1]

[1]Gonda Multidisciplinary Brain Research Center, Bar Ilan University, Ramat Gan 5290002, Israel, [2]Department of Psychology, University of Connecticut Storrs, Connecticut 06269, [3]Department of Psychiatry, Columbia University Medical Center, New York, New York 10032, [4]Translational Cognitive Neuroscience Program, Nathan Kline Institute, Orangeburg, New York 10962, [5]Department of Psychology, New York University, New York, New York 10003, and [6]Neuroscience Department, Max-Planck Institute for Empirical Aesthetics, Frankfurt, Germany 60322

Most humans have a near-automatic inclination to tap, clap, or move to the beat of music. The capacity to extract a periodic beat from a complex musical segment is remarkable, as it requires abstraction from the temporal structure of the stimulus. It has been suggested that nonlinear interactions in neural networks result in cortical oscillations at the beat frequency, and that such entrained oscillations give rise to the percept of a beat or a pulse. Here we tested this neural resonance theory using MEG recordings as female and male individuals listened to 30 s sequences of complex syncopated drumbeats designed so that they contain no net energy at the pulse frequency when measured using linear analysis. We analyzed the spectrum of the neural activity while listening and compared it to the modulation spectrum of the stimuli. We found enhanced neural response in the auditory cortex at the pulse frequency. We also showed phase locking at the times of the missing pulse, even though the pulse was absent from the stimulus itself. Moreover, the strength of this pulse response correlated with individuals’ speed in finding the pulse of these stimuli, as tested in a follow-up session. These findings demonstrate that neural activity at the pulse frequency in the auditory cortex is internally generated rather than stimulus-driven. The current results are both consistent with neural resonance theory and with models based on nonlinear response of the brain to rhythmic stimuli. The results thus help narrow the search for valid models of beat perception.

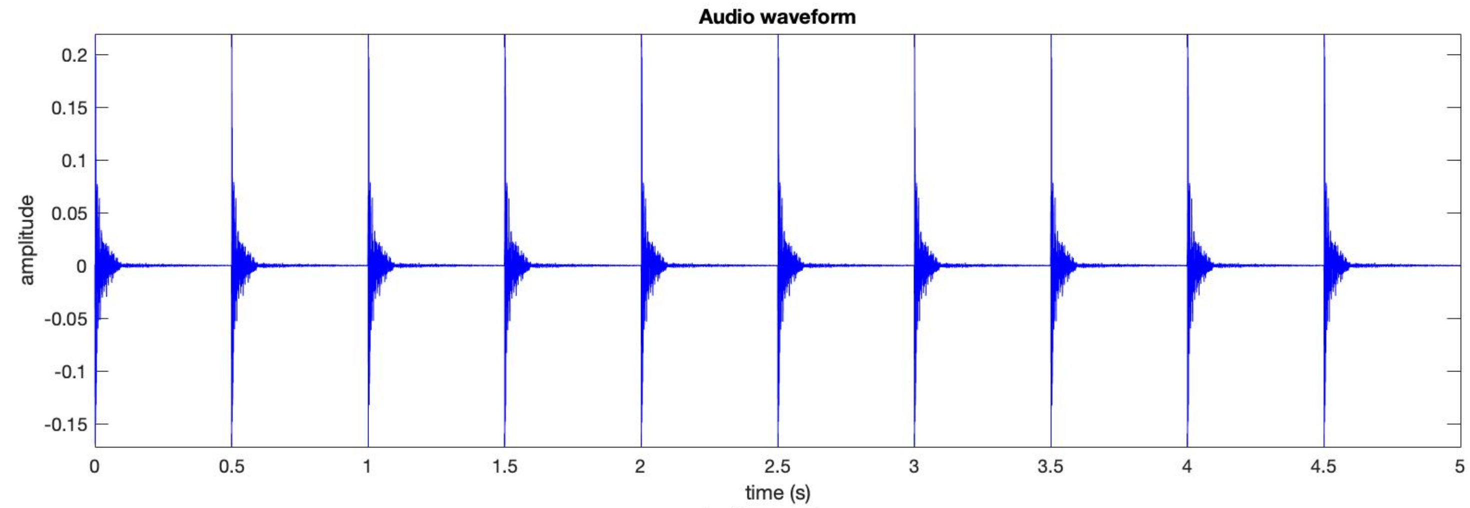
Key words: auditory rhythm; MEG; neural resonance theory; oscillations; pulse

Significance Statement

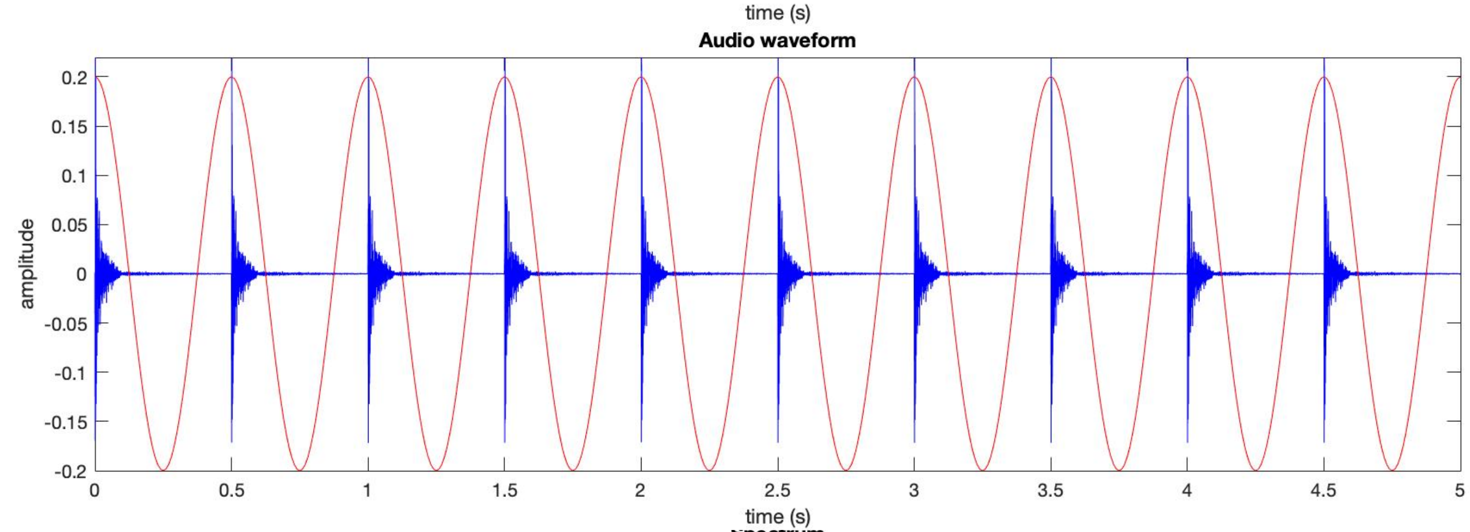
Humans perceive music as having a regular pulse marking equally spaced points in time, within which musical notes are temporally organized. Neural resonance theory (NRT) provides a theoretical model explaining how an internal periodic representation of a pulse may emerge through nonlinear coupling between oscillating neural systems. After testing key falsifiable predictions of NRT using MEG recordings, we demonstrate the emergence of neural oscillations at the pulse frequency, which can be related to pulse perception. These findings rule out alternative explanations for neural entrainment and provide evidence linking neural synchronization to the perception of pulse, a widely debated topic in recent years.

Fourier Transform

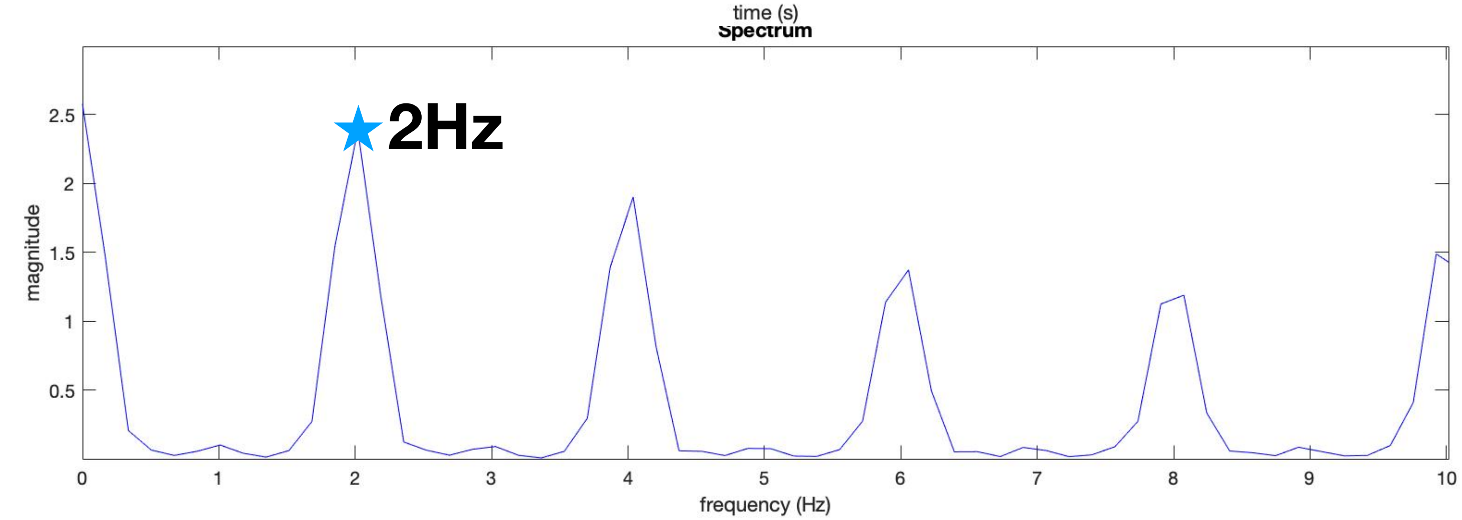
Signal to analyze:



2Hz sinusoid:

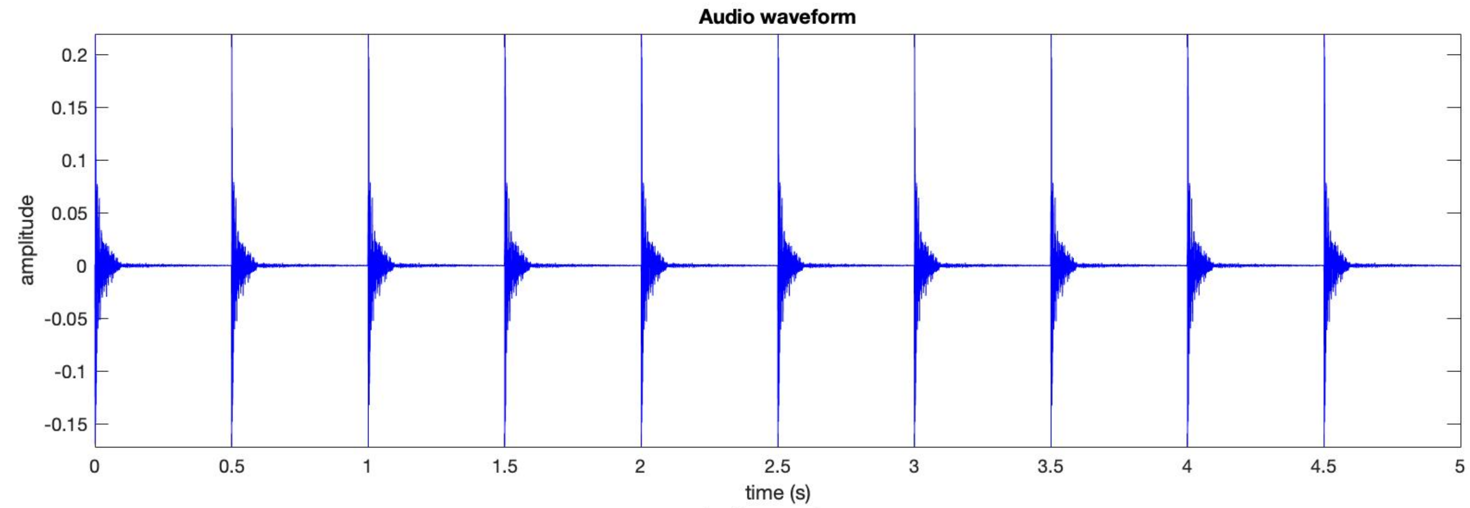


Fourier Transform:

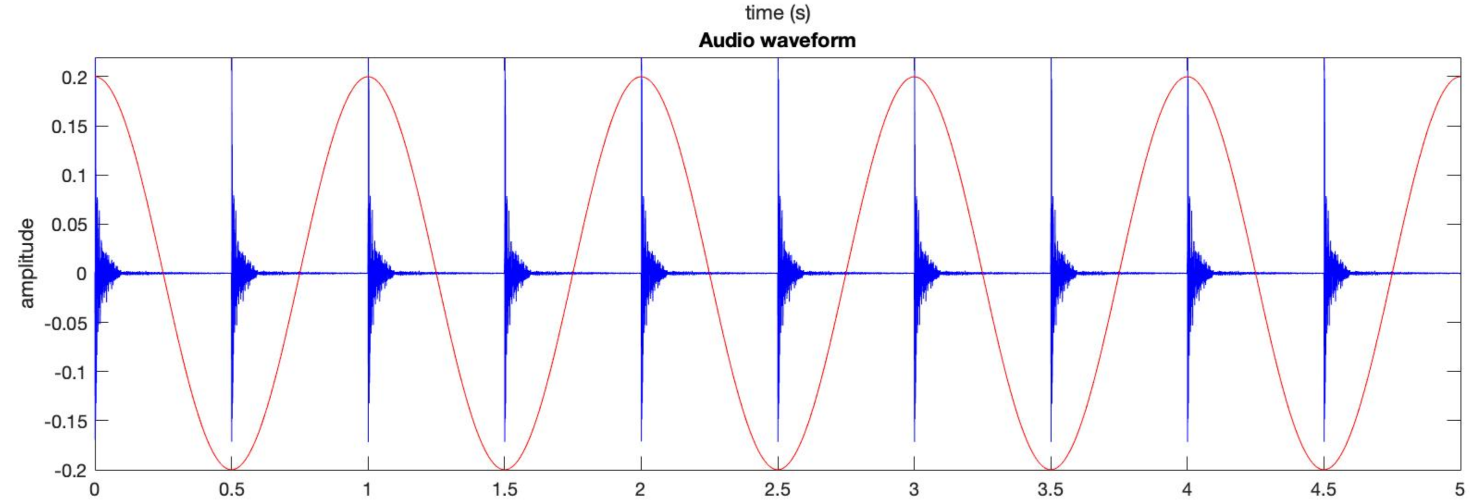


Fourier Transform

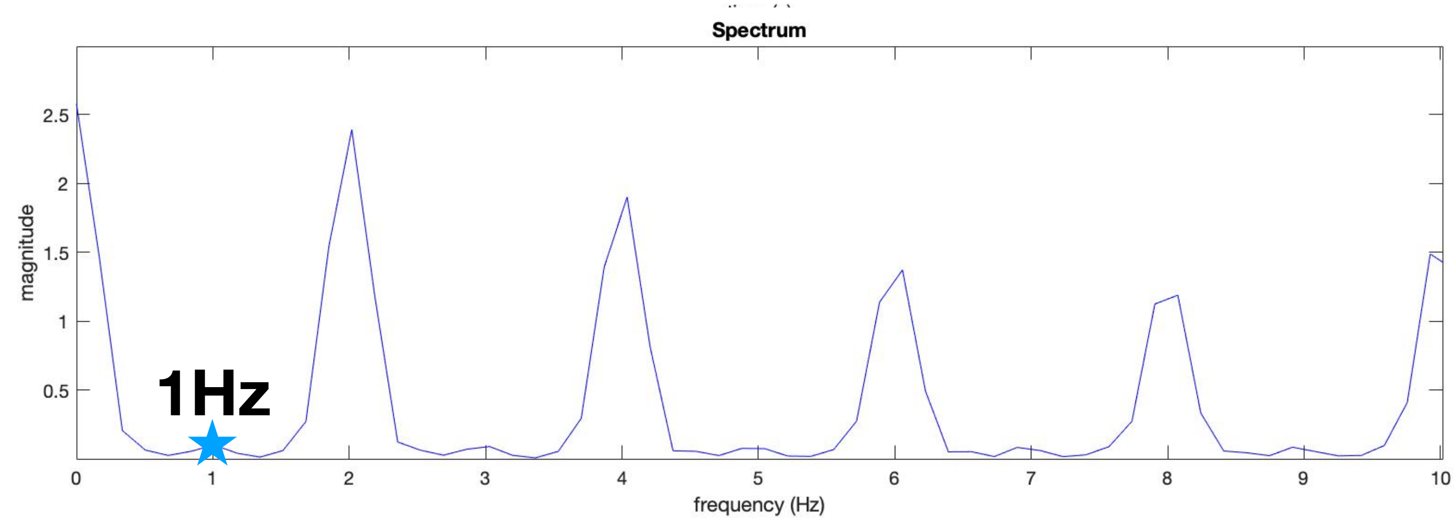
Signal to analyze:



1Hz sinusoid:

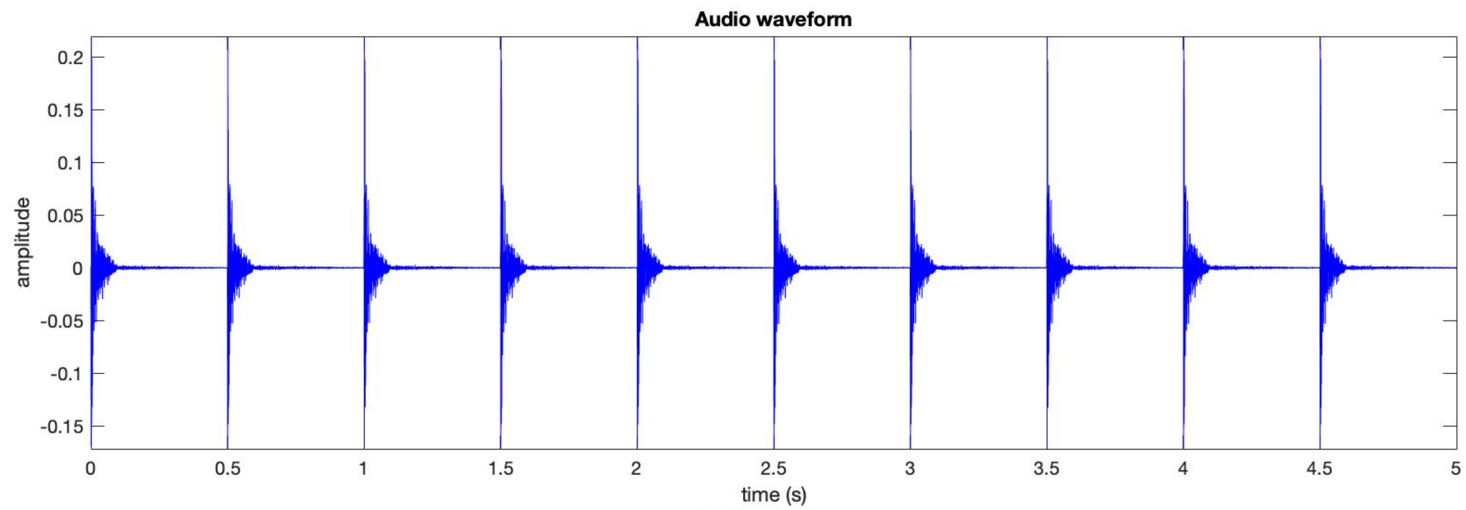


Fourier Transform:

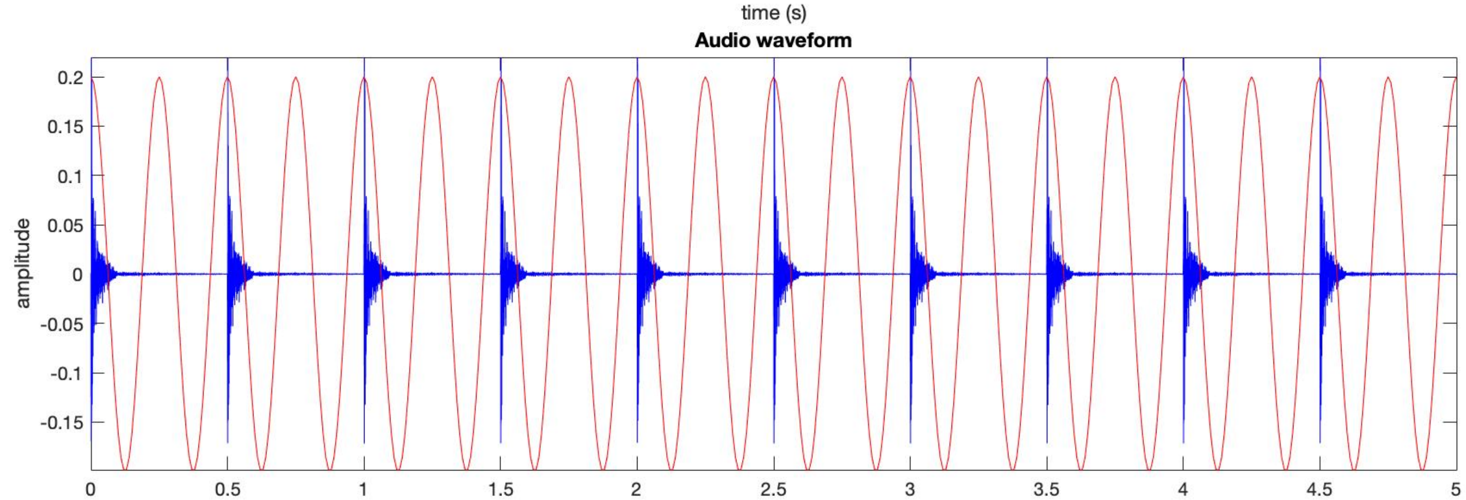


Fourier Transform

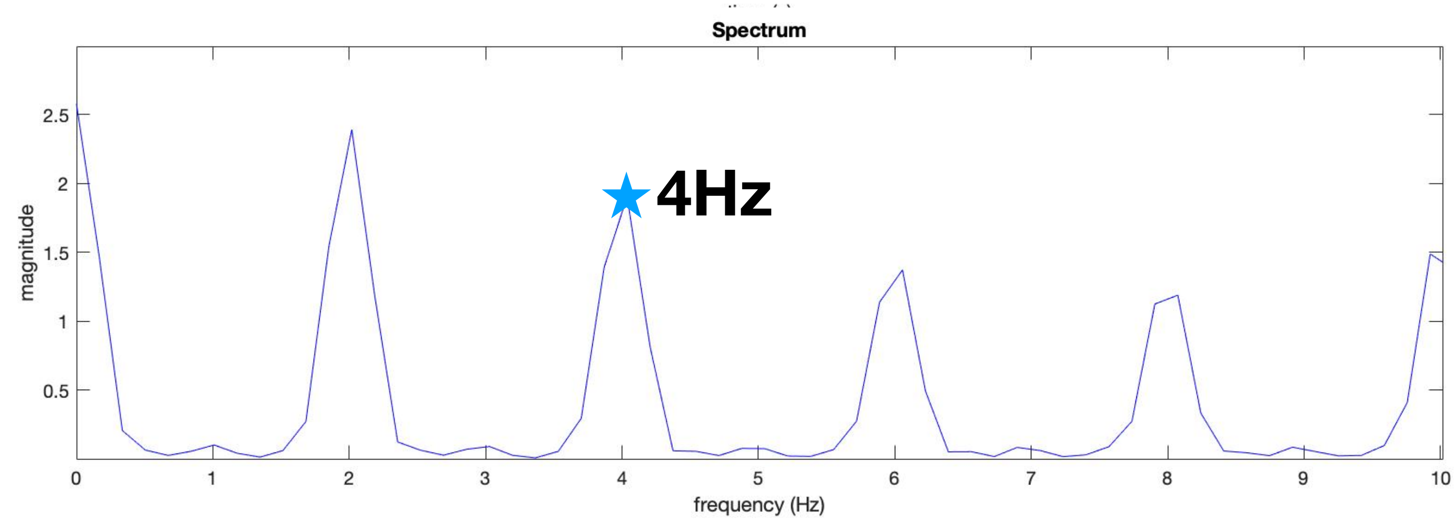
Signal to analyze:



4Hz sinusoid:

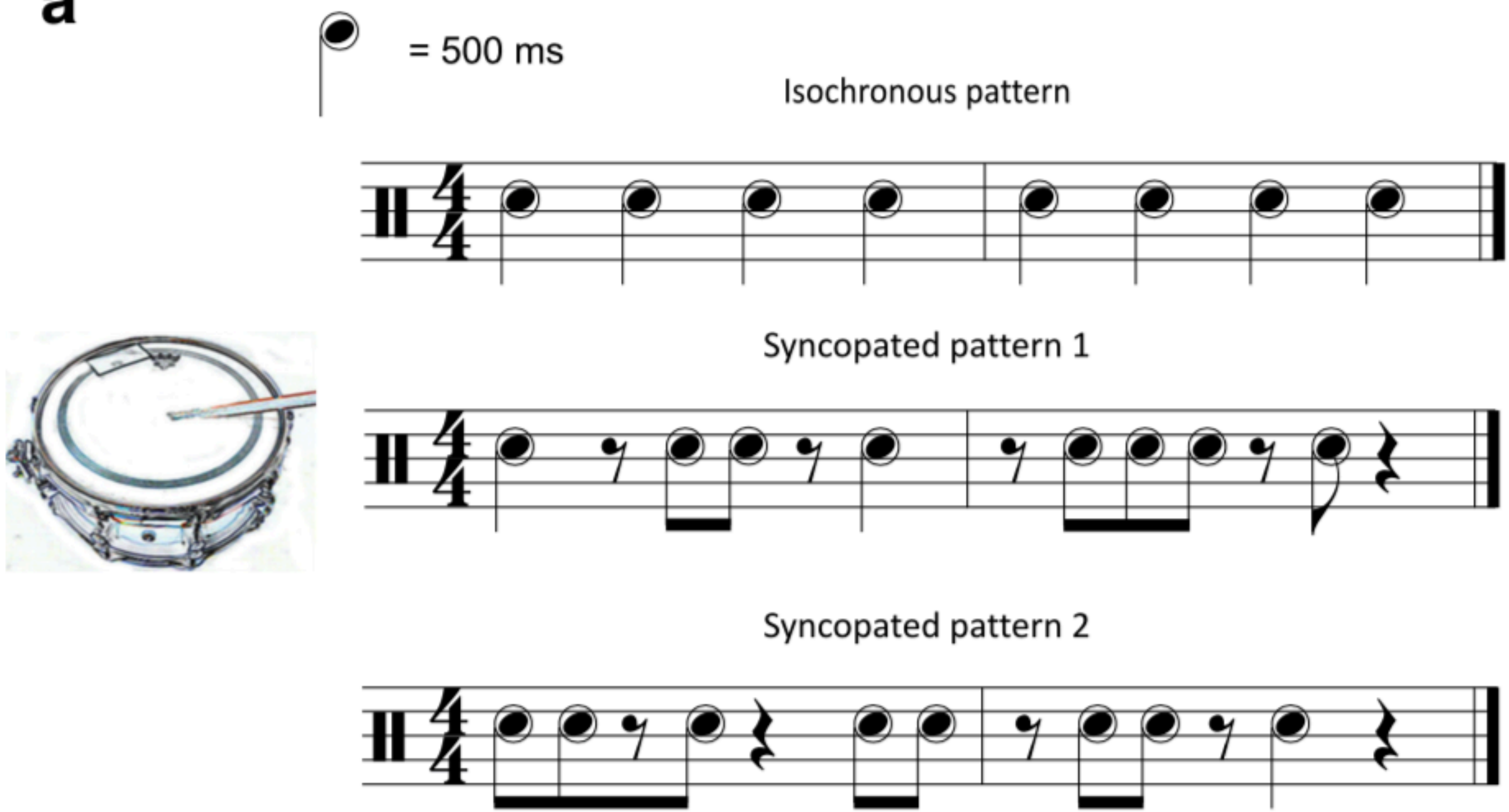


Fourier Transform:

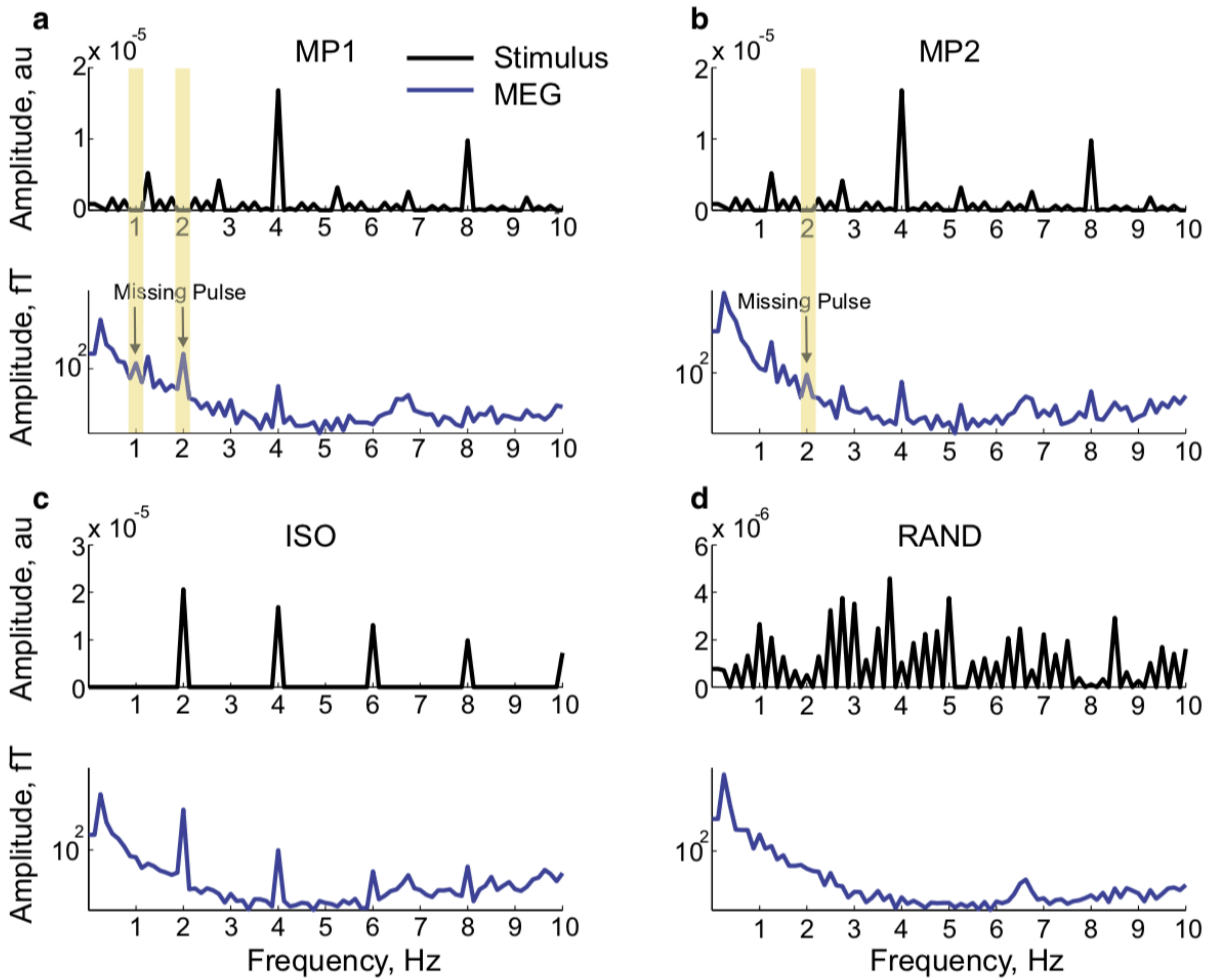


Fourier Transform

a



Fourier Transform



Fourier Transform

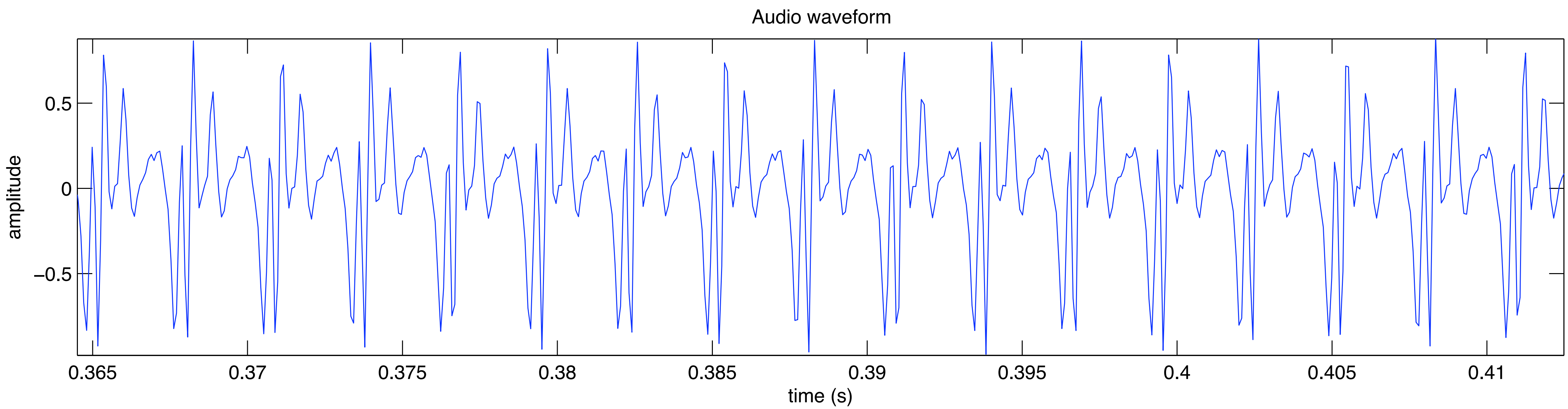
 = 500 ms

Malicious pattern

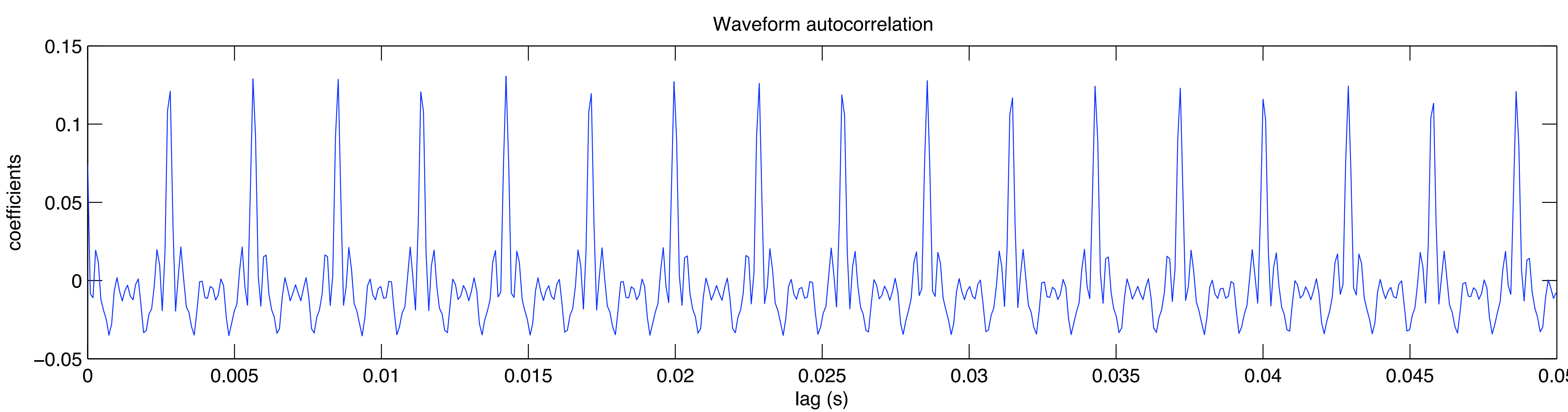
- **Exercice:** What is the Fourier Transformation coefficient corresponding to quarter-note (crochet) duration?

- Zero!

Autocorrelation function

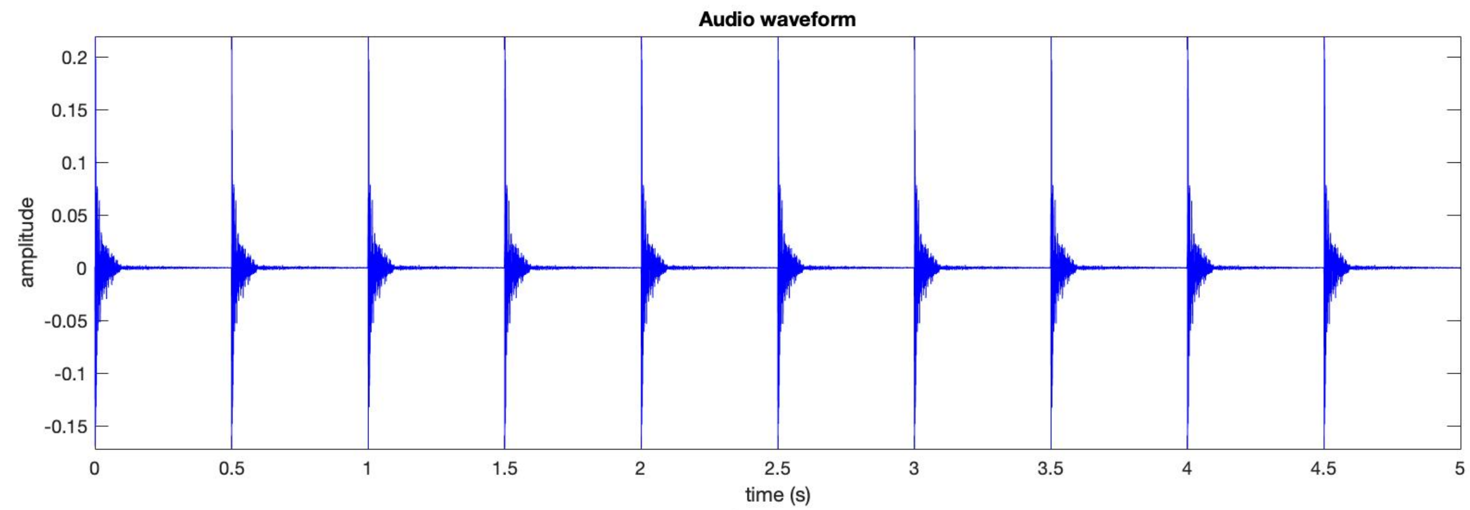


$$R_{xx}(j) = \sum_n x_n \overline{x}_{n-j} \ .$$

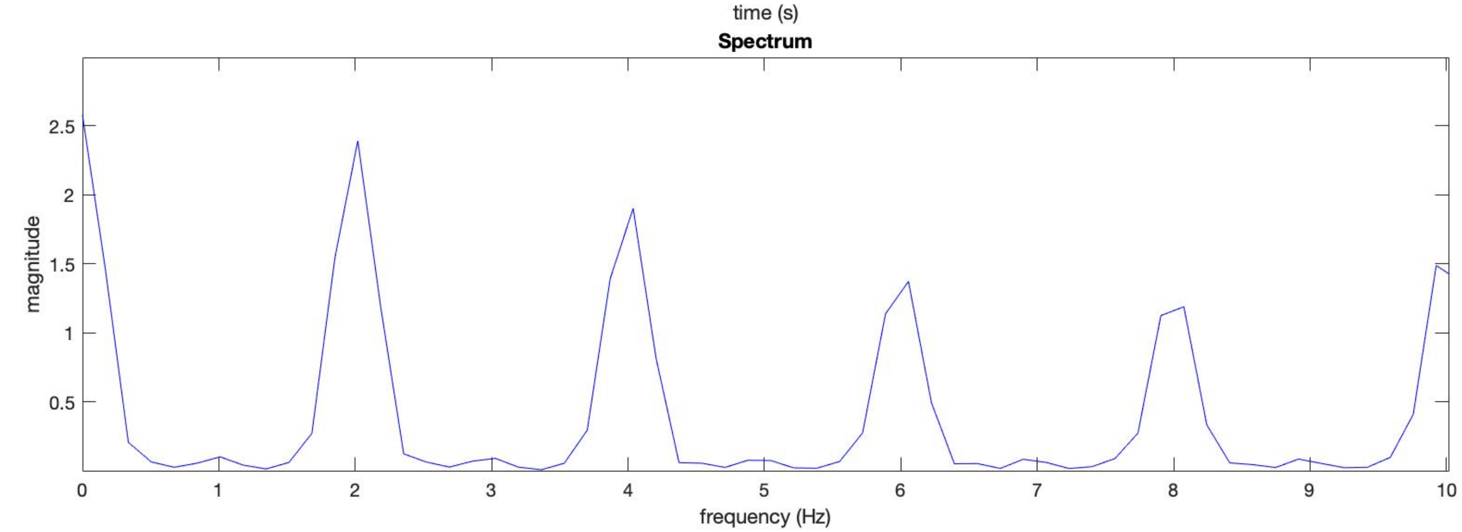


Autocorrelation function

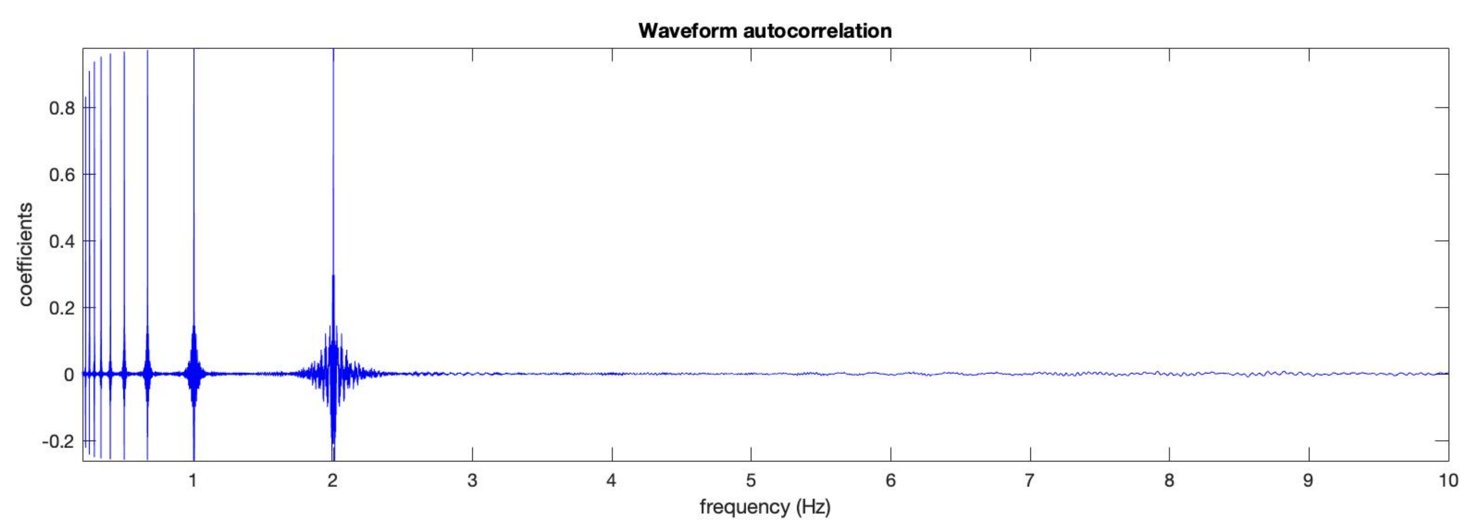
Signal to analyze:

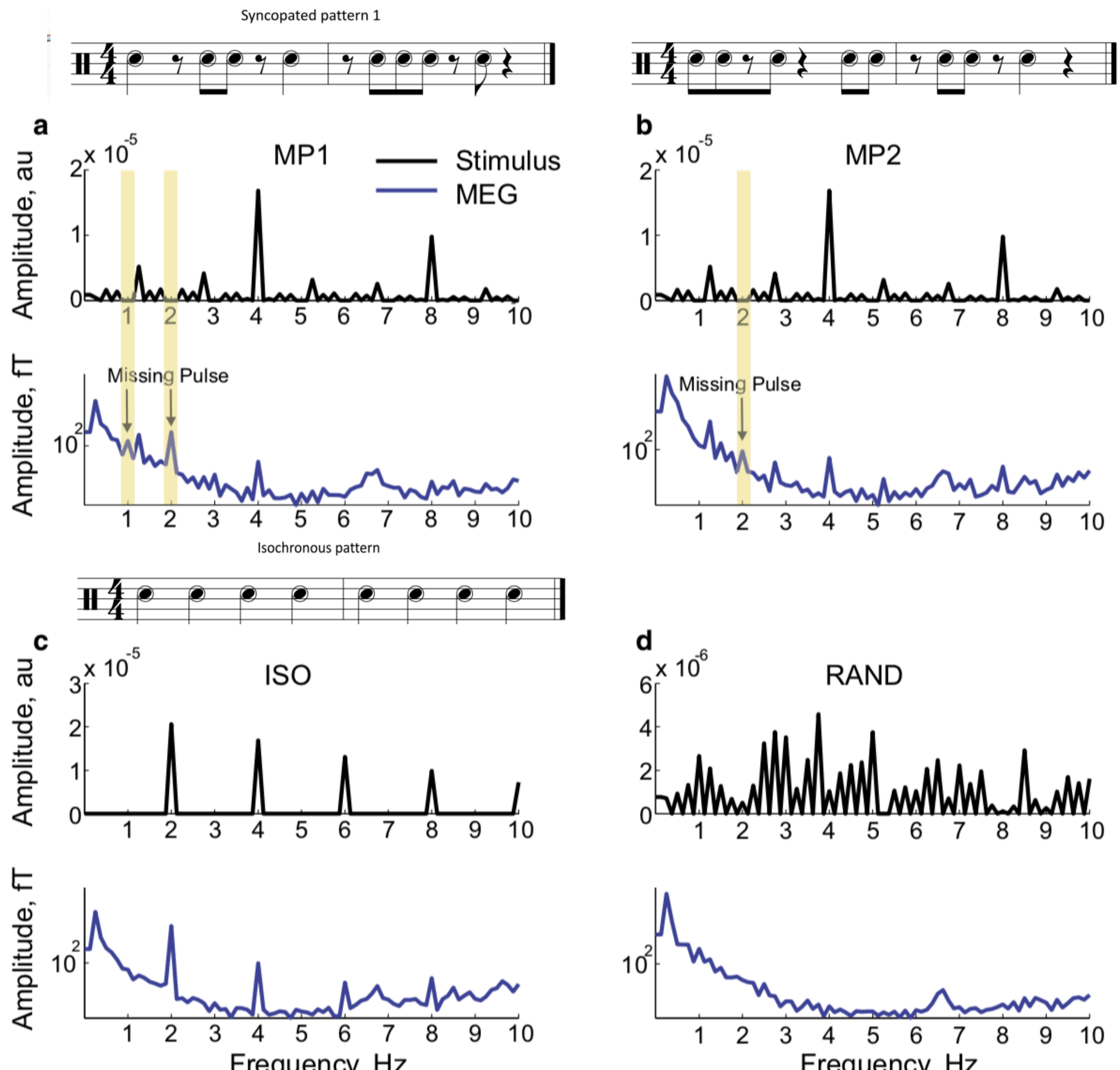


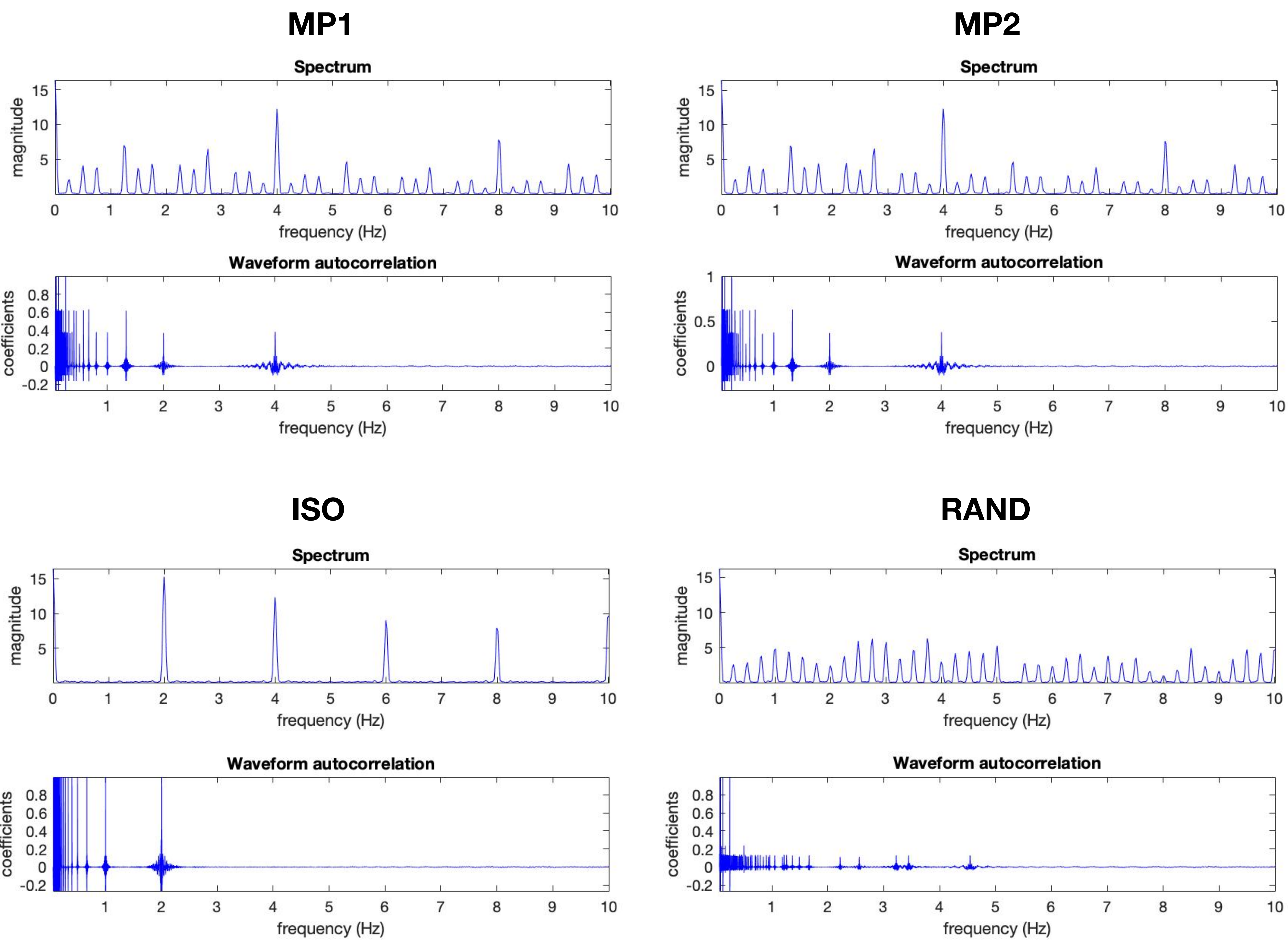
Fourier Transform:



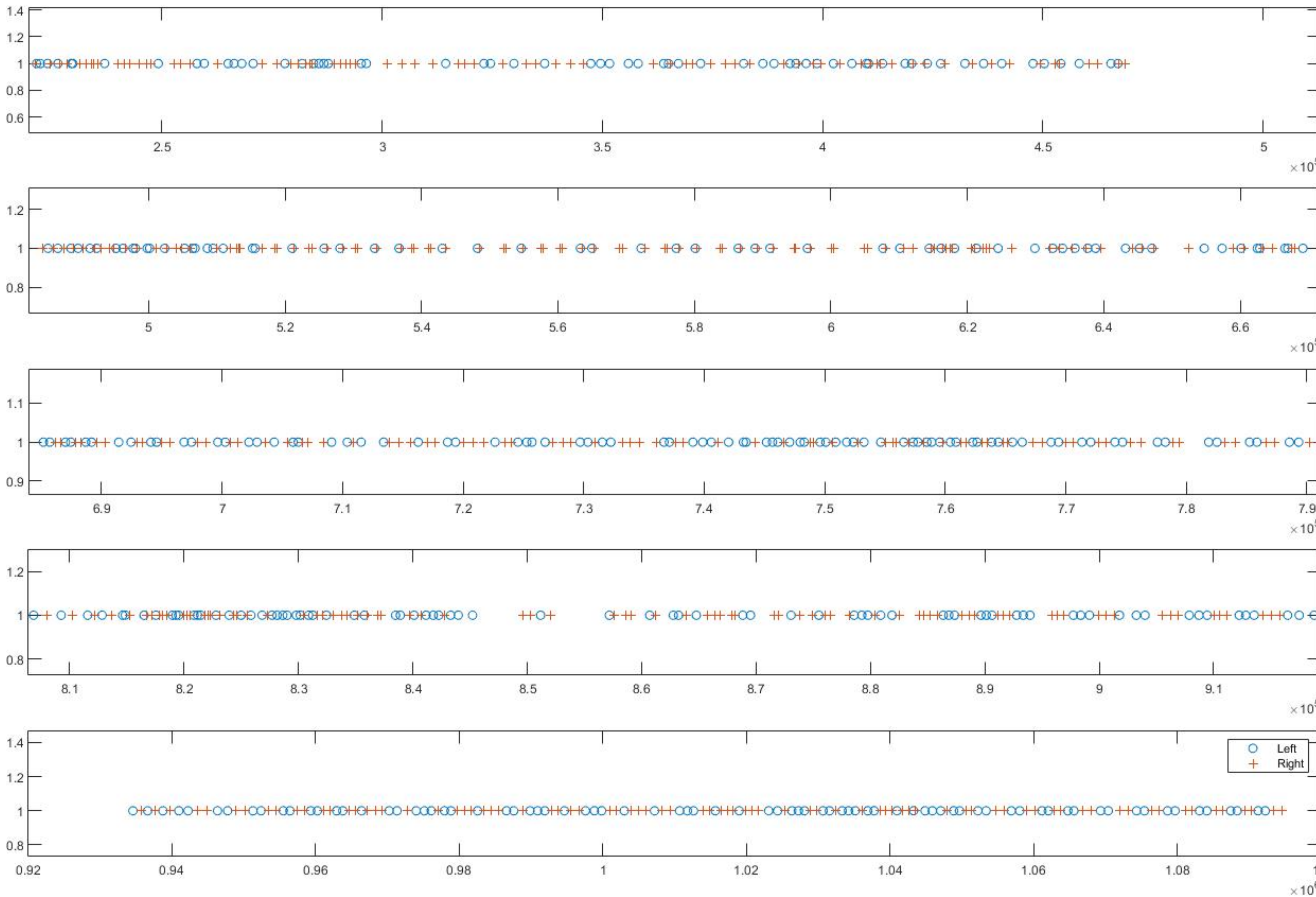
Autocorrelation function:



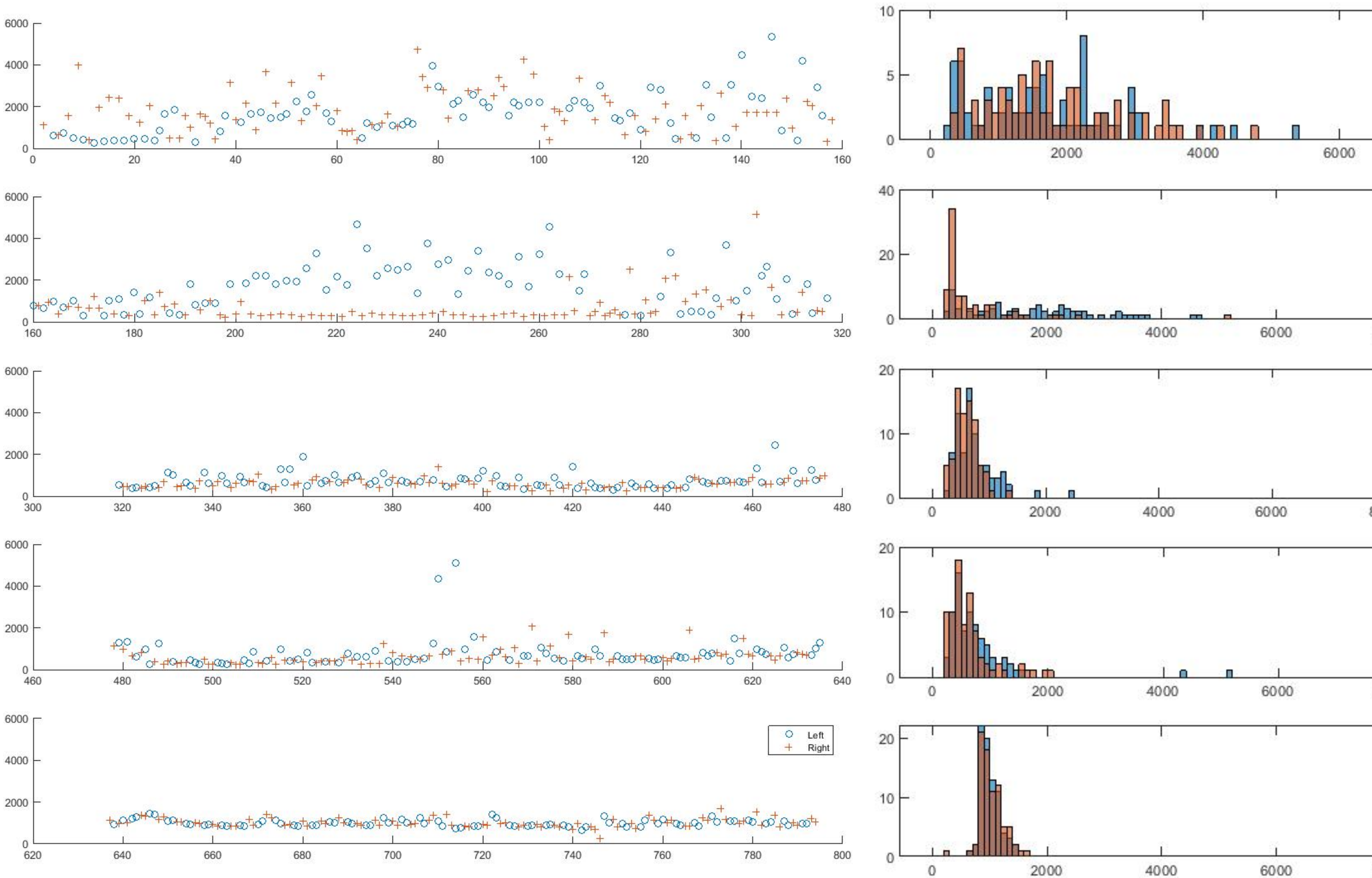

Syncopated pattern 1
Isochronous pattern
a
b
c
d
MP1
MP2
ISO
RAND
Stimulus
MEG
Missing Pulse
Amplitude, au
Amplitude, fT
Frequency, Hz


MP1
Spectrum
magnitude
frequency (Hz)
Waveform autocorrelation
coefficients
frequency (Hz)
MP2
Spectrum
magnitude
frequency (Hz)
Waveform autocorrelation
coefficients
frequency (Hz)
ISO
Spectrum
magnitude
frequency (Hz)
Waveform autocorrelation
coefficients
frequency (Hz)
RAND
Spectrum
magnitude
frequency (Hz)
Waveform autocorrelation
coefficients
frequency (Hz)

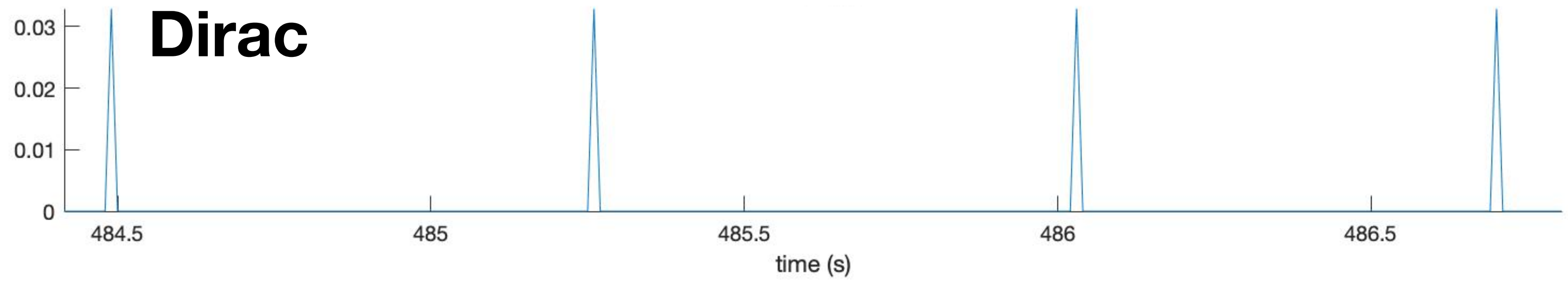
Sabine's data

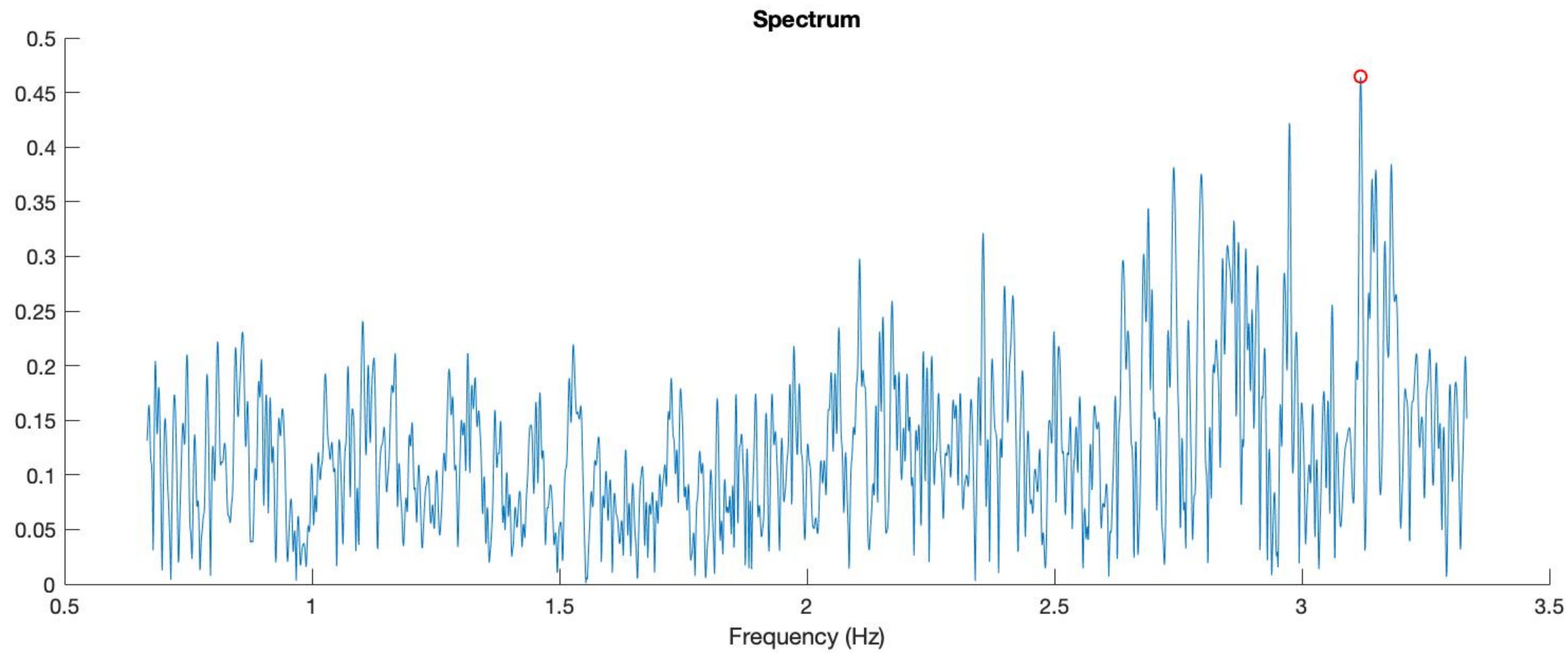


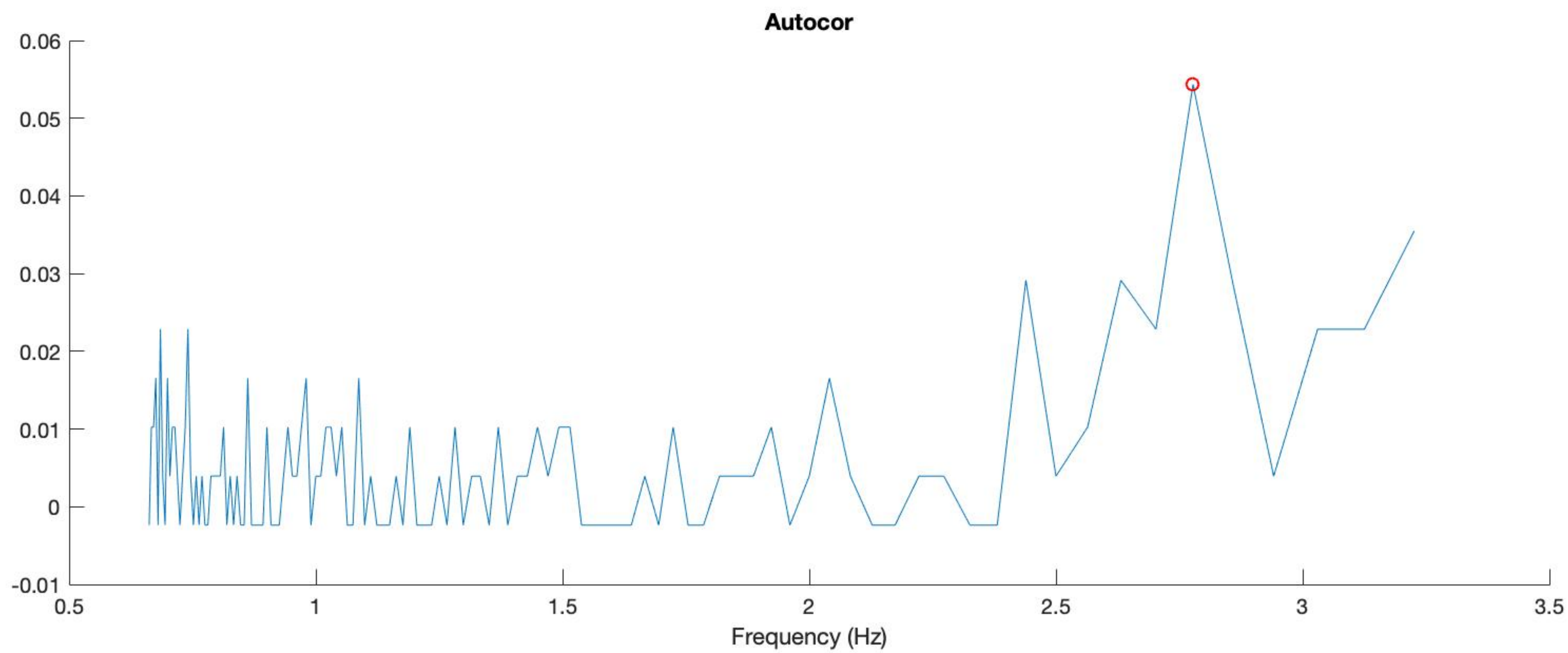
Sabine's data



Dirac **Bloc #2**

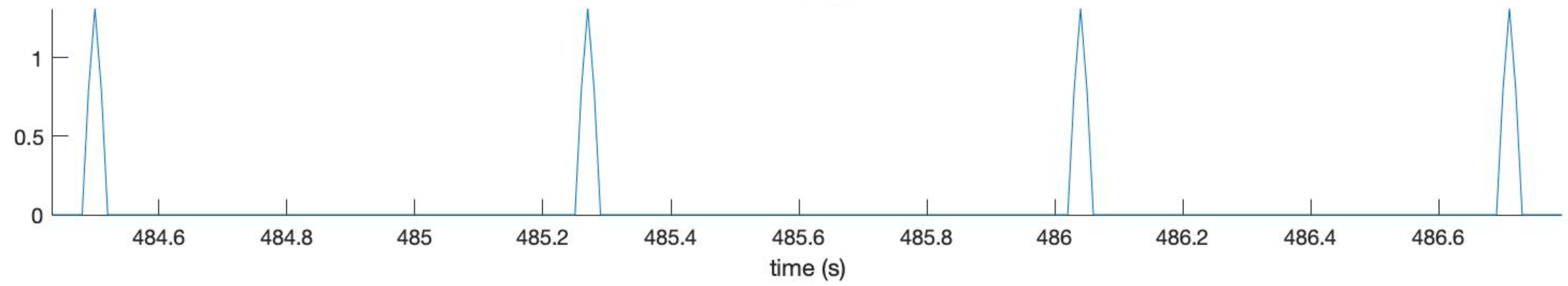

0.03
0.02
0.01
0
484.5
485
485.5
486
486.5
time (s)

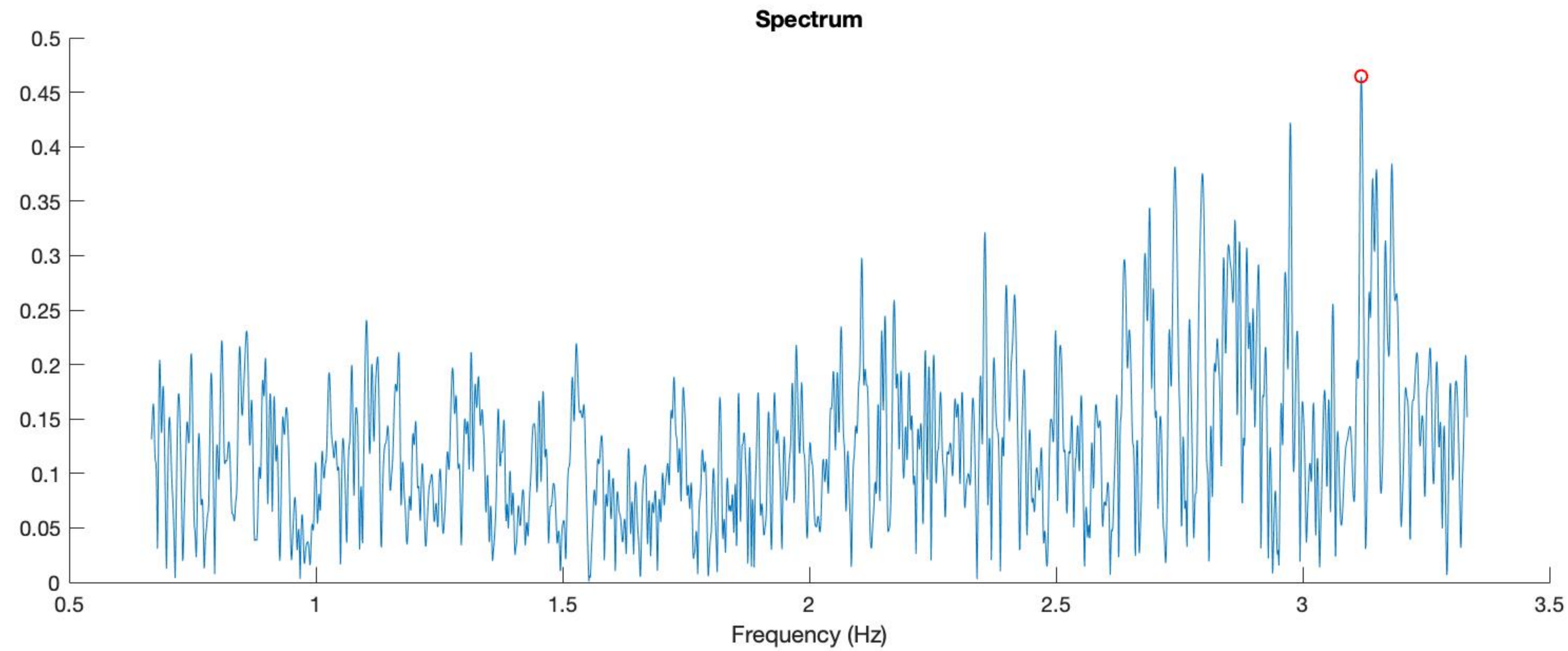

Spectrum
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.5
1
1.5
2
2.5
3
3.5
Frequency (Hz)

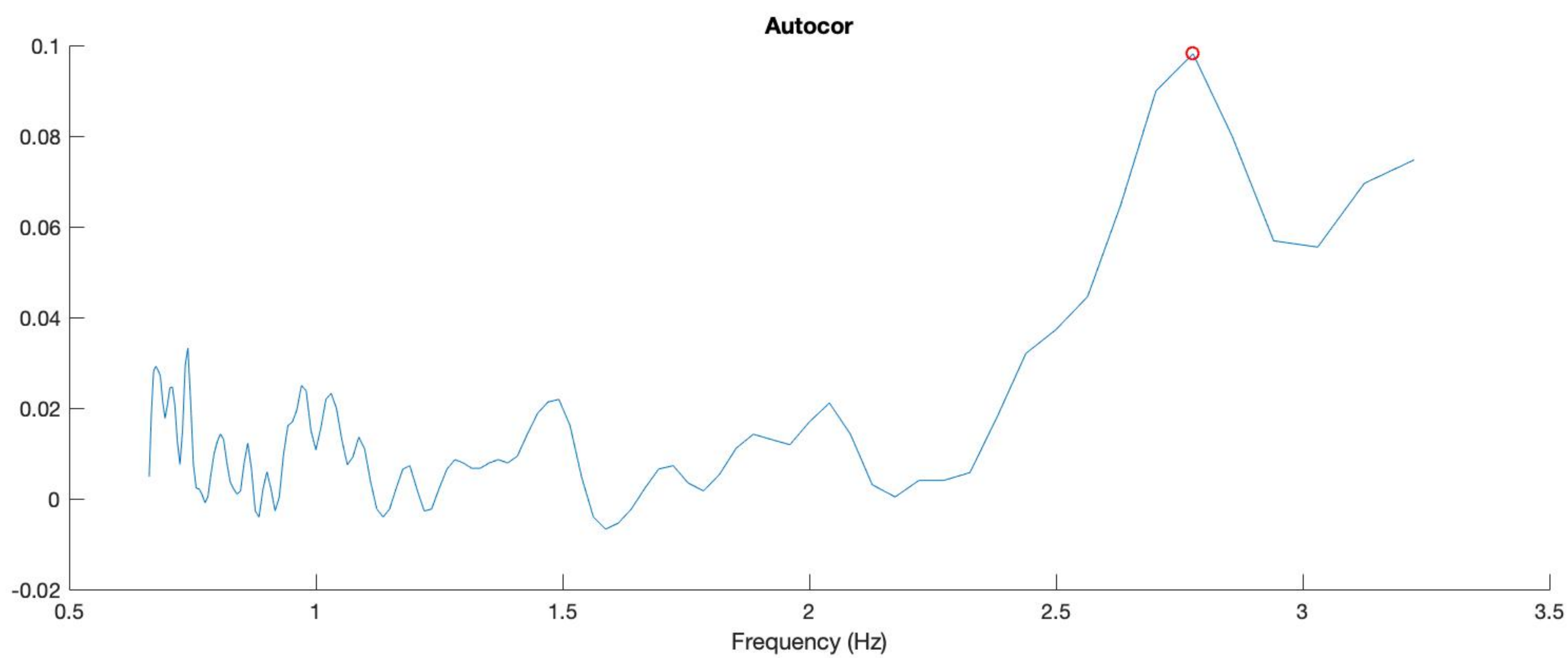

Autocor
0.06
0.05
0.04
0.03
0.02
0.01
0
-0.01
0.5
1
1.5
2
2.5
3
3.5
Frequency (Hz)

Bloc #2

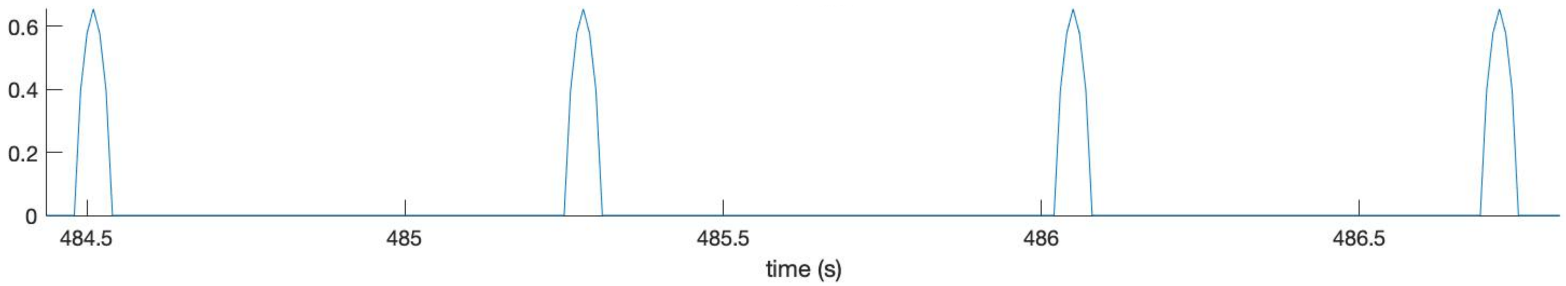
cf. Toiviainen & Snyder, 2003



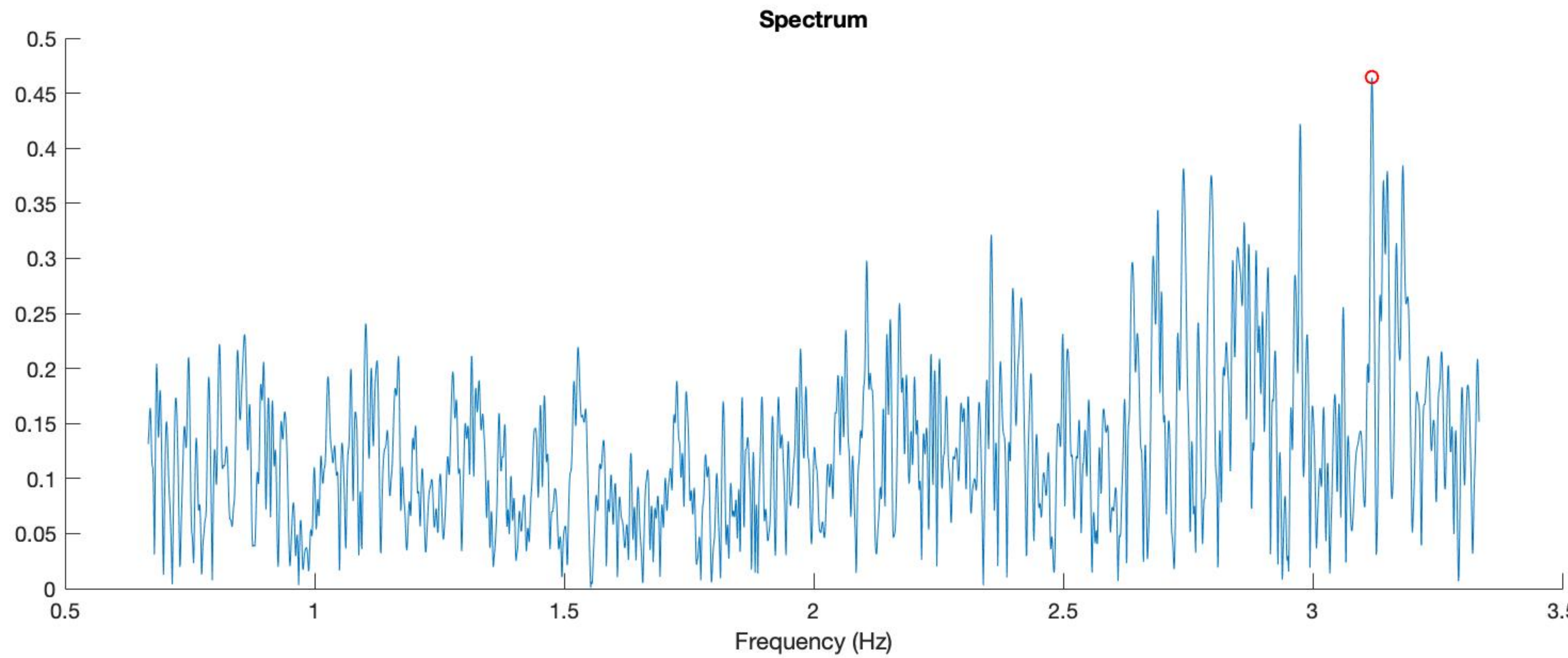


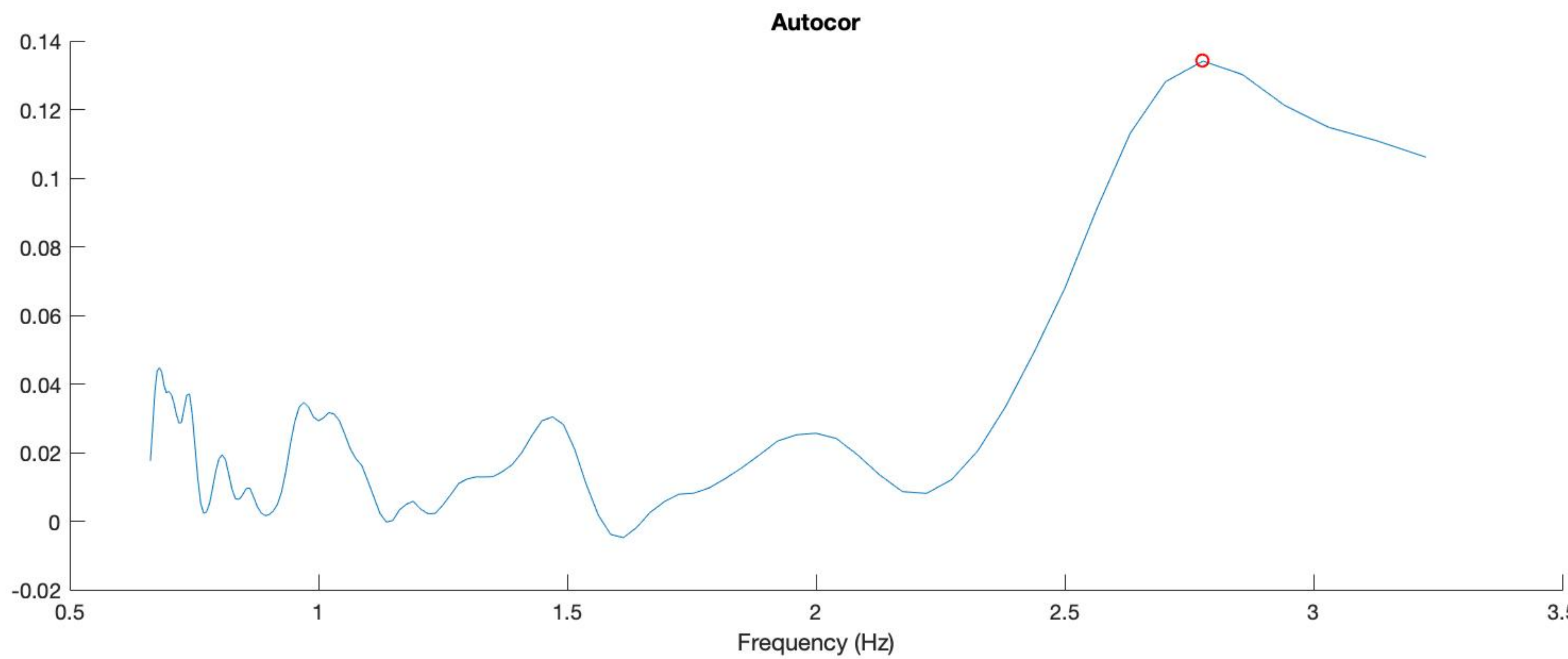


Bloc #2

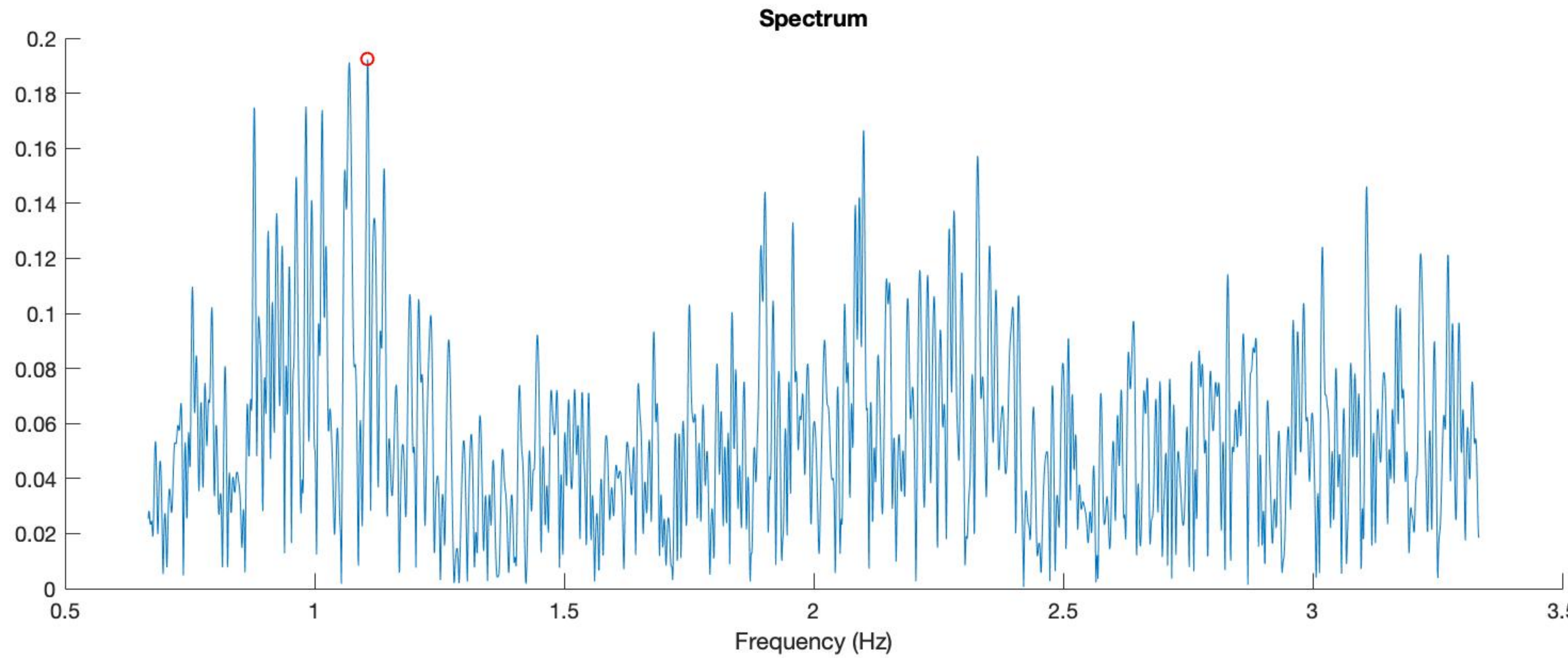


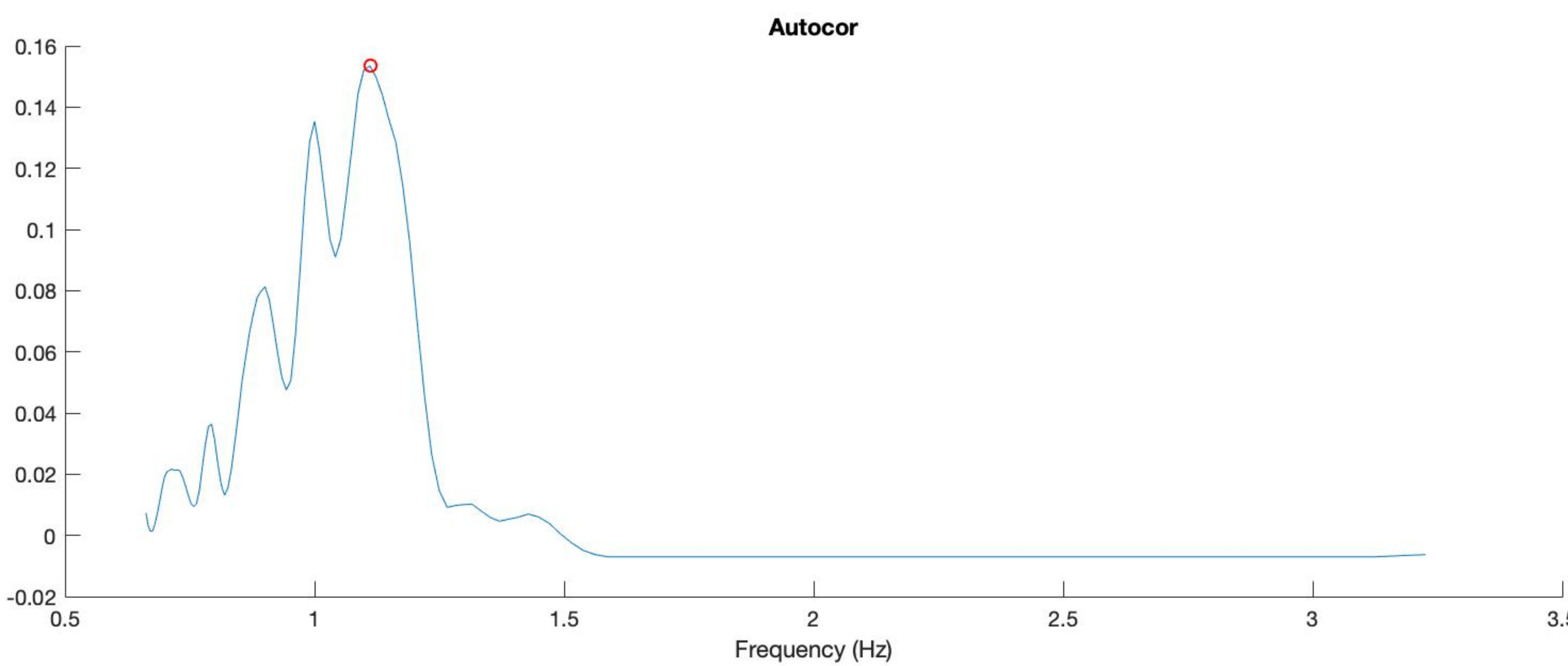
cf. Toiviainen & Snyder, 2003



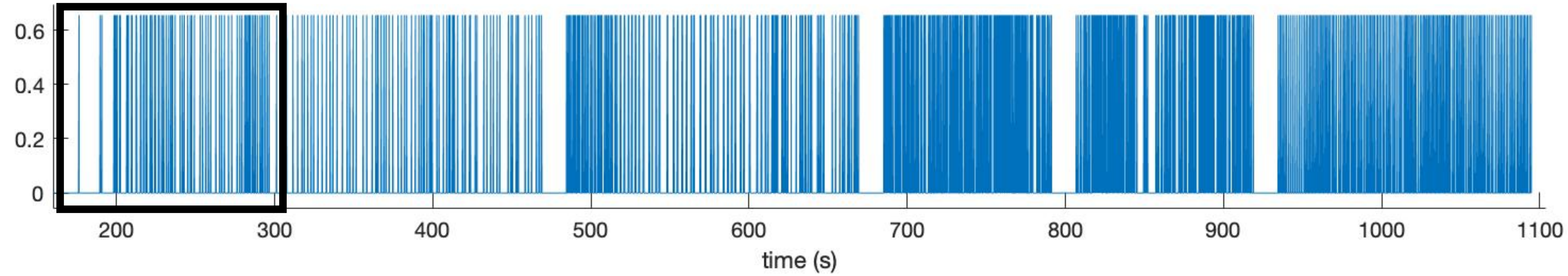


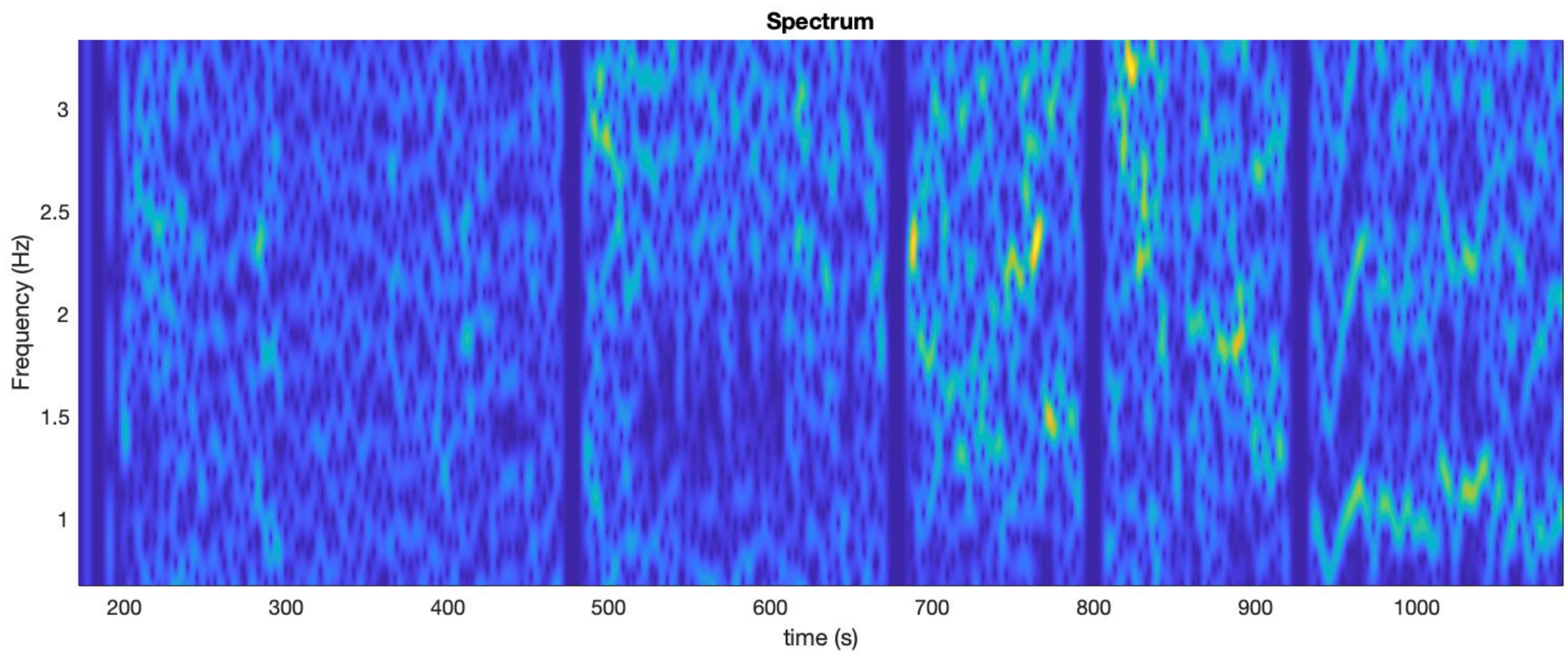
Bloc #5

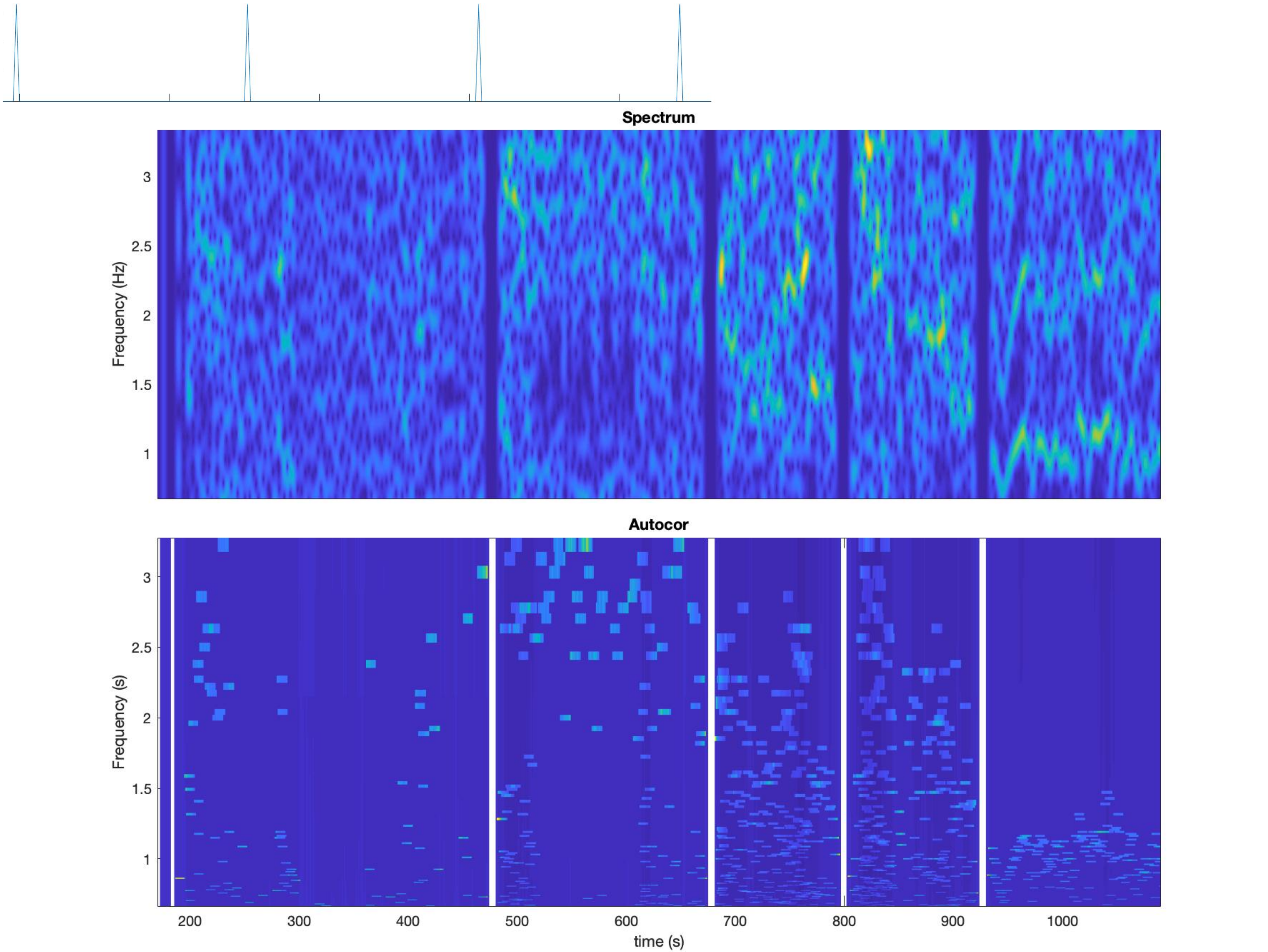

Spectrum
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0.5
1
1.5
2
2.5
3
3.5
Frequency (Hz)

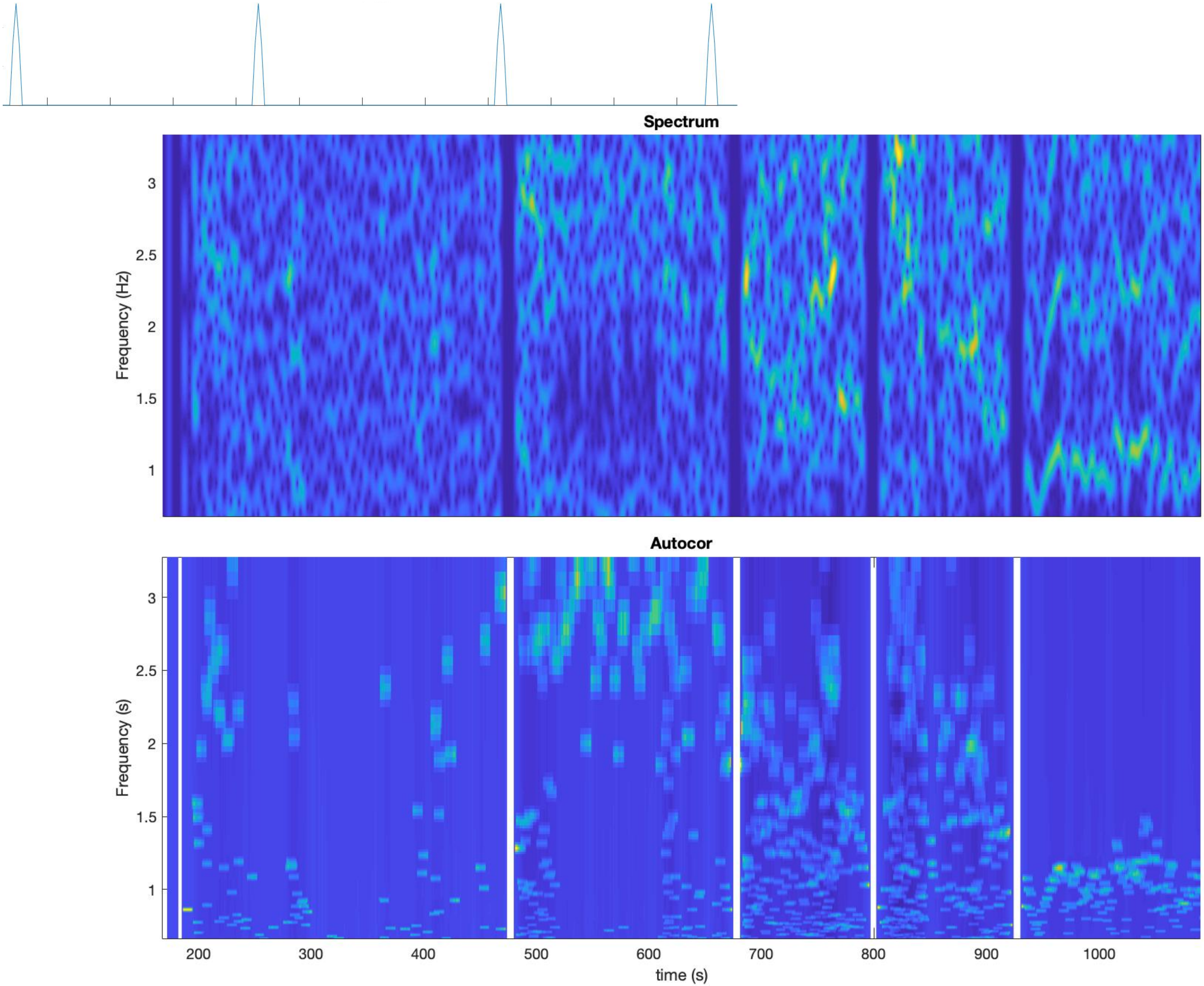

Autocor
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
-0.02
0.5
1
1.5
2
2.5
3
3.5
Frequency (Hz)

Short-Time Fourier Transform

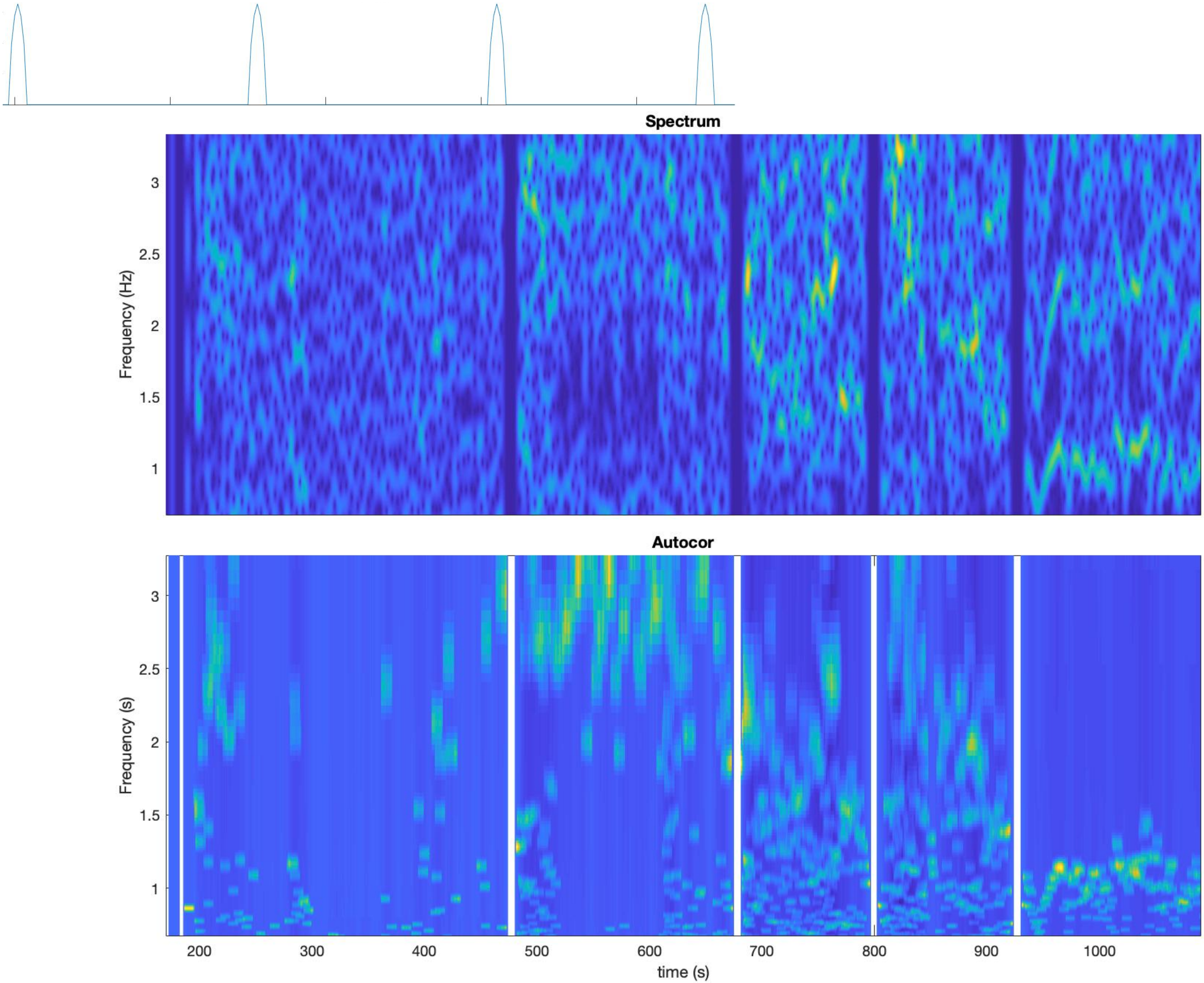





Spectrum
Frequency (Hz)
3
2.5
2
1.5
1
Autocor
Frequency (s)
3
2.5
2
1.5
1
200
300
400
500
600
700
800
900
1000
time (s)

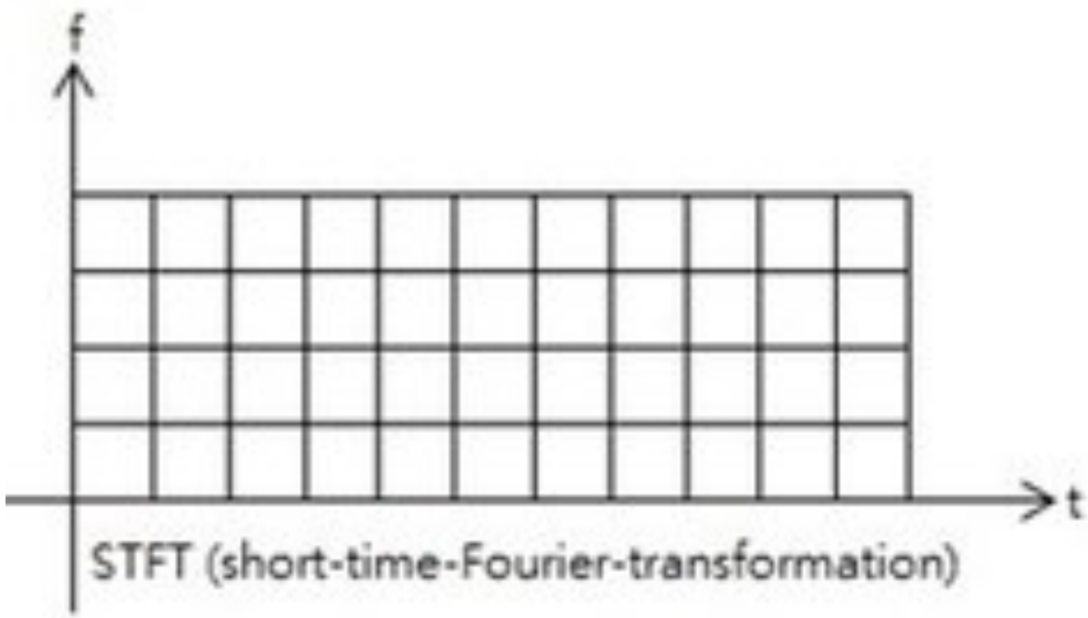


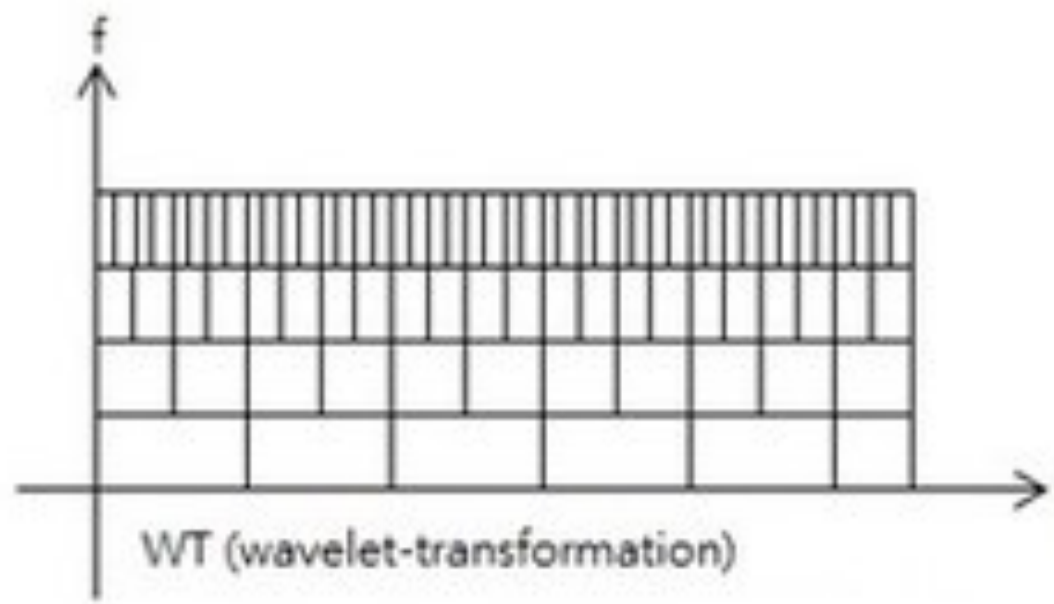
Spectrum
Frequency (Hz)
3
2.5
2
1.5
1
Autocor
Frequency (s)
3
2.5
2
1.5
1
200
300
400
500
600
700
800
900
1000
time (s)



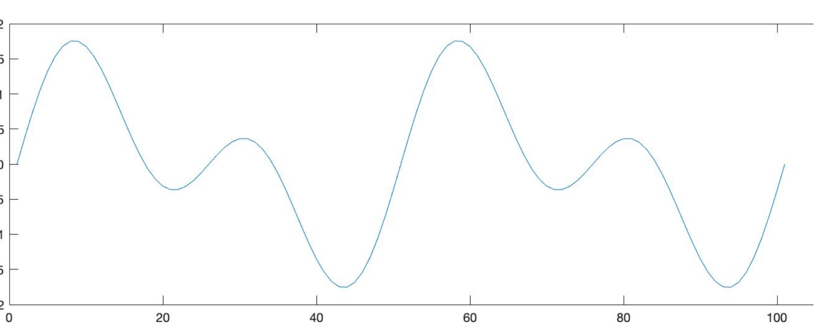
Spectrum
Frequency (Hz)
3
2.5
2
1.5
1
Autocor
Frequency (s)
3
2.5
2
1.5
1
200
300
400
500
600
700
800
900
1000
time (s)

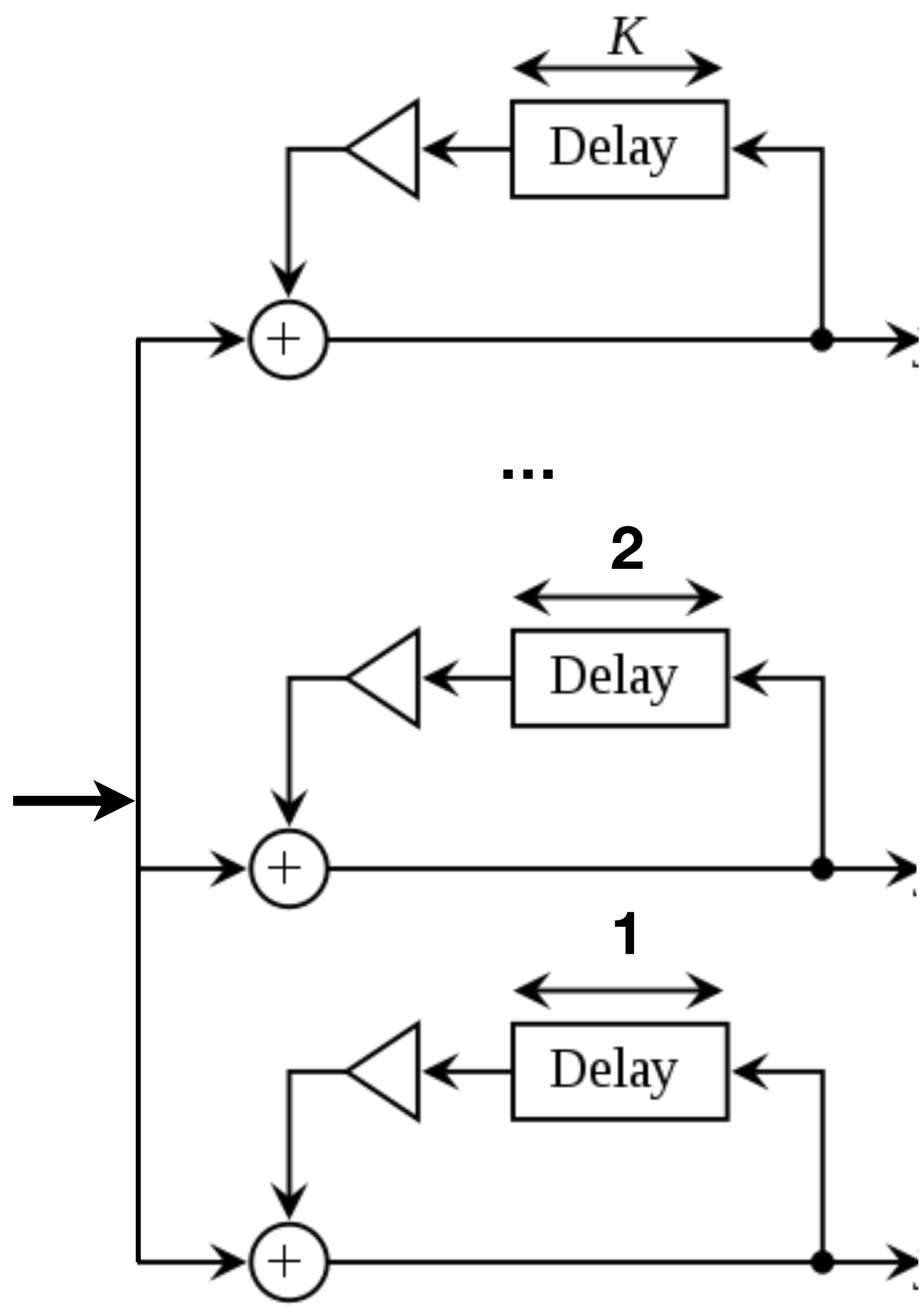
Wavelet


f
t
STFT (short-time-Fourier-transformation)


f
t
WT (wavelet-transformation)

Comb filter resonator





Contestant	SB	HS	EF	FW	GK	**OL**	AK	QH	NW	DP	ES	TL	GP	FK	CD	ZG	AD	SP	MD	DE	AP	PB	GT	CB	ZL	BD
Year 20..	15	18	13	15	11	**13**	06	14	10	06	10	10	12	12	13	11	06	11	14	06	06	06	10	13	18	14
Reference																										
P-score	.90	.88	.86	.83	.83	**.82**	.81	.80	.79	.78	.77	.76	.75	.75	.74	.73	.72	.71	.69	.67	.67	.63	.62	.61	.60	.54
1 tempo	.99	.98	.94	.95	.94	**.92**	.94	.92	.91	.93	.91	.89	.86	.85	.91	.82	.89	.93	.85	.79	.84	.79	.69	.85	.68	.64
both tempi	.69	.66	.69	.57	.62	**.57**	.61	.56	.50	.46	.55	.48	.61	.62	.55	.57	.46	.39	.47	.43	.48	.51	.51	.26	.46	.38

Table 1. Comparison of results from all contestants of MIREX Audio Tempo Extraction from 2006 to 2018. For each author, only the model yielding best P-score is shown. The model presented in this paper is shown in bold.

Lartillot, submission to SMC2019

ROBOD: a Real-time Online Beat and Offbeat Drummer

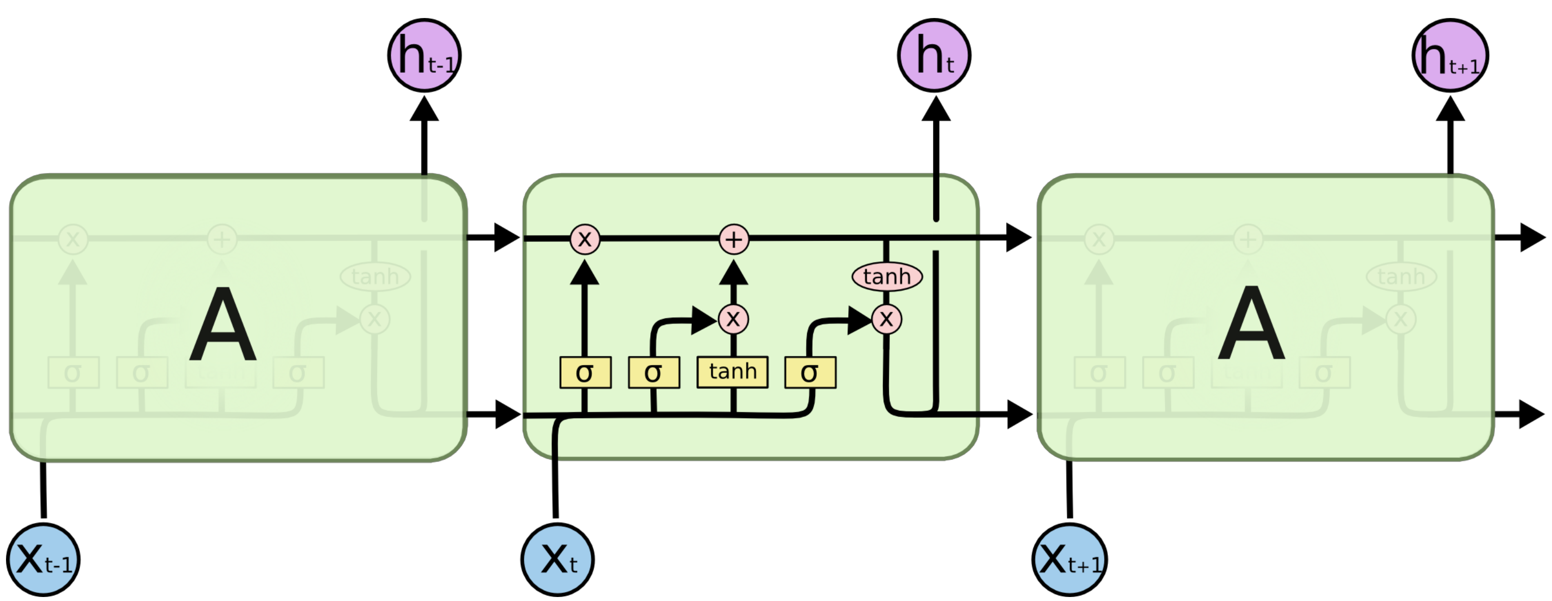
Sebastian Böck [1], *Florian Krebs* [12], *Amaury Durand* [3], *Sebastian Pöll* [1], *Raminta Balsyte* [1]

[1] Department of Computational Perception, Johannes Kepler University Linz, Austria
[2] Joanneum Research, Graz, Austria [3] Télécom ParisTech, Paris, France

Accurate Tempo Estimation based on Recurrent Neural Networks and Resonating Comb Filters

ISMIR 2015

Sebastian Böck, Florian Krebs and Gerhard Widmer
Department of Computational Perception
Johannes Kepler University, Linz, Austria
sebastian.boeck@jku.at



Long short-term memory (LSTM)

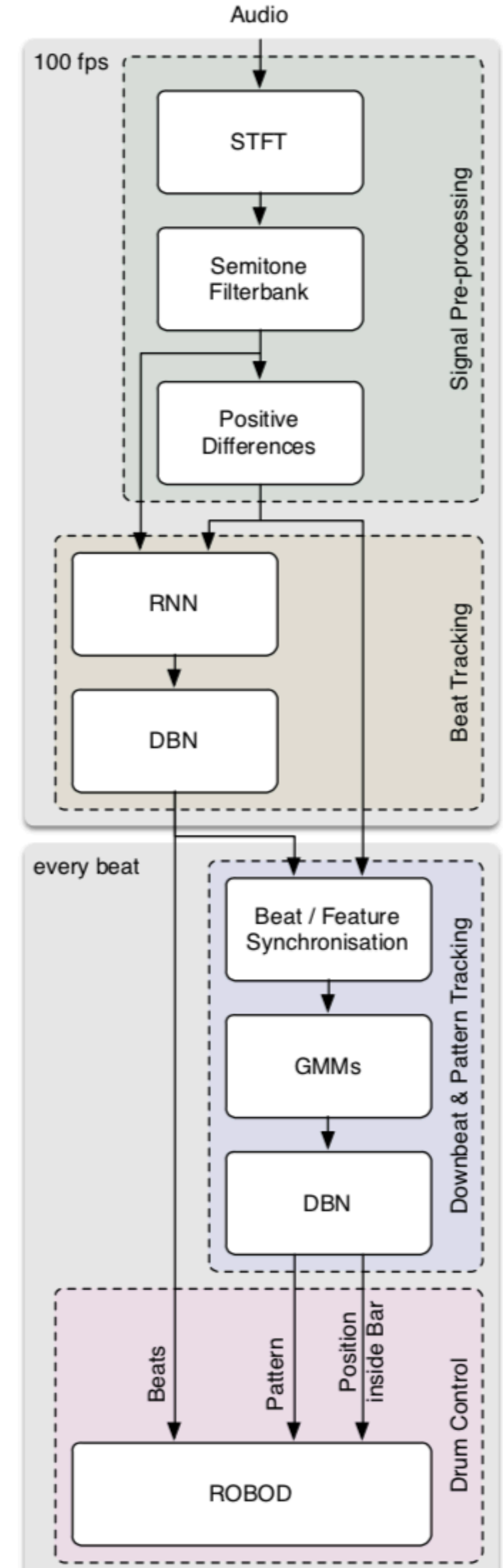


Fig. 1. ROBOD system overview.

MIREX 2012 AUDIO AUDIO TEMPO ESTIMATION EVALUATION: TEMPOKREB

Florian Krebs
Department of Computational Perception
Johannes Kepler University, Linz, Austria

Gerhard Widmer
Department of Computational Perception
Johannes Kepler University, Linz, Austria

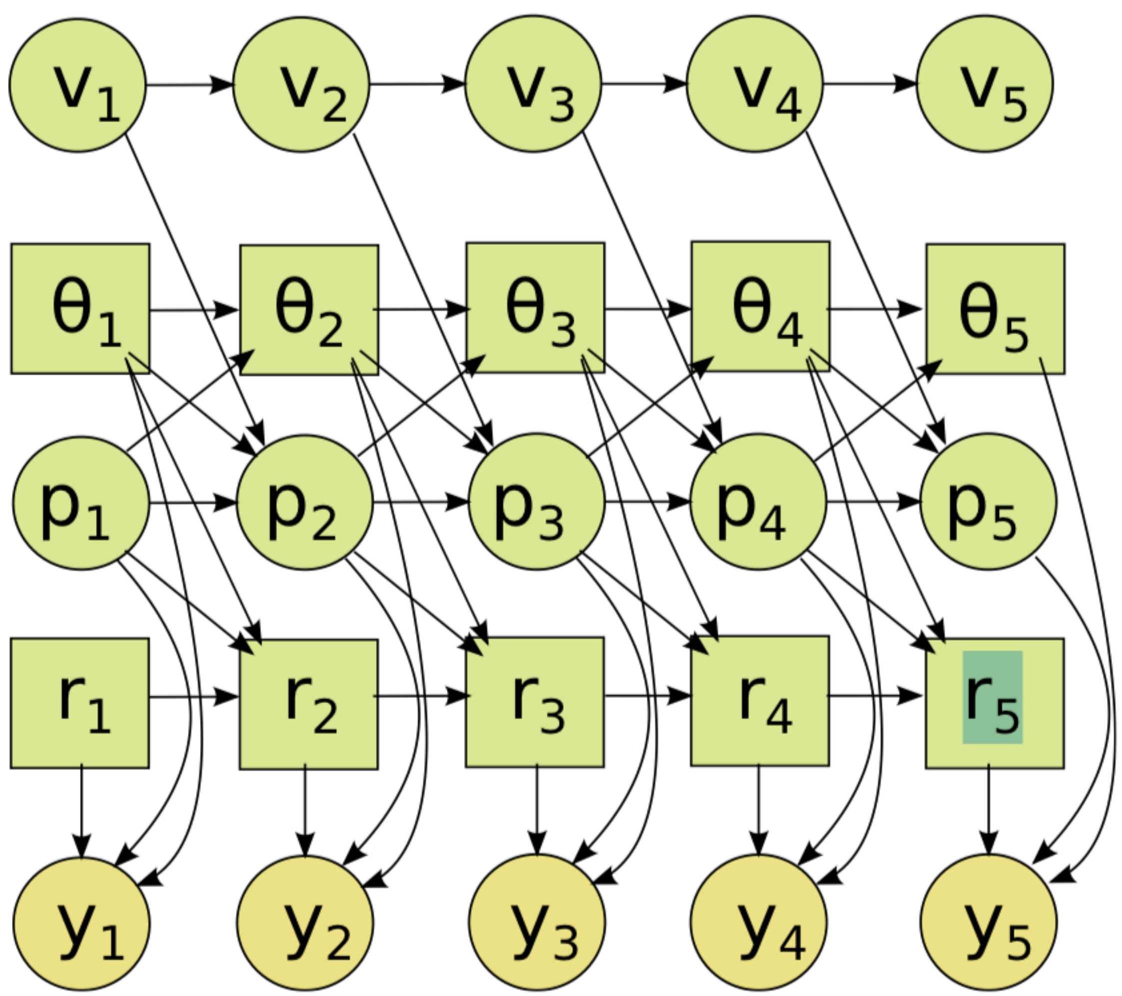


Figure 1. Dynamic Bayesian network

2. MODEL ARCHITECTURE

2.1 Dynamic Bar Pointer Model

Proposed in [7], the dynamic bar pointer model assigns each time instance k of an audio file to the hidden states:

1. current position inside a bar $p_k \in [0, 1)$;
2. current velocity of the bar pointer $v_k \in [v_{min}, v_{max}]$;
3. current meter $\theta_k \in \{\theta_1, \theta_2, ...\theta_{N_\theta}\}$; e.g., $\theta_k \in \{3/4, 4/4\}$, and
4. current rhythmic template $r_k \in \{r_1, r_2, ...r_{N_r}\}$;

Discussion

- Let’s write a reply to that Journal of Neuroscience 2017 article! Are those objections we found already discussed in the community?
- Other methods for periodicity analysis from symbolic representations, used in neuroscience (Alejandro) or in music (IOI histograms).
- Wavelets: concrete examples? Further study on this? But wavelet maybe not important for rhythm analysis due to short bandwidth anyway (Tor).
- Comb filter resonator: implementation for visualisation and listening different bands (Rolf Inge), link with cepstrum (Tor)
- cf. Sethares periodicity transform (Rolf Inge)
- (Machine learning) optimisation of parameters such as filter length (Alexander) or Gaussian width
- How time is modelled in RNN/DBN? (Alexander)
- Presentation about fractals? (Victor?)