UNIVERSITY of ROCHESTER

Music Information Retrieval and EEG-based Brain Decoding

Shadi Sartipi

Fall 2022

Decoding the music-listening brain

Music is a complex auditory stimulus.

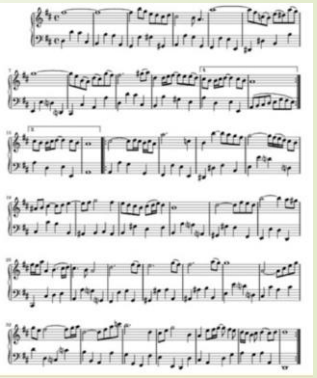
- Beat
- Harmony
- Genre
- Instrumentation and sound texture

How our brain process the music, specially real-world music, is not completely understood.





?



UNIVERSITY of ROCHESTER

Decoding the music-listening brain

Question:

Can we recover music information from brain responses (brain decoding)?

- Neuroscience approach explore brain responses
- Engineering approach for real-world applications



https://www.freethink.com/health/digital-therapy

Approaches:

- Signal-based analysis
- Machine learning/Deep learning

UNIVERSITY of ROCHESTER

Measuring Neural Activity with Electroencephalography (EEG)

- Records electrical activity of the brain.

Advantages:
- Ease of use
- Non-invasive
- High temporal resolution
- Inexpensive

Disadvantages:
- Low spatial resolution
- Low SNR (noisy data)

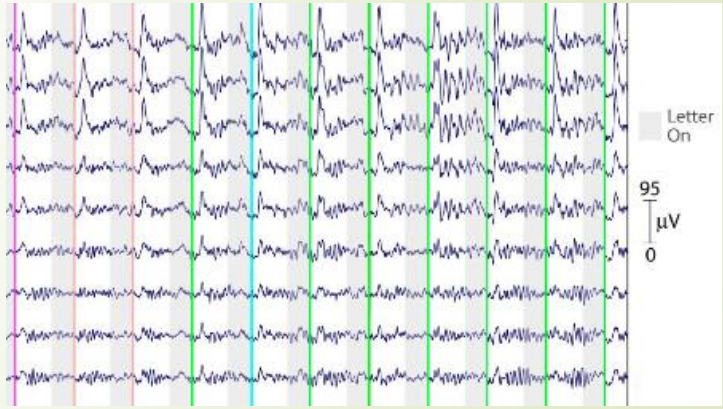


UNIVERSITY of ROCHESTER

Electroencephalography (EEG)

- Each EEG electrode produces a rime-series data
- Most of the information relies in the following frequency sub-bands
    - Delta (1-4 Hz)
    - Theta (4-8 Hz)
    - Alpha (8-12 Hz)
    - Beta (12-29 Hz)
    - Gamma (29-50 Hz)





- There is a 10-100 msec latency in stimulus-to response
- Data Can be analyzed in time and frequency domains

UNIVERSITY of ROCHESTER

Tempo Detection

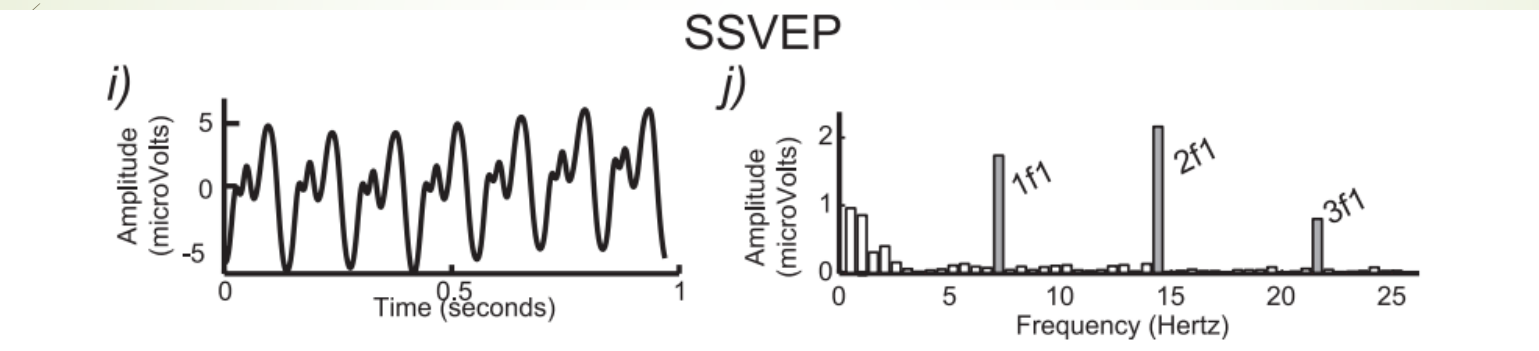
Losorelli, Steven, et al. "NMED-T: A Tempo-Focused Dataset of Cortical and Behavioral Responses to Naturalistic Music." *ISMIR*. Vol. 3. 2017.

EEG spectra reflects stimulus periodicities

Steady-State Paradigm:

- Stimulus periodicity can be seen in EEG response
- Periodicities in stimulus frequency and harmonics

Example:



Time domain with a stimulus frequency of 7.2 Hz

Response spectrum contains multiple harmonics

Norcia, Anthony M., et al. "The steady-state visual evoked potential in vision research: A review." *Journal of vision* 15.6 (2015): 4-4.

UNIVERSITY of ROCHESTER

Steady –state responses to music

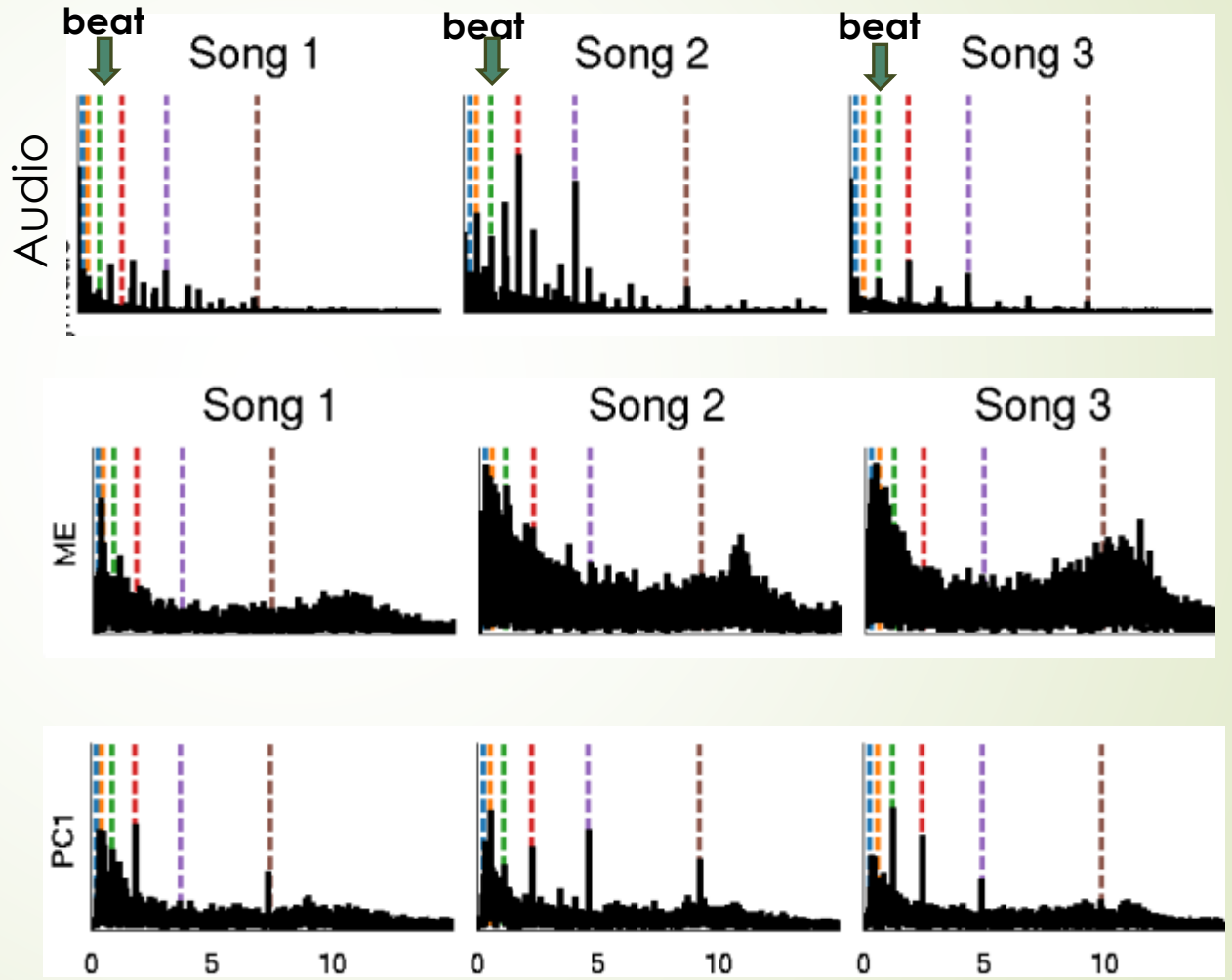
- Electronically musical beat can be considered a periodic stimulus attribute.
- Spectral peaks are observed in EEG responses to simple rhythms.
- Can we apply it to real-world music?
- Application: Tempo Detection

UNIVERSITY of ROCHESTER

NMED-T: A Tempo-Focused Dataset

FFT APPROACH

- Audio envelope contains spectral peaks at beat-related and other frequencies.
- EEG averaged across all electrodes: like noise floor
- EEG with spatial filter with PCA; Clear spectral peaks

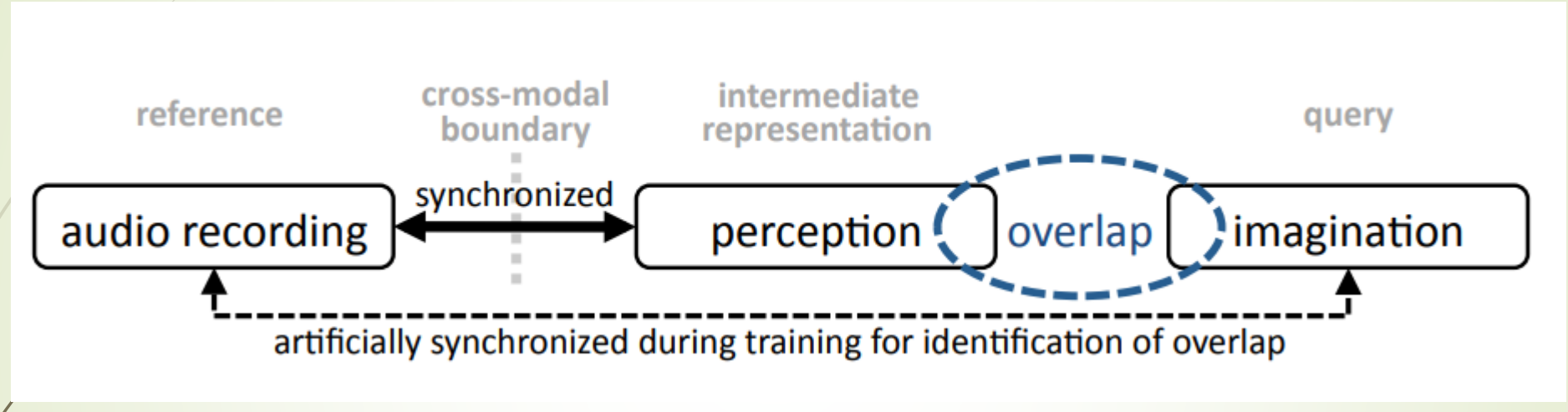


UNIVERSITY of ROCHESTER

An Open Dataset
Of Music Perception and Imagination Recordings

Stober, Sebastian, et al. "Towards Music Imagery Information Retrieval: Introducing the OpenMIIR Dataset of EEG Recordings from Music Perception and Imagination." *ISMIR*. 2015.

General Approach



Primary Goal:

- recognize individual pieces

Future Goal:

- Audio re-synthesis from EEG
- Cross-domain sync of EEG and audio

UNIVERSITY of ROCHESTER

Stimuli Selection

Table 1. Information about the tempo, meter and length of the stimuli (without cue clicks) used in this study.

ID	Name	Meter	Length	Tempo
1	Chim Chim Cheree (lyrics)	3/4	13.3s	212 BPM
2	Take Me Out to the Ballgame (lyrics)	3/4	7.7s	189 BPM
3	Jingle Bells (lyrics)	4/4	9.7s	200 BPM
4	Mary Had a Little Lamb (lyrics)	4/4	11.6s	160 BPM
11	Chim Chim Cheree	3/4	13.5s	212 BPM
12	Take Me Out to the Ballgame	3/4	7.7s	189 BPM
13	Jingle Bells	4/4	9.0s	200 BPM
14	Mary Had a Little Lamb	4/4	12.2s	160 BPM
21	Emperor Waltz	3/4	8.3s	178 BPM
22	Hedwig's Theme (Harry Potter)	3/4	16.0s	166 BPM
23	Imperial March (Star Wars Theme)	4/4	9.2s	104 BPM
24	Eine Kleine Nachtmusik	4/4	6.9s	140 BPM
	mean		10.4s	176 BPM

Different Classification problems can be defined…

UNIVERSITY of ROCHESTER

Experiment Setup

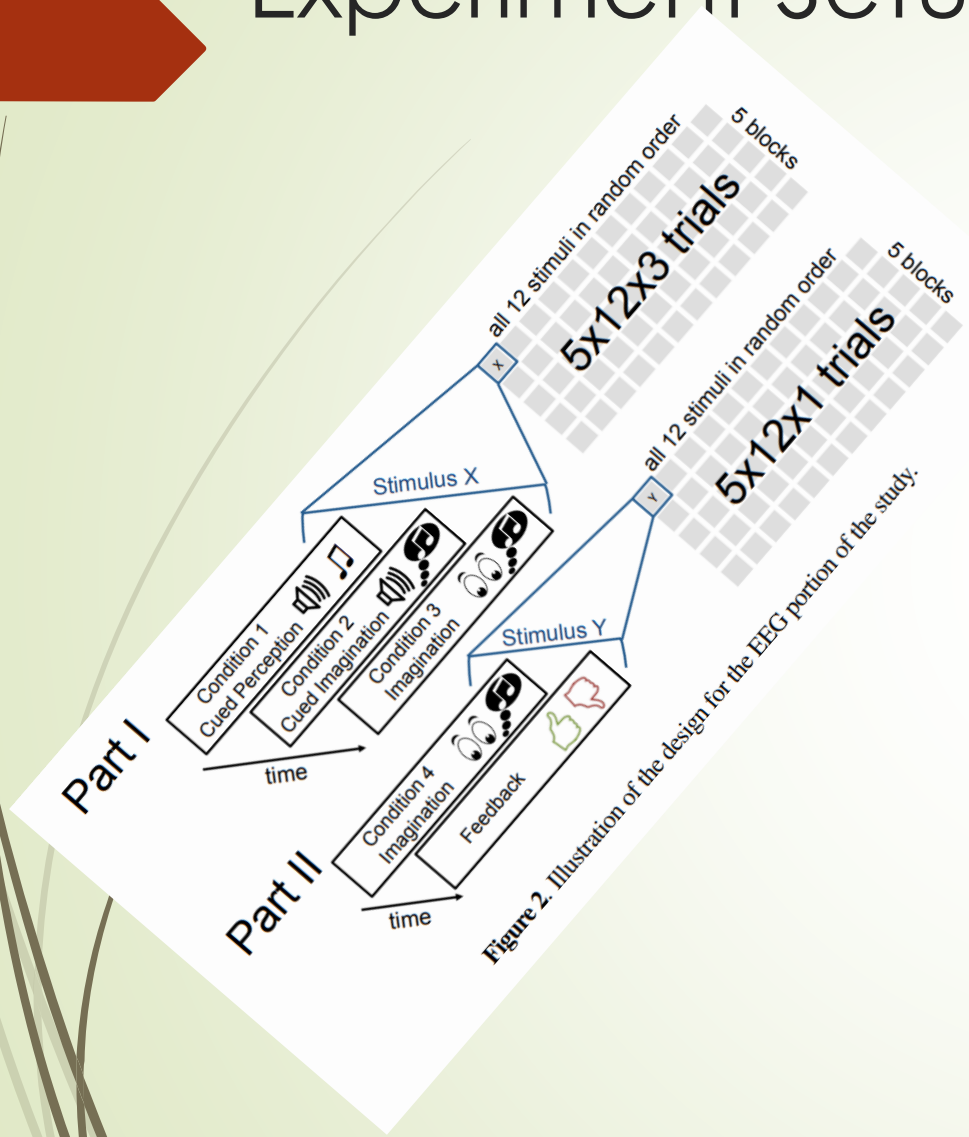
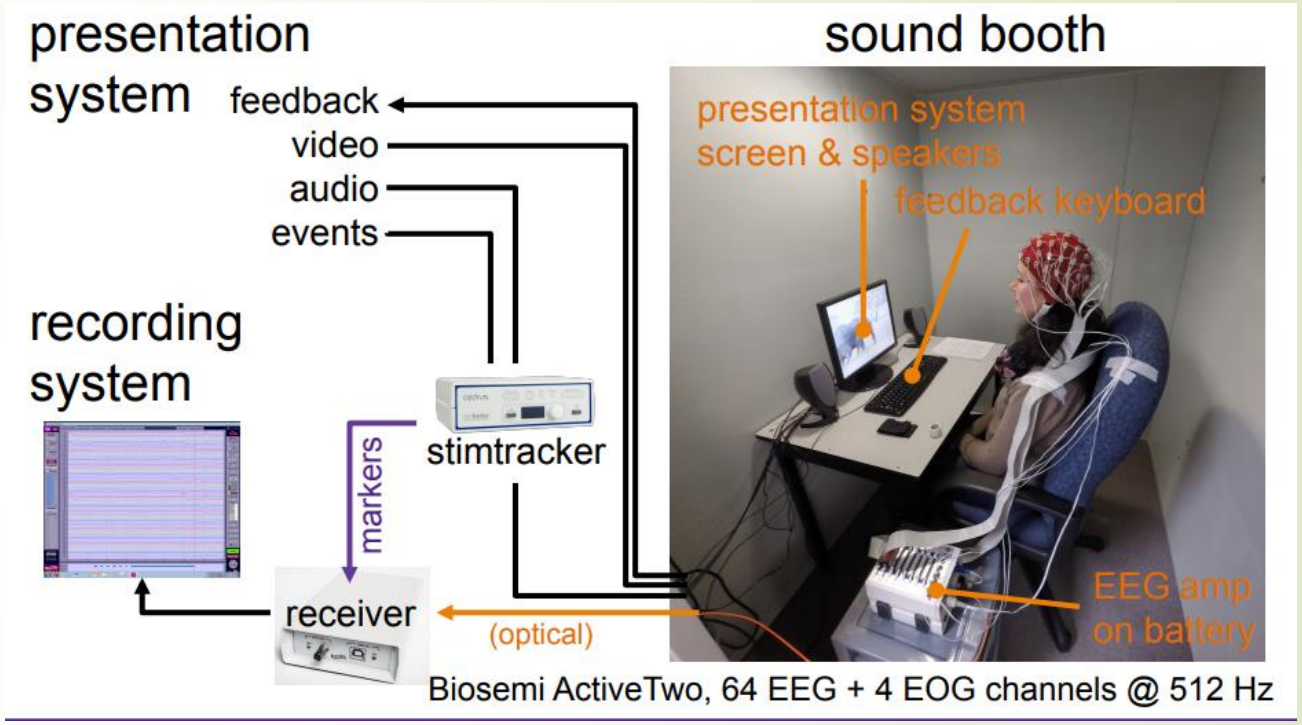


Figure 2. Illustration of the design for the EEG portion of the study.



UNIVERSITY of ROCHESTER

Ntalampiras, Stavros, and Ilyas Potamitis. "A statistical inference framework for understanding music-related brain activity." *IEEE Journal of Selected Topics in Signal Processing* 13.2 (2019): 275-284.

- Understanding the relationship between brain activity and music stimuli

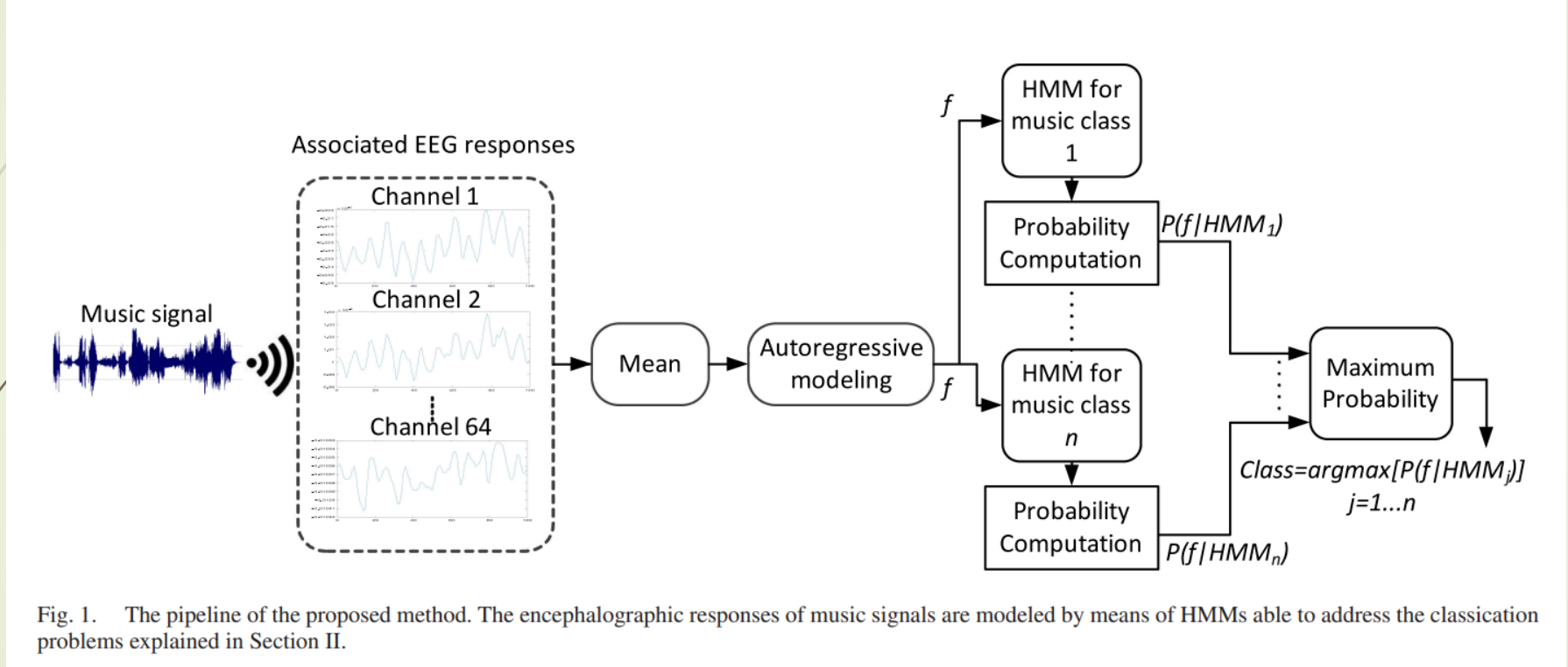


Fig. 1. The pipeline of the proposed method. The encephalographic responses of music signals are modeled by means of HMMs able to address the classication problems explained in Section II.

Three different classification tasks:

- Stimuli identification
- Groups
- Meter

UNIVERSITY of ROCHESTER

Ntalampiras, Stavros, and Ilyas Potamitis. "A statistical inference framework for understanding music-related brain activity." *IEEE Journal of Selected Topics in Signal Processing* 13.2 (2019): 275-284.

TABLE III

THE CONFUSION MATRIX OF TASK 1. THE HIGHEST CLASSIFICATION RATES PER CLASS ARE EMBOLDENED. THE AVERAGE CLASSIFICATION RATE IS 42.7%. HMMS COMPOSED OF 5 STATES AND 64 GAUSSIAN FUNCTIONS PROVIDED THE HIGHEST RECOGNITION RATES

Presented \ Responded	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Chim Chim Cheree (lyrics) 1	**34.1**	3.9	6.2	1.8	2	6	4	6.1	5.9	18.3	2	9.7
Take Me Out to the Ball game (lyrics) 2	2	**44**	2	1.5	2.5	6.1	3.9	2	10.3	5.7	6.2	13.8
Jingle Bells (lyrics) 3	2	10	**40.2**	1.8	2	9.5	2	6	6.5	7.7	-	12.3
Mary Had a Little Lamb (lyrics) 4	-	2	2	**43.7**	6.4	6.1	1.5	2	6	12.4	-	17.6
Chim Chim Cheree 5	2	12	4	2	**24**	10	2.5	7.5	4.1	14	1.9	16
Take Me Out to the Ball game 6	4	2	4	-	2	**52.5**	5.5	4.9	4	9.1	-	12
Jingle Bells 7	-	-	6	-	-	17.2	**49.5**	4	3.7	4	2	13.6
Mary Had a Little Lamb 8	-	8	2	4	2	8	2	**46**	2.1	11.9	2.9	11.1
Emperor Waltz 9	2	5.9	2.1	-	-	8.6	10	5.4	**50.8**	7.4	5.6	2
Hedwigs Theme (Harry Potter) 10	2	9.4	2.6	4	-	10.1	2	9.9	6	**42**	2.5	9.5
Imperial March (Star Wars Theme) 11	-	1.1	-	2	4.4	21.6	6	8.1	3.9	2	**36**	14
Eine Kleine Nachtmusik 12	4.1	5.9	4	4.1	3.9	4	7.7	8.3	2.5	3.5	2	**49.2**

TABLE V

CONFUSION MATRIX ACHIEVED BY THE PROPOSED APPROACH AS REGARDS TO THE GROUP CLASSIFICATION TASK. THE HIGHEST RATES ARE EMBOLDENED. THE AVERAGE CLASSIFICATION RATE IS 49.6%. HMMS COMPOSED OF 7 STATES AND 16 GAUSSIAN FUNCTIONS PROVIDED THE HIGHEST RECOGNITION RATES

Presented \ Responded	*1*	*2*	*3*
Songs recorded with lyrics 1	**52.3**	20.7	27
Songs recorded without lyrics 2	35	**42.7**	22.3
Instrumental pieces 3	22.8	23.2	**54**

TABLE VI

CONFUSION MATRIX ACHIEVED BY THE PROPOSED APPROACH AS REGARDS TO THE METER CLASSIFICATION TASK. THE HIGHEST RATES ARE EMBOLDENED. THE AVERAGE CLASSIFICATION RATE IS 68.7%. HMMS COMPOSED OF 6 STATES AND 2 GAUSSIAN FUNCTIONS PROVIDED THE HIGHEST RECOGNITION RATES

Presented \ Responded	*3/4*	*4/4*
3/4	**68.2**	31.8
4/4	30.9	**69.1**

Possible Future Works

- Explore temporal dependencies via recurrent neural networks
- Spatial features can be learned via CNNs
- Transformers are acting based on attention mechanism and can consider global temporal dependencies

UNIVERSITY of ROCHESTER