## Definition of anomalous signal

An anomalous signal is an abrupt variation from a normal network pattern.

## Anomaly detection in time series of graph

Commonly, a graph can be defined as:

$$G = (V, E)$$

In which V is the set of vertice and E is the set of edges. Thus, a time series of graphs can be defined as:

$$G(t) = (V, E(t))$$

In which the vertex set is fixed and the edges evolve as a function of time. There are two types of anomalous behavior in time series of graphs: 1) detecting anomalous time points in a time series and 2) detecting anomalous subsets of vertices.

## Simulation settings

*Random Dot Product Graphs*

Suppose for a certain distribution $X_1, X_2, \dots, X_n \sim F$, we have $X_i^T X_j \in [0,1]$ for all $X_i, X_j \in supp\ F \subset \mathbb{R}^d$. Then the distribution is a d-dimensional Inner Product Dstribution. Then if we sample the adjacency matrix on a Bernoulli distribution taking the form of:

$$A_{ij} | (X_i, X_j) \sim Bern(X_i^T X_j)$$

Then we have a randomly generated graph called a random dot product graph, with A being the adjacency matrix. Let $X = [X_1, X_2, \dots, X_n]^T$, then the rows of X gives the latent positions of the graph defined by the adjacency matrix A.

A time series of RDPGs can be defined as $G(t) = (V, E(t))$, and the vertex process can be noted as $\{X(t)\}_{t=1}^{\infty}$, which is an unobserved stochastic process.

*The anomaly detection problem*

If there is no anomalous time point in the whole time series, then the inner stochastic dynamics portrayed by a d-dimensional inner product distribution is identical. If there is a change-point at $t_1$ where the inner stochastic dynamics of a certain vertex i changes abruptly, then:

$$X_i(t_1 - 1) \neq X_i(t_1) \neq X_i(t_1 + 1) \neq \cdots$$

The anomaly detection task in this senario is to find both the time point and the changing vertex.

## Thoughts on how USPORF can be applied to the simulation

In the Euclidean based embeddings, anomalous time points and vertex are detected with the following mindset: firstly, a temporal gap of vertex i at the time point t is defined as:

$$y_i(t) = \left\| \widehat{X}_i(t-1) - \widehat{X}_i(t) \right\|_F$$

In which $\widehat{X_\iota}$ is the estimation of latent position matrix X obtained with two different Euclidean based embeddings which are OMNI and MASE. Then obviously, when $y_i(t)$ is big, there is an anomaly at time t, vertex i.

As for now, I can think of two ways of implementing USPORF on this problem. The first possible way is to use USPORF to get an estimation of the latent position matrices, and then use the matrices to determine the anomaly. The other way is to squeeze the adjacency matrice and run USPORF directly on that.

### Other thoughts and questions