# Edge selection preserving the topological features of brain network
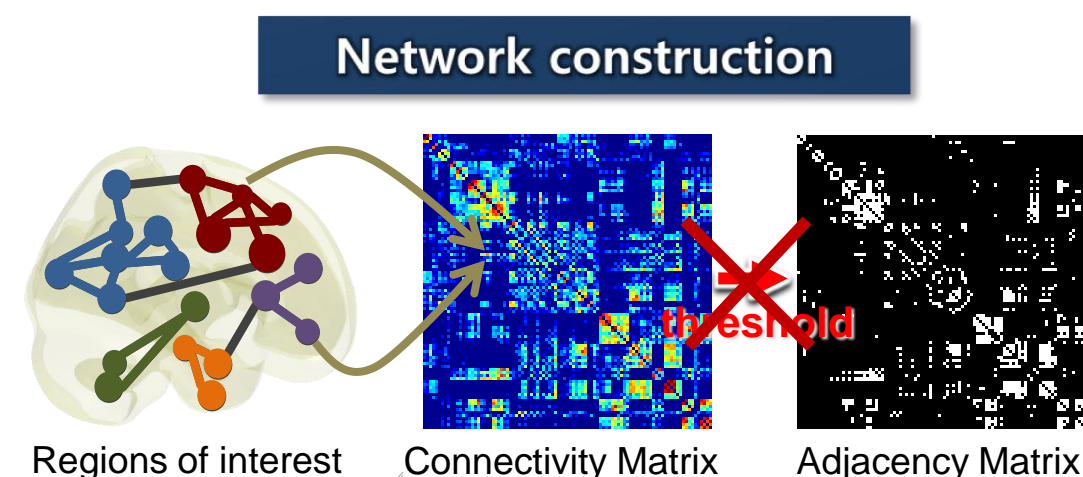
Hyekyoung Lee[1], Moo K. Chung[2], Hyejin Kang[1], Bung-Nyun Kim[1], Dong Soo Lee[1]

[1]Seoul National University, Seoul, Republic of Korea, [2]University of Wisconsin, Madison, WI

## Introduction

### Network construction



Regions of interest    Connectivity Matrix    Adjacency Matrix
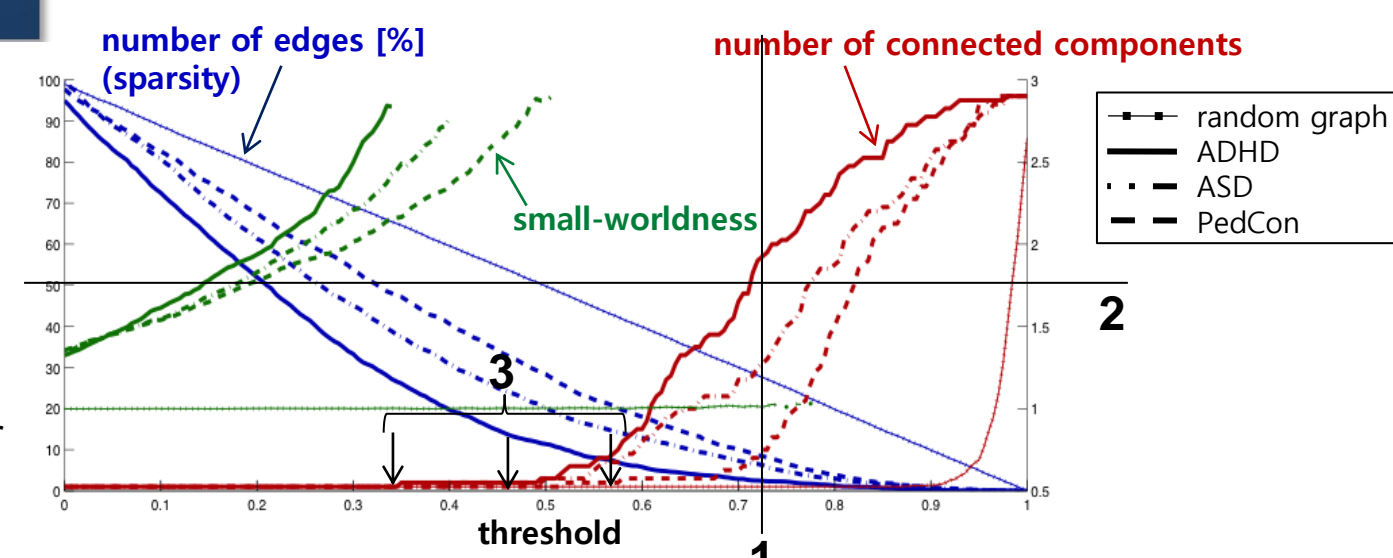
**Distance = 1 – positive correlation**

### Threshold

1. Fixing the threshold.
2. Fixing the sparsity.
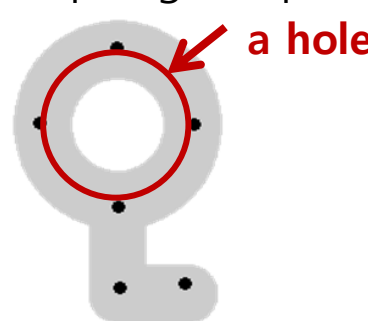3. Choosing the threshold when all nodes are connected, but with the minimum number of edges.

↓

There is no widely accepted rule for thresholding!



→ We propose the edge selection method which preserves the topological features of brain network based on the computational topology.

## Methods

### Betti number

Given the topological space,



a hole

- the zeroth Betti number, $\beta_0 = 1$ = number of connected components
- the first Betti number, $\beta_1 = 1$ = number of holes (Zomorodian and Carlsson, 2005)

### Filtration

Given a node set X consisting of p regions of interest (ROIs), the binary brain network, $B(X,\varepsilon)$ is constructed by connecting two ROIs i and j by an edge if their distance $d(i, j) = 1-\text{corr}(i,j)$ is smaller than $\varepsilon$. A sequence of binary networks for increasing $\varepsilon$ is called a filtration. During the filtration, the topological characteristics of binary network, $B(X,\varepsilon)$, is quantified by the 0th and 1st Betti numbers, $\beta_0$ and $\beta_1$, and visualized by the barcode as shown in Fig. 1 (b) and (c).

### Edge Selection Procedure

By increasing $\varepsilon$, the edge is added one by one and the binary network, $B(X, \varepsilon)$ and its Betti numbers, $\beta_0$ and $\beta_1$ are estimated. If $\beta_0$ and $\beta_1$ are changed when the edge is added, it is collected (See Fig. 1 (a)). Then, the set of selected edges can represent the topological characteristics of network as shown in Fig. 1 (d). And the shortest path lengths along the selected edges approximates the original distance matrix in (e-g).

### Validation

1. Estimate the number of selected edges and mean error and compare the existing methods, MST (Lee, H., et al., 2011) and thresholding (Basset, et al., 2006), to the proposed method.
2. Test the difference of number of holes between groups.
3. Estimate the Gromov-Hausdorff (GH) distance between the true shortest path length matrices (W), between the approximated W, and between single linkage matrices and the bottleneck distance of the zeroth and first Betti numbers, and cluster each matrices into 3 classes using the estimated GH distance.



**Figure 1. Toy example.**

Mean error = mean(abs((e) –(g)))
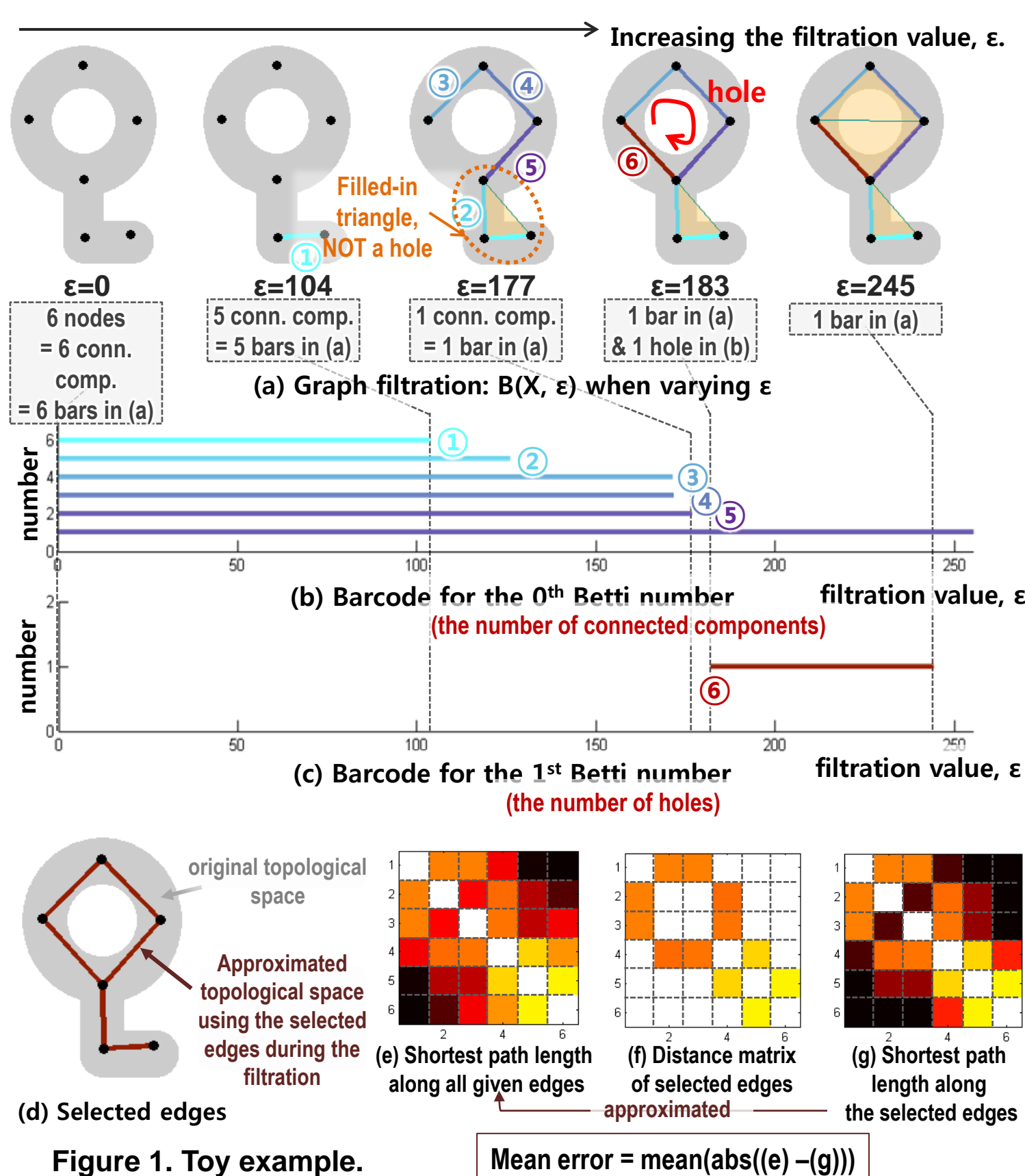
## Results

### Datasets

We used the FDG-PET data set: 24 attention deficit hyperactivity disorder (ADHD), 26 autism spectrum disorder (ASD) children and 11 pediatric control (PedCon) subjects.
FDG-PET images were preprocessed using Statistical Parametric Mapping package. After spatial normalization to the Korean standard template space, mean FDG uptake within 103 regions of interest were extracted (Lee, J.S., et al., 2004). The values of FDG uptake were globally normalized to the individual's total gray matter mean count. Using the general linear model, we factored out the effects of age.

### 1. Selected Edges

We illustrated the filtration and barcodes of ADHD, ASD and PedCon in Fig. 2. The total number of selected edges of ADHD, ASD and PedCon are 144, 136 and 148 in Fig. 3. The ADHD, ASD and PedCon networks have 42, 34 and 46 holes.
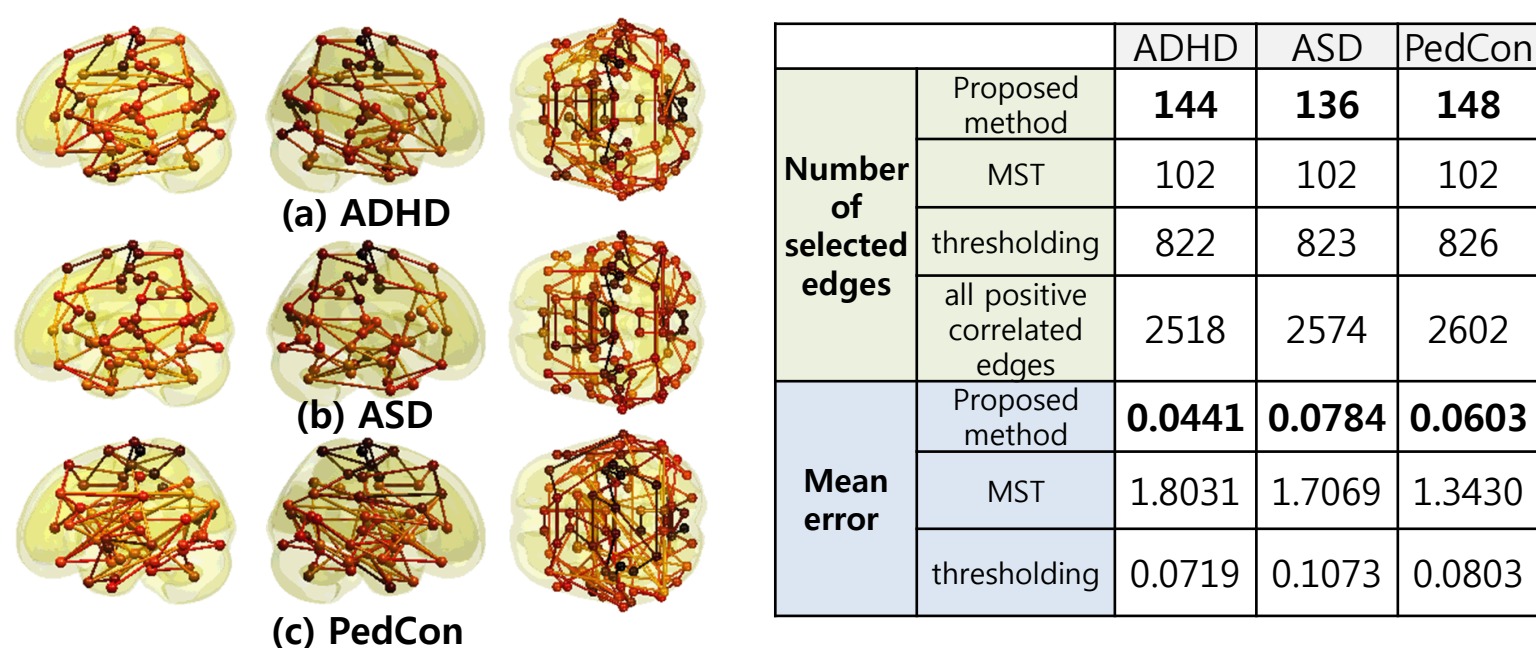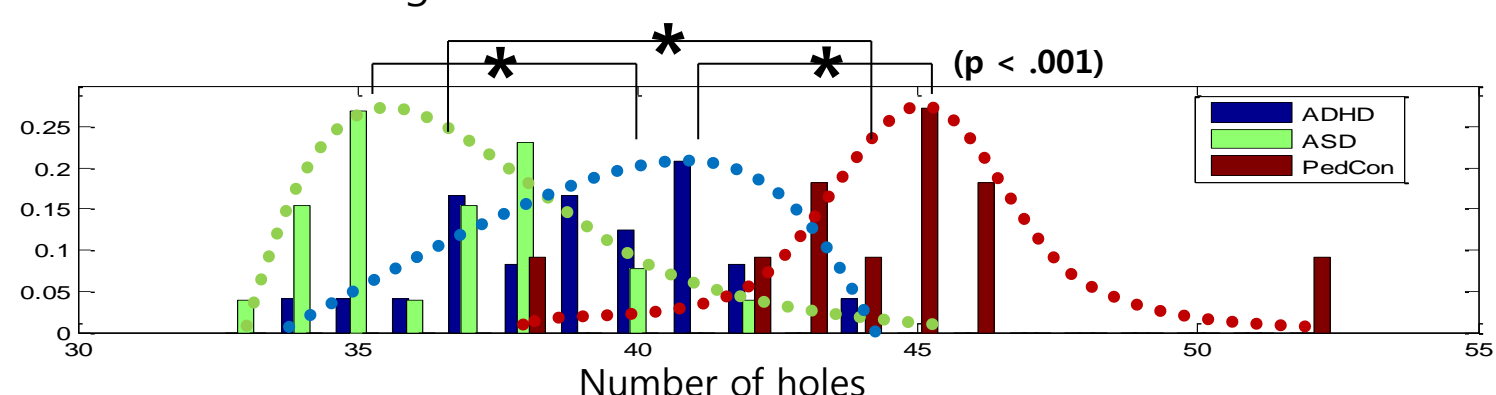


(a) ADHD
(b) ASD
(c) PedCon

| | | ADHD | ASD | PedCon |
|---|---|---|---|---|
| **Number of selected edges** | Proposed method | **144** | **136** | **148** |
| | MST | 102 | 102 | 102 |
| | thresholding | 822 | 823 | 826 |
| | all positive correlated edges | 2518 | 2574 | 2602 |
| **Mean error** | Proposed method | **0.0441** | **0.0784** | **0.0603** |
| | MST | 1.8031 | 1.7069 | 1.3430 |
| | thresholding | 0.0719 | 0.1073 | 0.0803 |

**Figure 3.** Comparison of the number of selected edges (as shown in left) and mean error to the given shortest path length matrix to the previous methods. The smaller error is, the better representation of the given distance matrix is. Our edge selection method finds the smaller number of edges with the smaller error.

### 2. Number of holes

We generate 24 ADHD, 26 ASD and 11 PedCon datasets using the jackknifed resampling method. We estimate the number of holes for each datasets and obtain the distribution of number of holes and the statistical significance of their differences using Wilcoxon rank sum test.
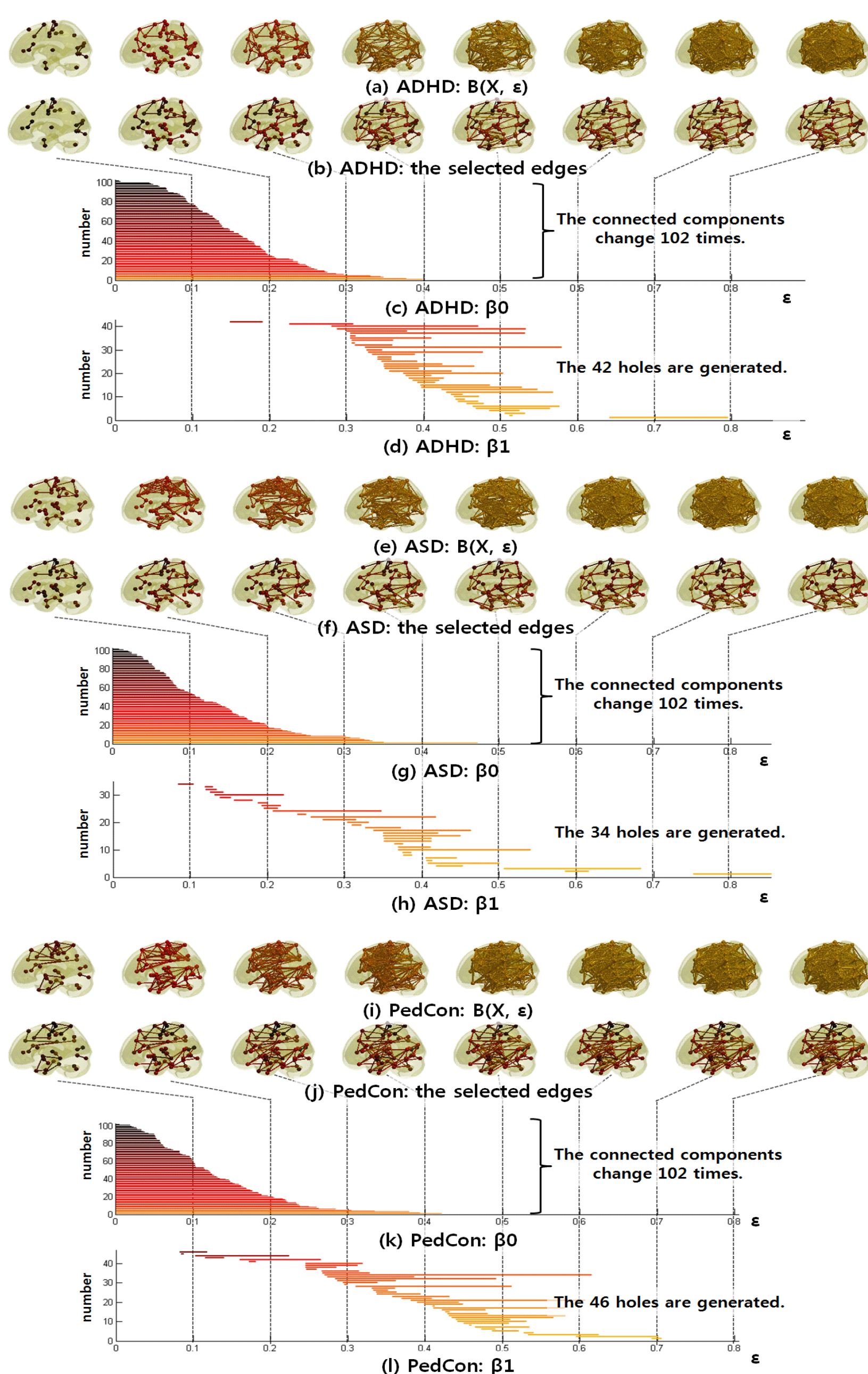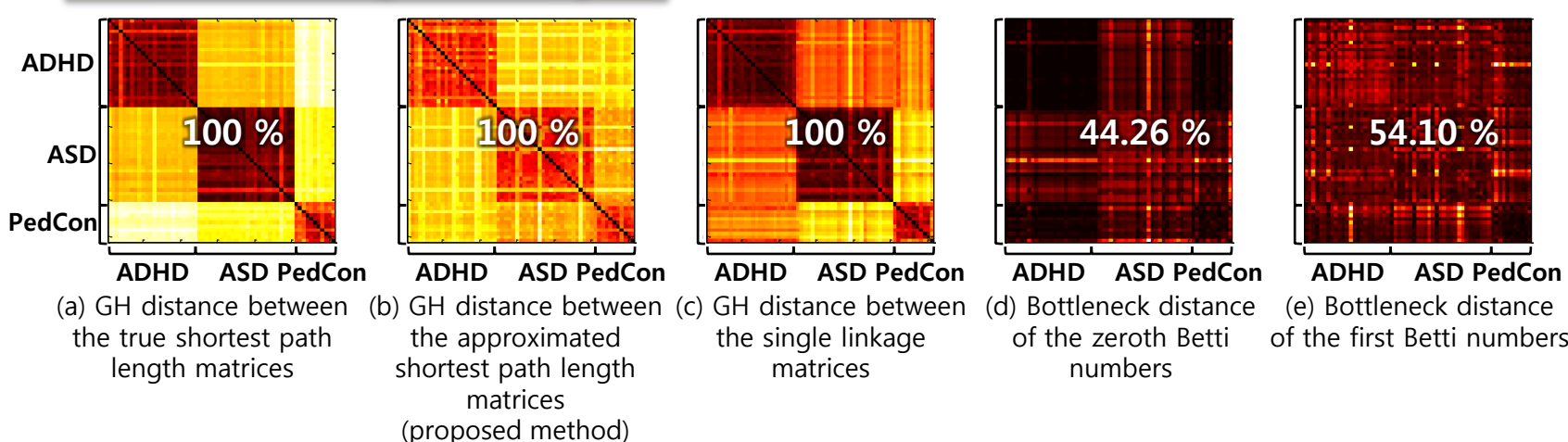


### 3. Clustering accuracy



(a) GH distance between the true shortest path length matrices
(b) GH distance between the approximated shortest path length matrices (proposed method)
(c) GH distance between the single linkage matrices
(d) Bottleneck distance of the zeroth Betti numbers
(e) Bottleneck distance of the first Betti numbers



(a) ADHD: $B(X, \varepsilon)$
(b) ADHD: the selected edges

The connected components change 102 times.

(c) ADHD: $\beta_0$

The 42 holes are generated.

(d) ADHD: $\beta_1$

(e) ASD: $B(X, \varepsilon)$
(f) ASD: the selected edges

The connected components change 102 times.

(g) ASD: $\beta_0$

The 34 holes are generated.

(h) ASD: $\beta_1$

(i) PedCon: $B(X, \varepsilon)$
(j) PedCon: the selected edges

The connected components change 102 times.

(k) PedCon: $\beta_0$

The 46 holes are generated.

(l) PedCon: $\beta_1$

**Figure 2. Graph filtration and barcode for edge selection of brain network**

### References

Bassett, D. S., (2006). 'Adaptive reconfiguration of fractal small-world human brain functional networks', Proceedings of the National Academy of Sciences, vol. 103, pp. 19518–19523.

Lee, J. S., . (2004). 'Quantification of brain images using Korean standard templates and structural and cytoarchitectonic probabilistic maps', Korean Journal of Nuclear Medicine, vol. 38, pp. 241–252.

Lee, H., (2011), 'Computing the shape of brain network using graph filtration and Gromov-Haudorff metric ', MICCAI, Lecture Notes in Computer Science (LNCS), vol. 6892, pp. 302-309.

Zomorodian, A. (2005), 'Computing persistent homology', Computing Persistent Homology, Discrete and Computational Geometry, vol. 33, pp. 249-274.