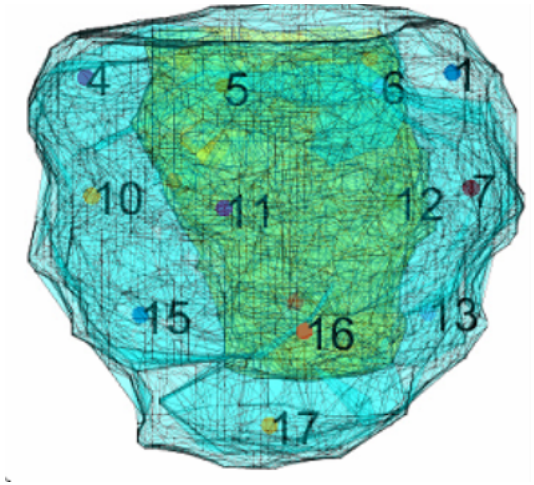
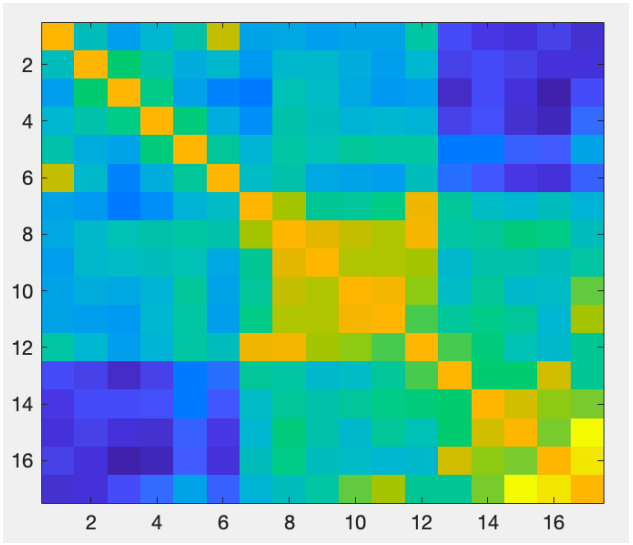
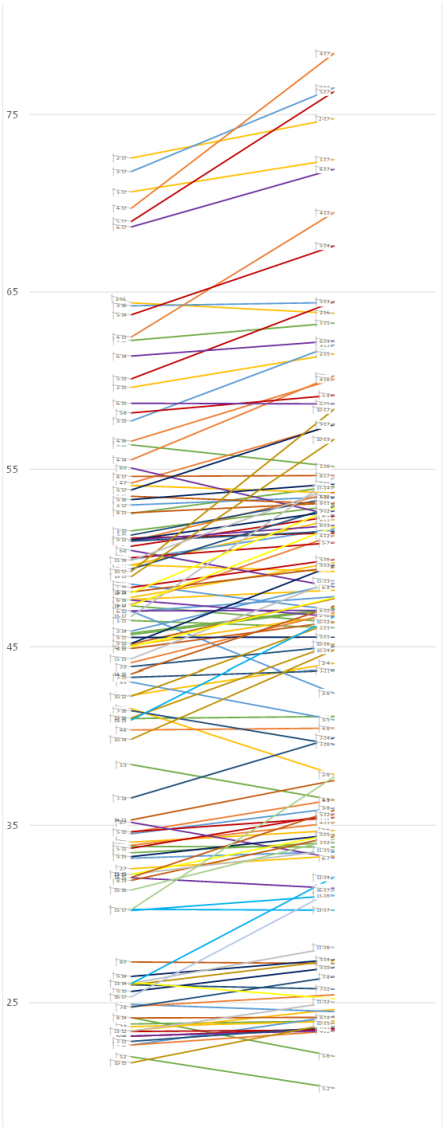
Graph Structure Inference on Cardiac Skeleton

Hypothesis: Heart Skeleton

the structural biomarker is associated with clinical outcomes– Worsen or Maintain function





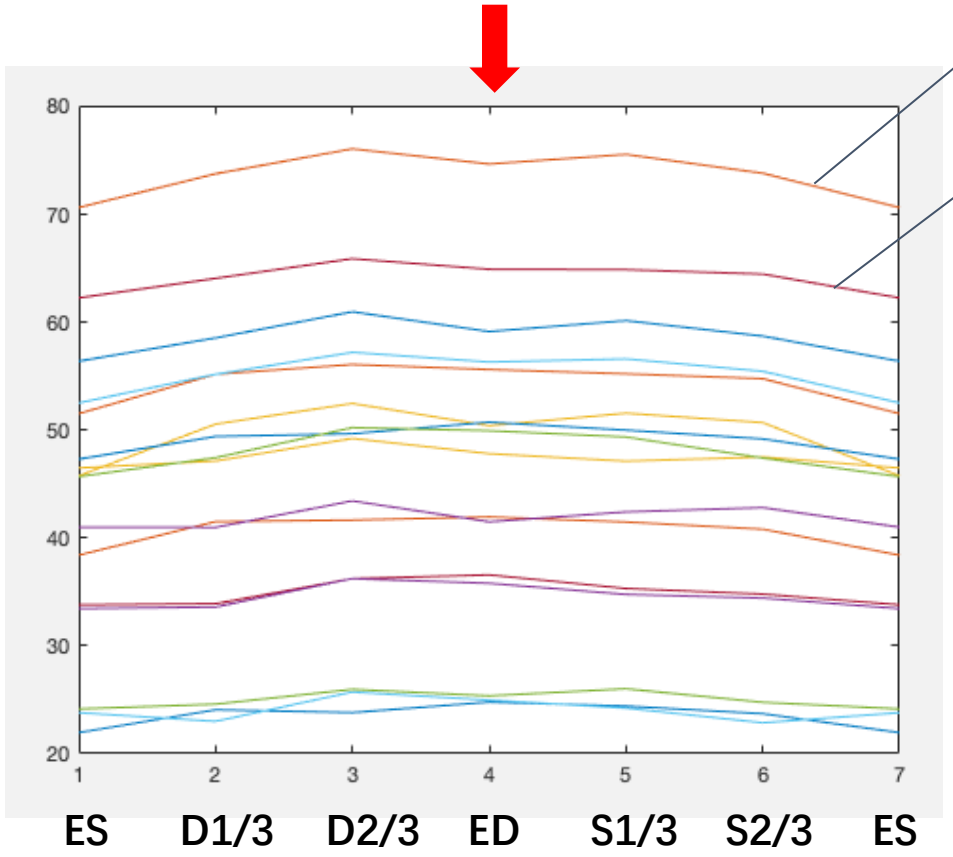


136 linkages

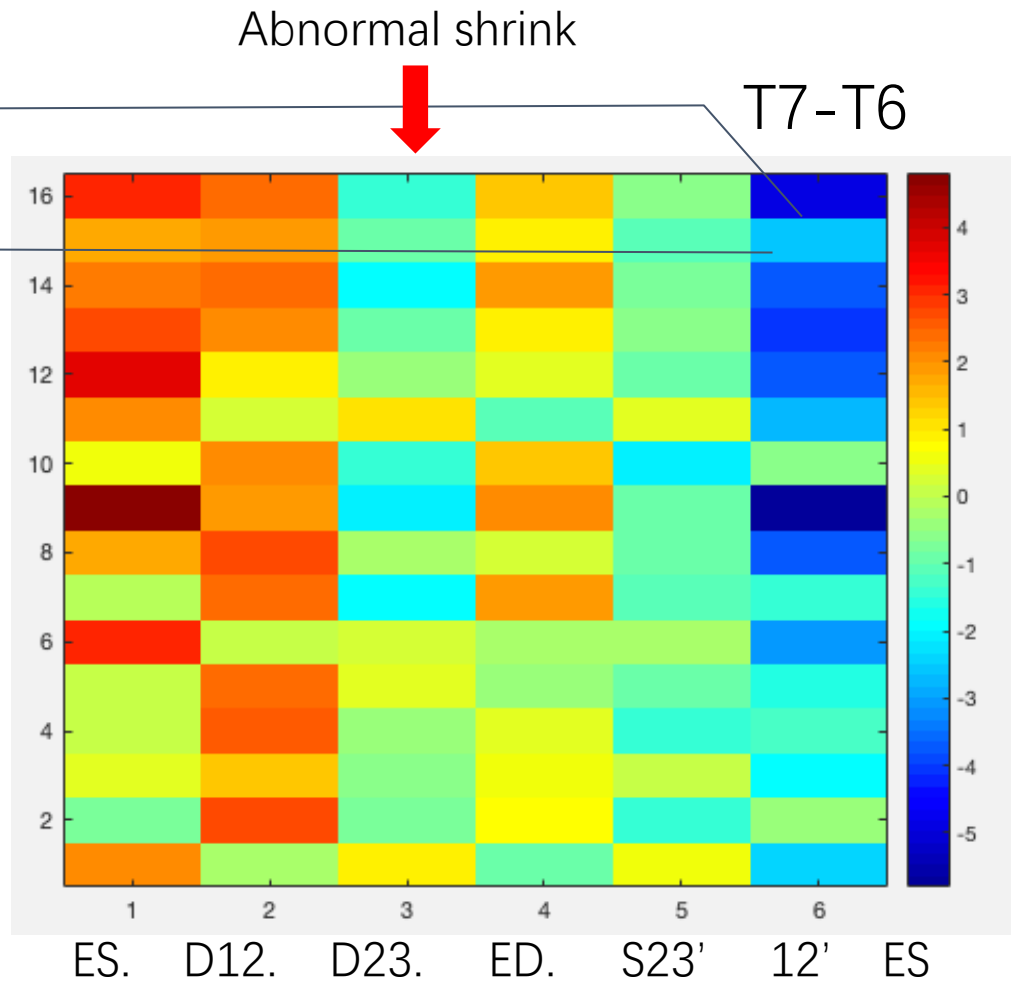
Skeleton was defined by linking the center points of 17 segments of the myocardium, which reflects the working relationship in between.

This system can be represent by a graph where nodes V=17 represent the key points, edges E = 136 represent the length of linkages between points.

An Worsen case:

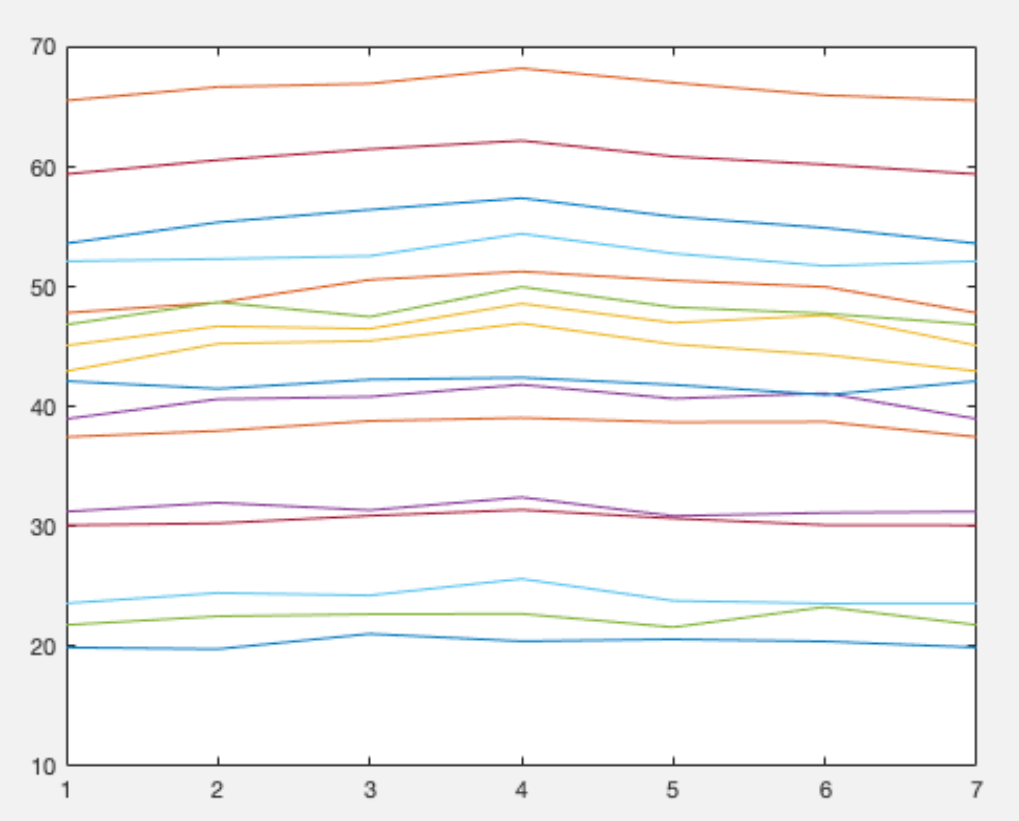


Segment 1 related edge values throughout the Cardiac Cycle



Δ Edge between neighbors on time frames

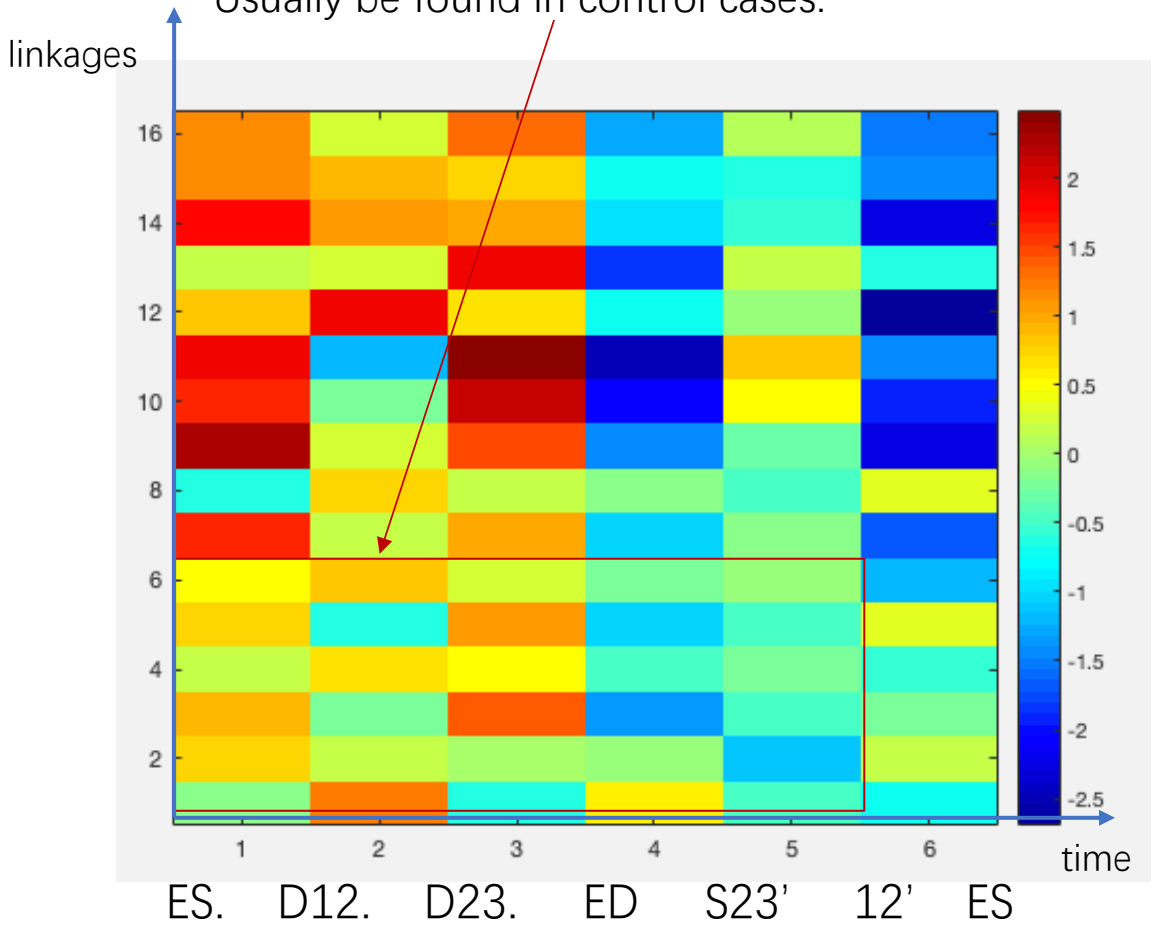
An Control Case



Segment 1 related edge values throughout the Cardiac Cycle

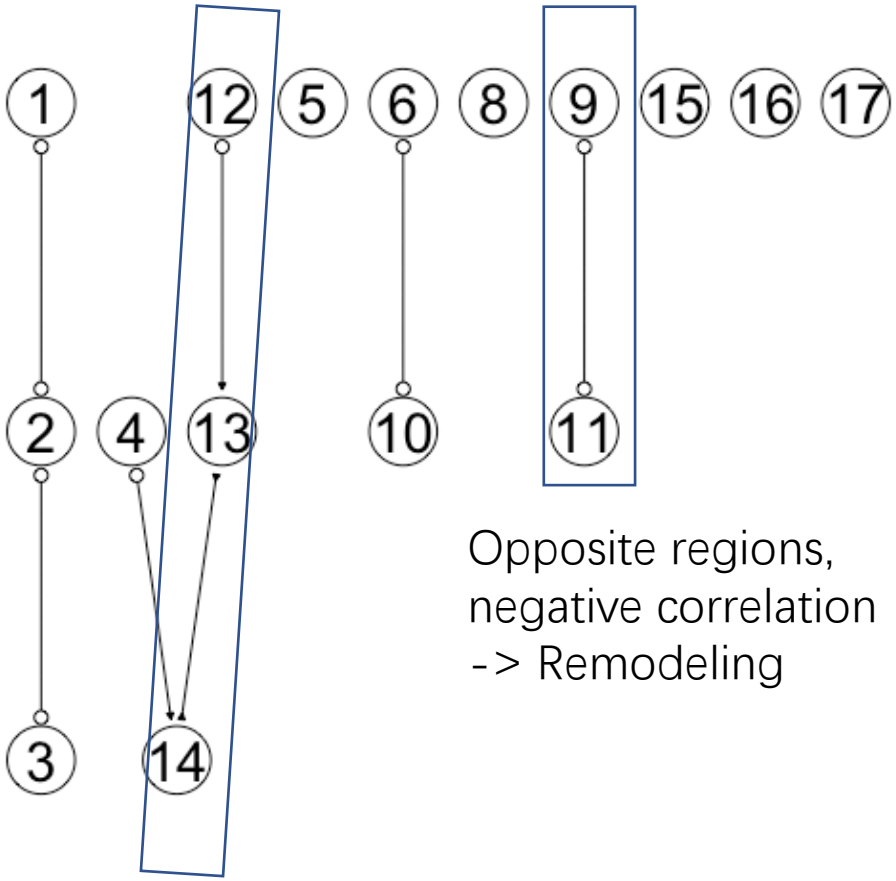
"Remodeling Compensation" : **negative correlation** between nearby regions throughout the cardiac cycle.

Usually be found in control cases.



Δ Edge between neighbors on time frames

Causality Extraction



Opposite regions,
negative correlation
-> Remodeling

Nearby regions,
positive correlation

The left figure showed the causality relationship between 16 linkages (time series difference, lag = 1) started from Region17 by using FCI algorithm. 25 time points were used. ("Causation, Prediction, and Search" , 1993, pp.140-145)

There are several problems about this model:

A. FCI does not work well on large number of features.

B. "Confounding" has no exact meanings in this cardiac system.

Fast Causal Inference Algorithm

A). Form the complete undirected graph Q on the vertex set **V**.
B). $n = 0$.
repeat
  repeat
    select an ordered pair of variables X and Y that are adjacent in Q such that **Adjacencies**(Q,X)\{Y} has cardinality greater than or equal to n, and a subset **S** of **Adjacencies**(Q,X)\{Y} of cardinality n, and if X and Y are d-separated given **S** delete the edge between X and Y from Q, and record **S** in **Sepset**(X,Y) and **Sepset**(Y,X)

  until all ordered variable pairs of adjacent variables X and Y such that **Adjacencies**(Q,X)\{Y} has cardinality greater than or equal to n and all subsets **S** of **Adjacencies**(Q,X)\{Y} of cardinality n have been tested for d-separation;
  $n = n + 1$;
until for each ordered pair of adjacent vertices X, Y, **Adjacencies**(Q,X)\{Y} is of cardinality less than n.
C). Let F' be the undirected graph resulting from step B). Orient each edge as o-o. For each triple of vertices A, B, C such that the pair A, B and the pair B, C are each adjacent in F' but the pair A, C are not adjacent in F', orient A *-* B *-* C as A *→ B ←* C if and only if B is not in **Sepset**(A,C).
D). For each pair of variables A and B adjacent in F', if A and B are d-separated given any subset **S** of **Possible-D-SEP**(A,B)\{A,B} or any subset **S** of **Possible-D-SEP**(B,A)\{A,B} in F remove the edge between A and B, and record **S** in **Sepset**(A,B) and **Sepset**(B,A).

The independence test is based on partial correlation.

$$\rho_{XY\cdot\mathbf{Z}} = \frac{\rho_{XY\cdot\mathbf{Z}\setminus\{Z_0\}} - \rho_{XZ_0\cdot\mathbf{Z}\setminus\{Z_0\}}\rho_{Z_0Y\cdot\mathbf{Z}\setminus\{Z_0\}}}{\sqrt{1-\rho^2_{XZ_0\cdot\mathbf{Z}\setminus\{Z_0\}}}\sqrt{1-\rho^2_{Z_0Y\cdot\mathbf{Z}\setminus\{Z_0\}}}}.$$

$$z(\hat{\rho}_{XY\cdot\mathbf{Z}}) = \frac{1}{2}\ln\left(\frac{1+\hat{\rho}_{XY\cdot\mathbf{Z}}}{1-\hat{\rho}_{XY\cdot\mathbf{Z}}}\right).$$

(for reference, from “Causation, Prediction, and Search”, 1993, pp.144)

Graph Convolutional Network with Causality Inference Inforamtion

- A. Graph convolutional network:

$$H^{(l+1)} = \sigma\left(\widetilde{D}^{-\frac{1}{2}}\tilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right)$$

Input: X- each row represents features of each subject.

H: Network layer.

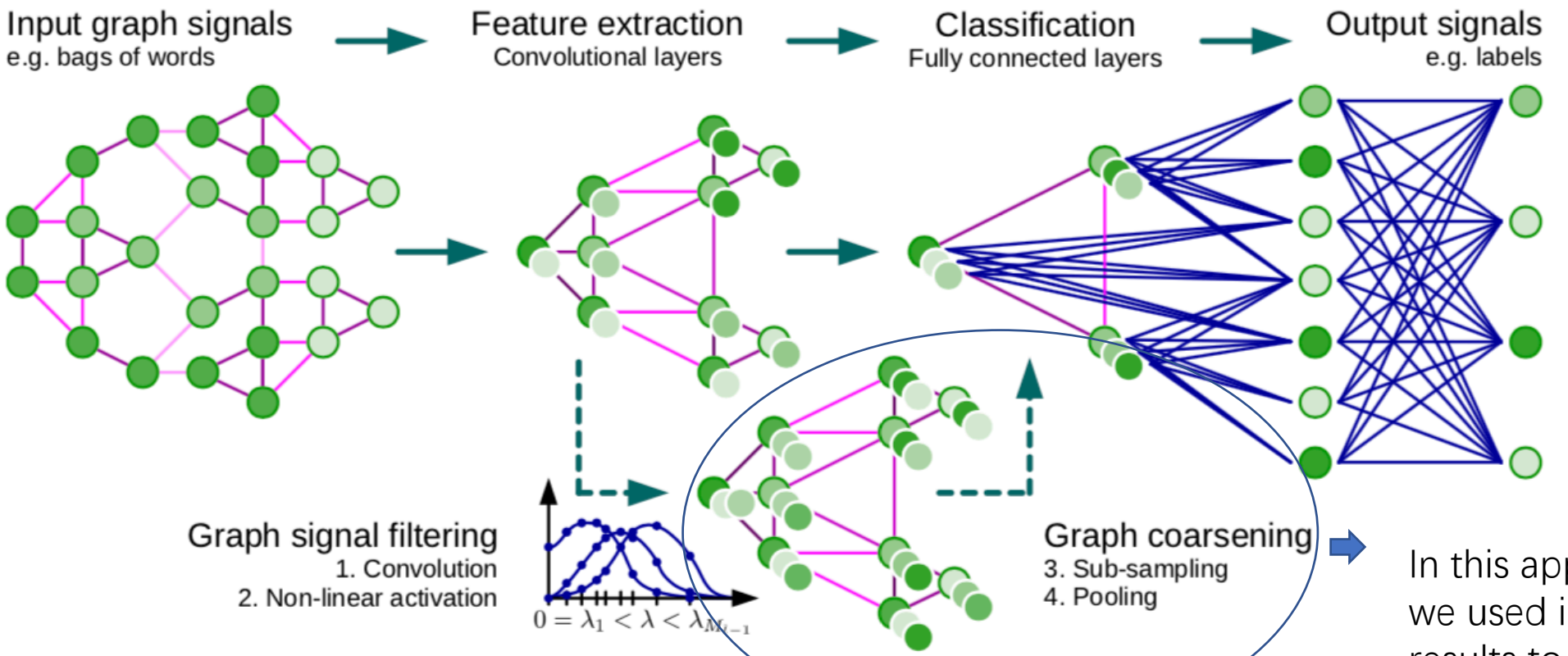
A: adjacent matrix of input X.

D: Diagonal(A).

W: Network weights.

σ: Relu function.

B. Spectral GCN structure



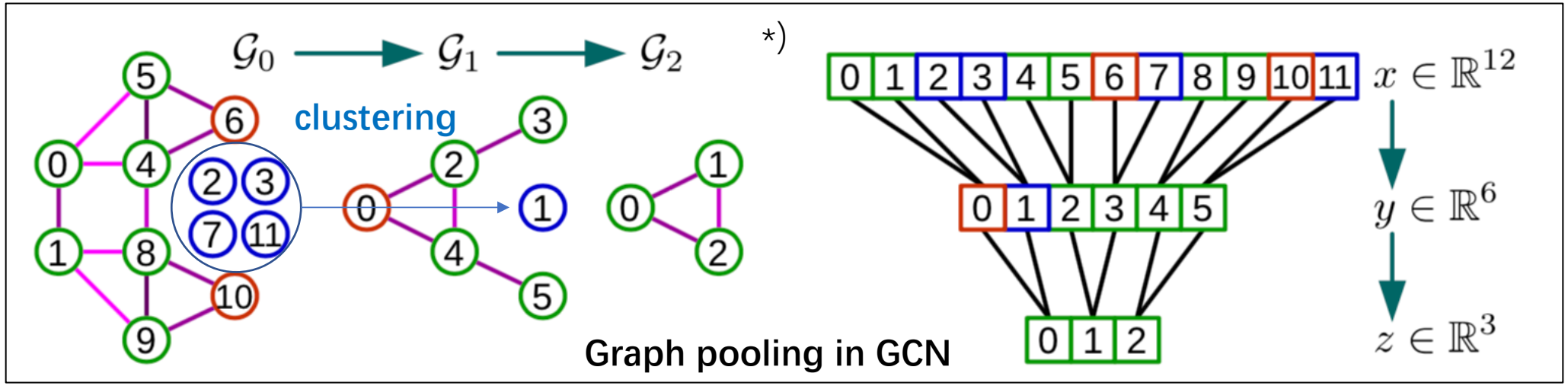
A. Cluster features together -> lower dimension.
B. 1D Pooling on binary tree.

In this application, we used inference results to modify the coarsening algorithm.

arXiv:1606.09375

- C. Motivation behind:

arXiv:1606.09375



This graph pooling algorithm starts from the biggest edges, thus smaller connections will be mixed together and biased.

In biomedical signal processing, the useful information is contained in functional edges-the connectivity inference. Therefore, setting tree roots in the causality inference as start points in the graph coarsening may be beneficial because more useful information would be preserved.

For reference:

Standard Graph Coarsening Algorithm

- Compute correlation / covariance map from features as a graph
- Compute Degree for each node
- **Degree= sum(Laplacian) at each row**

- Repeat
- Random start at V_i
- If V_i is unmarked:
- Compute weighted edge value between V_i and all the unmarked V_j
- **$W_{ij} = E_{ij} / Degree(V_i) + E_{ij} / Degree(V_j)$**
- Select (V_i, V_j) that has the biggest weighted edge value
- end if
- Sort all the subtrees based on the weights $W_{ij.}$
- (This gives subtrees with higher W_{ij} more chances to be selected earlier).
- Until all the nodes are marked

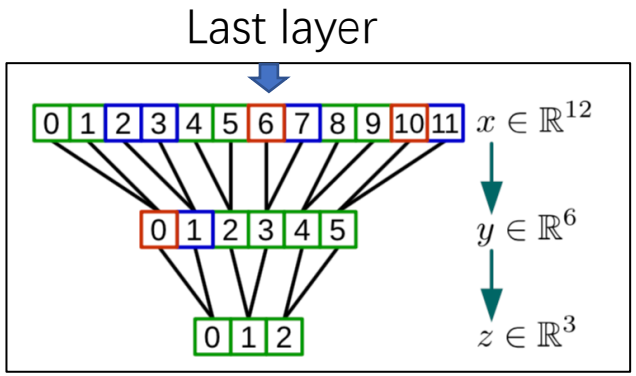
- Coarse original graph by combining edges started from the same cluster together.
- If number(V) in graph != 1:
- repeat (*);

Modified Graph Coarsening Algorithm with Causality Information

For an graph adjacency matrix W and the corresponding causality map C (without confounding edges)

W
[[0.03, 0.73, 0.93, 0.31, 0.23],
[0.73, 0.30, 0.86, 0.70, 0.29],
[0.93, 0.86, 0.95, 0.66, 0.22],
[0.31, 0.70, 0.66, 0.89, 0.43],
[0.23, 0.29, 0.22 , 0.43, 0.38]])

C
[[0 1 0 0 0]
[1 0 0 0 0]
[0 0 0 0 1]
[0 0 0 0 0]
[0 0 1 0 0]]



On the last layer of the graph coarsening tree, the goal of traditional graph coarsening algorithm is to find an adjacency matrix A which enables $\mathrm{min}(\sum_{i,k}(J_n - A)_{ik} \circ W_{ik})$, J_n is n x n matrix full of ones.

In reality, C_{ij} may be linked with more than 3 nodes, which makes it difficult to build a binary tree when we do graph pooling.

The following is a 'soft' way of combining causality information into the last layer:

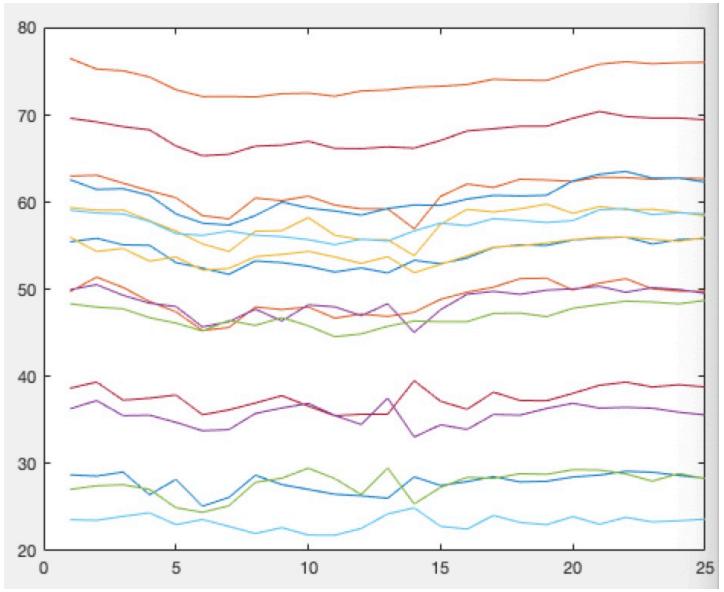
$$\begin{cases} \min G(F_1(A), F_2(A)) \\ F_1(\mathrm{A}) = \sum_{i,k}(J_n - A)_{ik} \circ W_{ik} \\ F_2(A) = \sum_{i,k}(J_n - C \circ A)_{ik} \circ W_{ik}) \\ \text{given} \sum_{i,k} A_{ik} = floor\left(\frac{n}{2}\right) * 2 \end{cases}$$

Note that only ∘ (element wise product) was used. Under the setting of linear scalarization, this problem can be solved in software like pulp (python) as an integer programing problem with n(n-1)/2 features.

Future work & Discussion:

We are still collecting baseline cardiac data.

A. After some experiments on small datasets, we found that the classification accuracy greatly depends on the outcomes of the causality inference results. We may need a better algorithm to deal with noisy data, like the case showed in the right figure.



B. Major problem: When we performed this algorithm on all the 136 linkages on the skeleton, the inference performance became quite unreasonable. Plus, the optimization algorithm for graph coarsening may not be practical on very large feature set.

We need to find a better framework to construct the causality map on all the features, which enables hierarchical structures or concatenating of smaller causality maps together in a DL architecture.

C. This work will be a part of our proposal to the AHA grant. https://www.heart.org/en/professional/institute/grants