INSTITUT POLYTECHNIQUE DE PARIS

Final Report

Pathological Gait Characterization using Frequency-domain and Time-Frequency-domain Clustering

Mochamad Ardiansyah NUGRAHA

IMA7515 Scientific Project Datapac

M1 Electrical Engineering for Communications and Information Processing – Datapac

Institut Polytechnique de Paris

2 July 2024

Introduction

Gait Analysis Importance

Understanding gait patterns is crucial for rehabilitation and treatment planning, particularly in populations with neurological conditions like hemiplegia, tetraplegia, and paraplegia.

Comparison Between Healthy and Pathological

Gait analysis allows for quantitative comparisons of biomechanical signals between healthy individuals and those with pathologies, highlighting deviations and impacts of specific conditions on walking mechanics.

Data

- Knee joint angle on sagittal plane
- Healthy Control: 52 person
- Pathological: 38 patients
    - Hemiplegia: 18
    - Paraplegia: 9
    - Tetraplegia: 11

Methodology

- Data pre-processing

Obtain knee cycles → Aggregate 3 cycles into 1 signal → Mean centering to remove DC offset

- Frequency-domain analysis and clustering

Pre-processed data → Compute Fast-Fourier Transform → Obtain 5 frequencies with highest magnitude → Plot histogram on frequency and magnitude, normalize → Fit Gaussian Mixture Model

Gaussian Mixture Model

- Incorporated the computation of Bayesian Information Criterion (BIC) to find the best number of Gaussian components with the lowest criterion value
- These criteria penalize the likelihood of the model based on the number of parameters, helping to avoid overfitting.
- BIC is simple, and a ready-to-use built-in method in scikit-learn's GaussianMixture.

The definition of BIC replace the constant 2 by $\log(N)$:

$$BIC = -2\log(\hat{L}) + \log(N)d$$

where N is the number of samples.

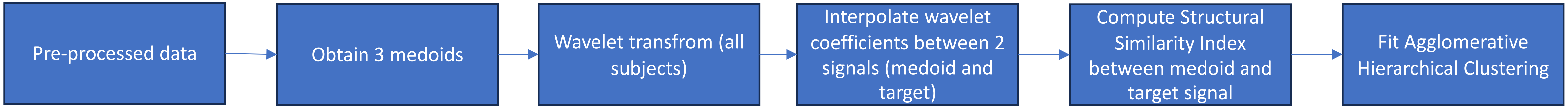
For a linear Gaussian model, the maximum log-likelihood is defined as:

$$\log(\hat{L}) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{2\sigma^2}$$

where σ^2 is an estimate of the noise variance, y_i and $\hat{y}_i$ are respectively the true and predicted targets, and n is the number of samples.

Methodology

- Time-Frequency-domain analysis and clustering



Wavelet transform

- Mother wavelet: Gaussian 1
- Wavelet scale: 1-64

Structural Similarity Index Measure (SSIM)

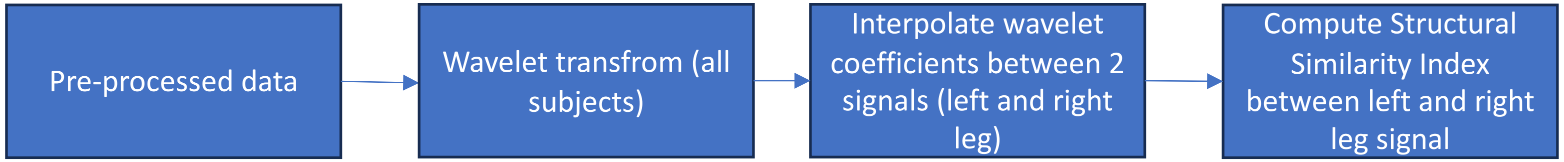
- Comparison between baseline and target signal
  - Baseline: 3 medoids (id_patient: 12, 23, 34), left legs only
  - Target: the rest (both Healthy Control and Pathological Subject)
- Before computing the SSIM, either baseline or target signal must be interpolated to have the same length/dimension.
- Output: for each target signal will have a 1-D SSIM vector [SSIM1, SSIM2, SSIM3]
  - SSIM1: SSIM between medoid 1 (id_patient 12) and target
  - SSIM2: SSIM between medoid 2 (id_patient 23) and target
  - SSIM3: SSIM between medoid 3 (id_patient 34) and target

Agglomerative Hierarchical Clustering

- Linkage: Ward
- n_cluster: 4 (Healthy Control, Hemiplegia, Paraplegia, Tetraplegia)

Methodology

- Time-Frequency-domain analysis and SSIM left-right leg symmetry



INSTITUT POLYTECHNIQUE DE PARIS

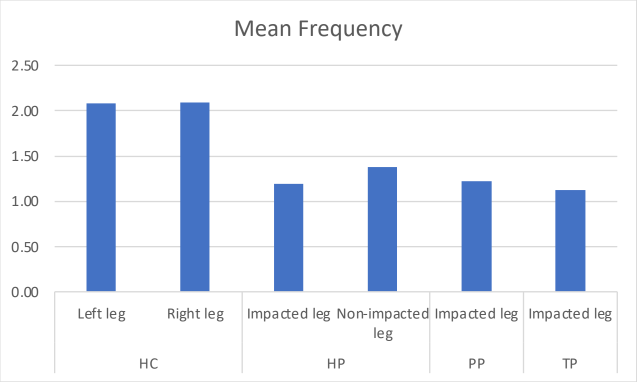
Result

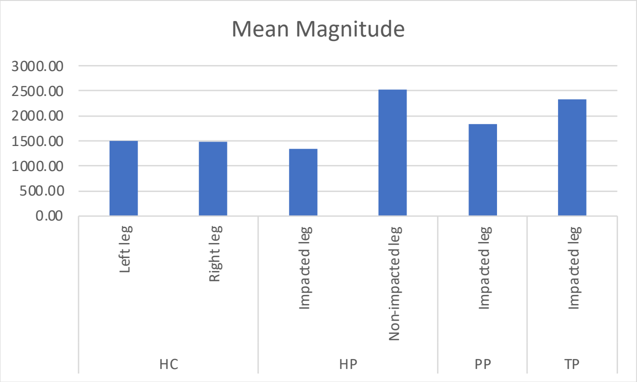
Frequency-domain Analysis

Top 5 Frequencies with Highest Magnitude

	top freq left	top mag left	top freq right	top mag right
Healthy Control				
mean	2.09	1492.25	2.09	1478.74
std	1.15	1297.47	1.16	1285.21
min	0.28	127.07	0.28	150.16
max	4.95	4358.82	5.00	4264.35

	top freq impacted	top mag impacted	top freq non-impacted	top mag non-impacted
Hemiplegia				
mean	1.19	1341.20	1.38	2526.31
std	0.89	1330.78	0.79	2028.41
min	0.07	156.26	0.06	283.82
max	3.53	7395.41	3.46	12373.76
Paraplegia				
mean	1.22	1841.41		
std	0.77	1928.81		
min	0.13	198.72		
max	3.21	7387.12		
Tetraplegia				
mean	1.13	2330.55		
std	0.77	2197.10		
min	0.09	200.23		
max	3.88	11235.62		





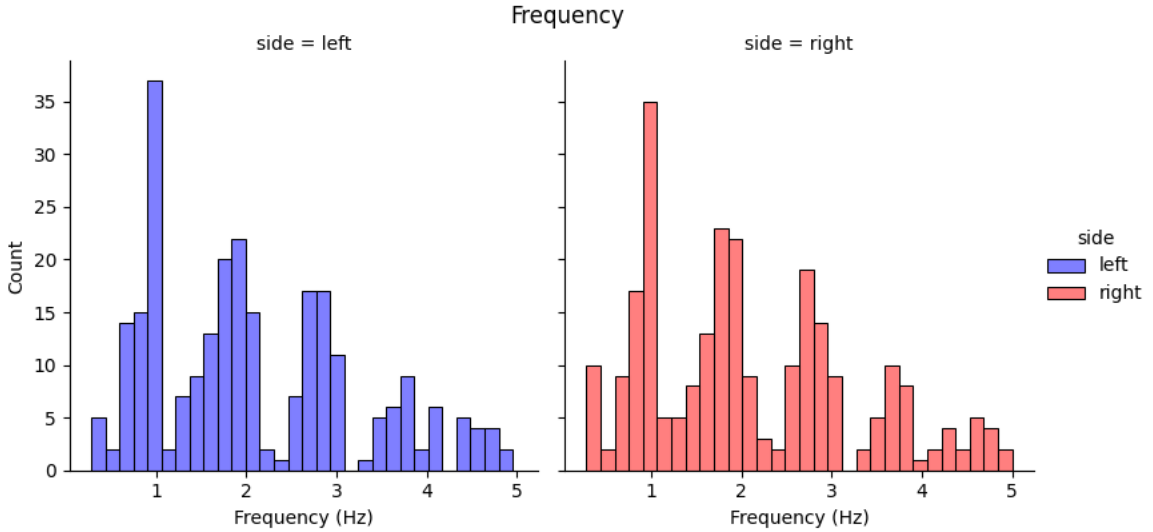
Frequency:

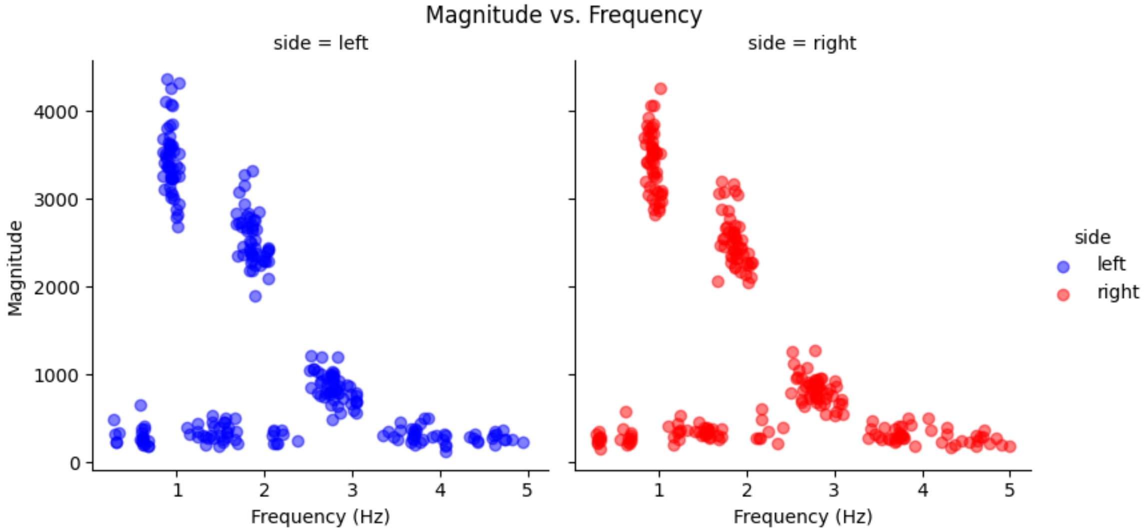
- All groups (hemiplegia impacted, hemiplegia non-impacted, paraplegia impacted, and tetraplegia impacted) show lower mean frequencies compared to healthy controls.
- This suggests that these conditions result in a decrease in the biomechanical performance.

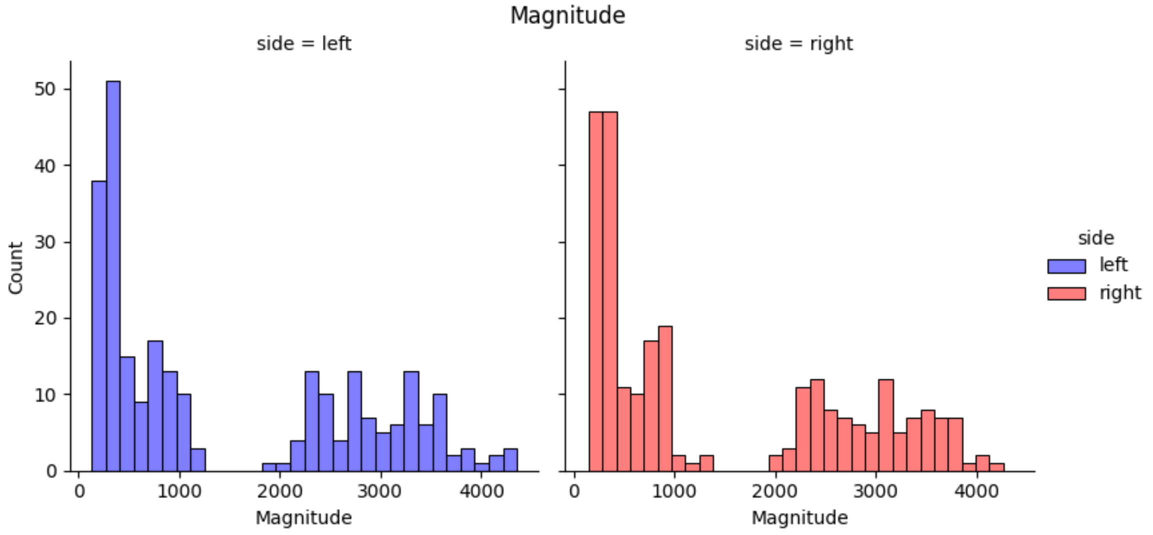
Magnitude:

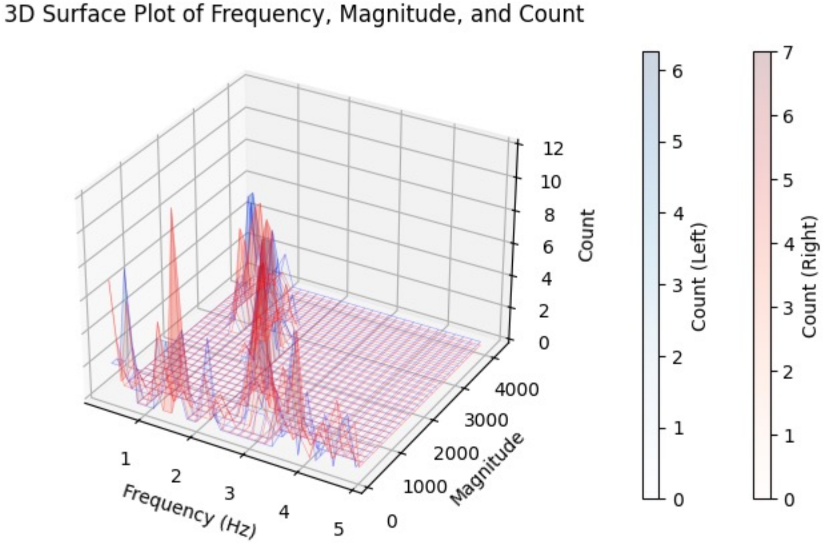
- Hemiplegia impacted individuals have slightly lower magnitudes compared to healthy controls.
- Hemiplegia non-impacted, paraplegia impacted, and tetraplegia impacted individuals have higher magnitudes compared to healthy controls.
- This might indicate a compensatory mechanism or altered biomechanical property in non-impacted sides (hemiplegia) or in cases of paraplegia and tetraplegia.

Healthy Control

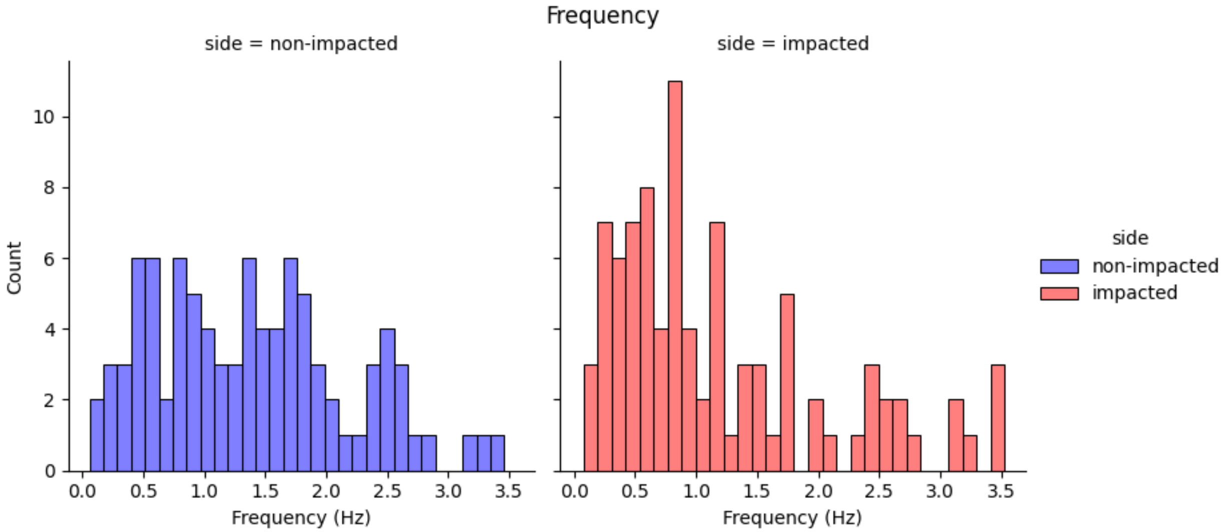


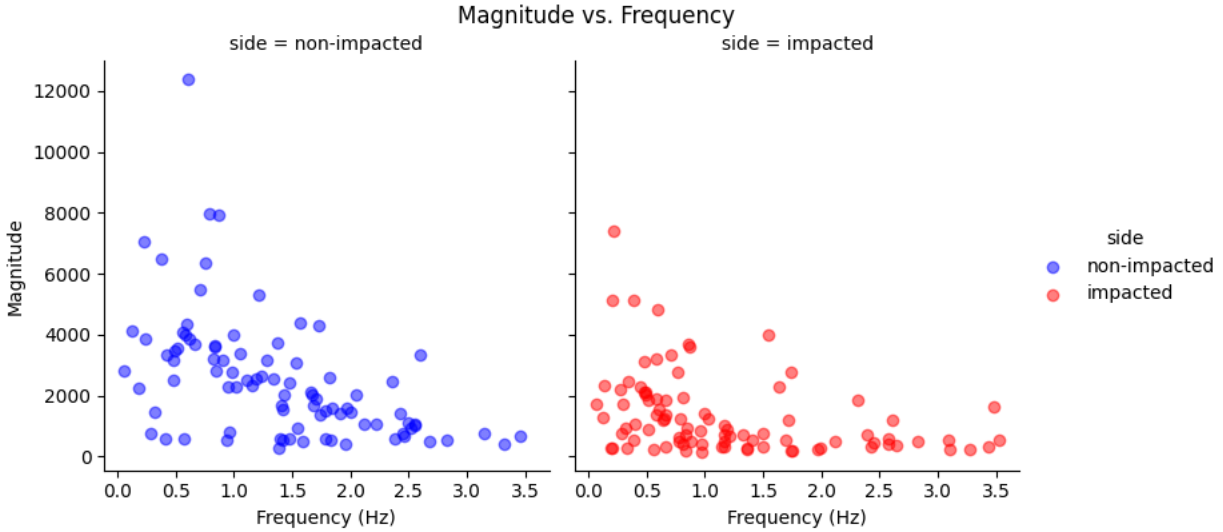


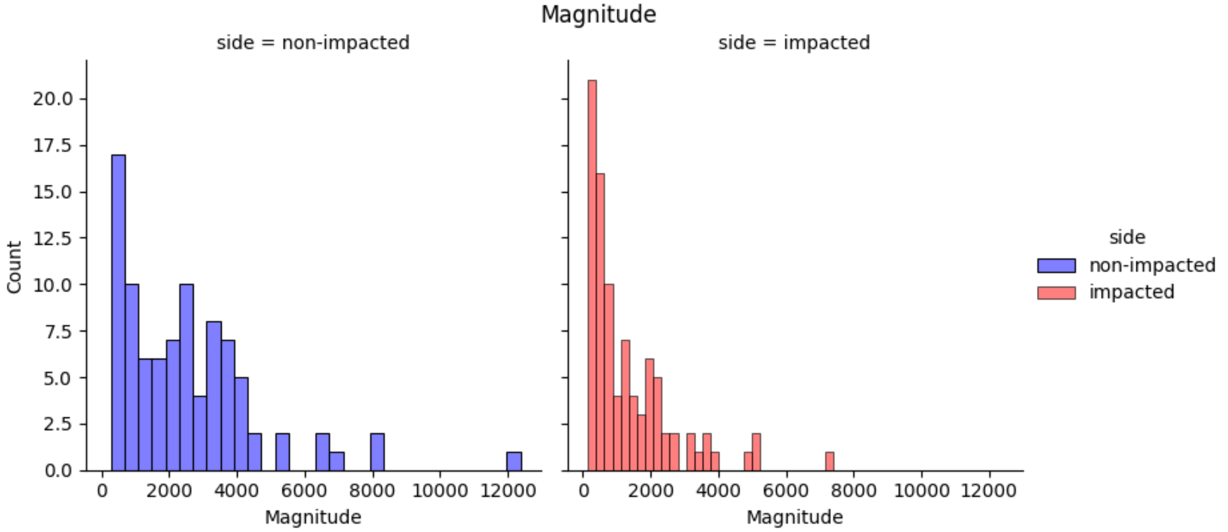


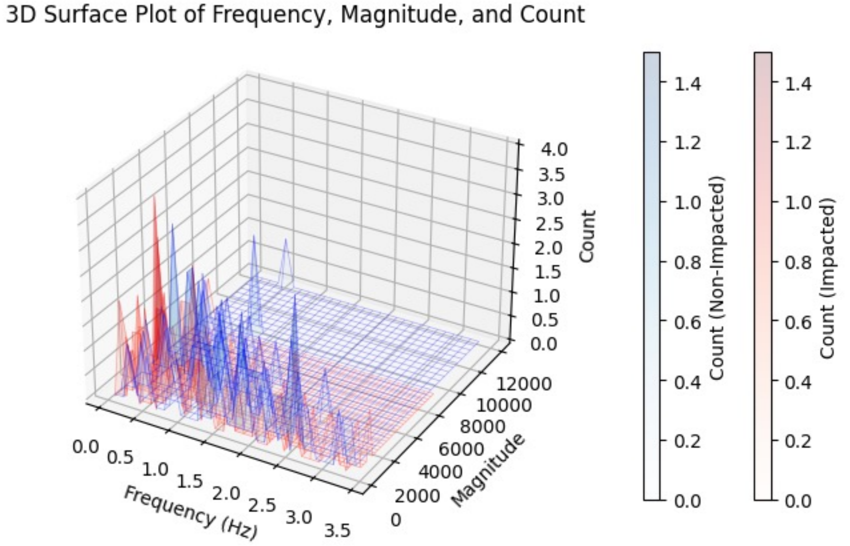


Pathological Subject – Hemiplegia

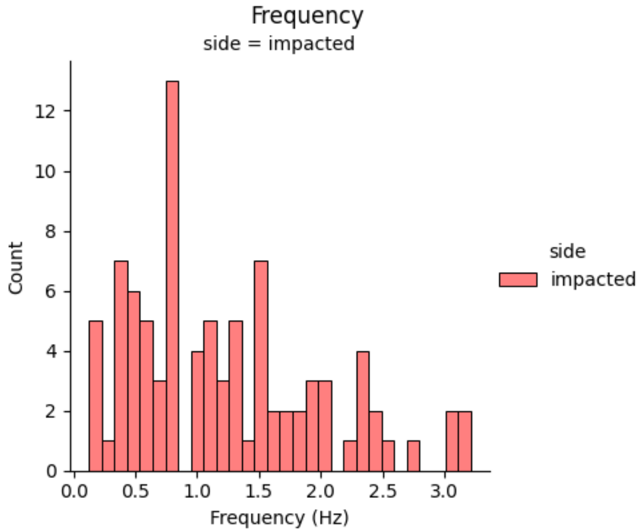


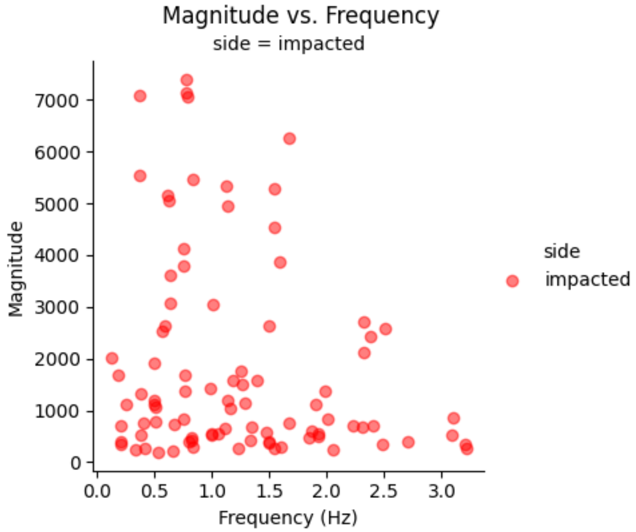


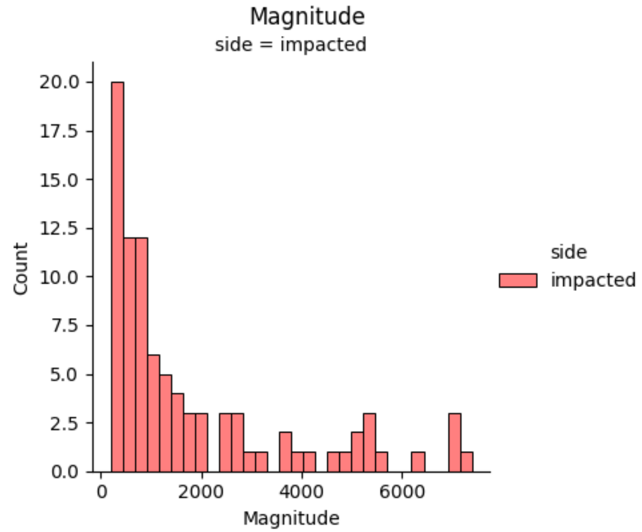


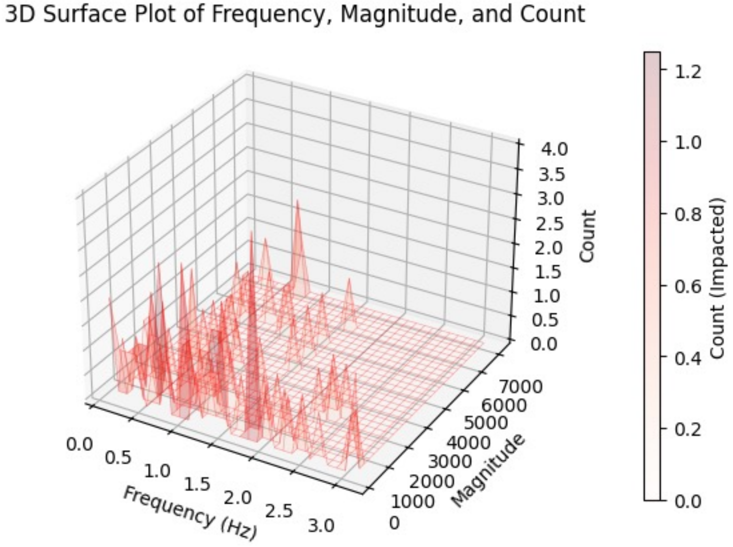


Pathological Subject – Paraplegia

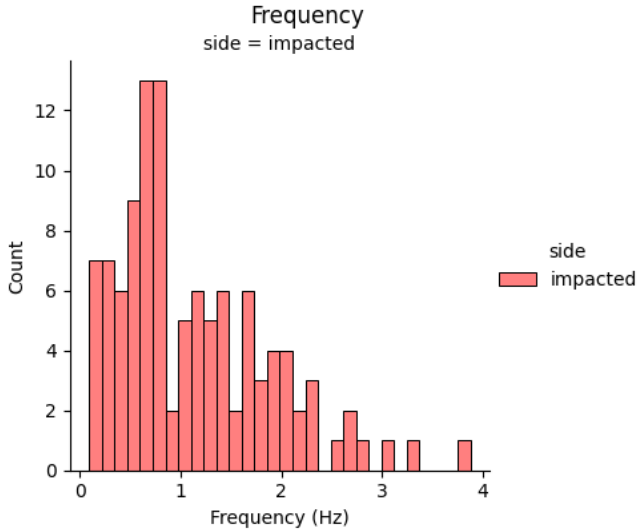


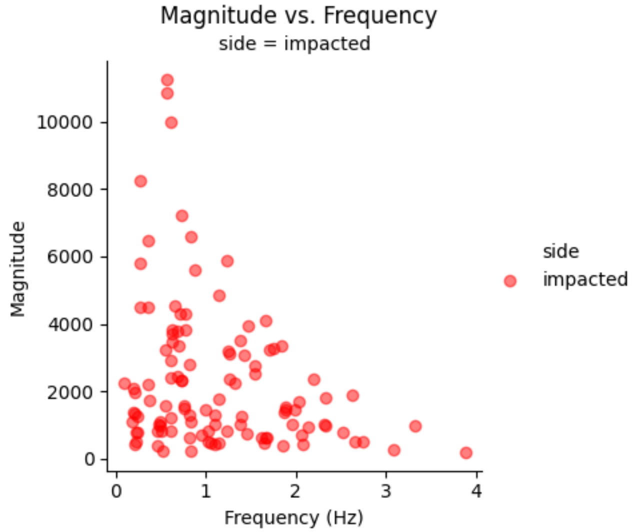


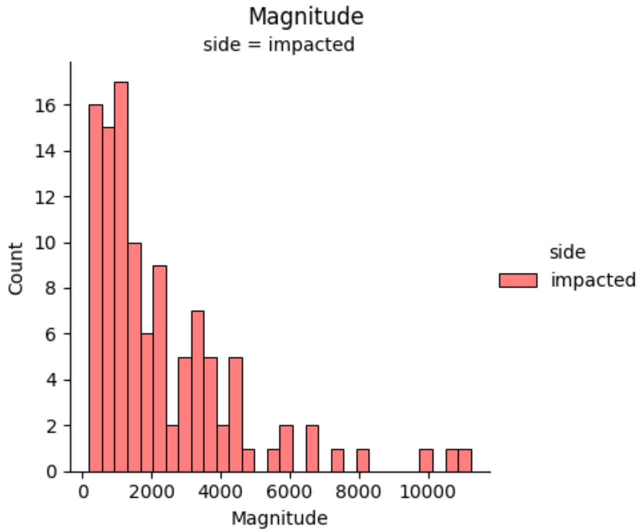


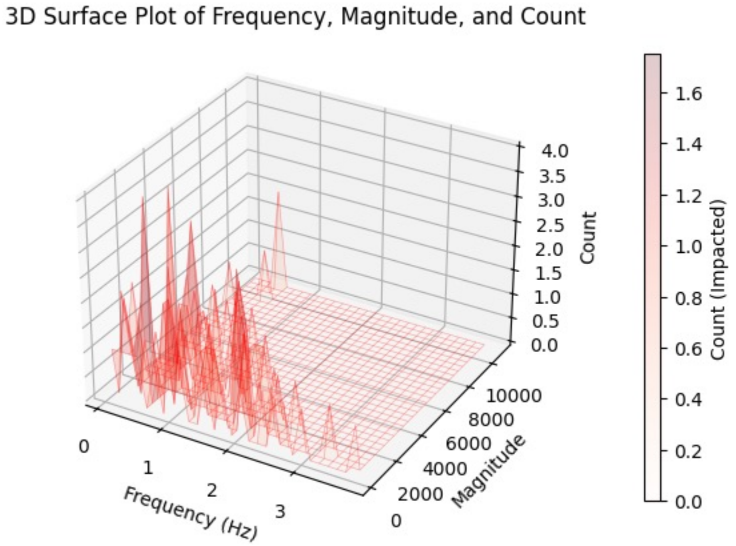


Pathological Subject – Tetraplegia





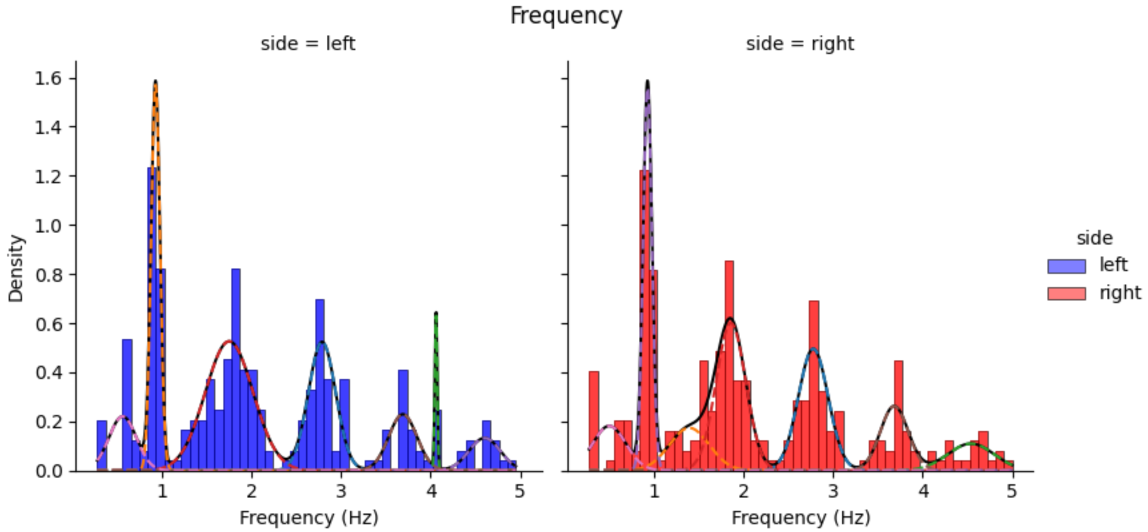


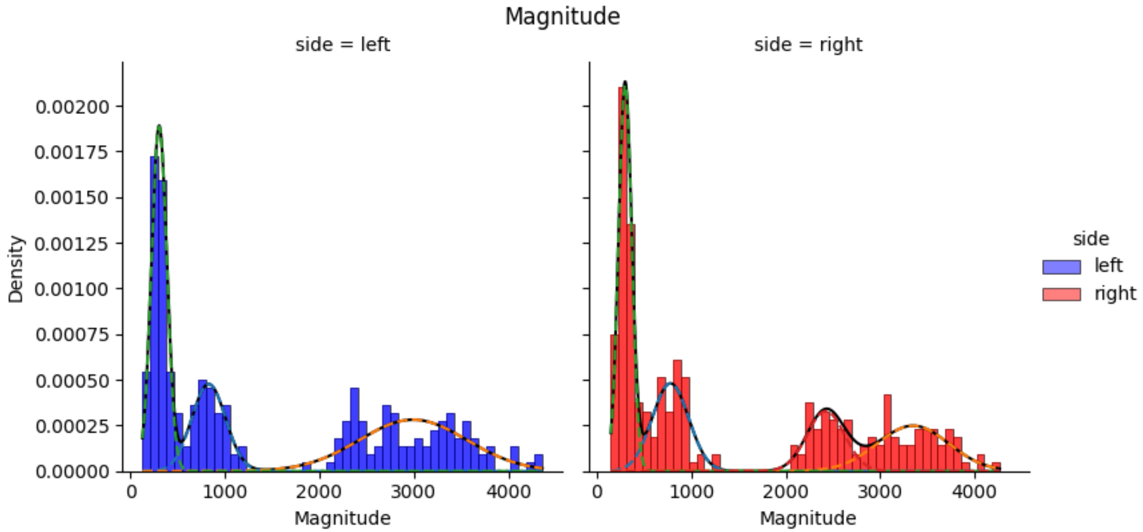


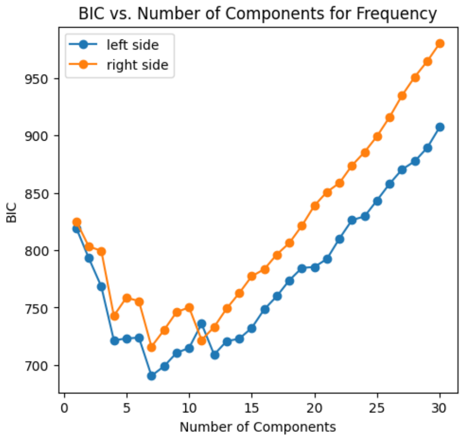
Result

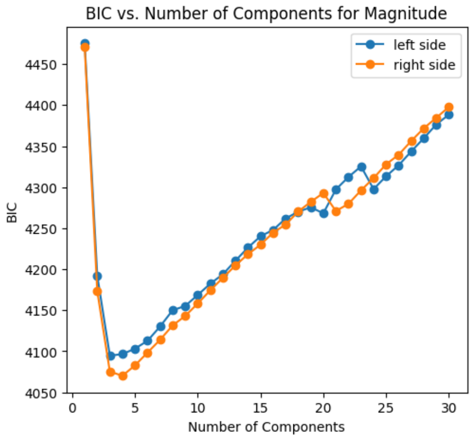
Gaussian Mixture Model

Healthy Control







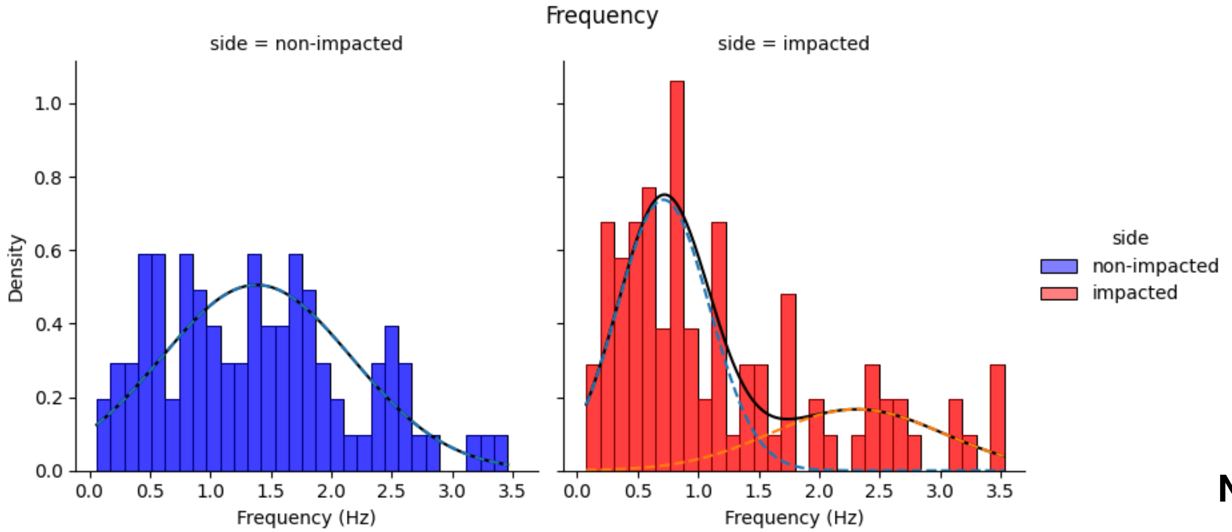


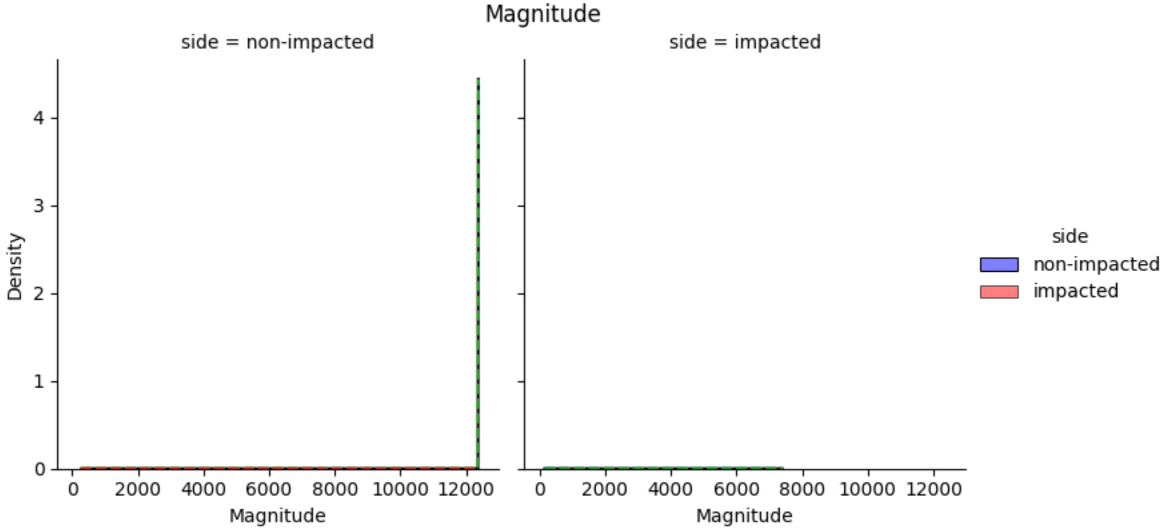
Number of Gaussian components

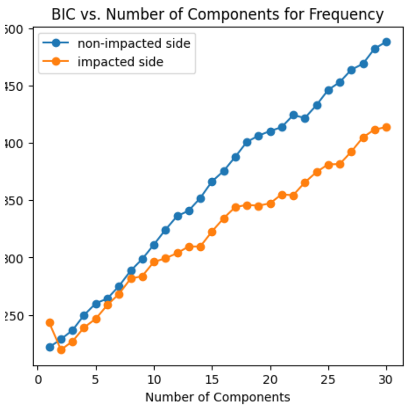
- Frequency
  - Left and right: 7
- Magnitude
  - Left: 3
  - Right: 4

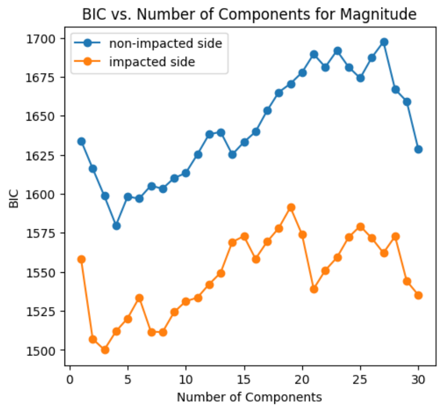
In Healthy Control population, right leg has more Gaussian components than left leg in terms of the magnitude.

Pathological Subject – Hemiplegia





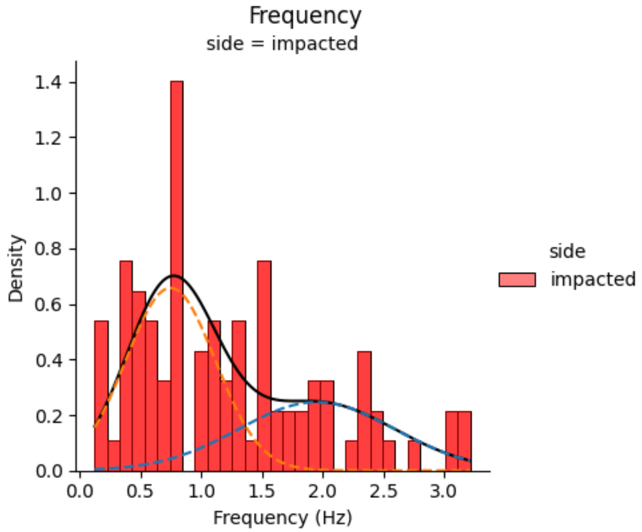


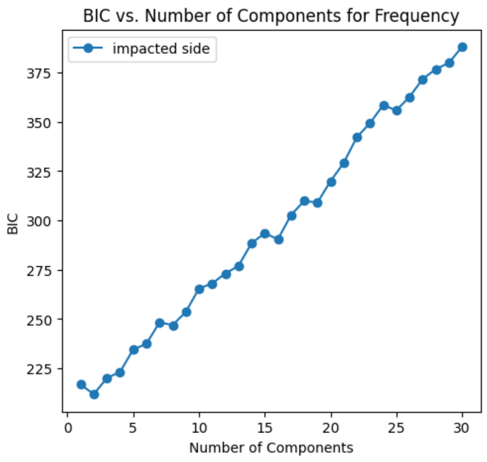


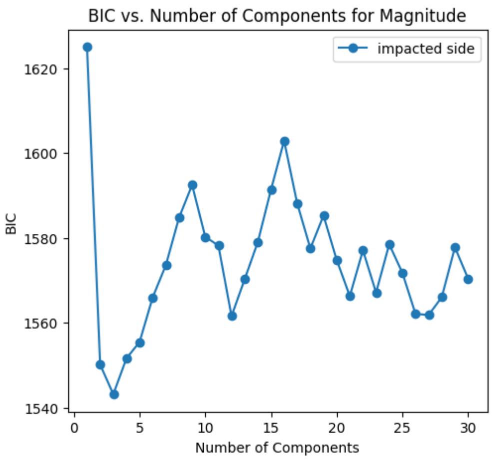
Number of Gaussian components

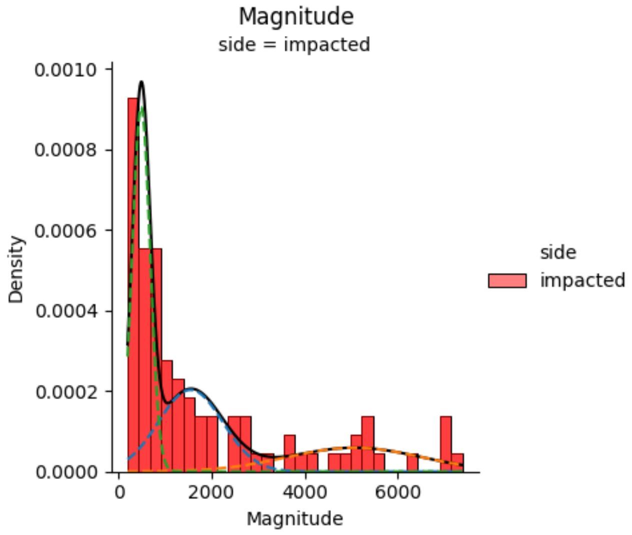
- Frequency
  - Non-impacted: 1
  - Impacted: 2
- Magnitude:
  - Non-impacted: 4
  - Impacted: 3

Pathological Subject – Paraplegia





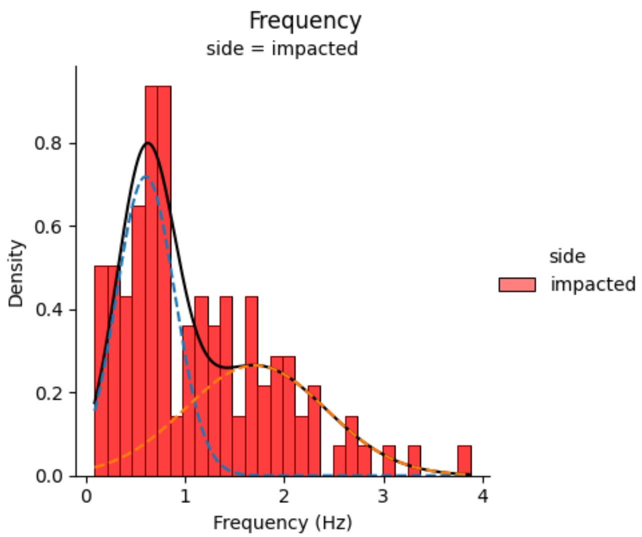


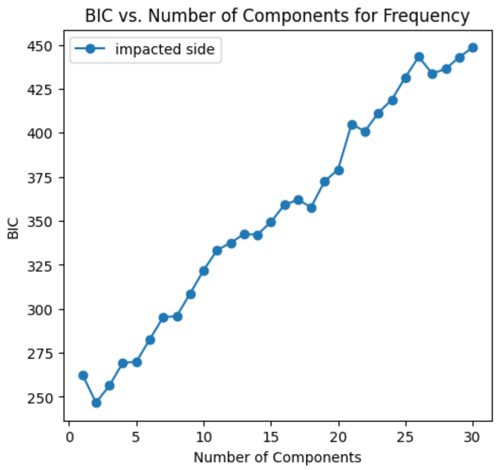


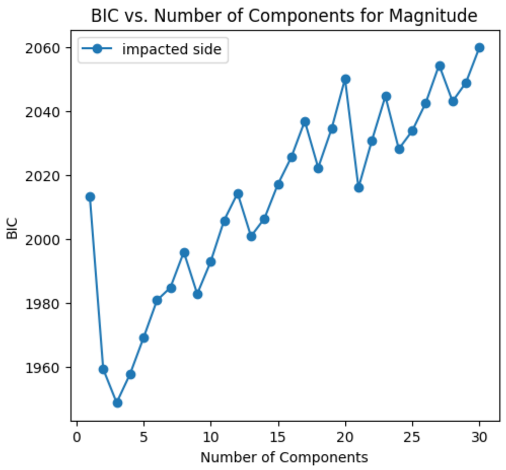
Number of Gaussian components

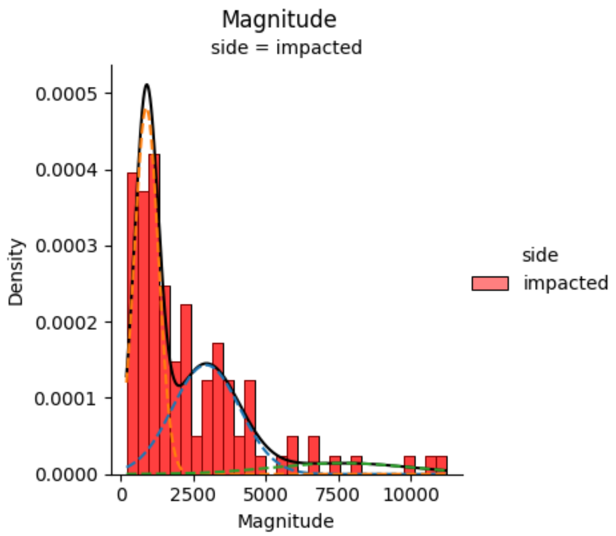
- Frequency impacted: 2
- Magnitude impacted: 3

Pathological Subject – Tetraplegia









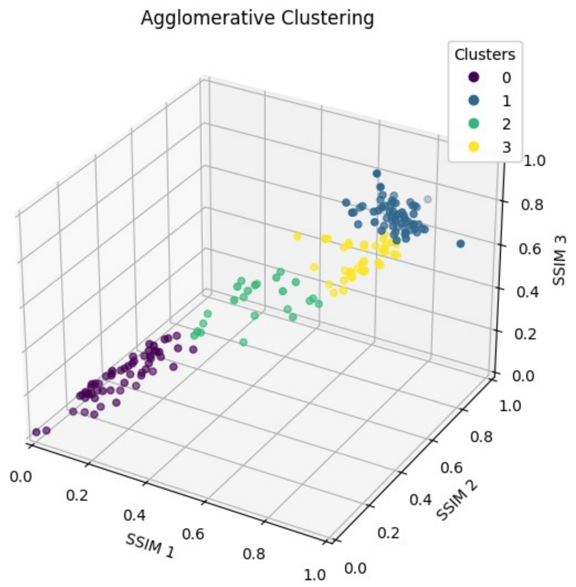
Number of Gaussian components

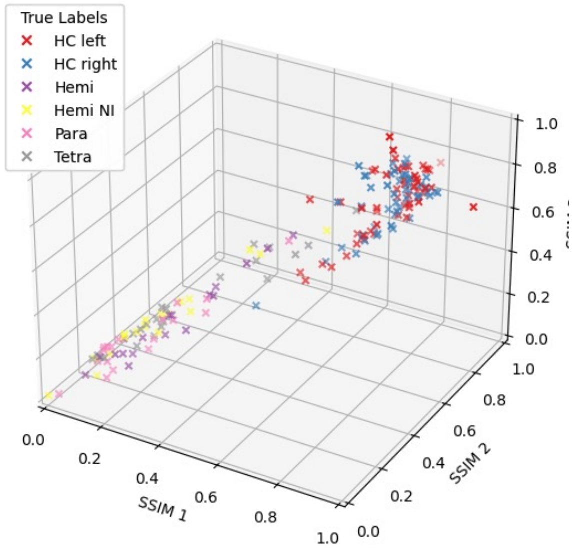
- Frequency impacted: 2
- Magnitude impacted: 3

Result

Time-Frequency-domain Anaysis and Agglomerative Clustering

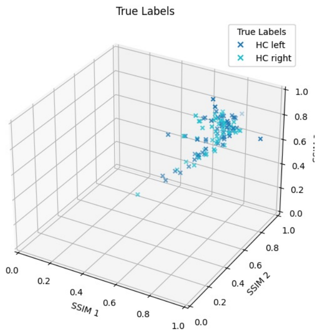
Agglomerative Hierarchical Clustering

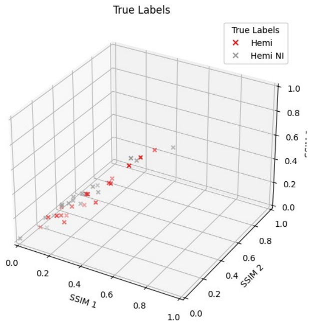


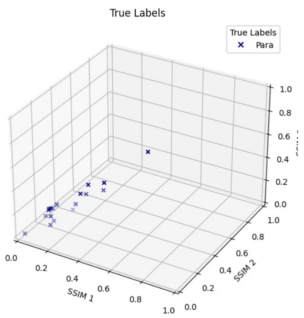


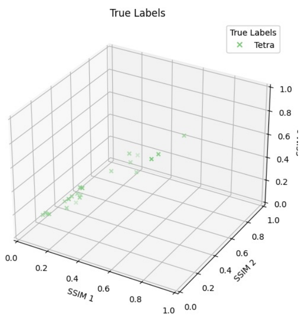
Cluster Analysis:

- Cluster 1: predominantly contains samples from both "HC left" and "HC right" groups, indicating that these samples are clustered together based on their similarity in SSIM values across three dimensions (SSIM 1, SSIM 2, SSIM 3).
- Cluster 3: a mix of "HC left" and "HC right" samples but with a higher proportion of "HC right" samples.
- Cluster 2: primarily contains samples from "Hemi" group, specifically AM, BM, BA, CB, GA, GM, and SV, indicating similarity among these samples.
- Cluster 0: includes samples from "Para" and "Tetra" groups, suggesting similarity in SSIM values for these samples.

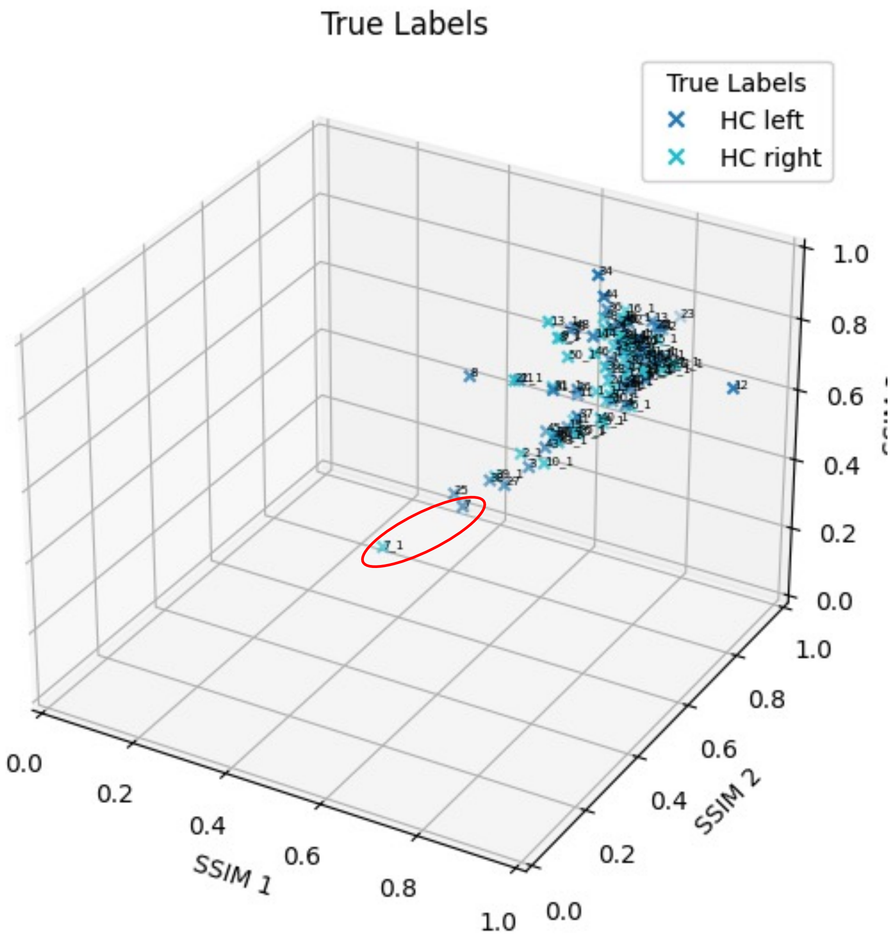




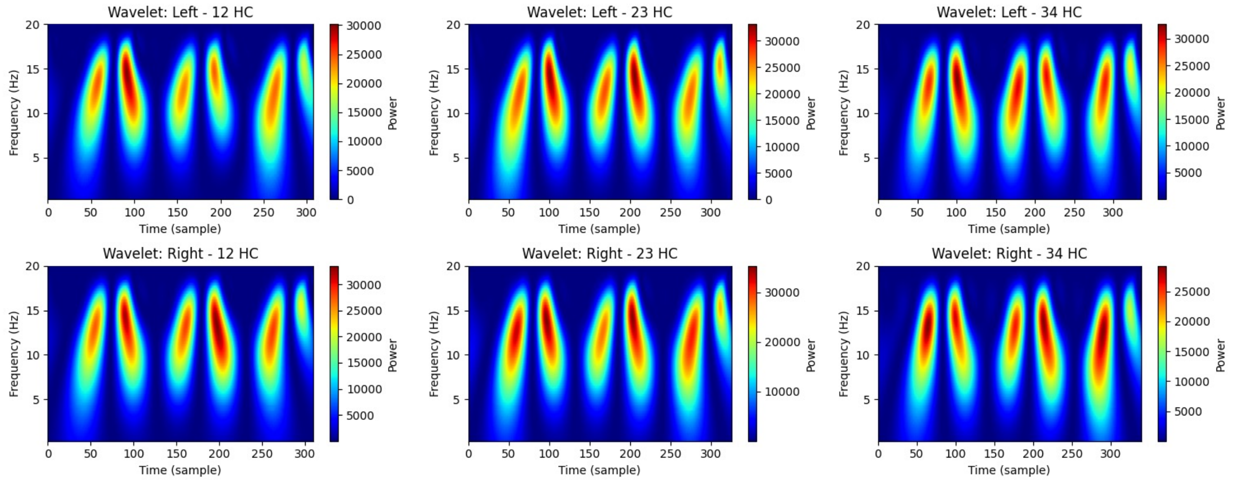




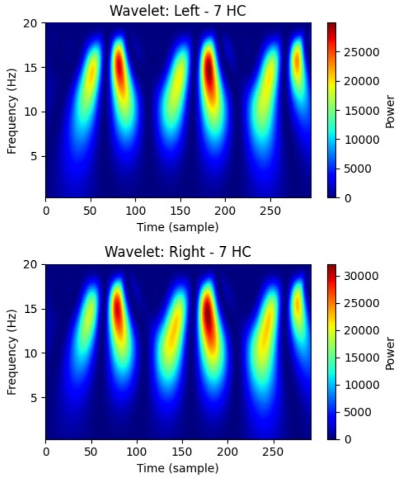
Healthy Control – id_patient = 7



Medoids



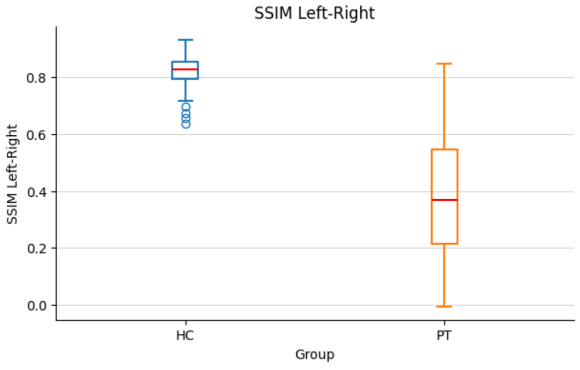
HC id = 7

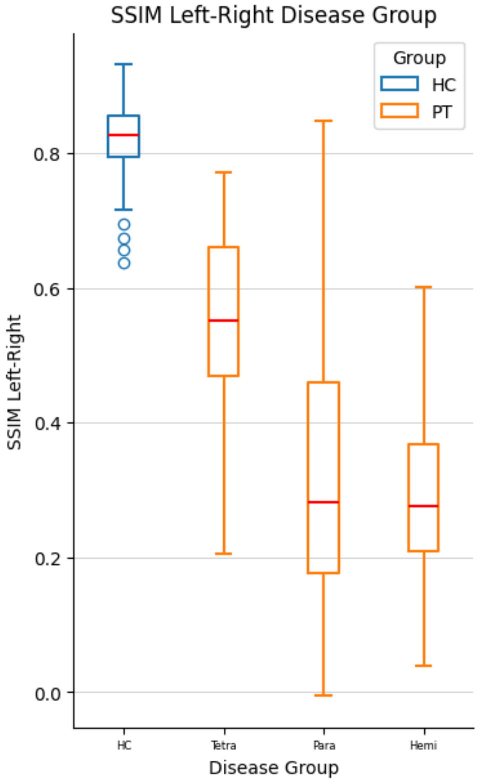


Result

SSIM on Left-Right Leg Symmetry

SSIM on Left-Right Leg Symmetry



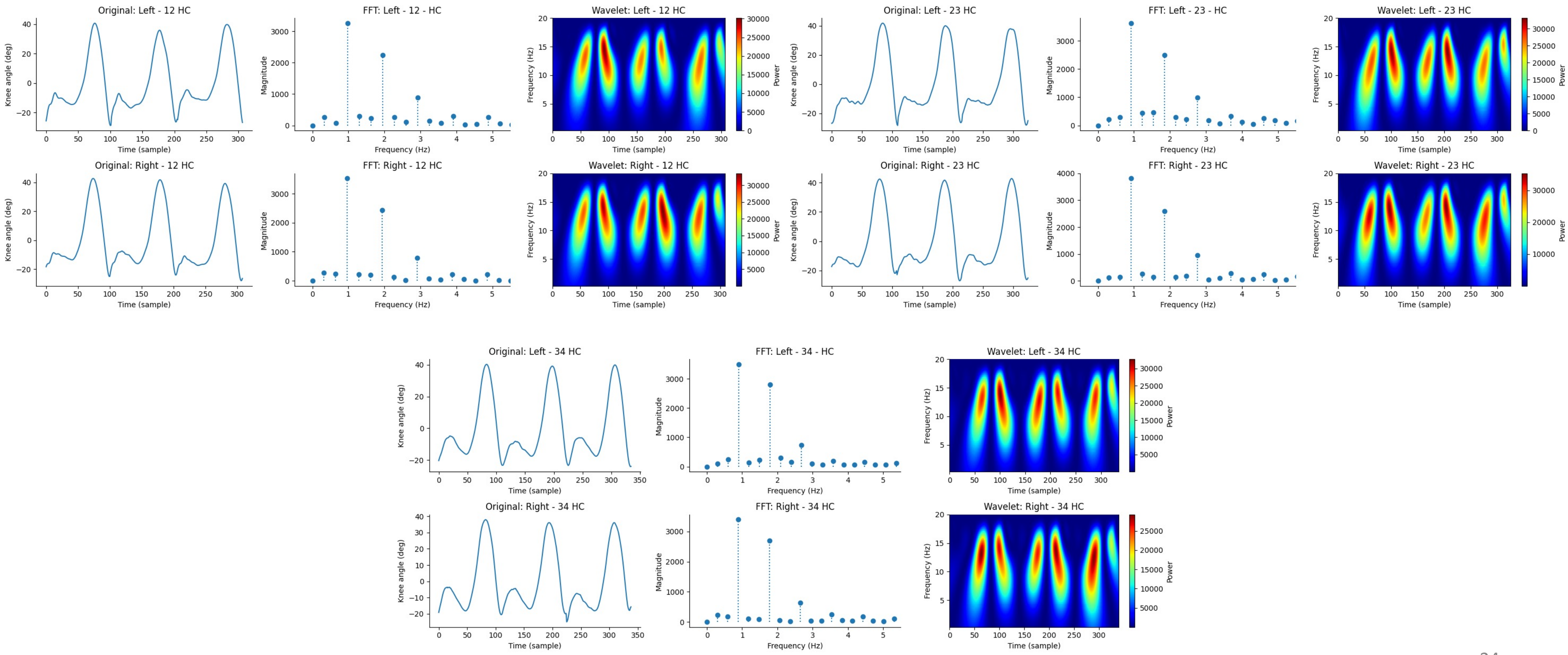


- Healthy control reported high SSIM values with low variance reflecting high symmetry in their structure and appearance
- Pathological subject exhibits a lower SSIM median that the left and right legs are less similar, suggesting asymmetry due to disease or injury

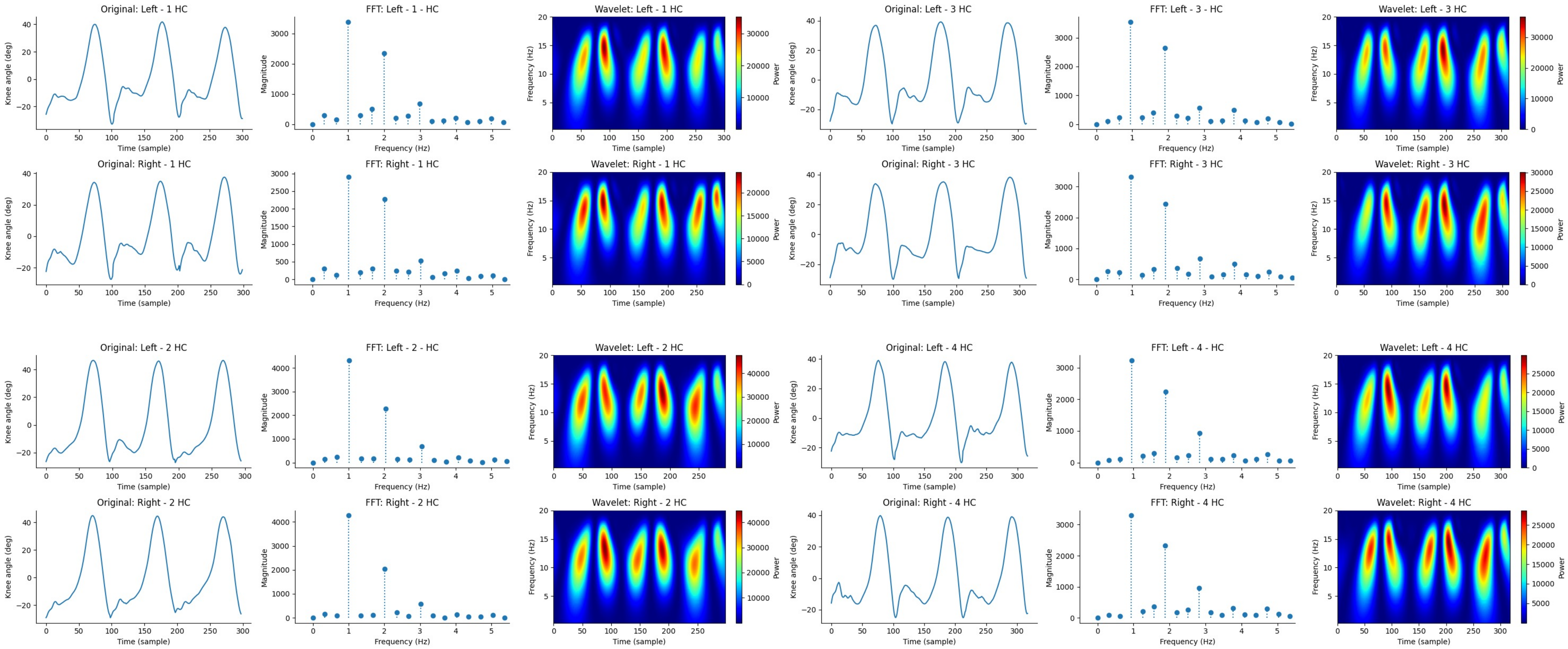
Conclusion

- The clustering suggests that SSIM values effectively differentiate between different disease groups, even within subcategories like "HC left" vs "HC right" and different types of hemiplegia.
- The clustering still hardly identify Paraplegia and Tetraplegia. Probably need more parameters to be considered.
- In general, Pathological Subjects have lower frequency and higher magnitude than Healthy Control due to biomechanical limitation and compensatory property.
- SSIM on left-right symmetry show a good result specifically on hemiplegia patients, highlighting the imbalance of the lower limbs.

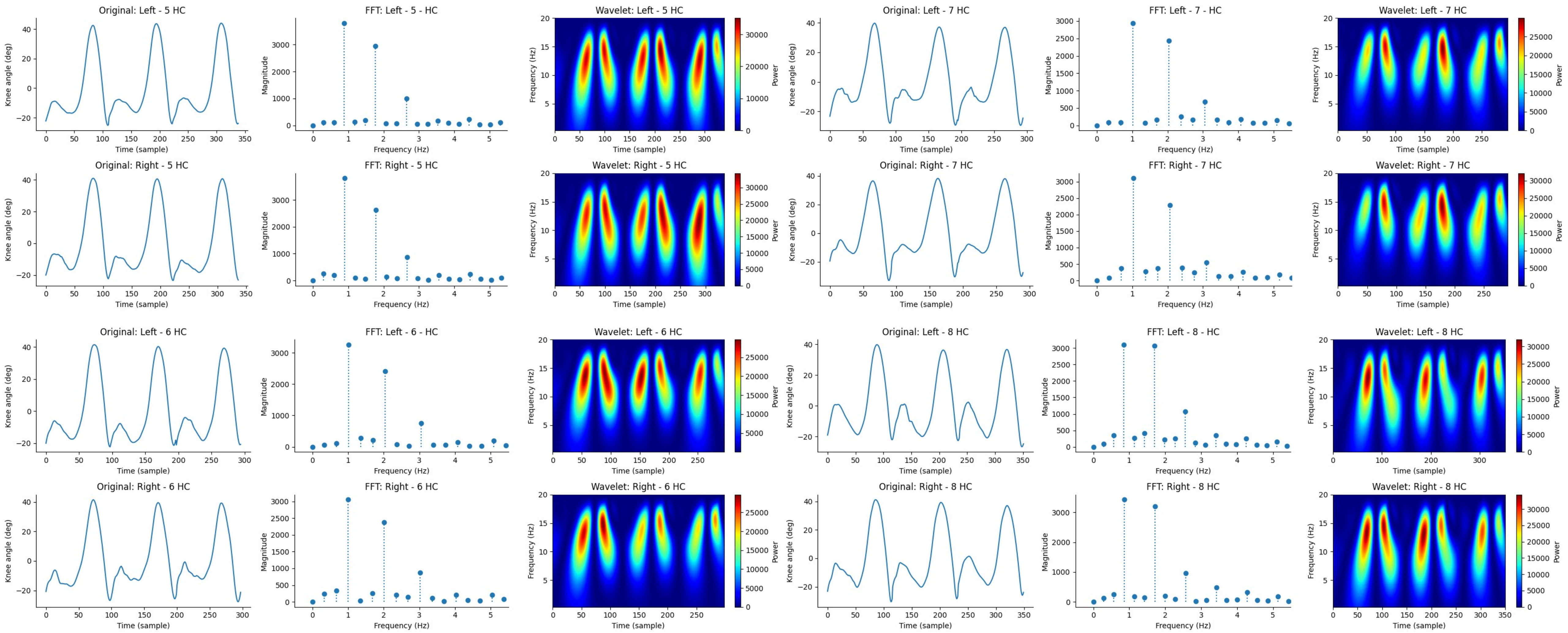
Appendix – Healthy Control – 3 Medoids



Appendix – Healthy Control – 12 samples (1/3)

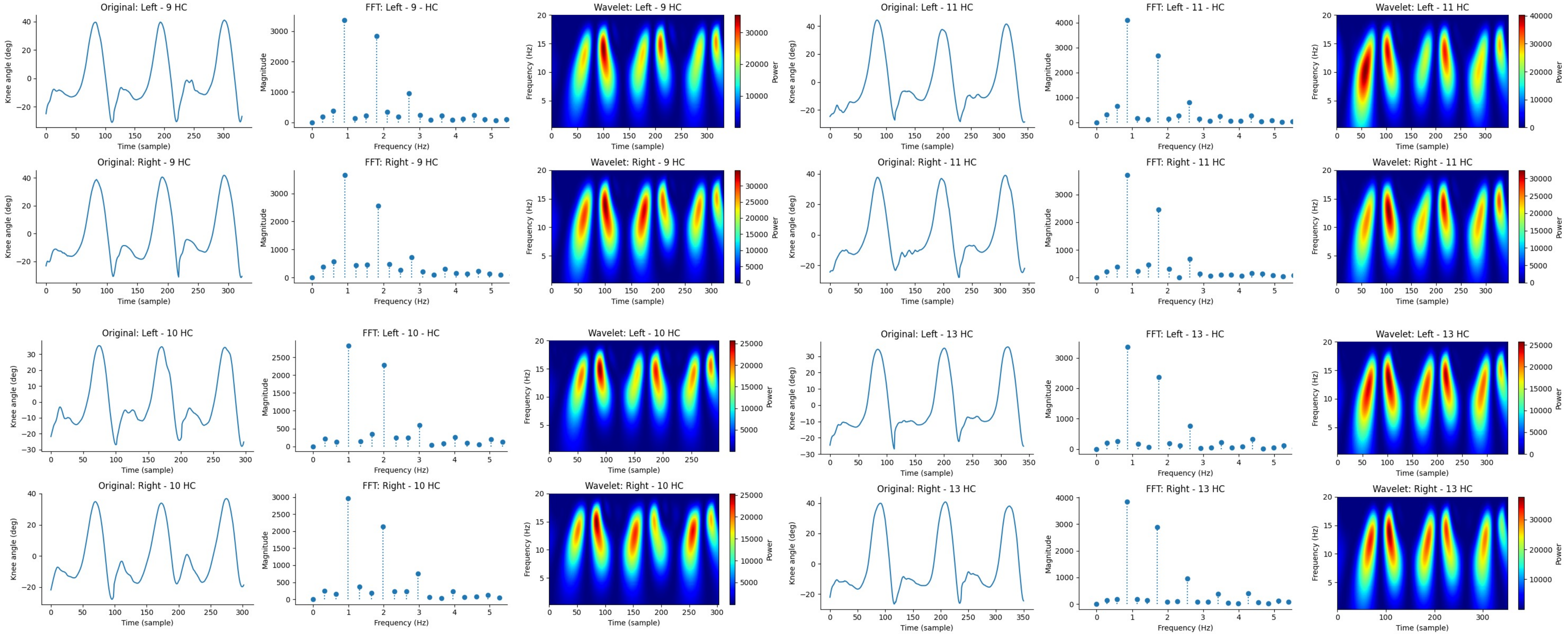


Appendix – Healthy Control – 12 samples (2/3)



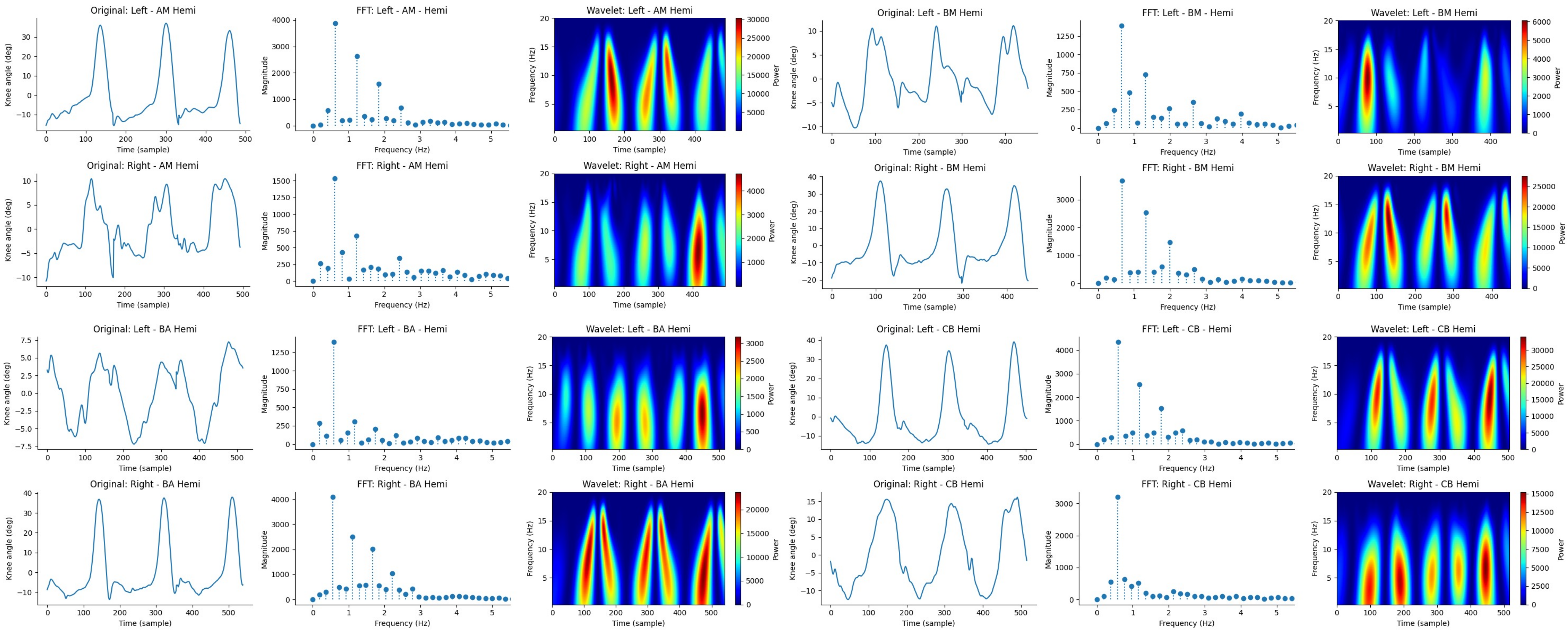
INSTITUT POLYTECHNIQUE DE PARIS

Appendix – Healthy Control – 12 samples (3/3)



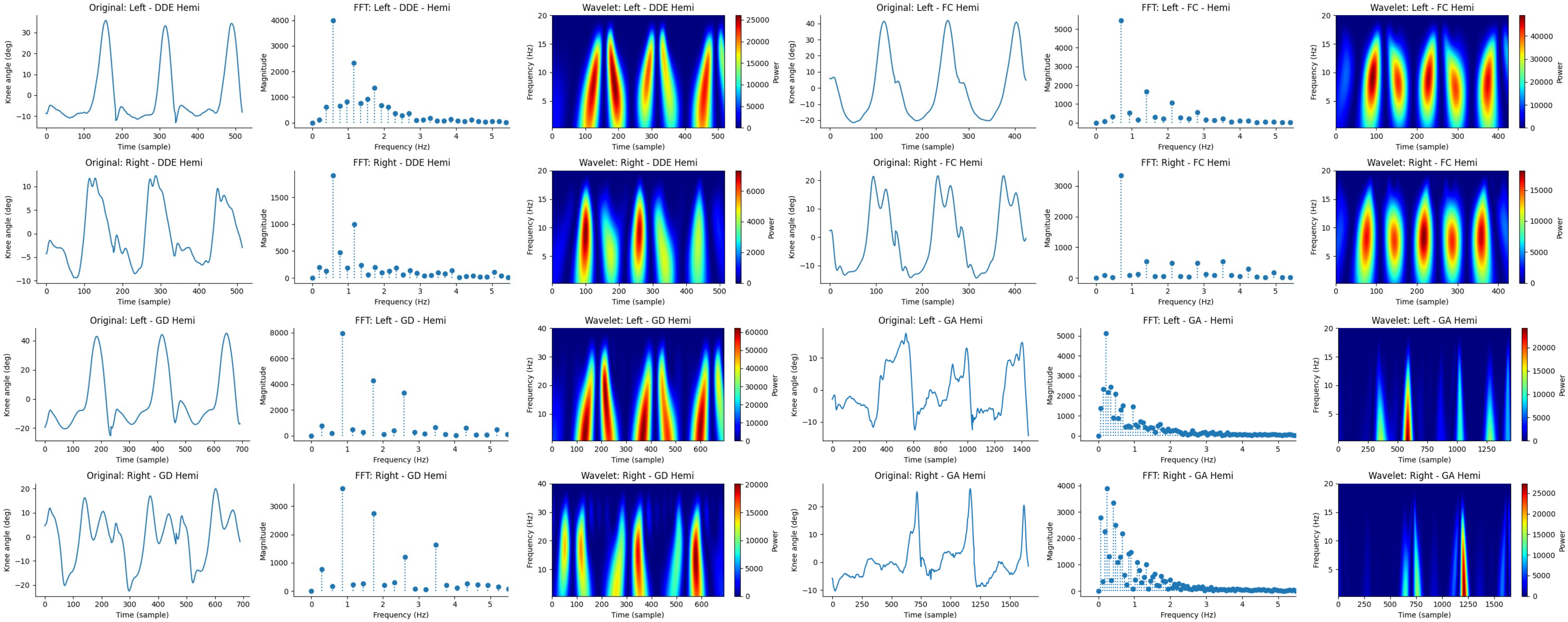
Appendix – Hemiplegia – 12 samples (1/3)

INSTITUT POLYTECHNIQUE DE PARIS

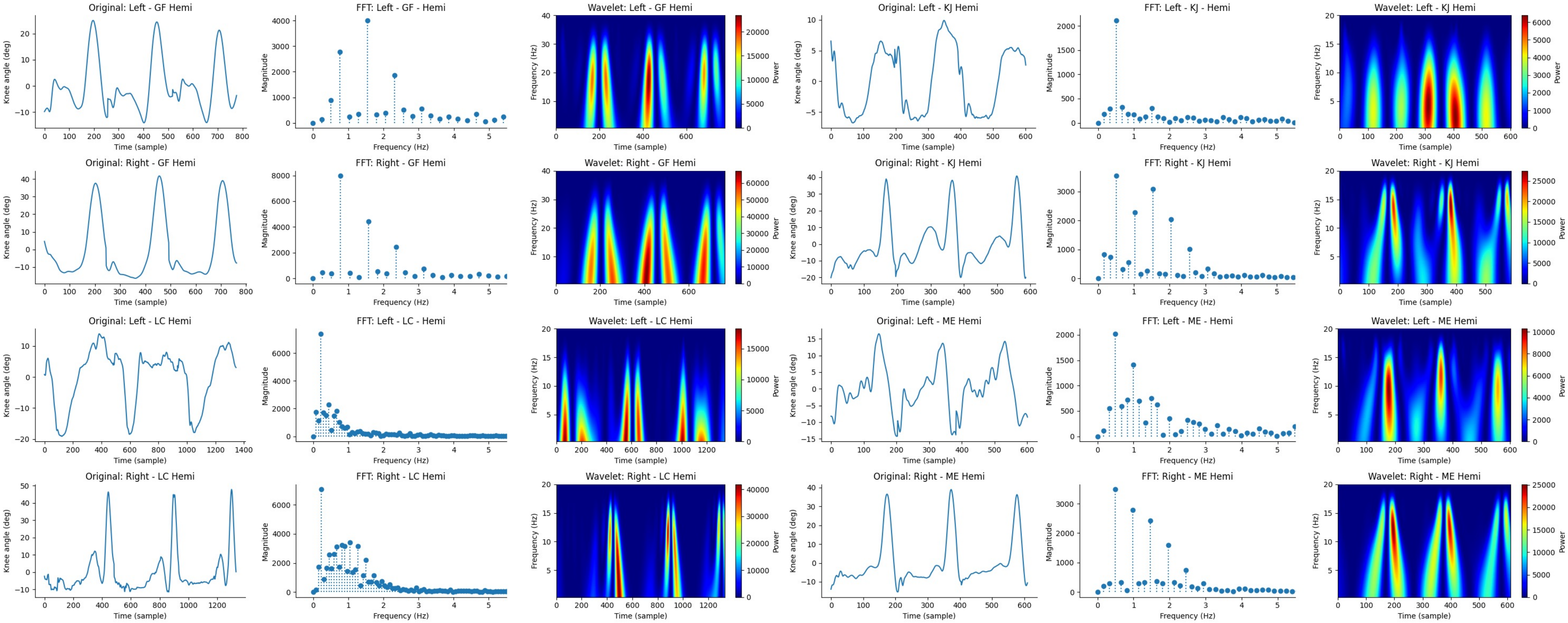


INSTITUT POLYTECHNIQUE DE PARIS

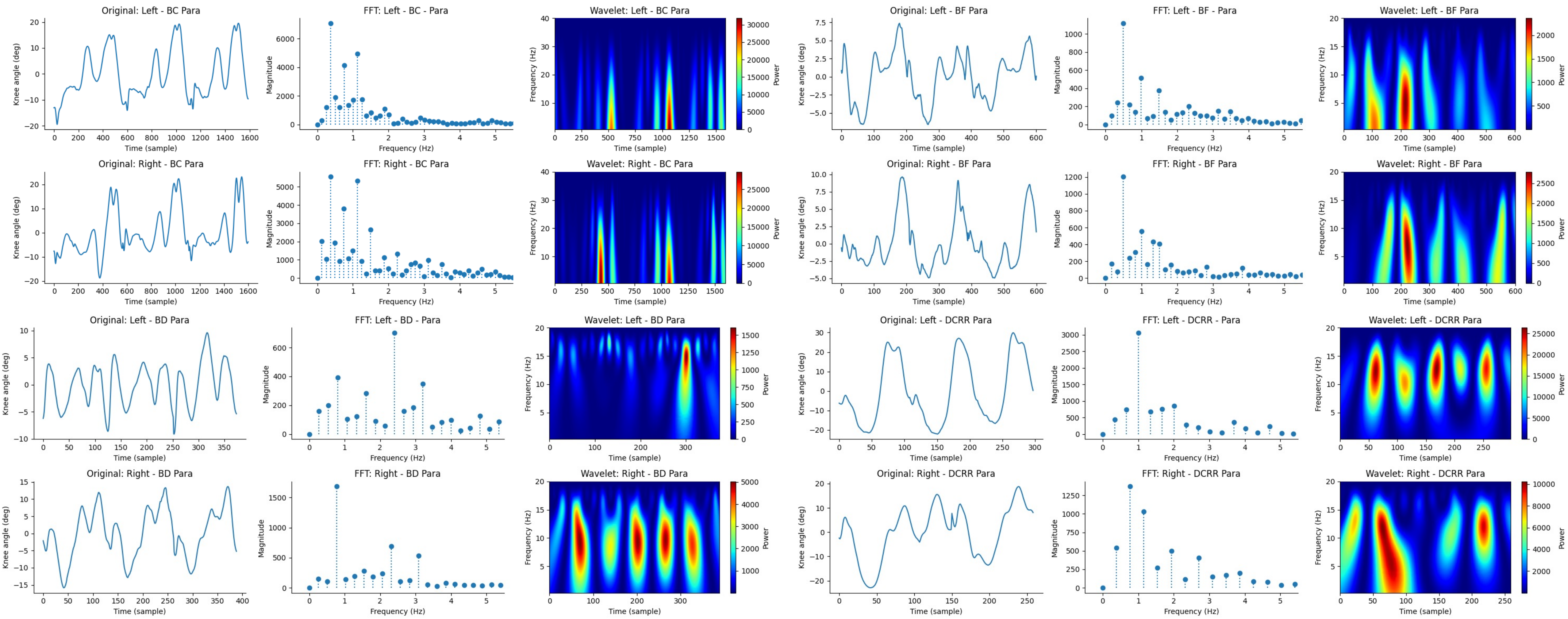
Appendix – Hemiplegia – 12 samples (2/3)



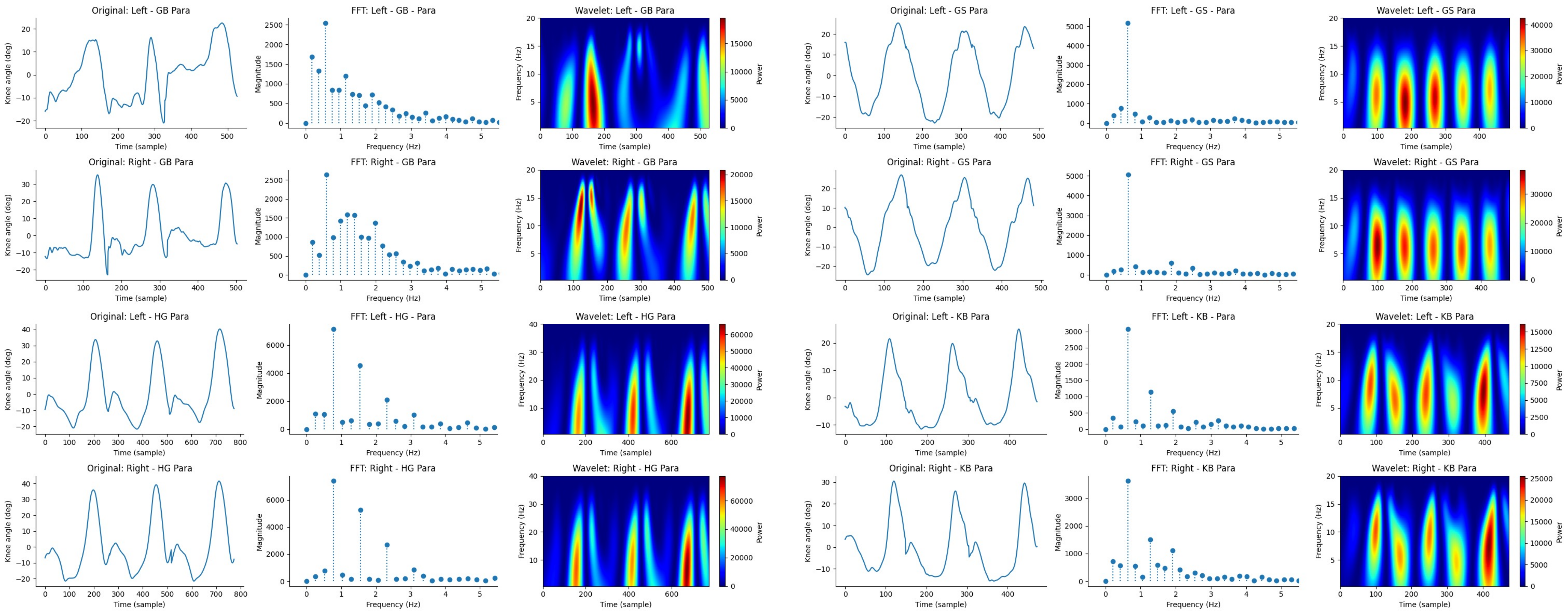
Appendix – Hemiplegia – 12 samples (3/3)



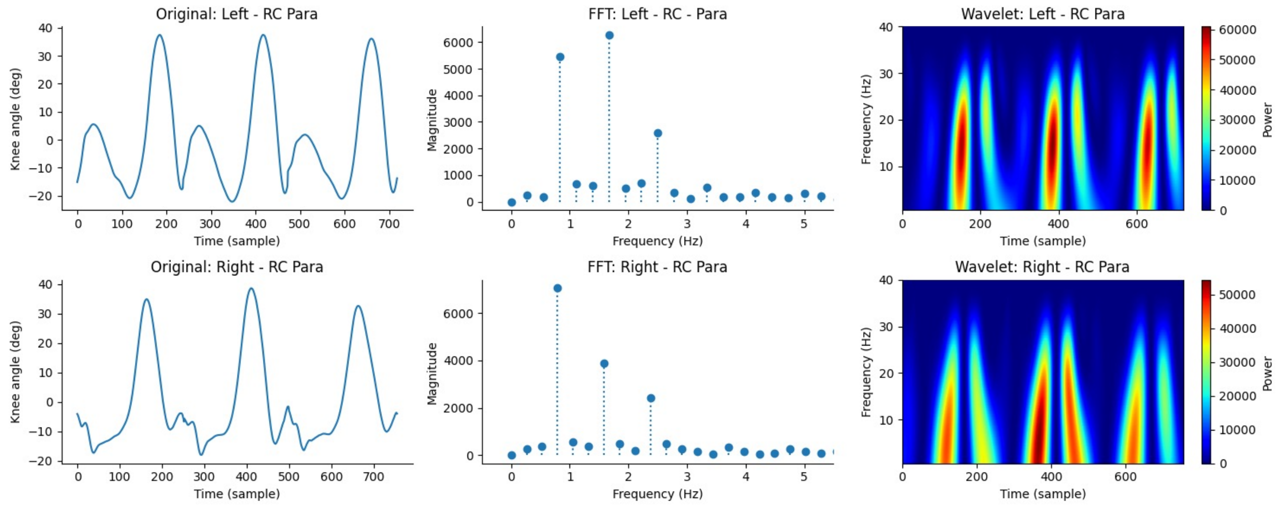
Appendix – Paraplegia (1/3)



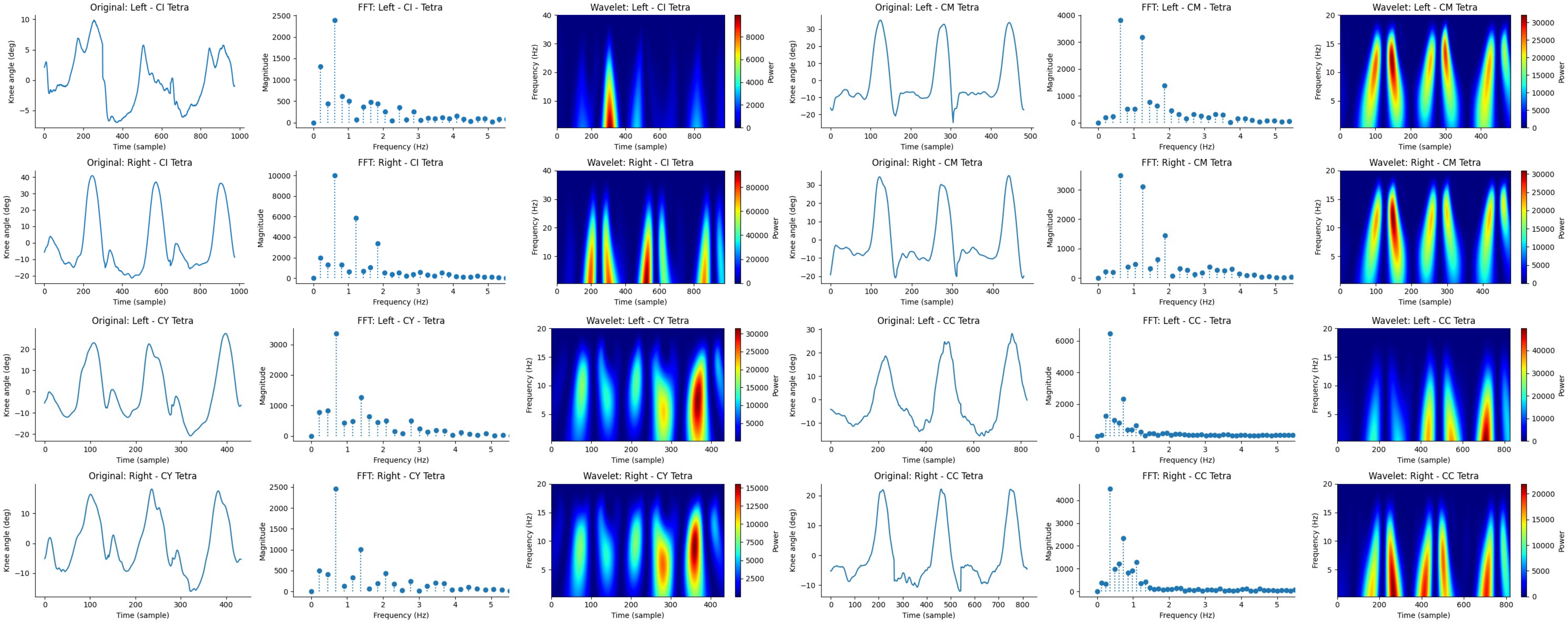
Appendix – Paraplegia (2/3)



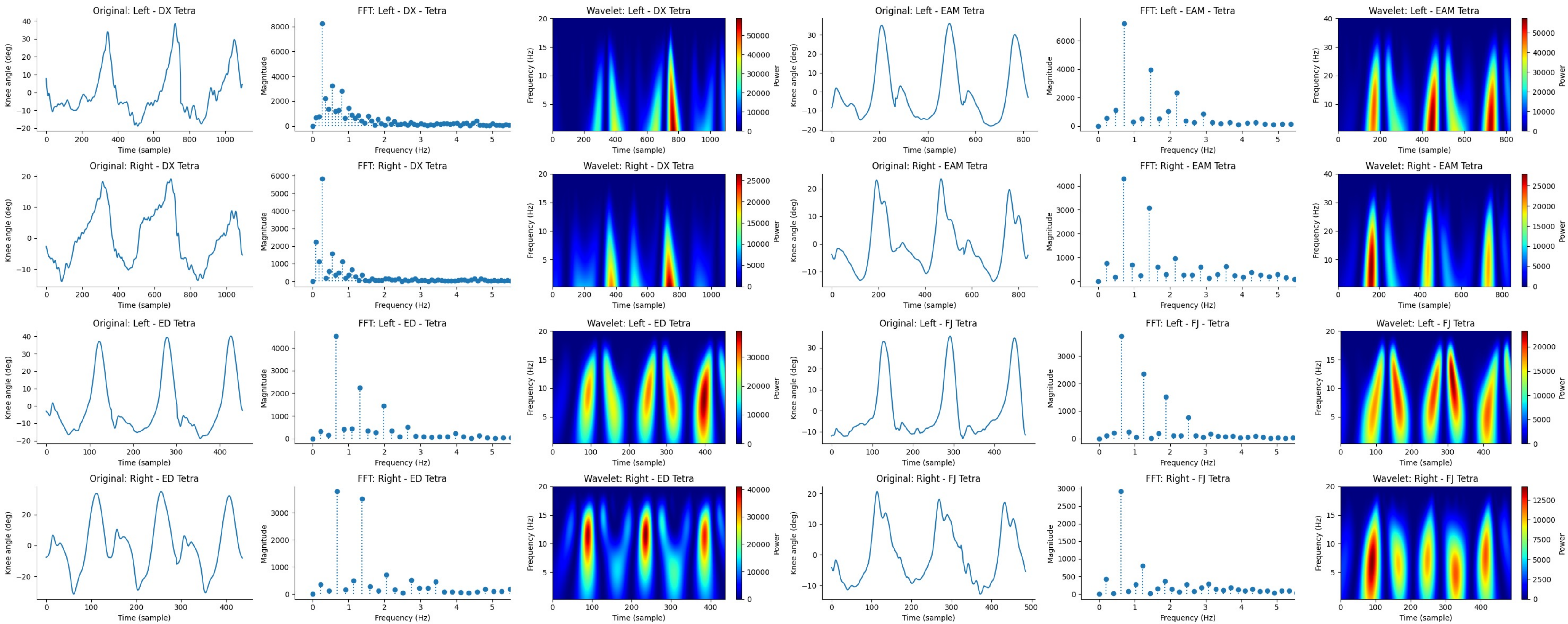
Appendix – Paraplegia (3/3)



Appendix – Tetraplegia (1/3)



Appendix – Tetraplegia (2/3)



INSTITUT POLYTECHNIQUE DE PARIS

Appendix – Tetraplegia (3/3)