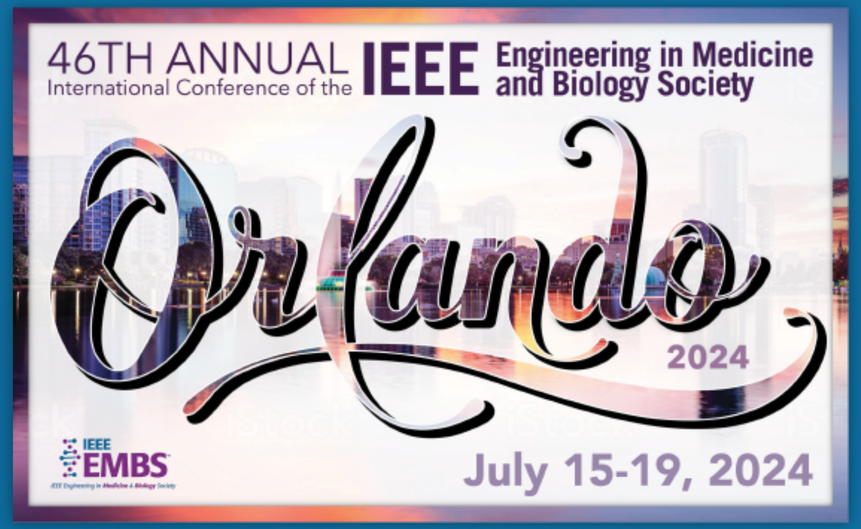
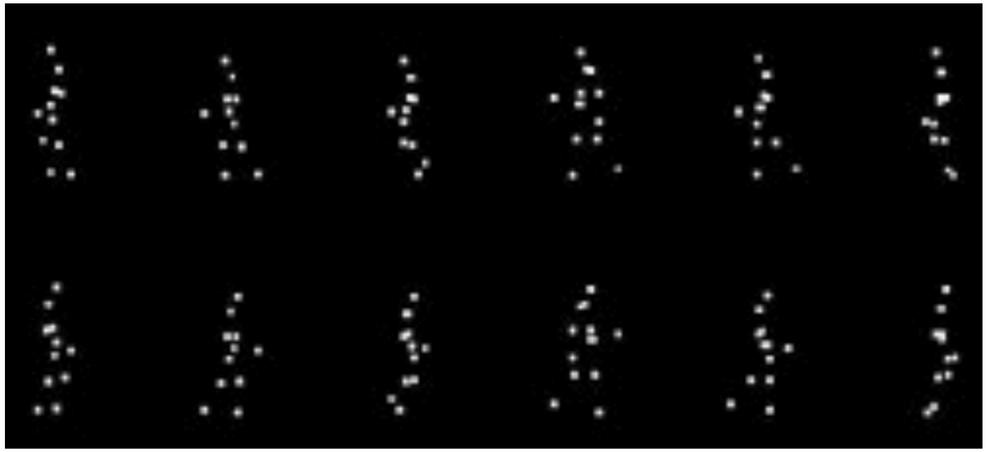


Graph dictionary learning for the study of human motion

Marion Chauveau, Antoine Mazarguil and Laurent Oudre

Université Paris-Saclay, Université Paris Cité, ENS Paris Saclay, CNRS, SSA,
INSERM, Centre Borelli, F-91190, Gif-sur-Yvette, France

Skeleton-based motion data



Film that illustrates the work of Gunnar Johansson on Motion perception, 1971, James Maas, Cornell University.

L. Feng et al, A comparative review of graph convolutional networks for human skeleton-based action recognition, 2022

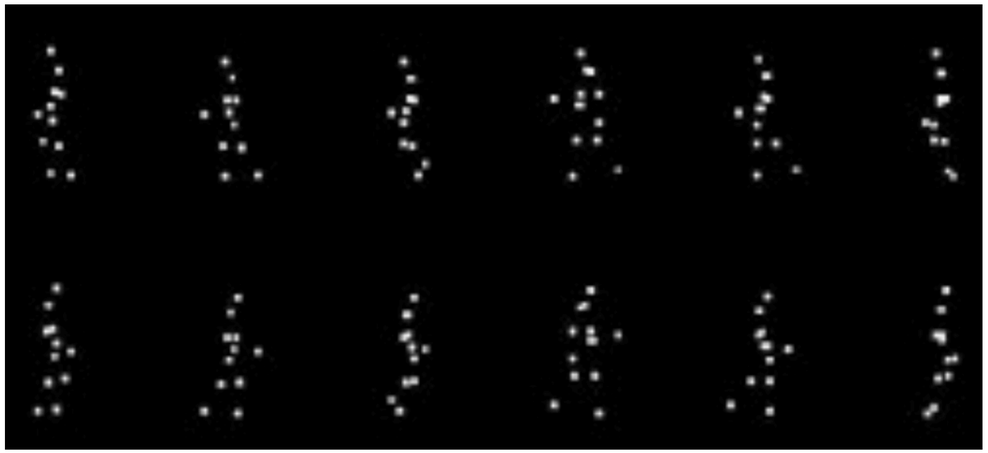
Q. Wang et al, Visual analysis of human motion: A survey on recent advances and applications, 2018

Z. Kertesz et al, 3d motion capture methods for pathological and non-pathological human motion analysis, 2006

S. Tanaka et al, Determination of human motion for rehabilitation based on time-scale transformation, 2007

G. Johansson, Visual perception of biological motion and a model for its analysis, 1973

Skeleton-based motion data



Film that illustrates the work of Gunnar Johansson on Motion perception, 1971, James Maas, Cornell University.

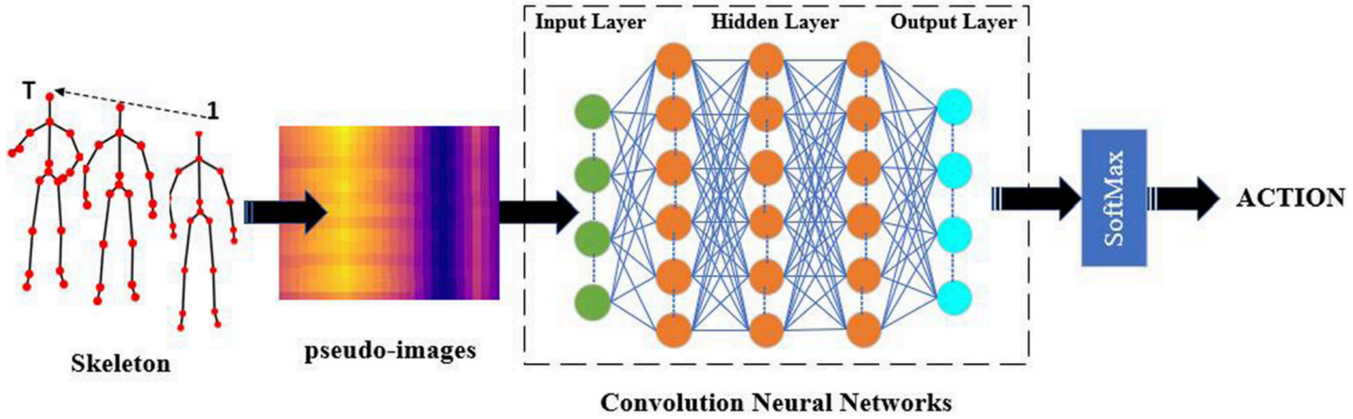


Illustration of a CNN-based human action recognition, L. Feng et al, 2022

L. Feng et al, A comparative review of graph convolutional networks for human skeleton-based action recognition, 2022

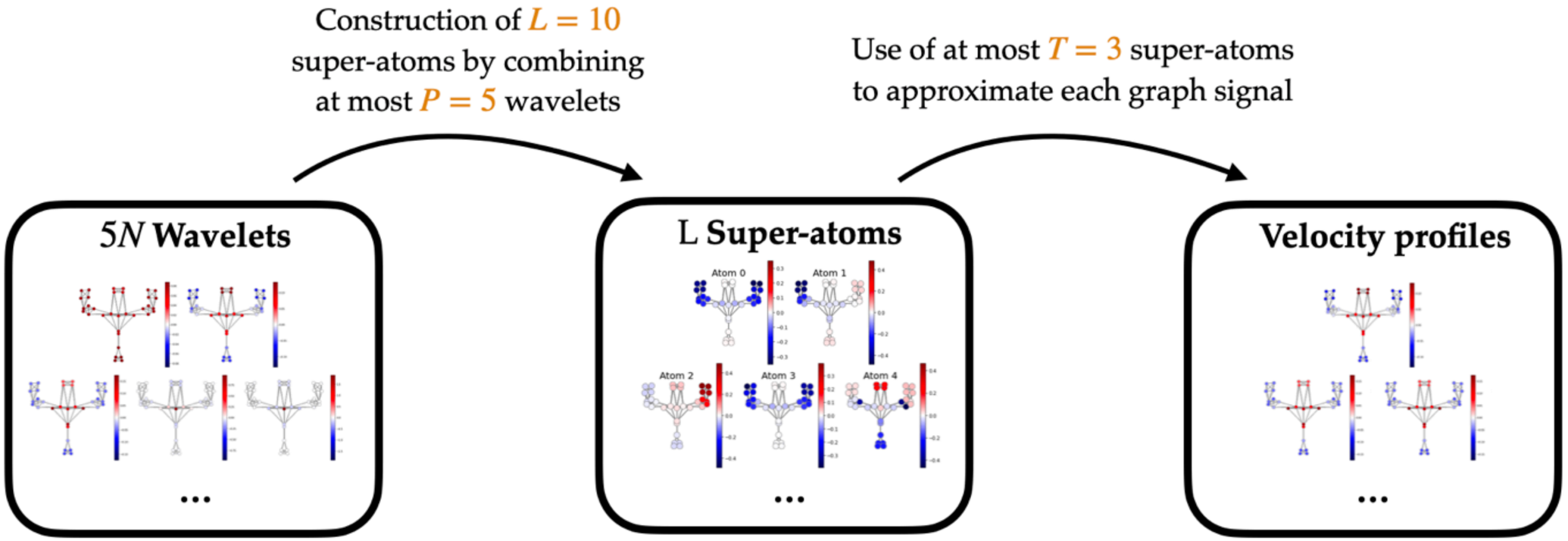
Q. Wang et al, Visual analysis of human motion: A survey on recent advances and applications, 2018

Z. Kertesz et al, 3d motion capture methods for pathological and non-pathological human motion analysis, 2006

S. Tanaka et al, Determination of human motion for rehabilitation based on time-scale transformation, 2007

G. Johansson, Visual perception of biological motion and a model for its analysis, 1973

Graph dictionary learning approach



Upper limbs movements: Arm-CODA

Subjects :  $\times 16$

 Healthy subjects

S. W. Combettes et al, Arm-CODA: A Data Set of Upper-limb Human Movement During Routine Examination, 2024

Upper limbs movements: Arm-CODA

Subjects : × 16

→ Healthy subjects

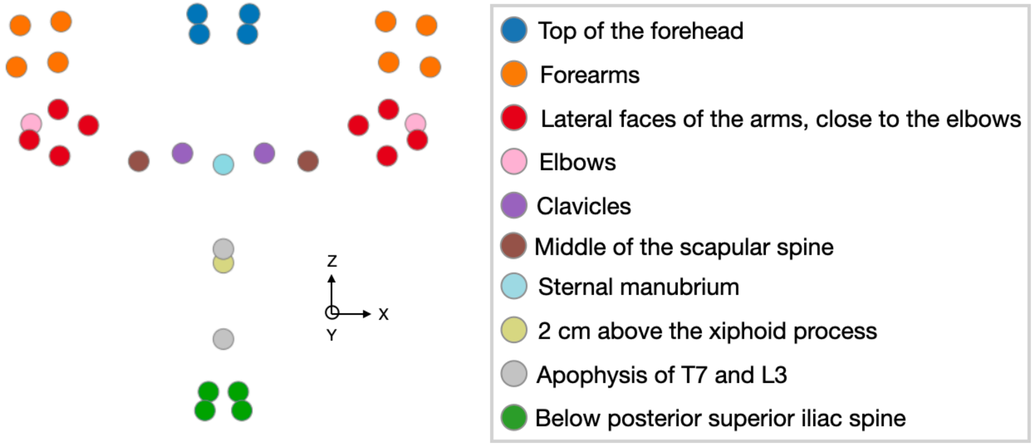
Sensors : × 34

→ Cartesian Optoelectronic Dynamic Anthropometer (CODA)

→ 3D positions data in millimeters at 100 Hz

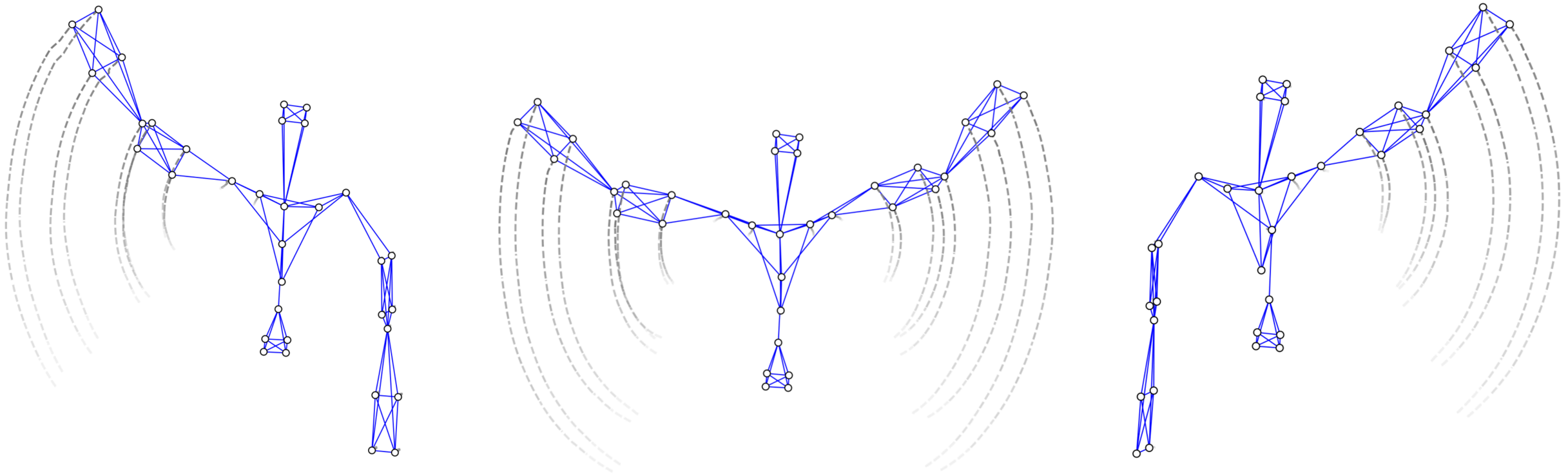
→ System of 6 depth Cameras

View in the Sagittal plane



S. W. Combettes et al, Arm-CODA: A Data Set of Upper-limb Human Movement During Routine Examination, 2024

Elevation movements in the scapular plane: right arm, left arm and bilateral elevation

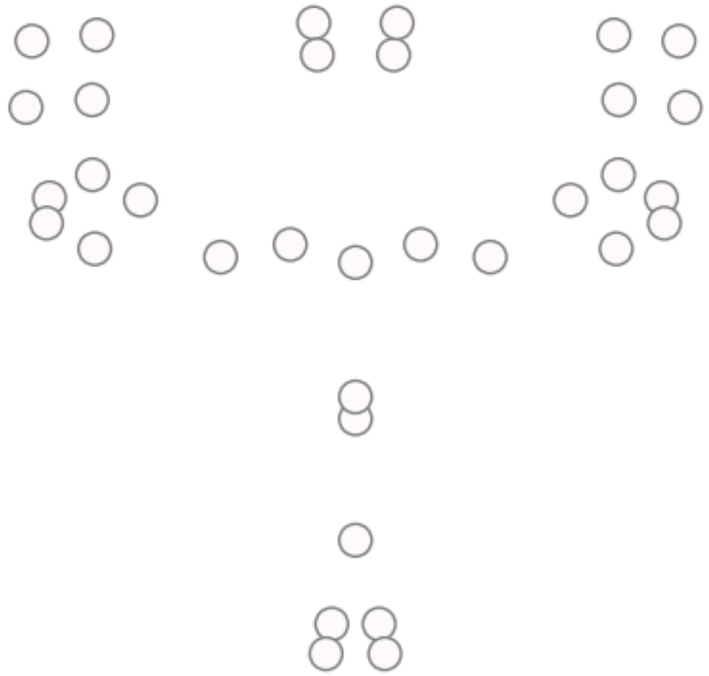


S. W. Combettes et al, Arm-CODA: A Data Set of Upper-limb Human Movement During Routine Examination, 2024

Graph

Graph signal

Nodes



A. Ortega et al, Graph signal processing: Overview, challenges and applications, 2017

X. Dong et al. Graph signal processing for machine learning: A review and new perspectives,

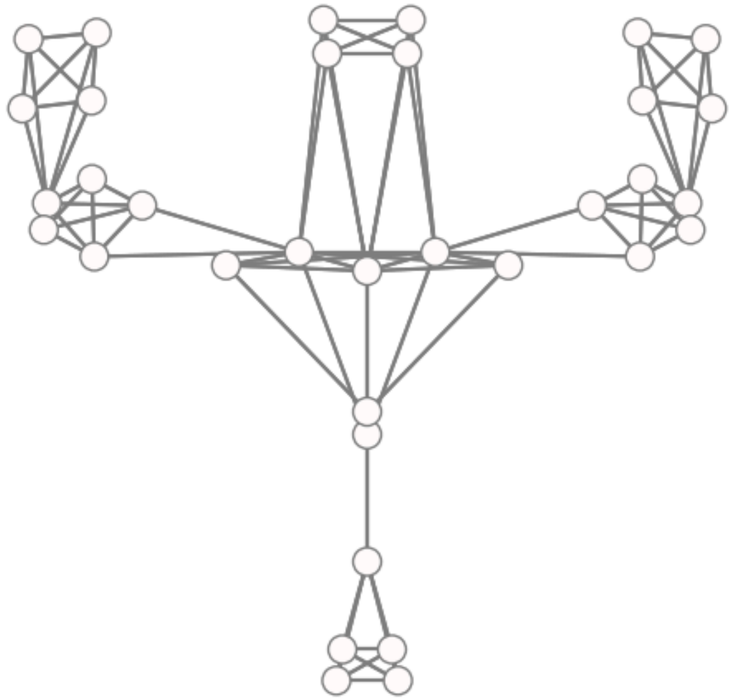
Graph

Graph signal

Nodes

+

Edges



A. Ortega et al, Graph signal processing: Overview, challenges and applications, 2017
X. Dong et al. Graph signal processing for machine learning: A review and new perspectives,

Graph

Graph signal

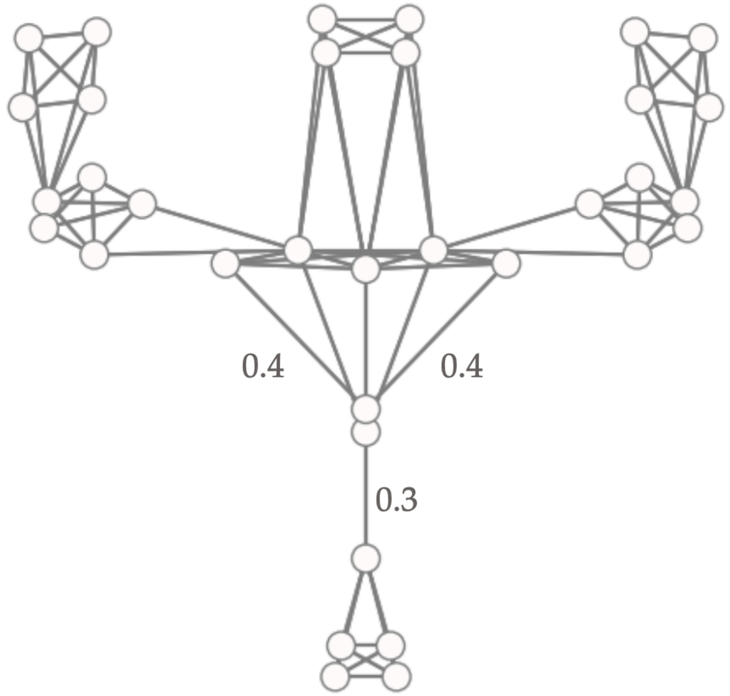
Nodes

+

Edges

+

Weights



A. Ortega et al, Graph signal processing: Overview, challenges and applications, 2017
X. Dong et al. Graph signal processing for machine learning: A review and new perspectives,

Graph

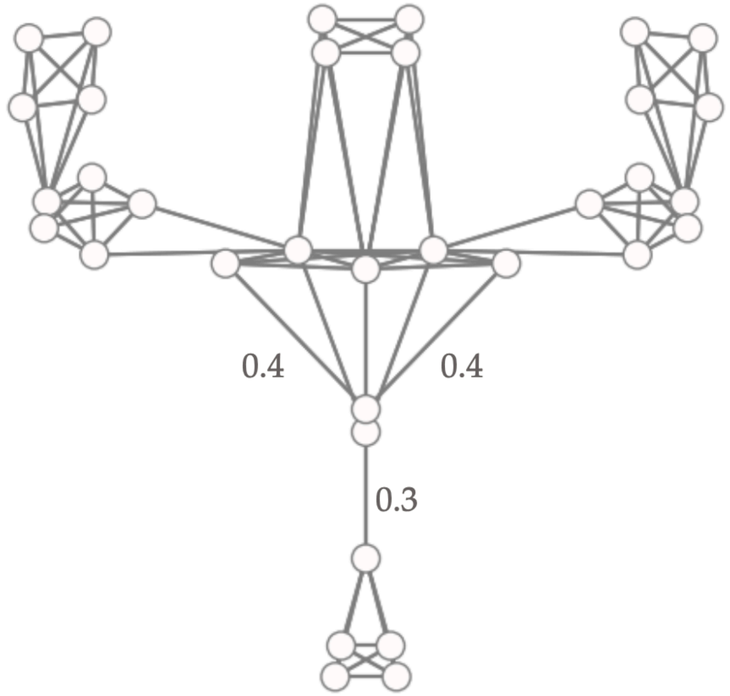
Nodes

+

Edges

+

Weights

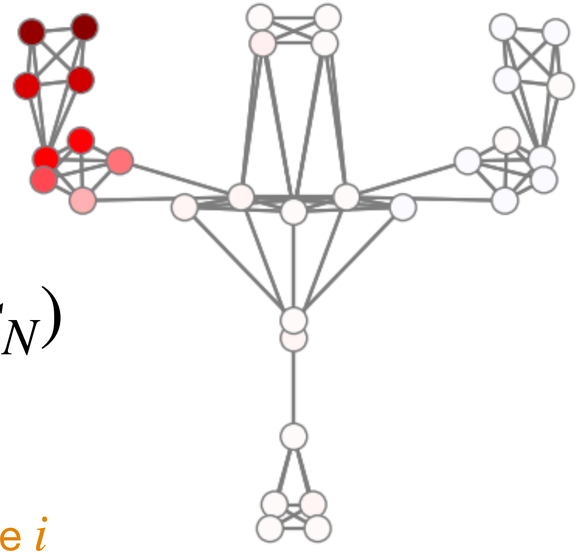


Graph signal

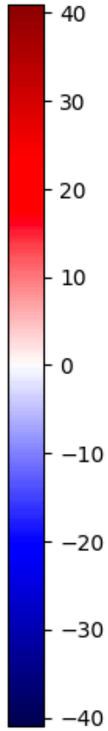
A value is associated to each node

$$\mathbf{s} = (s_1, \ldots, s_i, \ldots s_N)$$

Value associated to node i



N nodes



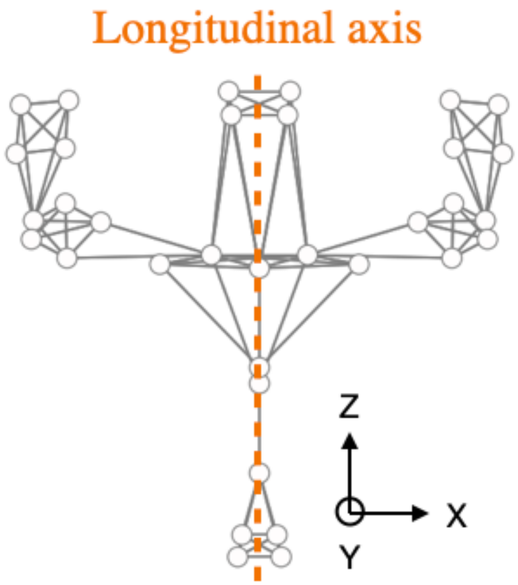
A. Ortega et al, Graph signal processing: Overview, challenges and applications, 2017

X. Dong et al. Graph signal processing for machine learning: A review and new perspectives,

Graph construction

Close nodes

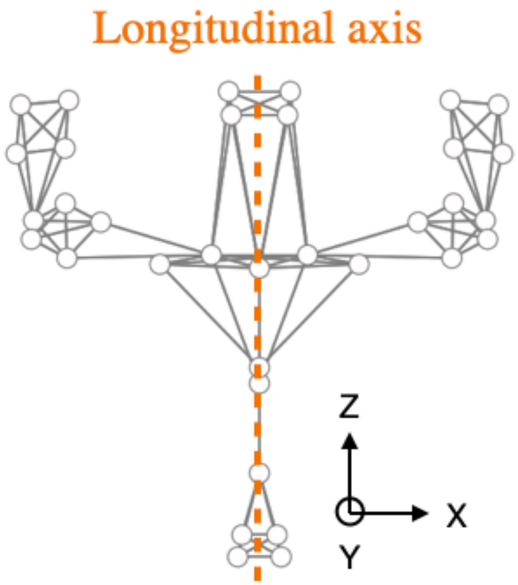
Close sensors
along the body



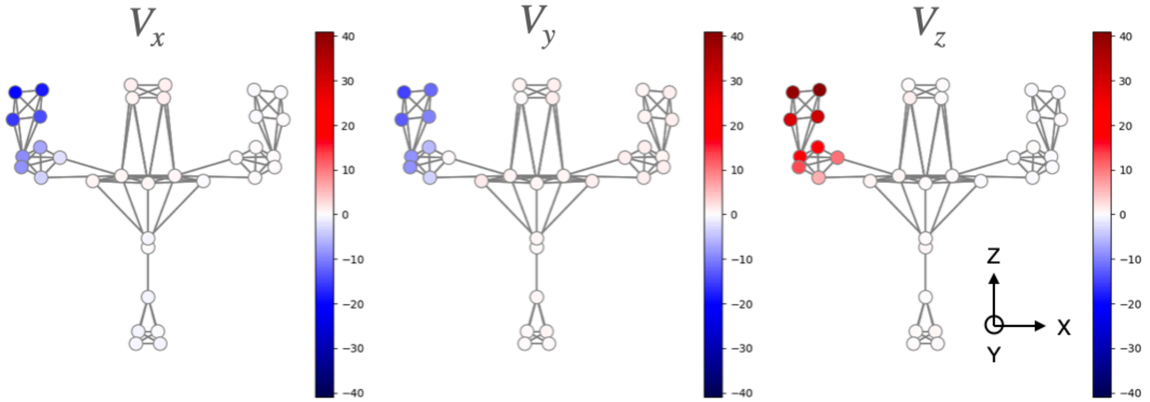
Graph construction

Close nodes

Close sensors along the body



Velocity profiles



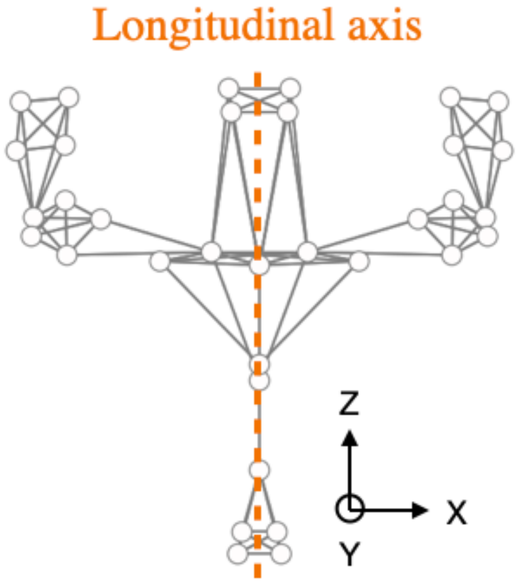
Velocity profile directed along the y-axis at a given instant

Graph construction

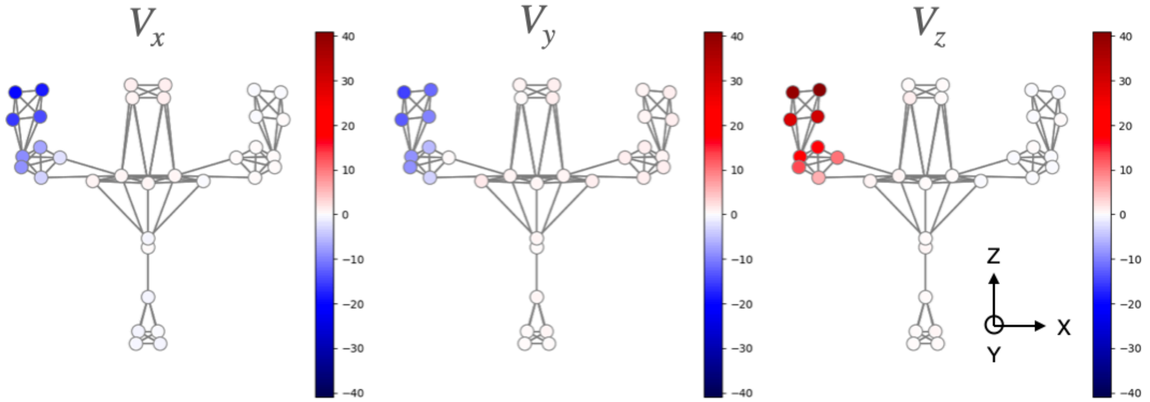
Close nodes

⬇

Close sensors along the body



Velocity profiles



Velocity profile directed along the y-axis at a given instant

Dictionary learning

$$\mathbf{y} \approx \sum_{l=1}^{L} x_l \, \mathbf{d}^l$$

Signal **Activations** **Atoms**

$$\underset{X,\, D}{\text{argmin}} \quad ||Y - DX||_F^2$$

$$\text{s.t.} \quad ||x_i||_0 \leq T \qquad \forall i$$

Y. Yankelevsky et al, Dictionary learning for high dimensional graph signals, 2018

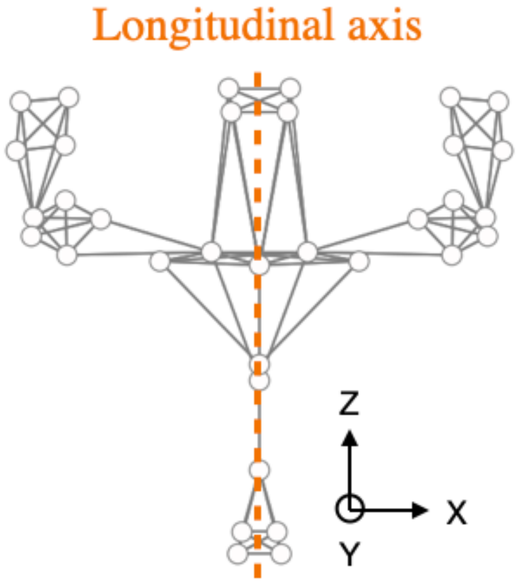
R. Rubinstein et al, Double sparsity: Learning sparse dictionaries for sparse signal approximation, 2010

Graph construction

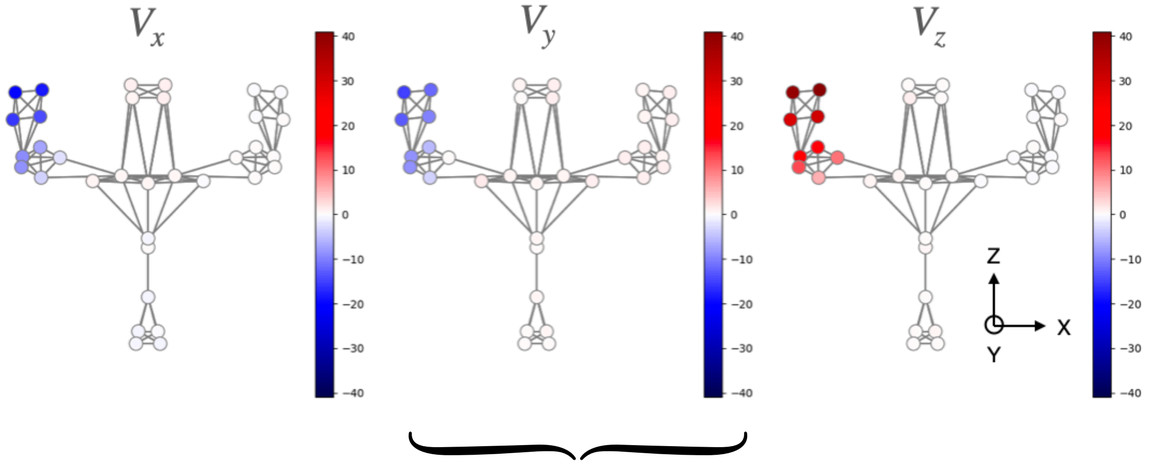
Close nodes

⬇

Close sensors along the body



Velocity profiles



Velocity profile directed along the y-axis at a given instant

Double sparsity

$$\mathbf{y} \approx \sum_{l=1}^{L} x_l \, \mathbf{d}^l$$

Signal — Activations — ~~Atoms~~ Super-atoms

$$\mathbf{d}^l \approx \sum_{k=1}^{K} a_k \, \phi^k$$

Super-atom — Atom

Y. Yankelevsky et al, Dictionary learning for high dimensional graph signals, 2018

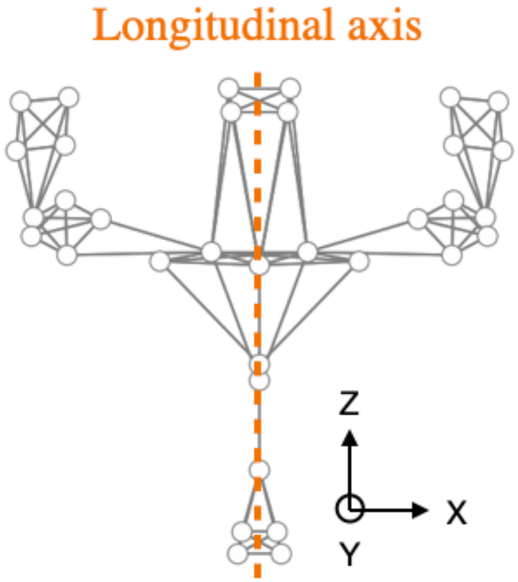
R. Rubinstein et al, Double sparsity: Learning sparse dictionaries for sparse signal approximation, 2010

Graph construction

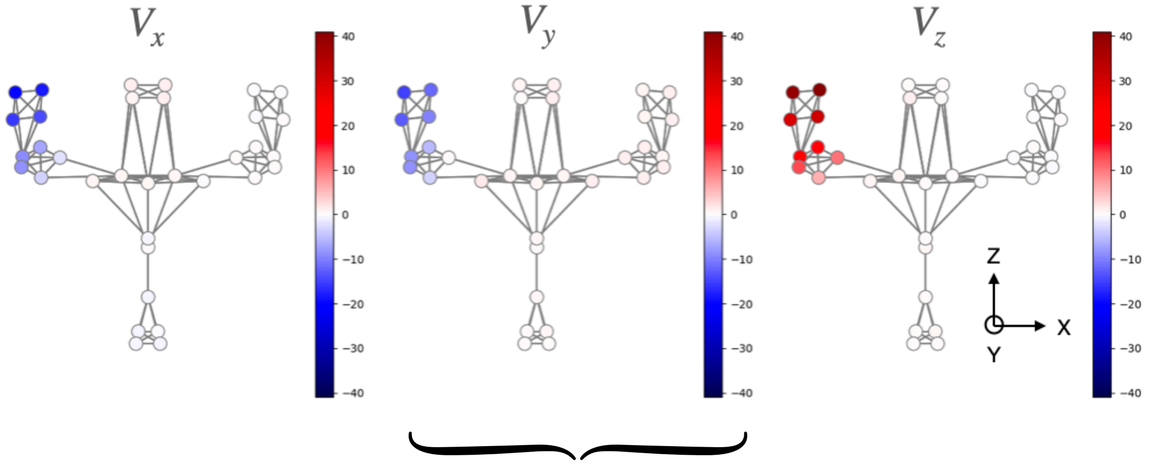
Close nodes

⬇

Close sensors along the body



Velocity profiles



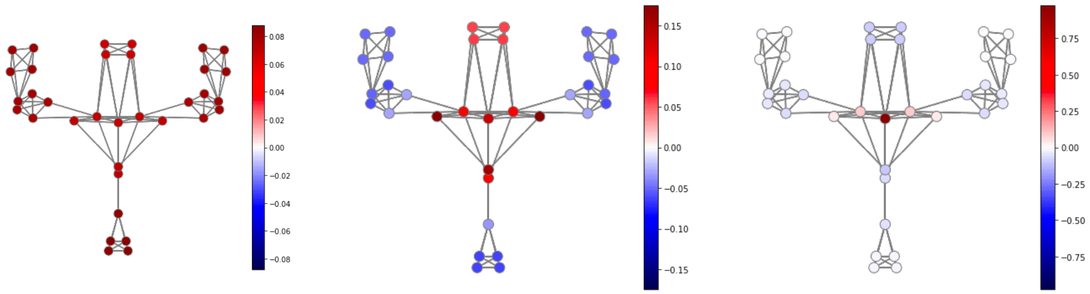
Velocity profile directed along the y-axis at a given instant

Double sparsity

$$\mathbf{y} \approx \sum_{l=1}^{L} x_l \, \mathbf{d}^l \qquad \mathbf{d}^l \approx \sum_{k=1}^{K} a_k \, \phi^k$$

Signal — Activations — ~~Atoms~~ Super-atoms — Super-atom — Atom

Mexican Hat wavelets



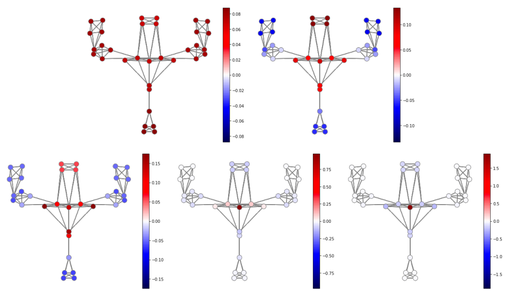
Each wavelet is characterized by a scale and a node on which it is centered

Y. Yankelevsky et al, Dictionary learning for high dimensional graph signals, 2018

R. Rubinstein et al, Double sparsity: Learning sparse dictionaries for sparse signal approximation, 2010

DSMH Method

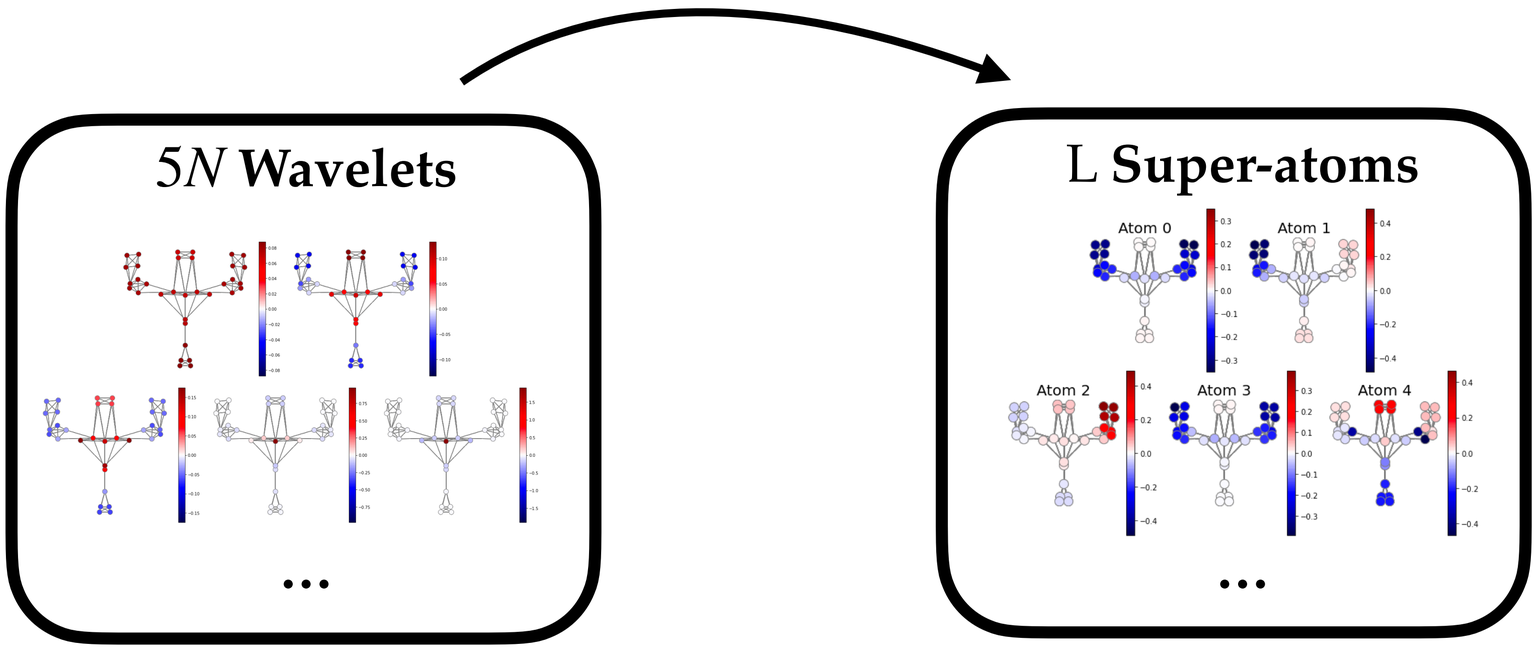
$5N$ Wavelets



...

DSMH Method

Construction of $L = 10$
super-atoms by combining
at most $P = 5$ wavelets

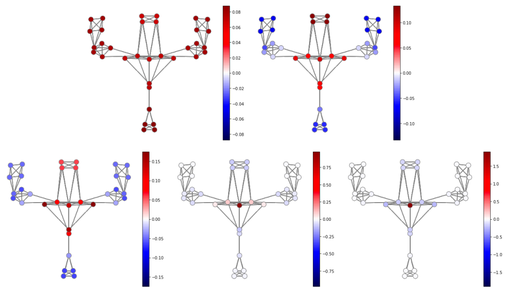


DSMH Method

Construction of $L = 10$ super-atoms by combining at most $P = 5$ wavelets

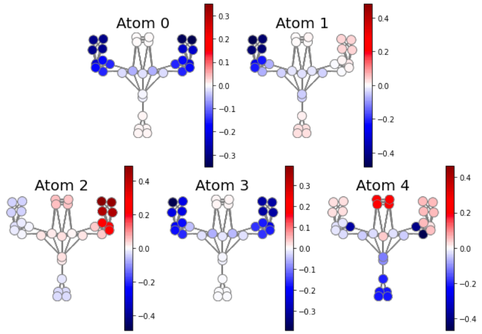
Use of at most $T = 3$ super-atoms to approximate each graph signal

$5N$ Wavelets



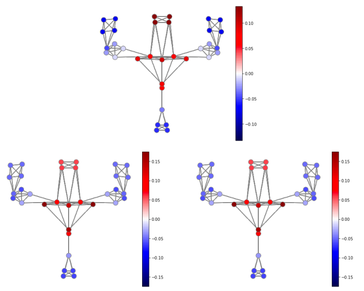
...

L Super-atoms



...

Velocity profiles



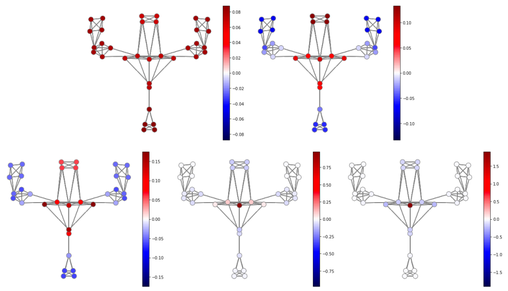
...

DSMH Method

Construction of $L = 10$ super-atoms by combining at most $P = 5$ wavelets

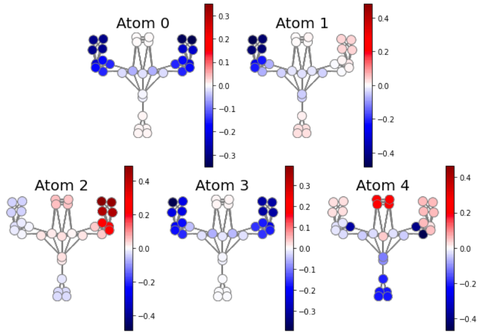
Use of at most $T = 3$ super-atoms to approximate each graph signal

$5N$ Wavelets



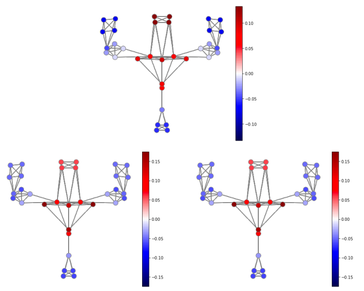
...

L Super-atoms



...

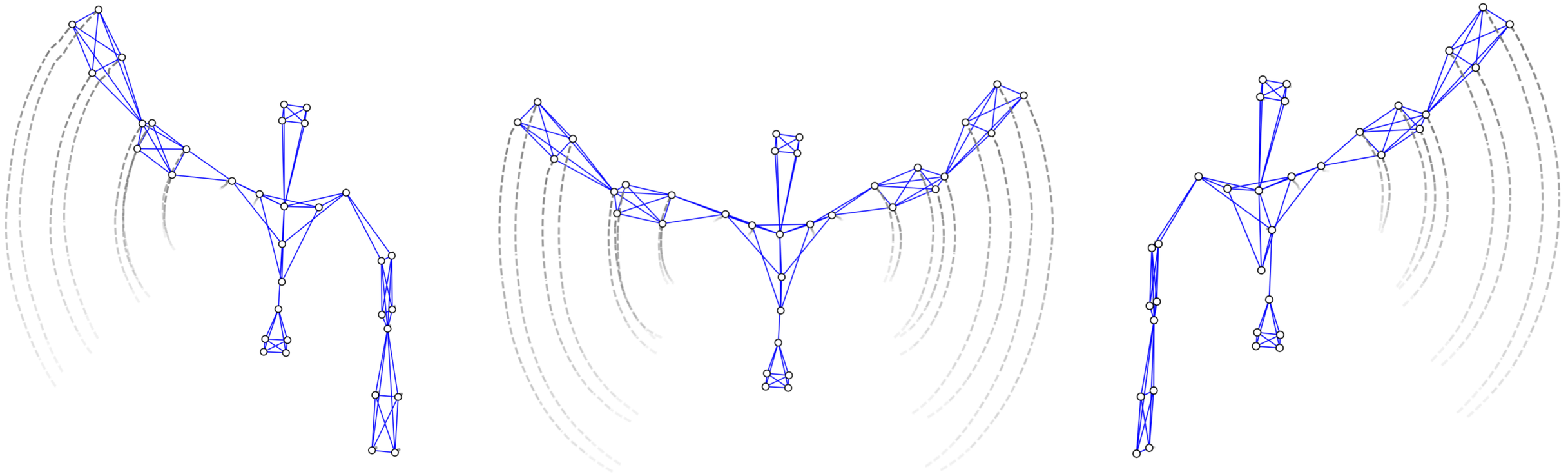
Velocity profiles



...

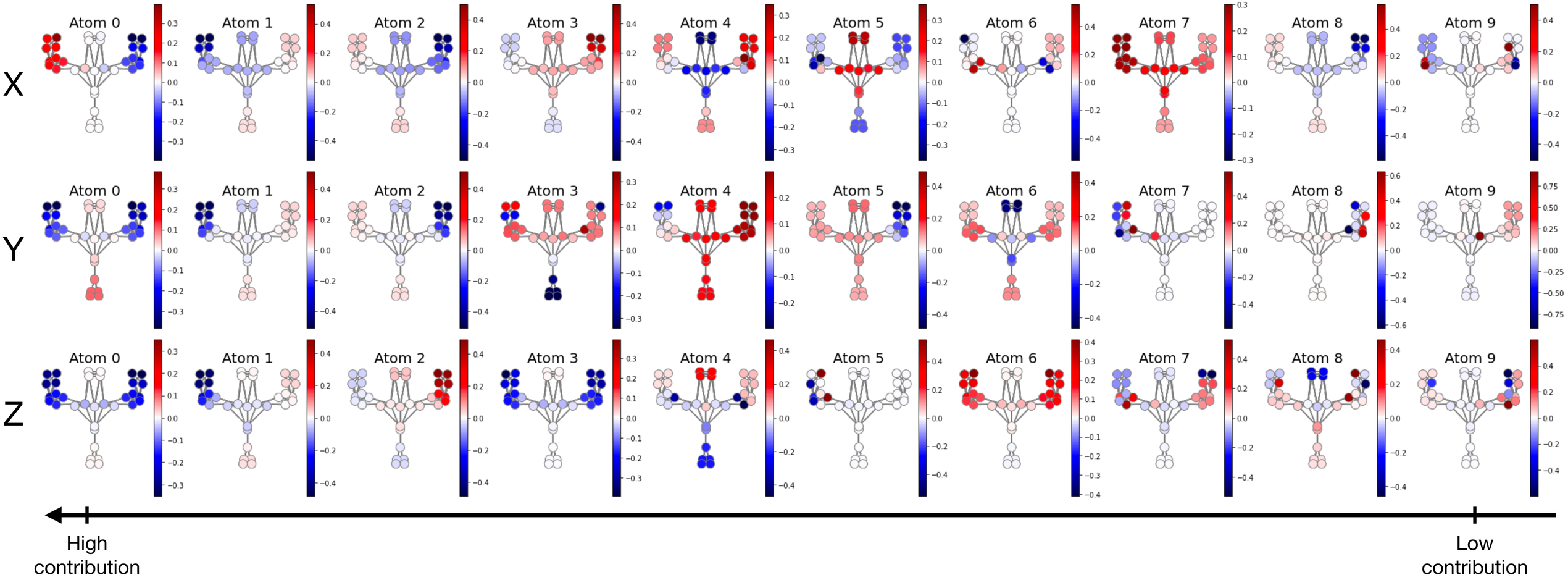
$L = 10 \quad P = 5 \quad T = 3$ → 80 % signal reconstruction

Elevation movements in the scapular plane: right arm, left arm and bilateral elevation

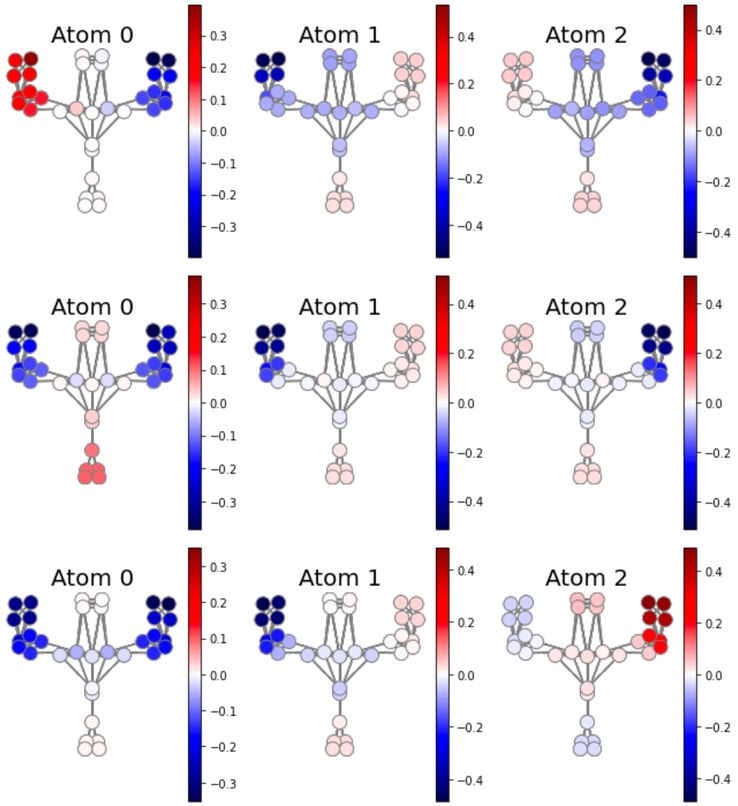


S. W. Combettes et al, Arm-CODA: A Data Set of Upper-limb Human Movement During Routine Examination, 2024

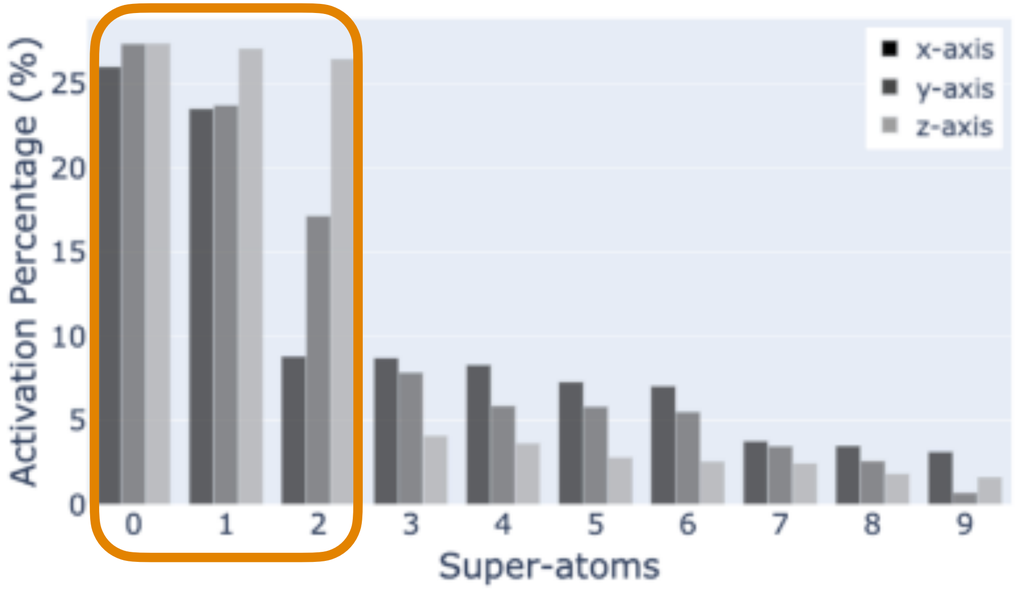
DSMH dictionary



DSMH dictionary



Elevation movements in the scapular plane : right arm, left arm and bilateral elevation



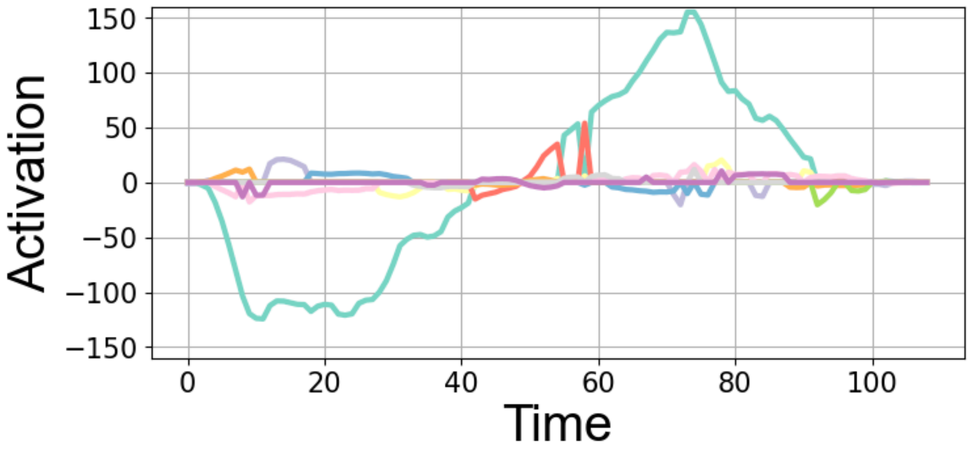
High contribution

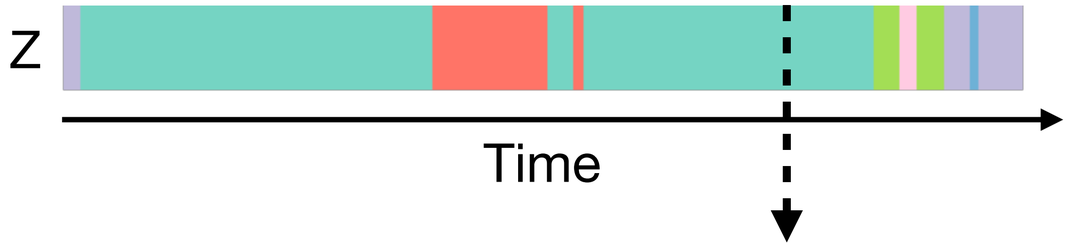
Low contribution

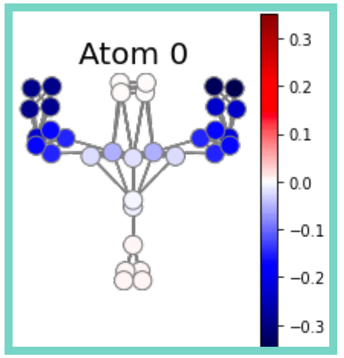
Timelines

Most activated super-atom over time

Bilateral elevation, Z-axis





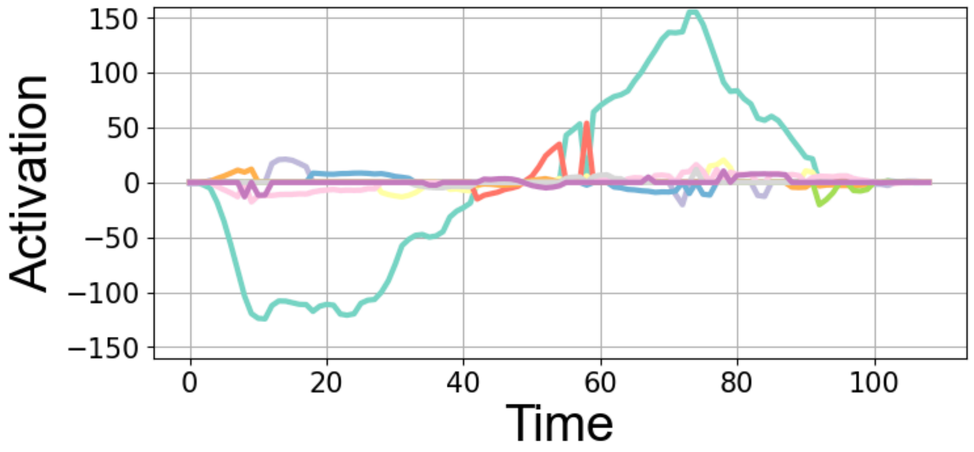


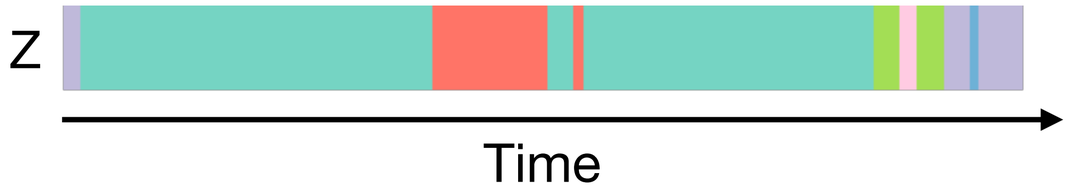
Super-atoms

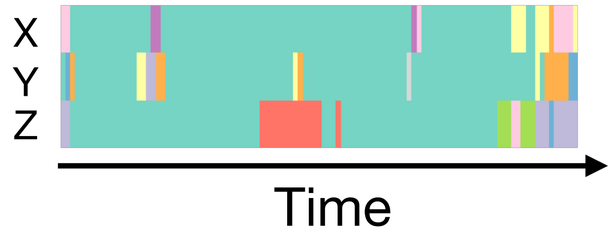
Timelines

Most activated super-atom over time

Bilateral elevation, Z-axis

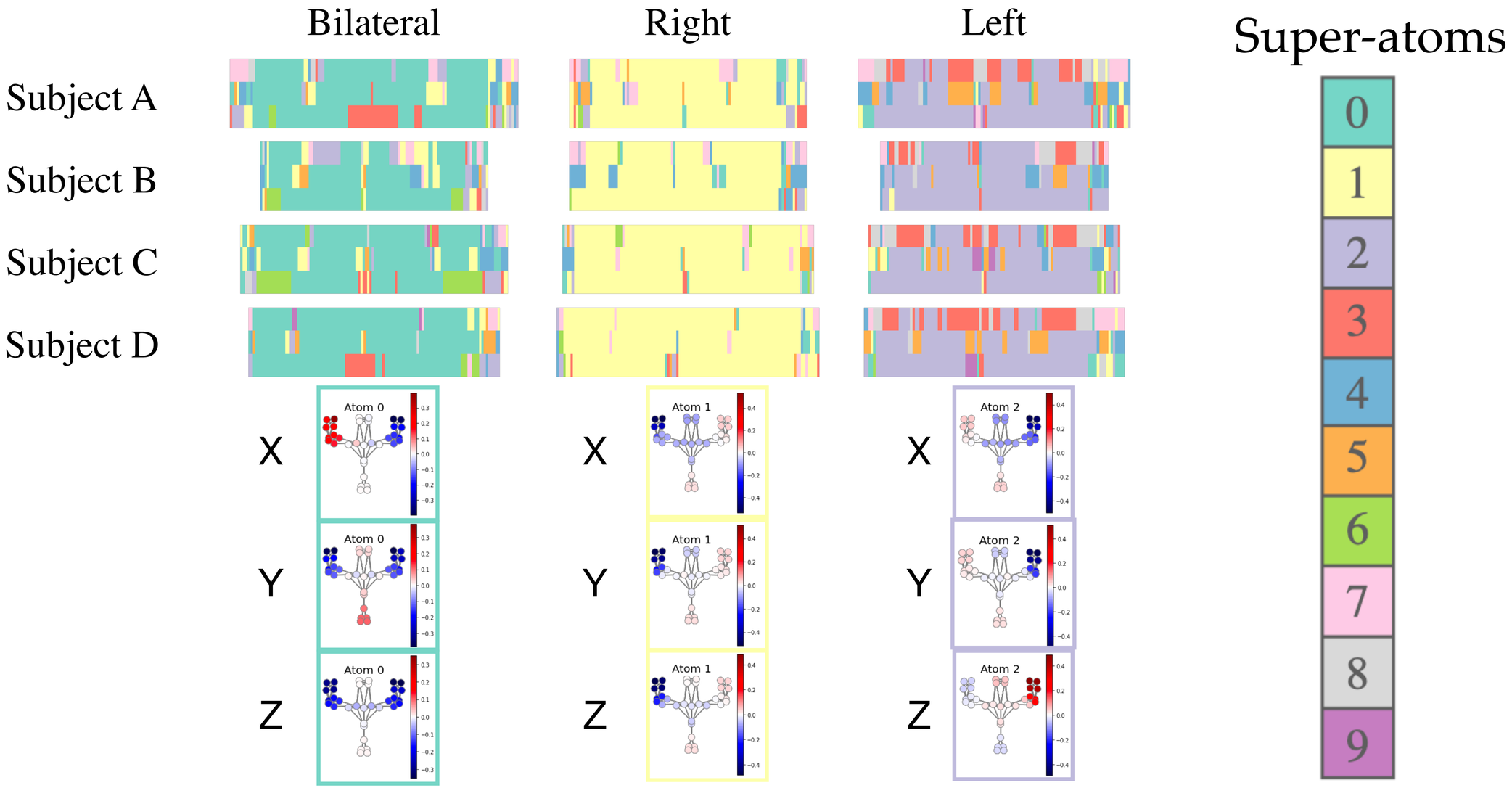






Super-atoms

Timelines 1st activation



Zoom on 1st activation Timelines

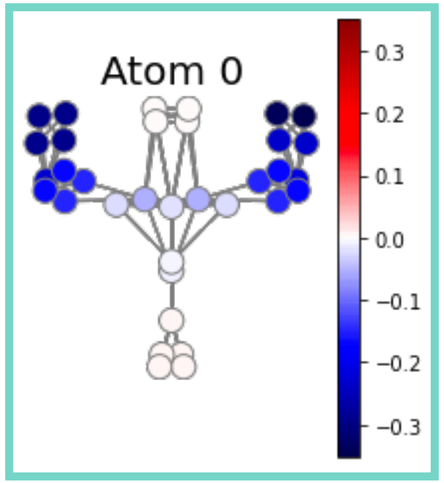
Bilateral, z-axis

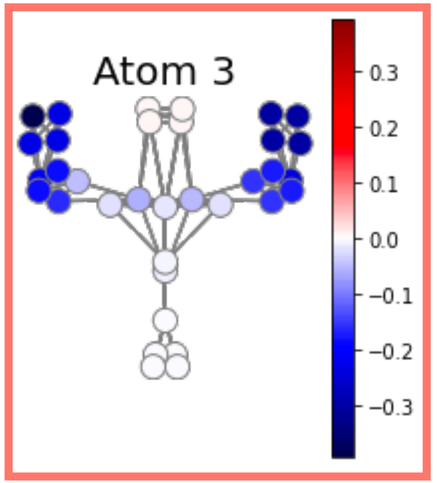
Subject A

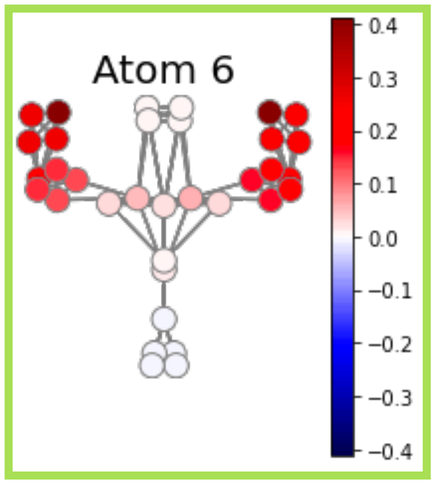
Subject B

Subject C

Subject D



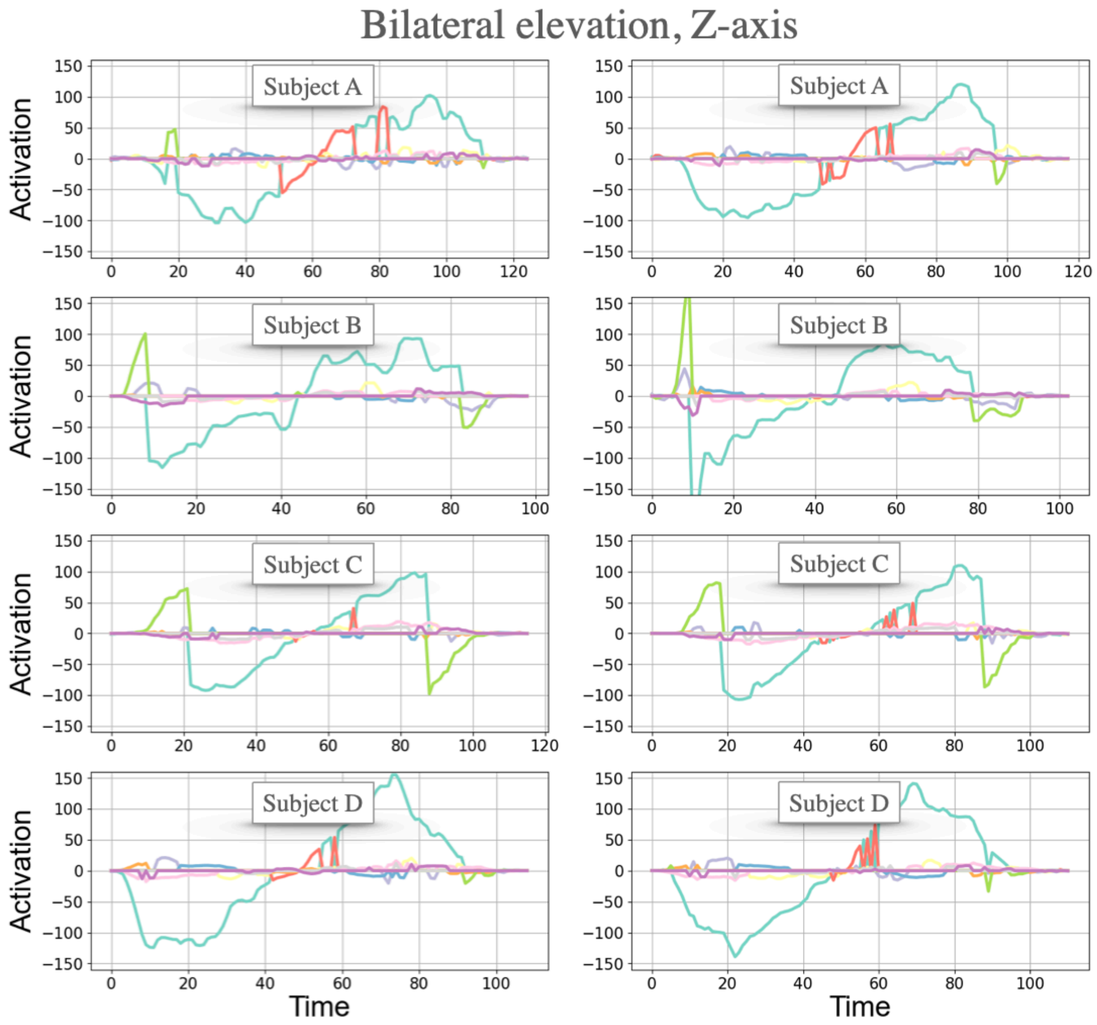




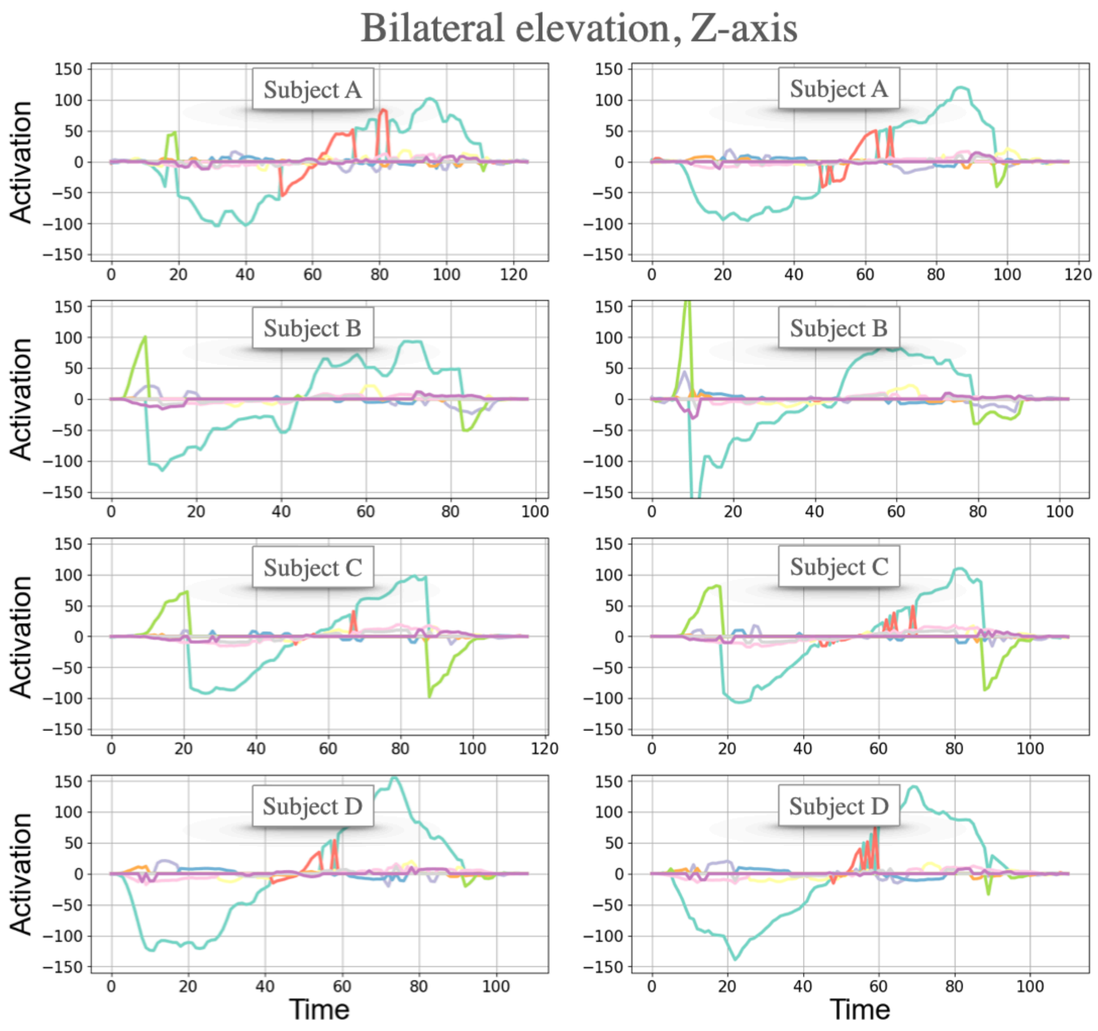
→ Different movement strategies

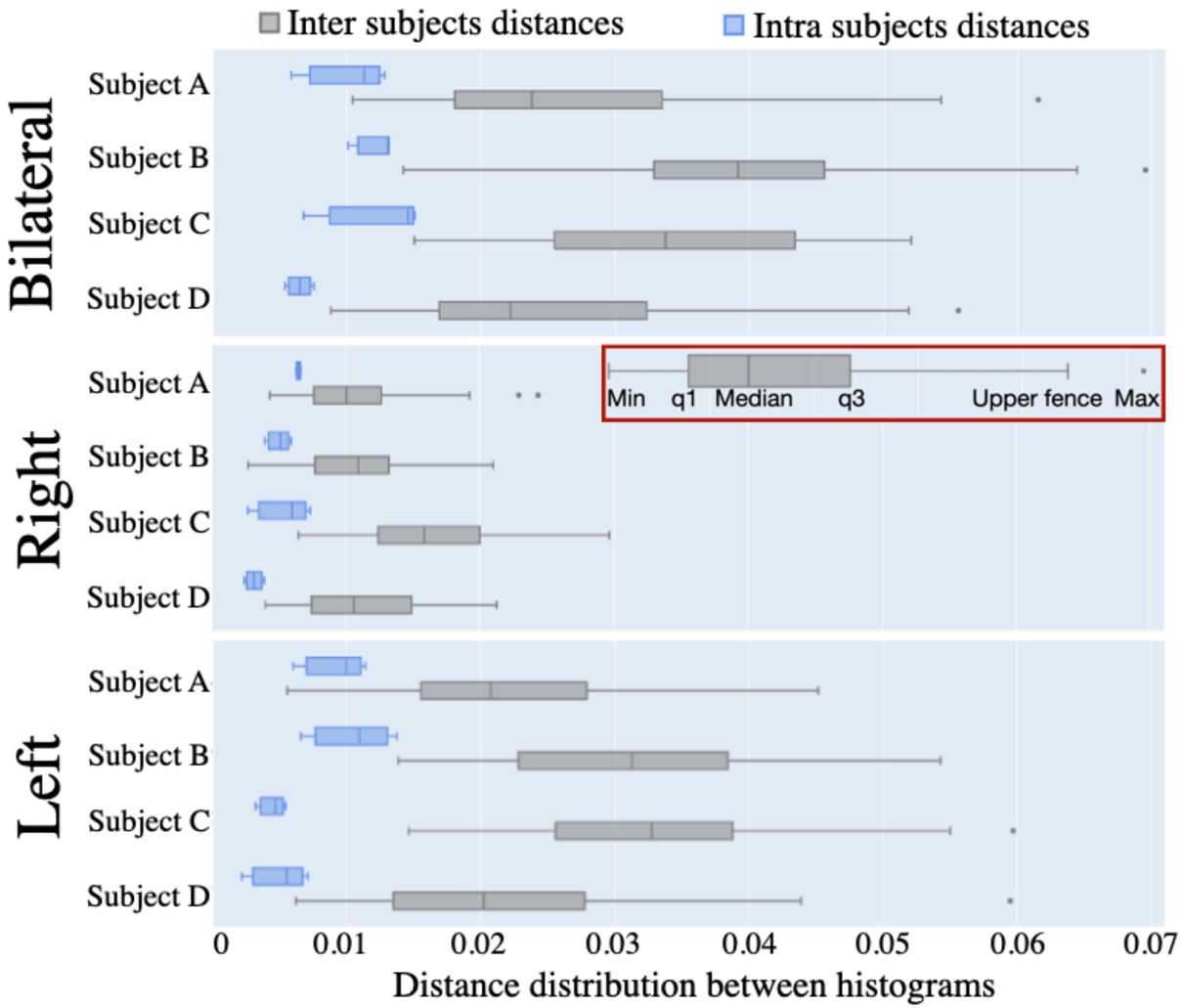
Super-atoms

Inter / Intra subject comparison

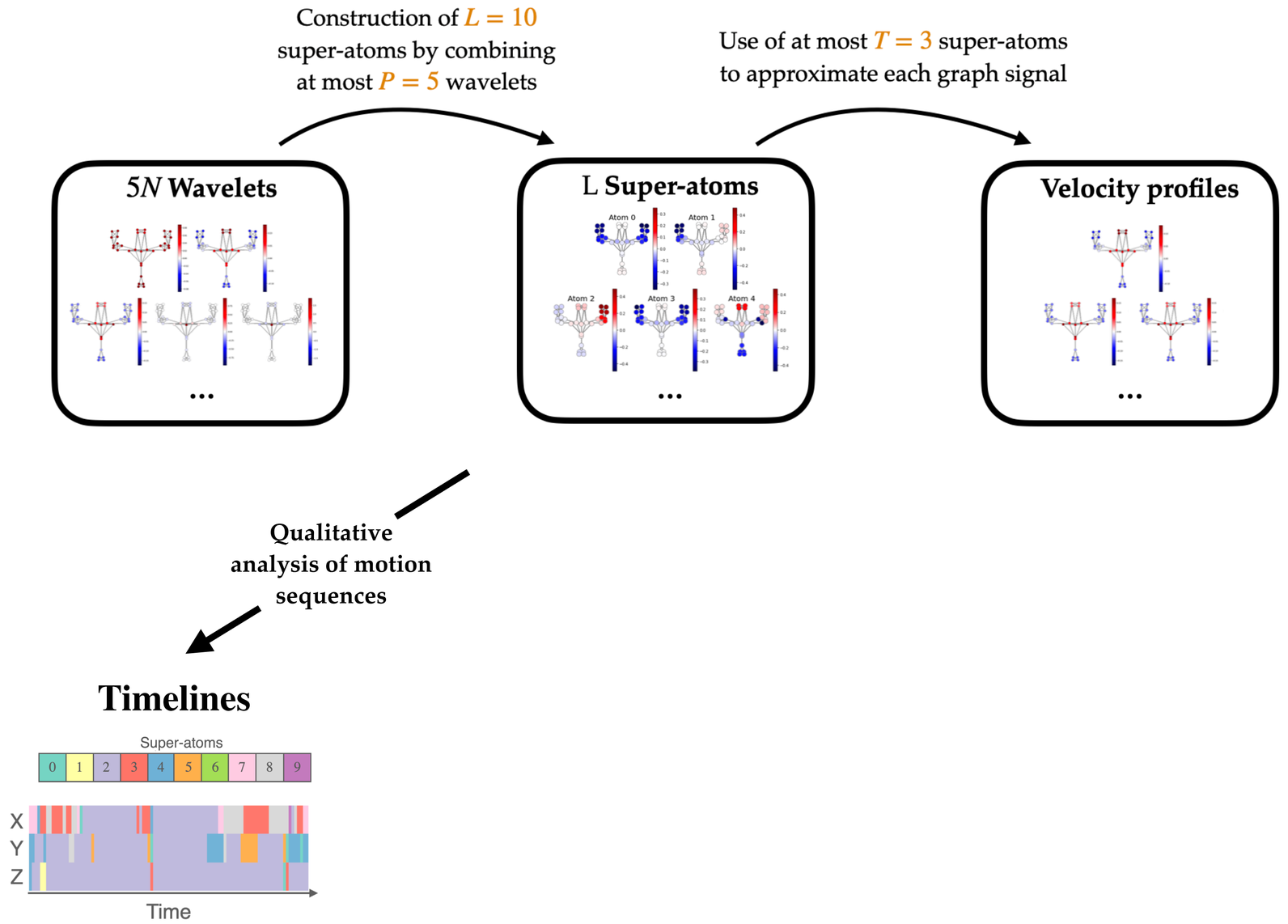


Inter / Intra subject comparison

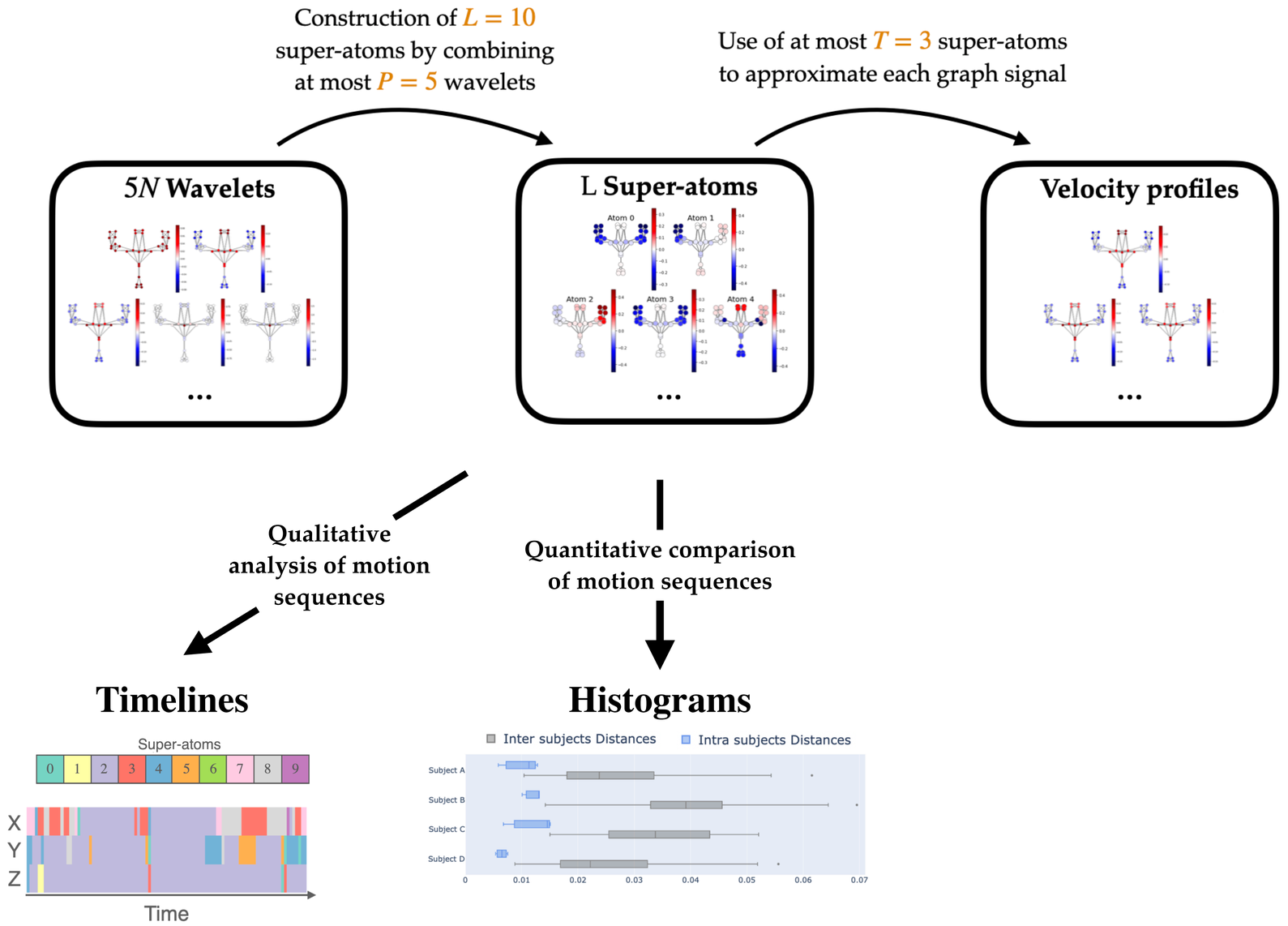




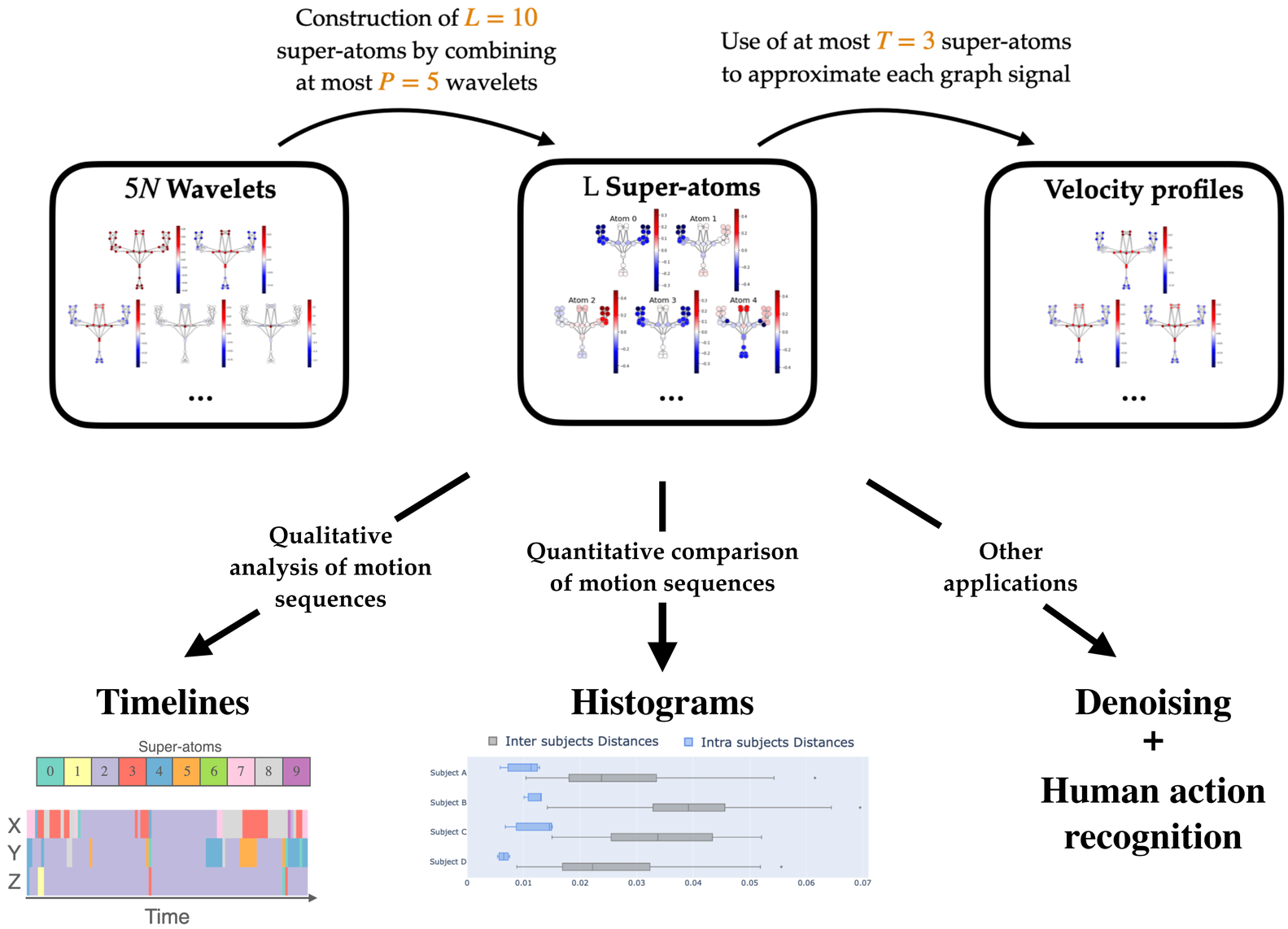
DSMH Method


Construction of $L = 10$ super-atoms by combining at most $P = 5$ wavelets
Use of at most $T = 3$ super-atoms to approximate each graph signal
5N Wavelets
L Super-atoms
Atom 0
Atom 1
Atom 2
Atom 3
Atom 4
Velocity profiles
Qualitative analysis of motion sequences
Timelines
Super-atoms
0 1 2 3 4 5 6 7 8 9
X
Y
Z
Time

DSMH Method


Construction of $L = 10$ super-atoms by combining at most $P = 5$ wavelets
Use of at most $T = 3$ super-atoms to approximate each graph signal
5N Wavelets
L Super-atoms
Atom 0
Atom 1
Atom 2
Atom 3
Atom 4
Velocity profiles
Qualitative analysis of motion sequences
Quantitative comparison of motion sequences
Timelines
Histograms
Super-atoms
0 1 2 3 4 5 6 7 8 9
X
Y
Z
Time
Inter subjects Distances
Intra subjects Distances
Subject A
Subject B
Subject C
Subject D
0 0.01 0.02 0.03 0.04 0.05 0.06 0.07

DSMH Method


Construction of $L = 10$ super-atoms by combining at most $P = 5$ wavelets
Use of at most $T = 3$ super-atoms to approximate each graph signal
5N Wavelets
L Super-atoms
Atom 0
Atom 1
Atom 2
Atom 3
Atom 4
Velocity profiles
Qualitative analysis of motion sequences
Quantitative comparison of motion sequences
Other applications
Timelines
Super-atoms
0 1 2 3 4 5 6 7 8 9
X
Y
Z
Time
Histograms
Inter subjects Distances
Intra subjects Distances
Subject A
Subject B
Subject C
Subject D
0 0.01 0.02 0.03 0.04 0.05 0.06 0.07
Denoising
+
Human action recognition

Thanks for your attention

Marion Chauveau, Antoine Mazarguil and Laurent Oudre

Contact: marion.chauveau@univ-grenoble-alpes.fr

Université Paris-Saclay, Université Paris Cité, ENS Paris-Saclay, CNRS,
SSA, INSERM, Centre Borelli, F-91190, Gif-sur-Yvette, France

CENTRE BORELLI

école normale supérieure paris-saclay

université PARIS-SACLAY

46TH ANNUAL IEEE Engineering in Medicine and Biology Society
Orlando
EMBS 2024
July 15-19, 2024

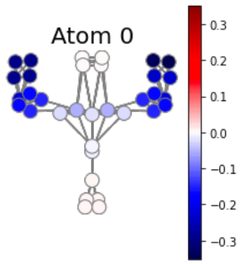
Annexes

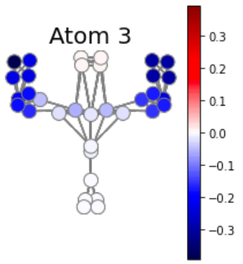
DSMH dictionary

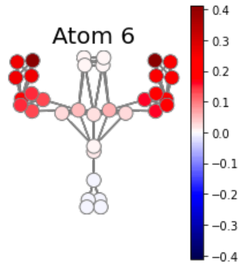
X

Y

Z







High contribution

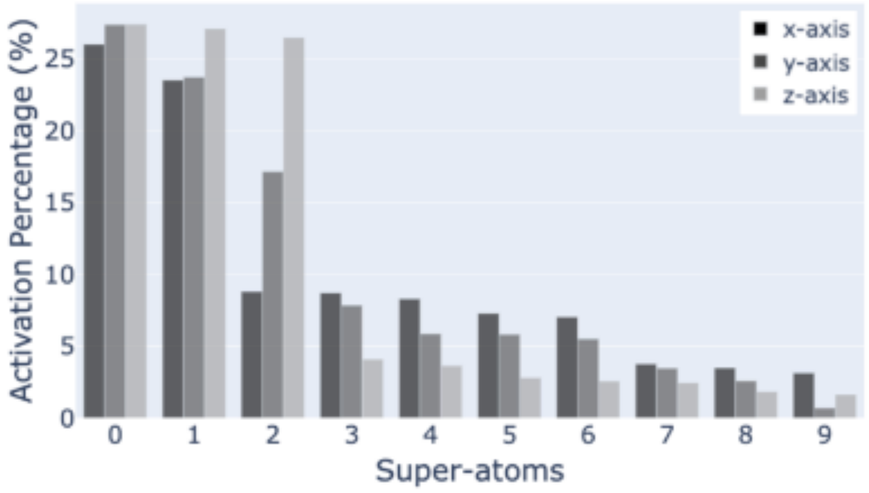
Low contribution

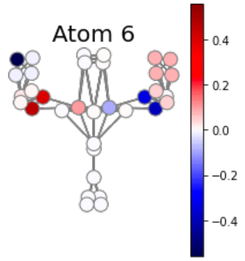
DSMH dictionary

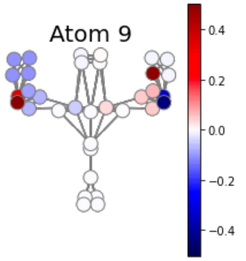
X

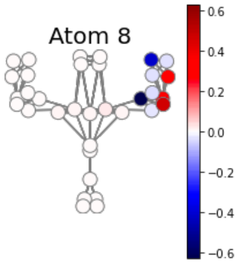
Y

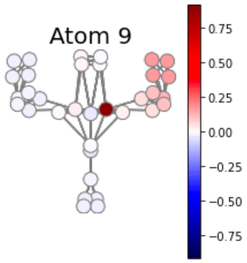
Z

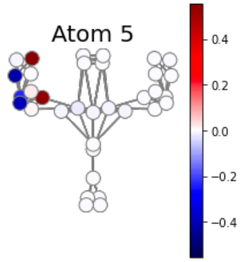


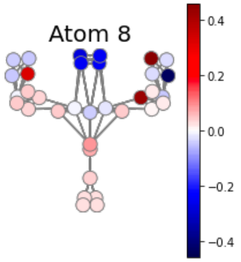


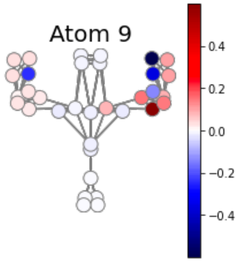












High contribution

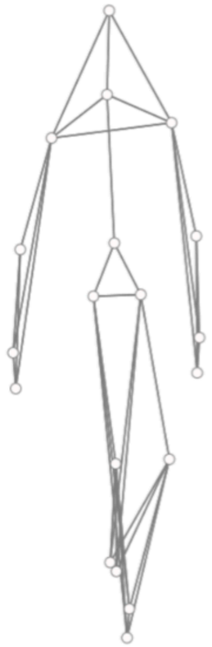
Low contribution

Human action recognition

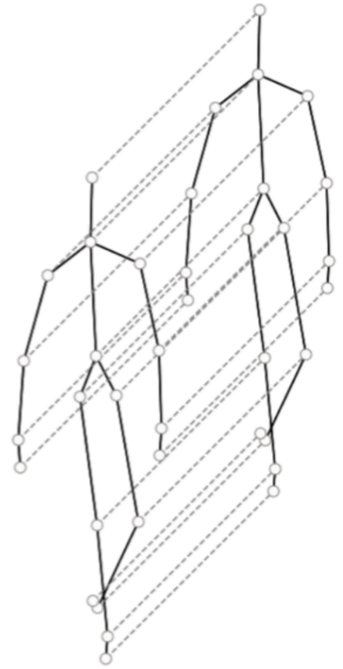
Recognition Method	UTK	MSR	F3D
Spatio-temporal Graph + GFT	95 %	71,45 %	**82,63** %
Spatio-temporal Graph + DSMH	95,5 %	70,14 %	78,12 %
Weighted Graph + GFT	94,97 %	70,78 %	79,76 %
Weighted Graph + DSMH	**96** %	**71,94** %	82,38 %

Recognition Method	UTK [14]	MSR [15]	F3D [16]	ntu_cs_mini [5] [17]
Kao et al. [8]	95.00	71.45	82.63	-
ST-GCN [1], [5], [18]	-	27.64 (CS)	-	71.53 (CS)
GR-GCN [3]	98.5*	-	98.4*	-
Deep STGC$_K$ [2]	-	-	99.1*	-
shift-GCN [4], [5]	-	-	-	60.00 (CS)
Our method	96.00	71.94	82.38	66.25

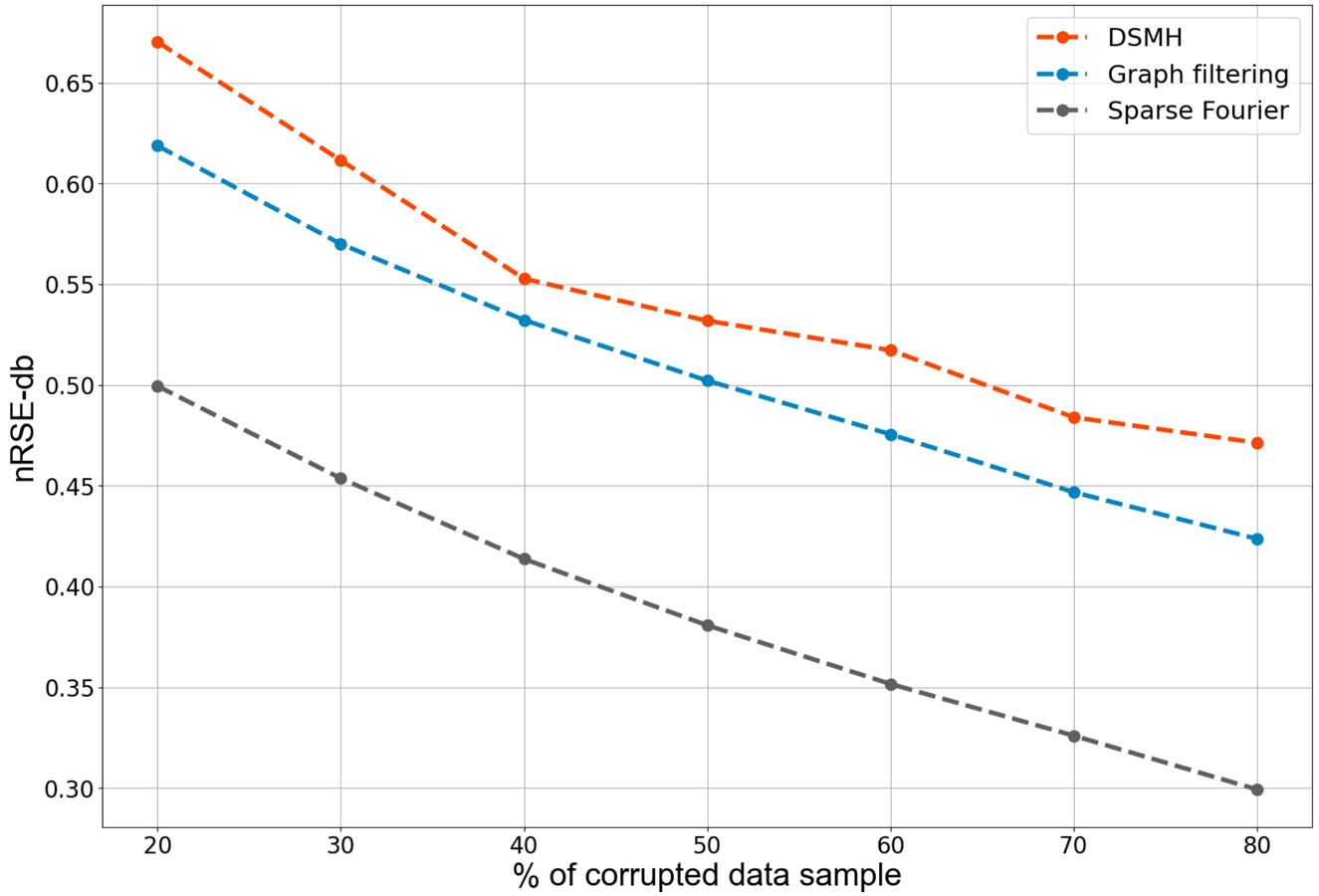
Weighted graph



Spatio-temporal graph



Denoising



$$\text{nRSE-db} = -\log_{10}\left(\frac{||Y - Y_{denoised}||_F}{||Y||_F}\right)$$

DSMH method:

$$Y_{denoised} = \Phi A X$$

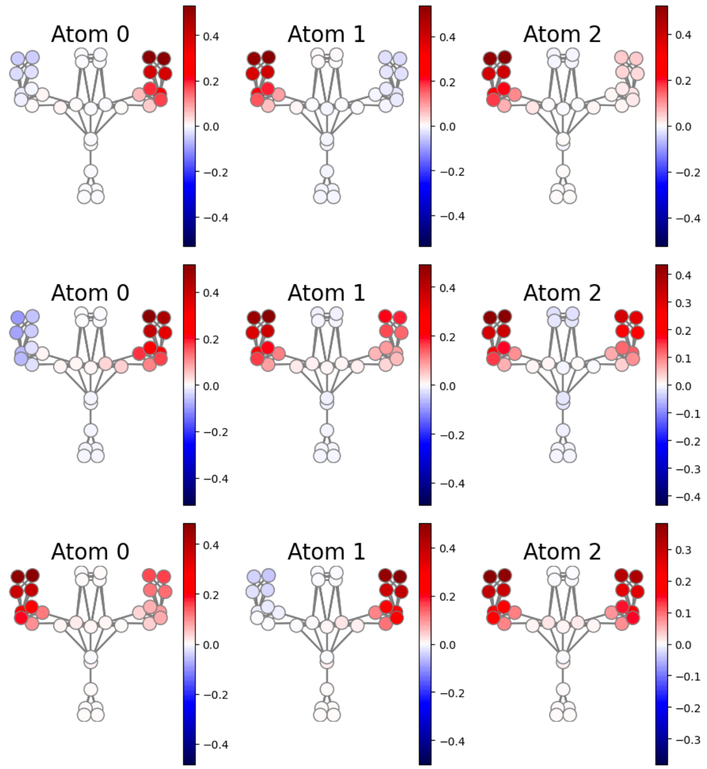
Graph filtering method:

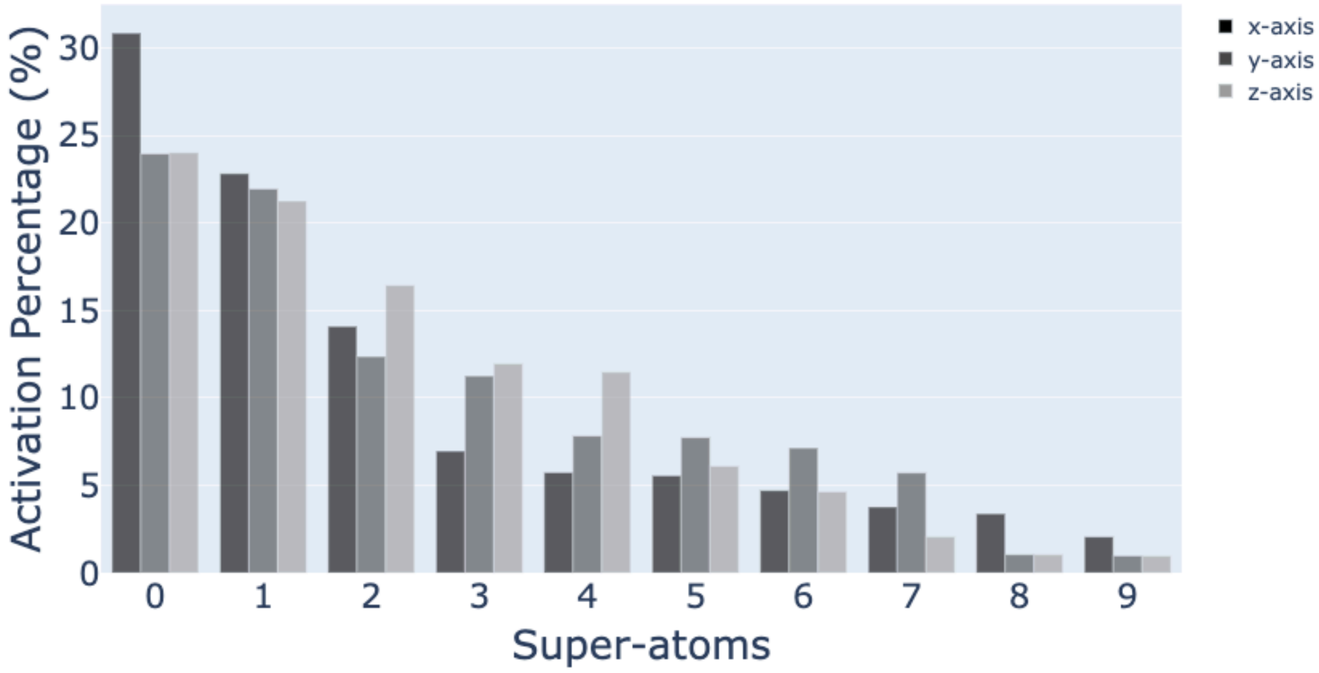
$$Y_{denoised} = (I + \alpha(I - A)^*(I - A))^{-1}Y$$

Method	20% of corr. data	50% of corr. data
Noisy data	0.49	0.29
Graph filtering [45]	0.62	0.50
Sparse Fourier [44]	0.50	0.38
DSMH	**0.72**	**0.55**

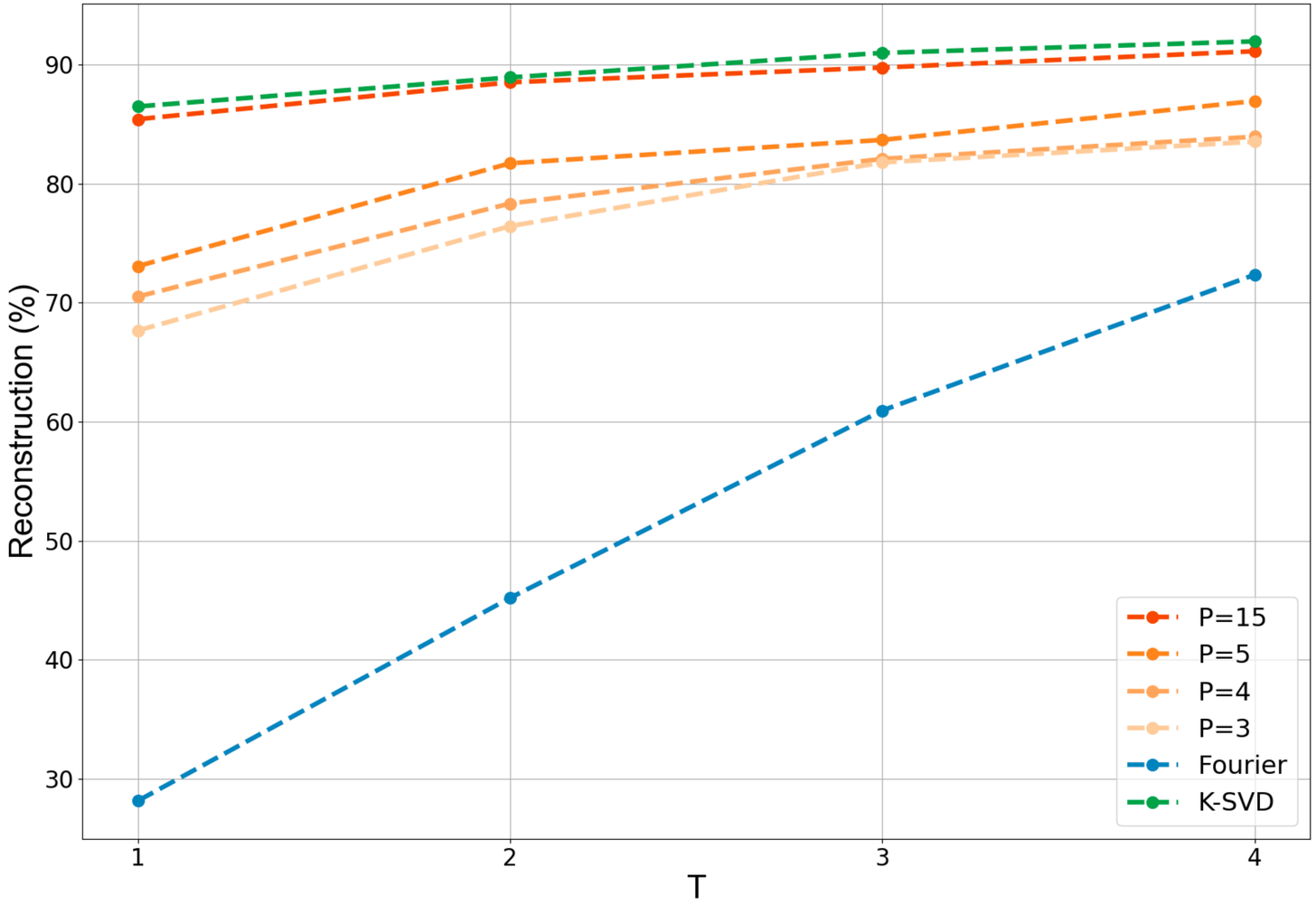
TABLE I. nRSE-db obtained for 20% and 50% of corrupted samples with the following 3 methods: the double sparsity method ($s_1 = 5$, $s_2 = 7$, $L = 10$), the graph filtering method ($\alpha = 0.39$), and the sparse Fourier method ($s_1 = 5$).

Unconstrained dictionary

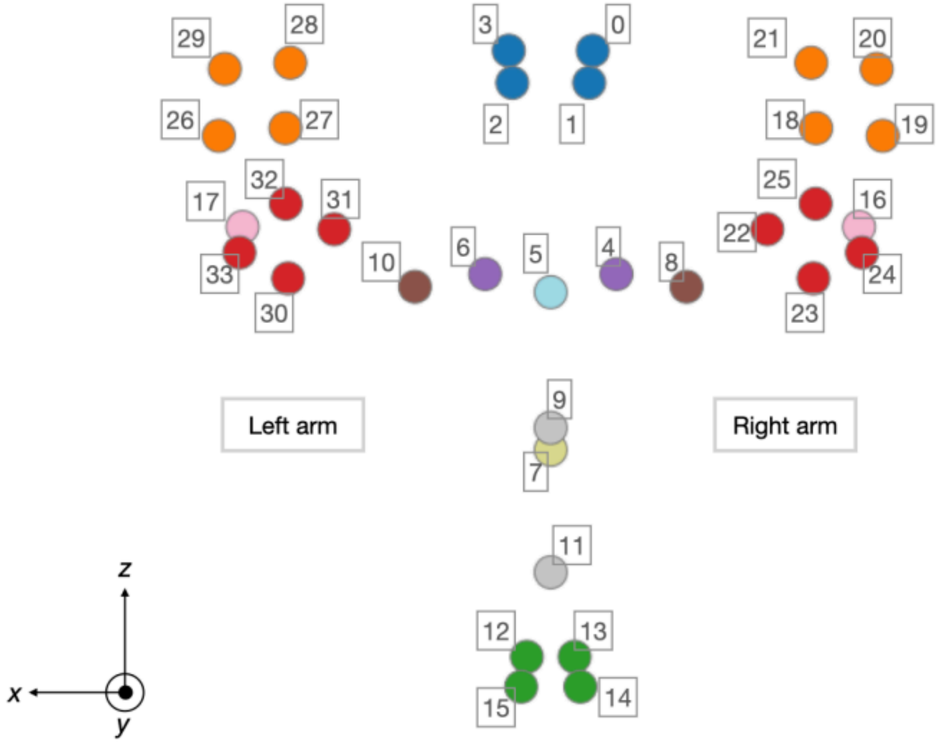


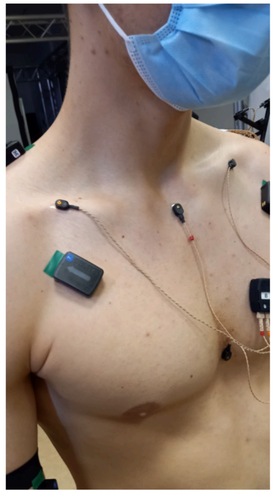


Signal reconstruction



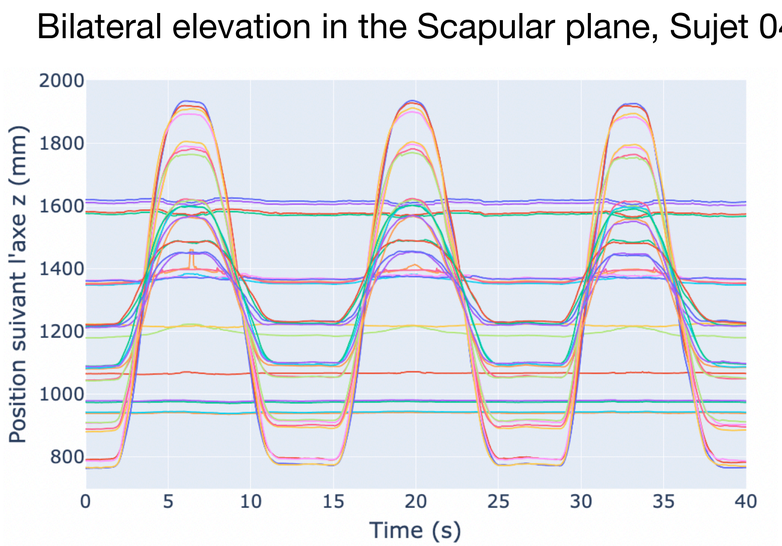
Arm-CODA





(a) CODA simple marker placed on the front side of the subject

(b) CODA cluster placed on the subject's forearm



Marker index	Marker type	Position
0-3	Cluster	Top of the forehead
4	Simple marker	Middle of the right clavicle
5	Simple marker	Sternal manubrium, just below the jugular incisure
6	Simple marker	Middle of the left clavicle
7	Simple marker	2 cm above the xiphoid process
8	Simple marker	Middle of the right scapular spine
9	Simple marker	Apophysis of T7
10	Simple marker	Middle of the left scapular spine
11	Simple marker	Apophysis of L3
12-15	Cluster	Below posterior superior iliac spine
16	Simple marker	Right elbow, lateral epicondyle of the humerus
17	Simple marker	Left elbow, lateral epicondyle of the humerus
18-21	Cluster	Right forearm
22-25	Cluster	Lateral face of the right arm, close to the elbow
26-29	Cluster	Left forearm
30-33	Cluster	Lateral face of the left arm, close to the elbow

S. W. Combettes et al, Arm-CODA: A Data Set of Upper-limb Human Movement During Routine Examination, 2024

Arm-CODA

	Age	Height (cm)	Weight (kg)	BMI (kg/m^2)
mean ($\pm$ std)	44.2 ($\pm$ 14.1)	173.7 ($\pm$ 8.4)	73.7 ($\pm$ 11.8)	24.4 ($\pm$ 3.1)
min	23.0	156.0	51.0	19.1
max	65.0	188.0	95.0	29.0

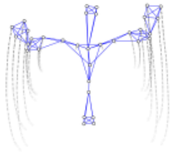
Table 1: Summary of the participants' characteristics.

34 Cartesian Optoelectronic Dynamic Anthropometer (CODA) motion system 3D position markers
Active sensors emit an optical signal received by six depth cameras
Sensor fusion algorithm merges the measurements of the different cameras and obtain estimate of the position of each sensor
Data position of the 34 sensors is measured at 100 Hz

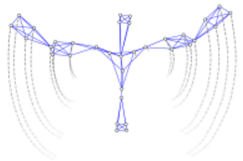
- Simple markers (see Figure 5(a)), attached to the subject's skin by an adhesive strip, powered by a battery placed on the subject's body (each battery powers up to 4 markers).
- Markers associated with a rectangular cluster (see Figure 5(b)), including a battery, attached to the subject with a strap. Each cluster was shaped like a rectangle with 4 markers positioned in its corners, allowing it to locate the fixed structure in space and orientation.

S. W. Combettes et al, Arm-CODA: A Data Set of Upper-limb Human Movement During Routine Examination, 2024

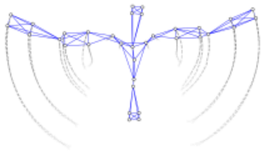
Arm-CODA



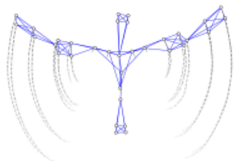
Mvt 0: Sagittal plane elevation (Seated)



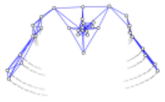
Mvt 1: Scapular plane elevation (Seated)



Mvt 2: Frontal plane elevation (Seated)



Mvt 3: Scapular plane elevation with lateral wrist orientation (Seated)



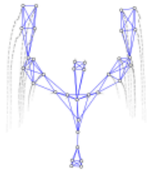
Mvt 4: Lateral elbow rotation (Seated and seen from above)

Mvt 5: Hair combing, right arm (Seated)

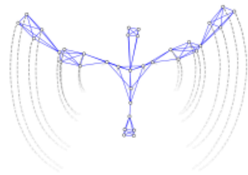
Mvt 6: Hair combing, left arm (Seated)

Mvt 7: Low back washing, right arm (Seated)

Mvt 8: Low back washing, left arm (Seated)



Mvt 9: Sagittal plane elevation (Standing)

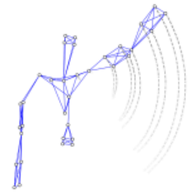


Mvt 10: Scapular plane elevation (Standing)

Mvt 11: Sagittal right arm elevation (Standing)

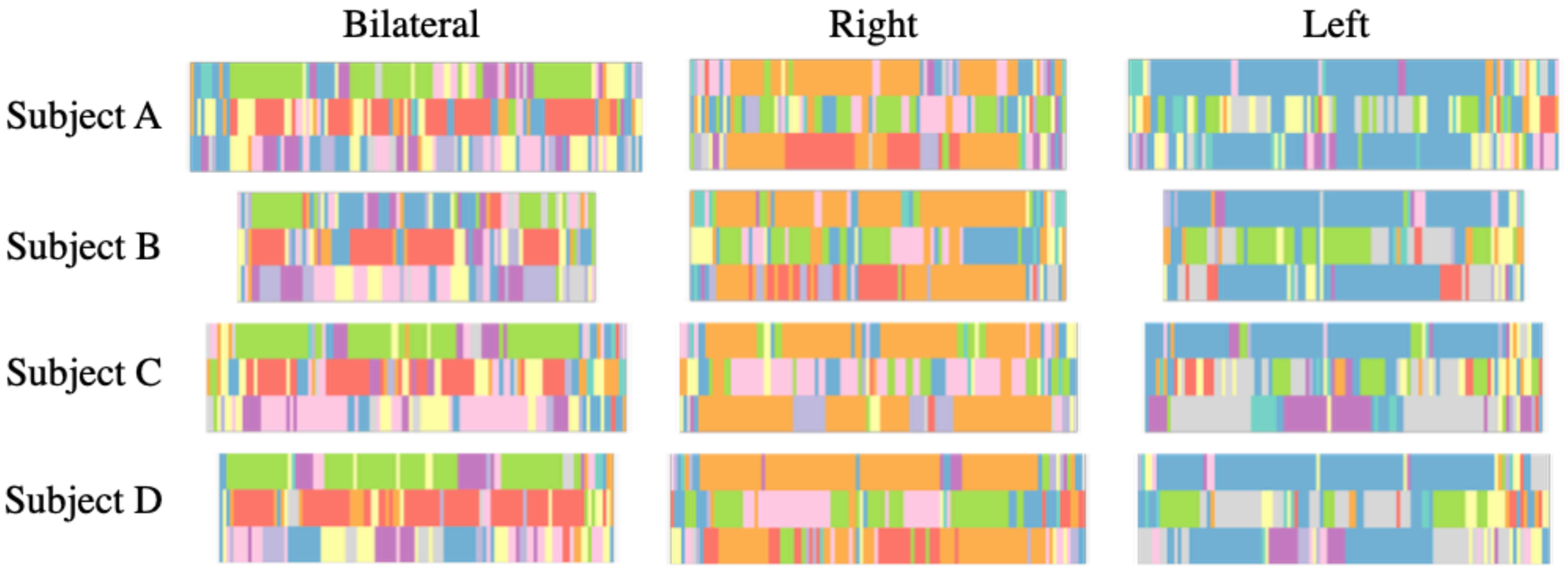
Mvt 12: Sagittal left arm elevation (Standing)

Mvt 13: Scapular right arm elevation (Standing)



Mvt 14: Scapular left arm elevation (Standing)

Timelines 2nd activation

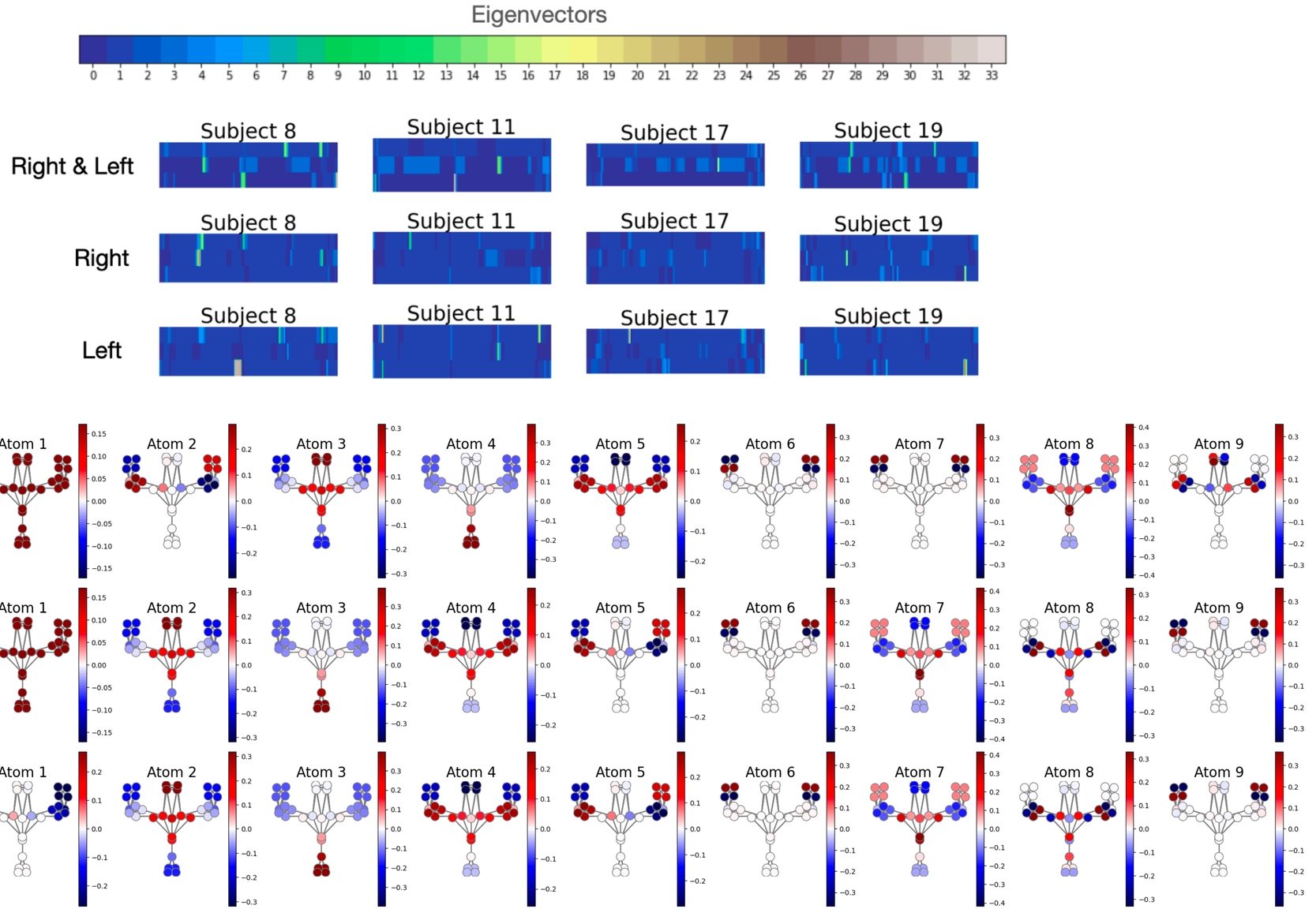


Super-atoms

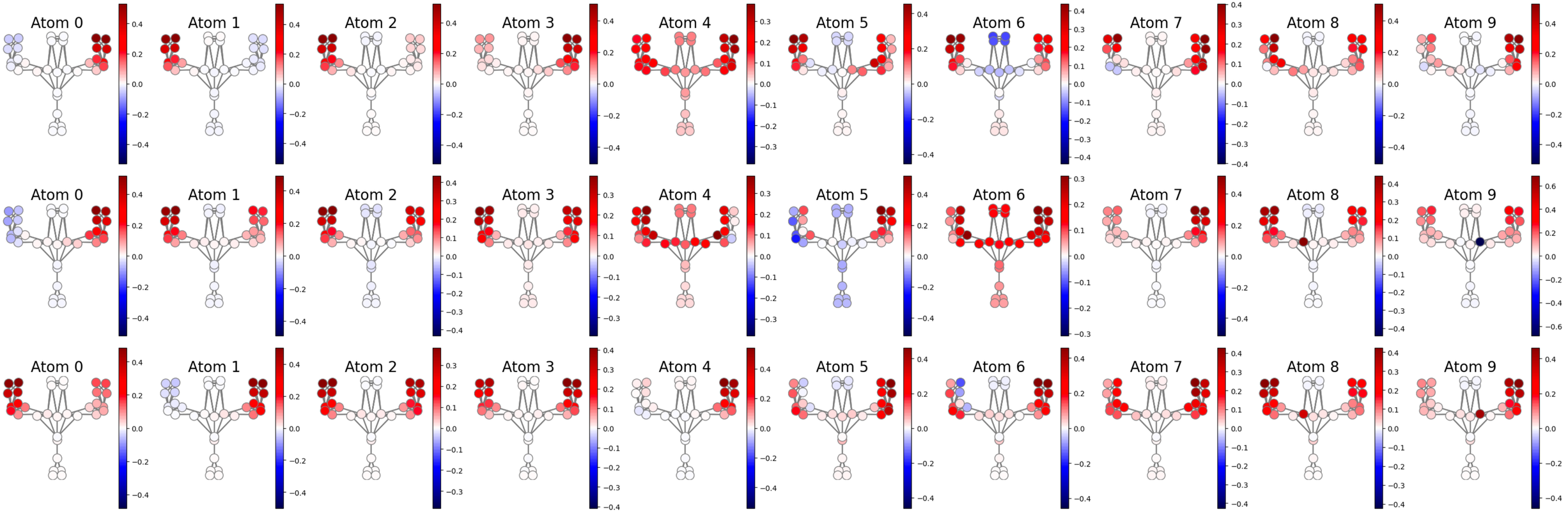
More variations over time & more differences between subjects
BUT two different movements can easily be distinguished

This second level of activation still allows to capture phenomena specific to signals

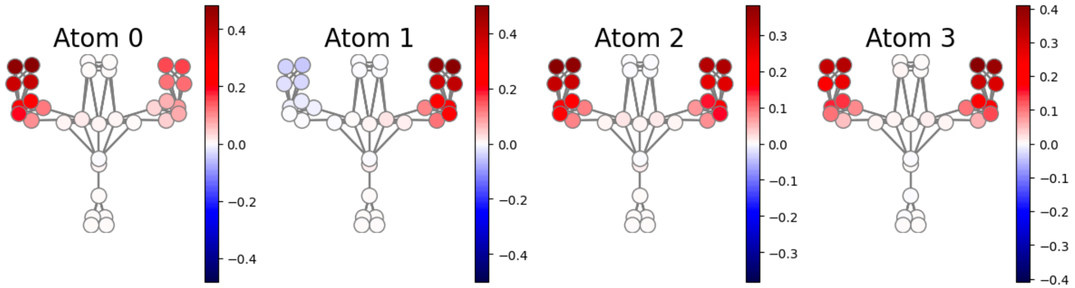
Representations with Fourier basis



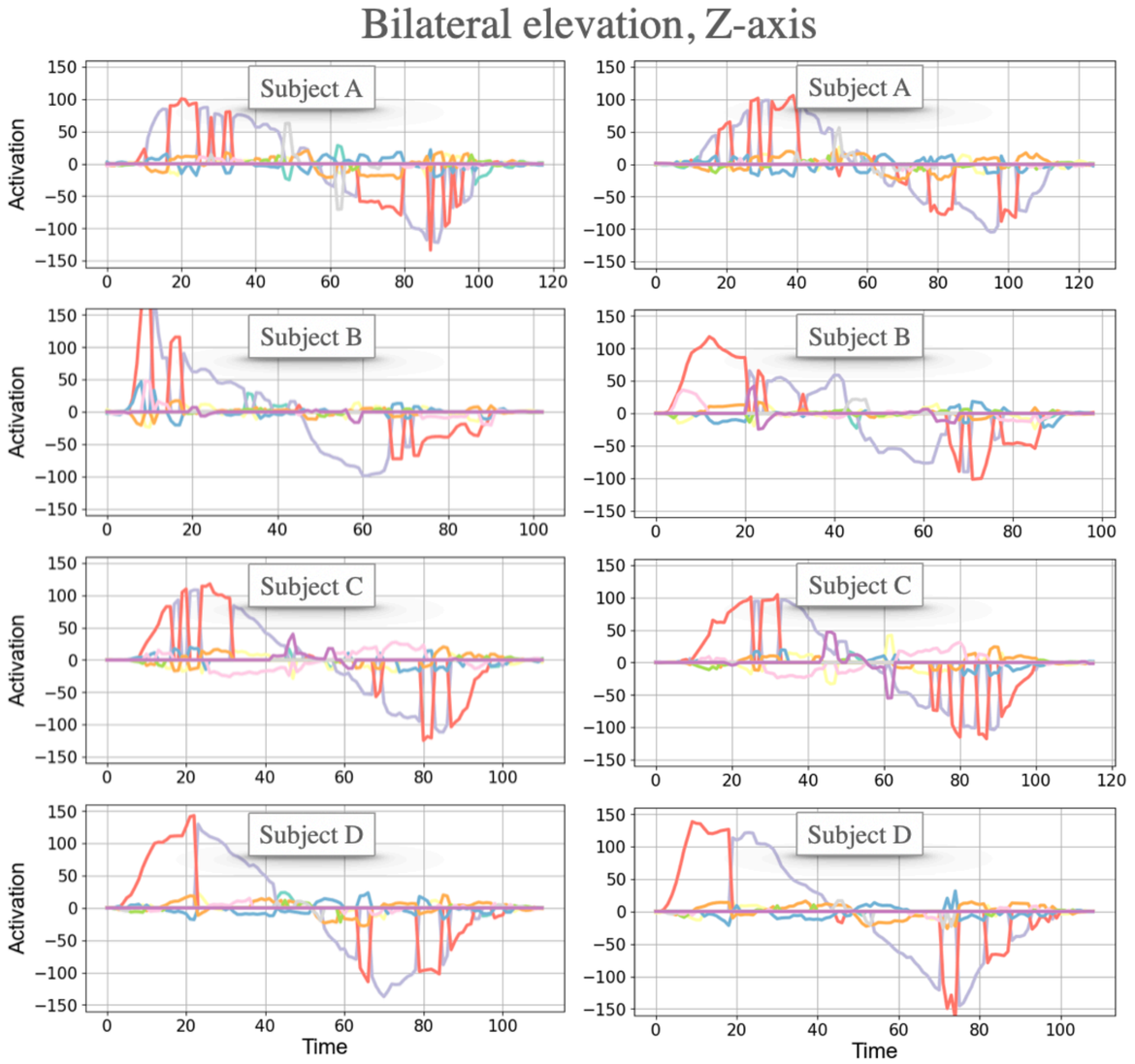
Unconstrained Dictionary



Unconstrained Dictionary

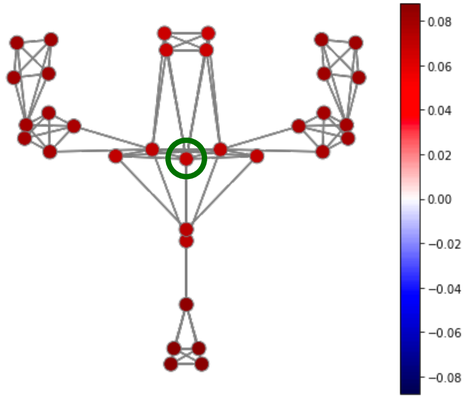


Super-atoms

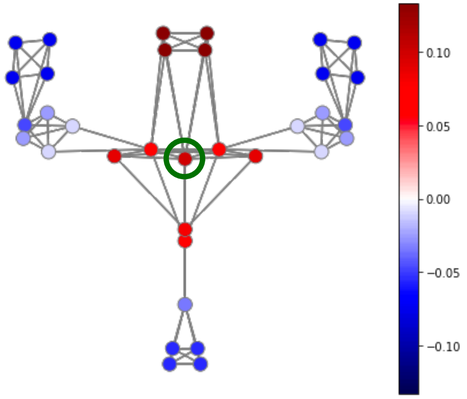


Mexican Hat Wavelets

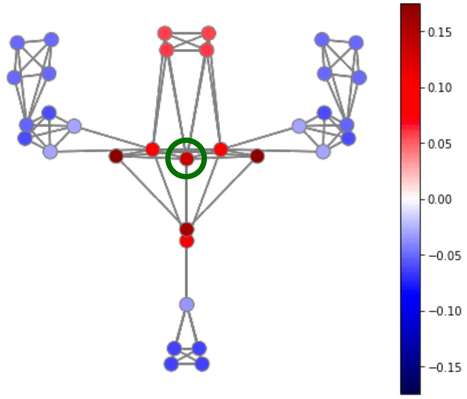
Wavelet $h(x)$



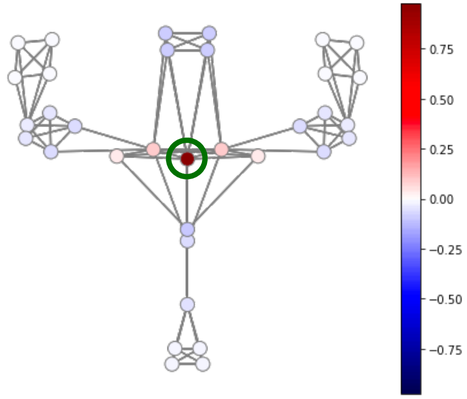
Wavelet $g(t_1x)$



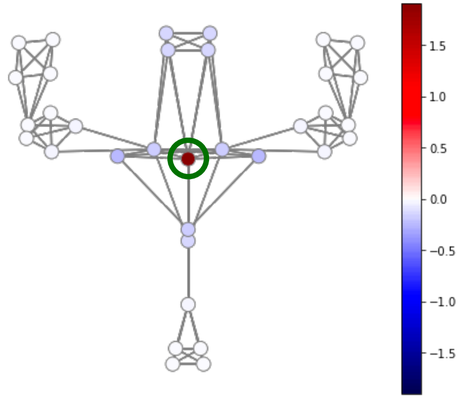
Wavelet $g(t_2x)$



Wavelet $g(t_3x)$



Wavelet $g(t_4x)$



$$h(x) = \gamma e^{-(\frac{20}{0.4\lambda_{max}}x)^4} \qquad g(tx) = tx \times e^{-tx} \qquad \hat{T_g\delta_n}(l) = g(\lambda_l)\hat{\delta}_n(l)$$