Understanding developing networks with stochastic block models*

Zitong Zhang[†,‡,??], **Shizhe Chen**[§,??]

Address of the First and Second authors
Usually a few lines long
e-mail: zztzhang@ucdavis.edu

Abstract: In neuroscience, it is not well understood how newborn neurons form a mature nervous system, due to the lack of observations. A recent study (Wan et al. 2019) made available a dataset of this functional maturation process on zebrafish. This novel data, however, introduce inherent challenges for the analysis of the formation of the nervous system. First, the formation process is transient by nature. The non-stationarity of the process makes the amount of data pale in comparison to the size of the neural network. Moreover, combining observations on multiple subjects are not straightforward since the neural circuits are not identical across subjects. In this talk, we propose a model for describing the emergence of a coordinated network from isolated nodes. To this end, we adapt and generalize the stochastic block model for random graphs. The proposed method learns the transferable features across subjects, while allowing for individual variabilities. Briefly, the proposed method classifies nodes into different functional groups by identifying typical connecting behavior. We further employ the shape invariant models to handle the nodal delays due to neuron-specific delays. We establish the consistency and minimax optimality of the proposed estimator. We demonstrate the performance of our algorithm on simulation experiments, and on the real zebrafish dataset.

MSC 2010 subject classifications: Primary 60K35, 60K35; secondary 60K35.
Keywords and phrases: sample, LaTeX 2_ε.

1. Introduction

Dynamic networks emerge in many area, such as the neuronal network in the brain during disease[] or learning tasks[], the social network in a time period[], etc. A lot of work has been done in order to study the dynamic networks, but most of these research focus on analyzing the dynamics in a well-developed network. In this paper, we concentrate on the growing networks where isolated nodes develop into a mature functioning network. To the best of our knowledge, there is no existing study about the growing networks, partly because of the lack of data. Fortunately, the data collected by Wan et al. [7] provides us the possibility to pursue this study.

*Footnote to the title with the "thankstext" command.
†Some comment
‡First supporter of the project
§Second supporter of the project

In this paper, we focus on the neural data provided by [7]. Since the neuronal network is complicated, we try to break it down and propose to model it by identifying the typical roles of individual neurons. As supported by [7], neurons in a growing network play different roles in terms of active time and connecting patterns. The connecting pattern of a neuron can be described as the occurrence time of edges that include this neuron. Neurons with the same roles perform similar activities and thus have similar connecting patterns. Learning these roles can help us understand the development procedure of growing networks.

However, being able to identify the typical roles in a single growing network is infeasible because the growing network is transient — there is only one measurement of a growing network due to its non-stationarity. Due to this limit, we combine networks of other subjects as additional samples. This is not trivial, because the neurons in different networks are not one-to-one mapped, and the roles identified from different subjects are not transferable . This constrains our ability to study the common features across subjects. For this reason, we propose to use the stochastic block model as it allows to combine multiple networks and hence resolve the above problems. [SBM is.... It has been applied to ...]

There are, nevertheless, some unique properties of the data that are beyond the scope of the stochastic block model. First, the connection between two neurons is measured over time. Second, the connecting pattern between two neurons are determined not only by their roles but also by their active time, which varies from node to node. Third, the connection is also effected by the spatial distance between neurons — connection cannot occur if two neurons are too away from each other. To incorporate such uniqueness, we propose a generalized stochastic block model in this paper. [Note that these properties also appear in other problems, e.g. venmo...]

Related work

The stochastic block model is first proposed by Holland, Blackmond and Leinhardt [3]. It has many dynamic extensions, Yang et al. [10], Xu and Hero [9], Matias and Miele [4], Xu [8] use the Markov chain to model the time-varying connecting probabilities and/or the clustering matrix. EM algorithm or iterative optimization algorithm is commonly used for inference.

Matias, Rebafka and Villers [5] adapt the stochastic block model to the context of recurrent interaction events in continuous time, where the recurrent events are modeled by Poisson processes with intensities determined by the nodes' group memberships. The maximum likelihood estimator is proposed, but no theoretical analysis is available in the paper.

Optimal rate of convergence is also studied. Gao, Lu and Zhou [2] provides an optimal rate under the mean squared error for the stochastic block model. Pensky [6] derives a penalized least square estimator in a dynamic network setting, and shows that the estimator satisfies an oracle inequality and attains the minimax lower bound for the risk.

Contribution

In this paper, we propose a method for analyzing the growing networks. Our method is able to identify the roles of individual nodes and the connecting patterns. We adapt the stochastic block model to the growing networks context by generalizing the connecting probabilities to intensities of point processes. In addition, we incorporate the time delay of each node so that our model is able to handle the network where nodes become active over time. We derive a least square estimator and show that the estimator converges [in a certain rate]. Finally, an algorithm combining the k-means method and the shape invariant method is proposed for estimation.

Future work

Future working directions include (but not limited to) (i) identifying clusters with similar connecting pattern but different active time phase or different vertex degree, (ii) incorporating the movement of nodes, (iii) seeking for a convex relaxation method that convexify over both clustering matrix and time lags (convex relaxation can also be adapted to solve the penalized least square problem in Pensky [6]), (iv) try other clustering methods.

Organization

The rest of this paper is organized as follows. In Section 2, we review the stochastic block model and introduce the proposed dynamic generalization of the stochastic block model. We introduce the least square estimator and the estimation algorithm in Section 3. Theoretical results are provided in Section 4. Section 5 shows the numerical experiments.

2. Model

2.1. Stochastic block model

A set of n nodes $\Gamma = \{1, \cdots, n\}$ is partitioned into k clusters $\Gamma_1, \cdots, \Gamma_k$. The cluster of node i is represented by $z_i \in \{1, \cdots, k\}$, and the vector of clusters is $\mathbf{z} = (z_i)_{i=1}^n$. Define the adjacency matrix $\mathbf{A} \in \{0,1\}^{n \times n}$ where $A_{i,j} = 1$ if an edge is observed between node i and node j and $A_{i,j} = 0$ otherwise. We set $A_{i,i} \equiv 0$ for any $i = 1, \cdots, n$, and assume that $A_{i,j}$'s are conditionally independent given the cluster vector $\mathbf{z}$:

$$A_{i,j} | z_i = q, z_j = l \overset{ind}{\sim} \text{Bernoulli}(C_{q,l}), \qquad i \neq j,$$

where $\mathbf{C} \in [0,1]^{k \times k}$ denotes the connecting probability matrix.

2.2. Dynamic generalization of the stochastic block model

In a growing network where the edges appear over time, the edge between a pair of nodes i and j can be represented by $N_{i,j}(\cdot)$ with intensify function

$$\lambda_{i,j}(t) = \lambda_{z_i,z_j}(t), \qquad t \in [0,T], \quad i,j = 1,\cdots,n,$$

where $[0,T]$ is overall time period, $\lambda_{z_i,z_j}(\cdot)$ is the connecting intensity function between cluster z_i and z_j. Similar to the stochastic block model, we set $\lambda_{i,i}(\cdot) \equiv 0$ for $i = 1,\cdots,n$.

In practice, there are usually more restrictions to the network. Motivated by the neuronal network (see [7] for detail) where neurons become active in different time and only connect with neighbor neurons, we incorporate the time delay and spatial location of each node and propose the following model:

$$\lambda_{i,j}(t) = \lambda_{z_i,z_j}(t - \tau_{i,j}) \cdot \mathbf{1}_{\{d_{i,j} \le d^*\}}, \qquad t \in [0,T], \quad i,j = 1,\cdots,n,$$

where T and λ_{z_i,z_j} is defined as before, $\tau_{i,j}$ is the time delay caused by both node i and node j, $d_{i,j}$ is the spatial distance between node i and j, and the node i and j can be connected only if $d_{i,j} \le d^*$.
For the convenience of estimation, we make the following assumptions.

Assumption 1. *$\tau_{i,j}$ only depends on node i, that is, $\tau_{i,j} = \tau_i$ for all $j \ne i, i = 1,\cdots,n$.*

With Assumption 1, we may consider the integrated point process $N_i(\cdot) := \sum_{j \ne i} N_{i,j}(\cdot)$ and write its intensity function as

$$\begin{aligned}\lambda_{N_i}(t) &= \sum_{l=1}^{k} \left(\lambda_{z_i,l}(t - \tau_i) \cdot \sum_{j \in \Gamma_l, j \ne i} \mathbf{1}_{\{d_{i,j} \le d^*\}} \right) \\ &=: \sum_{l=1}^{k} \lambda_{z_i,l}(t - \tau_i) \cdot w_{i,l}.\end{aligned}$$

Here $w_{i,l}$ is the number of nodes from cluster l that is in the neighborhood of node i.
We also assume that the nodes are distributed uniformly in the sense that $w_{i,l}$ and $w_{j,l}$ are identically distributed for any nodes i,j from the same cluster. More formally, we have the following assumption.

Assumption 2. *For any $l = 1,\cdots,k$, $\{w_{i,l}\}_{i \in \Gamma_l}$ are i.i.d. with mean $\bar{w}_{z_i,l}$ and variance $\sigma^2 < \infty$.*

By Assumption 2, $\lambda_{N_i}(t+\tau_i) \overset{d}{=} \lambda_{N_j}(t+\tau_j)$ for node i,j with $z_i = z_j$. And hence we can define the mean intensify function for each group $\lambda_l(t) \overset{\triangle}{=} \mathbb{E}\lambda_{N_i}(t+\tau_i), i \in \Gamma_l$. We aim at estimating the clusters $\mathbf{z}$, the mean intensity functions $\{\lambda_l(\cdot)\}_{l=1}^{k}$, and the time delays $\{\tau_i\}_{i=1}^{n}$.

3. Method

3.1. Shift-invariant distance between curves

We aim to minimize the following loss function

$$\min_{\mathbf{z}} \sum_{i=1}^{n} \left(\sum_{q=1}^{k} \alpha_{i,q} \cdot d(\hat{f}_{i,q}, \hat{f}_{z_i,q}) \right), \tag{3.1}$$

where $\alpha_{i,q}$ is the weight measuring the reliability of the distance $d(\hat{f}_{i,q}, \hat{f}_{z_i,q})$. The distance function $d(f, g)$ is defined as

$$d(f,g) = \inf_{\tau \in [-T,T]} \left(\int_{-T}^{2T} |f(t-\tau) - g(t)|^2 \mathrm{d}t \right)^{1/2}, \tag{3.2}$$

which is invariant to shifts of curves and measures the difference between shapes of curves.

3.2. Aligning curves

We use the shape-invariant method proposed by Bigot and Gendre [1] to evaluate the distance defined in (3.2). The idea is to find the optimal shift that aligns f and g best in the Fourier domain. Let θ_j and γ_j, $j = -(N-1)/2, \cdots, (N-1)/2$ be the discrete Fourier coefficients of f and g, where N is the length of discretized f and g. The time shift parameter τ can be estimated by solving the following problem by gradient descent

$$\hat{n} = \operatorname*{arg\,min}_{|n| \le (N-1)/2} \sum_{|j| \le (N-1)/2} \left| \theta_j e^{\mathrm{i}2\pi jn/N} - \gamma_j \right|^2, \tag{3.3}$$

where $n = N \cdot \tau / 2T$.

The objective function (3.3) may have a large plane when the supports of f and g are far away from each other, in which case a good initialization is crucial. To address this problem, we compute and align the corresponding cumulative distribution functions, and the $\hat{\tau}$ leading to the best alignment is then used as the initialization of the gradient descent for solving (3.3).
[Write down the loss function for aligning cdf?]

3.3. Algorithm

We propose to iteratively update the group membership $\mathbf{z}$ and the connecting patterns $\{\hat{f}_{i,q}\}_{i \in [n], q \in [k]}$ and $\{\hat{f}_{l,q}\}_{l,q \in [k]}$.

Given current clustering $\mathbf{z}$ a The algorithm is summarized as follows. Initialization step will be discussed in the next section.

Algorithm 1: [name of algorithm]

Set $s = 0$;
Initialize $\mathbf{z}^{[0]}$ and $\tau^{[0]}$;
while *not converge* **do**
  Update $\{\hat{f}_{i,q}^{[s+1]}\}_{i\in[n],q\in[k]}$ and $\{\hat{f}_{l,q}^{[s+1]}\}_{l,q\in[k]}$ with $\mathbf{z} = \mathbf{z}^{[s]}$ and $\tau = \tau^{[0]}$;
  Compute $\{\alpha_{i,q}\}_{i\in[n],q\in[k]}$ with $\mathbf{z} = \mathbf{z}^{[s]}$;
  Update $\mathbf{z}^{[s+1]}$ with $\{\hat{f}_{i,q}^{[s+1]}\}_{i\in[n],q\in[k]}$, $\{\hat{f}_{l,q}^{[s+1]}\}_{l,q\in[k]}$ and $\{\alpha_{i,q}\}_{i\in[n],q\in[k]}$;
  Evaluate the stopping criterion;
  $s \leftarrow s + 1$;
end
Update $\tau^{[1]}$ with $\mathbf{z}^{[\mathbf{s}]}$ and $\{\hat{f}_{i,q}^{[s]}\}_{i\in[n],q\in[k]}$;
Output: $\mathbf{z}^{[\mathbf{s}]}$, $\{\hat{f}_{l,q}^{[s]}\}_{l,q\in[k]}$.

3.4. Initialization: spectral clustering

Spectral clustering algorithm is used to obtain the initial clustering and time shifts.

4. Theory

Assume $\mathbf{F} = (F_i)_{i=1}^n$ is from the parameter space

$$\mathcal{F}_k = XXX.$$

Theorem 4.1. *For any constant $C' > 0$, there is a constant $C > 0$ only depending on C', such that*

$$\frac{1}{n}\sum_{i=1}^{n}\left\|\hat{F}_i - F_i\right\|^2 \leq C\left(XXX\right),$$

with probability at least $1 - \exp\left(-C'XXX\right)$, uniformly over $\mathbf{F} \in \mathcal{F}_k$.

Proof. This is a sketch of proof and is based on the proof in [2].

We denote the true value by $\theta_i^* = F_{z_i^*}^*(\cdot - \tau_i^*)$. For the estimated $\hat{z}$, define $\tilde{\theta} = \underset{\theta\in\Theta_k(\hat{z})}{\arg\min}\|\theta^* - \theta\|^2$. By the definition of the estimator, we have

$$L(\hat{F}, \hat{Z}, \hat{\tau}) \leq L(F^*, Z^*, \tau^*),$$

which can be rewritten as

$$\|\hat{\theta} - F^{obs}\|^2 \leq \|\theta^* - F^{obs}\|^2. \tag{4.1}$$

The left-hand side of (4.1) can be decomposed as

$$\|\hat{\theta} - \theta^*\|^2 + 2\langle \hat{\theta} - \theta^*, \theta^* - F^{obs} \rangle + \|\theta^* - F^{obs}\|^2. \tag{4.2}$$

Combining (4.1) and (4.2), we have

$$\|\hat{\theta} - \theta^*\|^2 \leq 2\langle \hat{\theta} - \theta^*, F^{obs} - \theta^* \rangle. \tag{4.3}$$

The right-hand side of (4.3) can be bounded as

$$\begin{aligned}
\langle \hat{\theta} - \theta^*, F^{obs} - \theta^* \rangle &= \langle \hat{\theta} - \tilde{\theta}, F^{obs} - \theta^* \rangle + \langle \tilde{\theta} - \theta^*, F^{obs} - \theta^* \rangle \\
&\leq \|\hat{\theta} - \tilde{\theta}\| \left| \left\langle \frac{\hat{\theta} - \tilde{\theta}}{\|\hat{\theta} - \tilde{\theta}\|}, F^{obs} - \theta^* \right\rangle \right| \qquad (4.4) \\
&\quad + \left(\|\tilde{\theta} - \hat{\theta}\| + \|\hat{\theta} - \theta^*\| \right) \left| \left\langle \frac{\tilde{\theta} - \theta^*}{\|\tilde{\theta} - \theta^*\|}, F^{obs} - \theta^* \right\rangle \right|. \qquad (4.5)
\end{aligned}$$

Using Lemmas XXX, the following three terms:

$$\|\hat{\theta} - \tilde{\theta}\|, \quad \left| \left\langle \frac{\hat{\theta} - \tilde{\theta}}{\|\hat{\theta} - \tilde{\theta}\|}, F^{obs} - \theta^* \right\rangle \right|, \quad \left| \left\langle \frac{\tilde{\theta} - \theta^*}{\|\tilde{\theta} - \theta^*\|}, F^{obs} - \theta^* \right\rangle \right| \tag{4.6}$$

can all be bounded by XXX with probability at least

$$XXX.$$

Combining these bounds with (4.4), (4.5) and (4.3), we get

$$\|\hat{\theta} - \theta^*\|^2 \leq XXX$$

with probability at least XXX. □

Now we present the lemmas, which bound the three terms in (4.6), respectively.

Lemma 1. *For any constant $C' > 0$, there exists a constant $C > 0$ only depending on C', such that*

$$\|\hat{\theta} - \tilde{\theta}\| \leq CXXX,$$

with probability at least XXX.

Proof of Lemma 1.
Step 1: Control $\mathbb{E}\|\hat{\tau} - \tilde{\tau}\|^2$.

Fix a group $\hat{\Gamma}_k$, denote the group size by $J = |\Gamma_k|$. For notation simplicity, relabel the nodes in group $\hat{\Gamma}_k$ with $\{1, \cdots, J\}$. Let

$$M(\tau_1, \cdots, \tau_J) = \frac{1}{J}\sum_{j=1}^{J}\sum_{|k|\le k_0}\left|c_{j,k}e^{\mathrm{i}2\pi k\tau_j} - \frac{1}{J}\sum_{j'=1}^{J} c_{j',k}e^{\mathrm{i}2\pi k\tau_{j'}}\right|^2,$$

where $c_{j,k}, k \in \mathbb{Z}$, are the Fourier coefficients of the distribution function $F^*_{z^*_j}(\cdot - \tau^*_j)$. Let

$$\hat{M}(\tau_1, \cdots, \tau_J) = \frac{1}{J}\sum_{j=1}^{J}\sum_{|k|\le k_0}\left|\hat{c}_{j,k}e^{\mathrm{i}2\pi k\tau_j} - \frac{1}{J}\sum_{j'=1}^{J} \hat{c}_{j',k}e^{\mathrm{i}2\pi k\tau_{j'}}\right|^2,$$

where $\hat{c}_{j,k}, k \in \mathbb{Z}$, are the Fourier coefficients of the empirical distribution function F^{obs}_j. Then $\hat{\tau}$ is the minimizer of $\hat{M}$, and $\tilde{\tau}$ is the minimizer of M.
By Proposition 3.1 in Bigot and Gendre [1], under proper assumptions of F and distribution of τ,

$$\frac{1}{J}\|\hat{\tau} - \tilde{\tau}\|^2 \le C^{-1} \cdot (M(\hat{\tau}_1, \cdots, \hat{\tau}_J) - M(\tilde{\tau}_1, \cdots, \tilde{\tau}_J)).$$

Note that $M(\hat{\tau}) - M(\tilde{\tau}) = M(\hat{\tau}) - \hat{M}(\hat{\tau}) + \hat{M}(\hat{\tau}) - M(\tilde{\tau}) \le 2\ \sup_\tau |M(\tau) - \hat{M}(\tau)|$, so it suffices to control $\mathbb{E}\sup_\tau |M(\tau) - \hat{M}(\tau)|$.

$\mathbb{E}\sup_\tau |M(\tau) - \hat{M}(\tau)|$ is controlled by $F^{obs}_j - F^*_j$ or $\hat{c}_{j,k} - c_{j,k}$?

Step 2: Bound $\|\hat{\theta} - \tilde{\theta}\|$.

□

Lemma 2. *For any constant $C' > 0$, there exists a constant $C > 0$ only depending on C', such that*

$$\left|\left\langle \frac{\tilde{\theta} - \theta^*}{\|\tilde{\theta} - \theta^*\|}, F^{obs} - \theta^* \right\rangle\right| \le CXXX,$$

with probability at least XXX.

Proof. Note that

$$\tilde{\theta}_i - \theta^*_i = \tilde{F}_{\hat{z}_i}(\cdot - \tilde{\tau}_i) - F^*_{z^*_i}(\cdot - \tau^*_i)$$

is a function of the partition $\hat{\mathbf{z}}$, then we have

$$\left|\sum_i \left\langle \frac{\tilde{\theta}_i - \theta^*_i}{\sqrt{\sum_i \|\tilde{\theta}_i - \theta^*_i\|^2}}, F^{obs}_i - F^*_{z^*_i}(\cdot - \tau^*_i)\right\rangle\right| \le \max_{\mathbf{z}\in\mathcal{Z}_{n,k}}\left|\sum_i \left\langle \gamma_i(\mathbf{z}), F^{obs}_i - F^*_{z^*_i}(\cdot - \tau^*_i)\right\rangle\right|$$

where

$$\gamma_i(\mathbf{z}) \propto \tilde{F}_{z_i}(\cdot - \tilde{\tau}_i) - F^*_{z^*_i}(\cdot - \tau^*_i)$$

satisfies $\sum_i \|\gamma_i(\mathbf{z})\|^2 = 1$. By [some inequality similar to Hoeffding's inequality] and union bound, we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{\mathbf{z}\in\mathcal{Z}_{n,k}}\left|\sum_i \left\langle \gamma_i(\mathbf{z}), F_i^{obs} - F^*_{z_i^*}(\cdot - \tau_i^*)\right\rangle\right| > t\right)\\
\leq &\sum_{z\in\mathcal{Z}_{n,k}} \mathbb{P}\left(\left|\sum_i \left\langle \gamma_i(\mathbf{z}), F_i^{obs} - F^*_{z_i^*}(\cdot - \tau_i^*)\right\rangle\right| > t\right)\\
\leq &|\mathcal{Z}_{n,k}| \exp\left(-C_1 t^2\right),
\end{aligned}$$

for some universal constant $C_1 > 0$. Choosing $t \propto \sqrt{n\log k}$, the proof is complete.

□

Lemma 3. *For any constant $C' > 0$, there exists a constant $C > 0$ only depending on C', such that*

$$\left|\left\langle \frac{\hat\theta - \tilde\theta}{\|\hat\theta - \tilde\theta\|}, F^{obs} - \theta^*\right\rangle\right| \leq CXXX,$$

with probability at least XXX.

Proof. Note that

$$\hat\theta_i - \tilde\theta_i = \hat F_{\hat z_i}(\cdot - \hat\tau_i) - \tilde F_{\hat z_i}(\cdot - \tilde\tau_i)$$

is a function of both the partition $\hat{\mathbf{z}}$ and the observations F^{obs}. For each $\mathbf{z}\in\mathcal{Z}_{n,k}$, define the set $\mathcal{B}_{\mathbf{z}}$ by

$$XXX$$

Thus, we have the bound

$$\left|\sum_i \left\langle \frac{\tilde\theta_i - \theta_i^*}{\sqrt{\sum_i \|\tilde\theta_i - \theta_i^*\|^2}}, F_i^{obs} - F^*_{z_i^*}(\cdot - \tau_i^*)\right\rangle\right| \leq \max_{\mathbf{z}\in\mathcal{Z}_{n,k}} \sup_{c\in\mathcal{B}_{\mathbf{z}}} \left|\sum_i \left\langle c_i, F_i^{obs} - F^*_{z_i^*}(\cdot - \tau_i^*)\right\rangle\right|$$

If set $\mathcal{B}_{\mathbf{z}}$ is not too large, applying union bound (and Hoeffding-like inequality) completes the proof.

□

5. Simulation

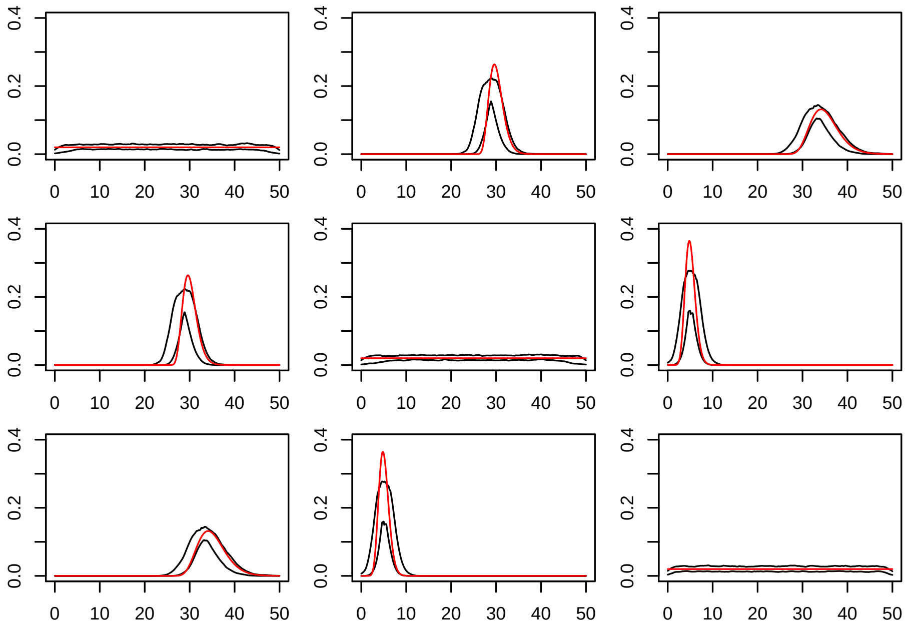


Fig 1: Connecting patterns.

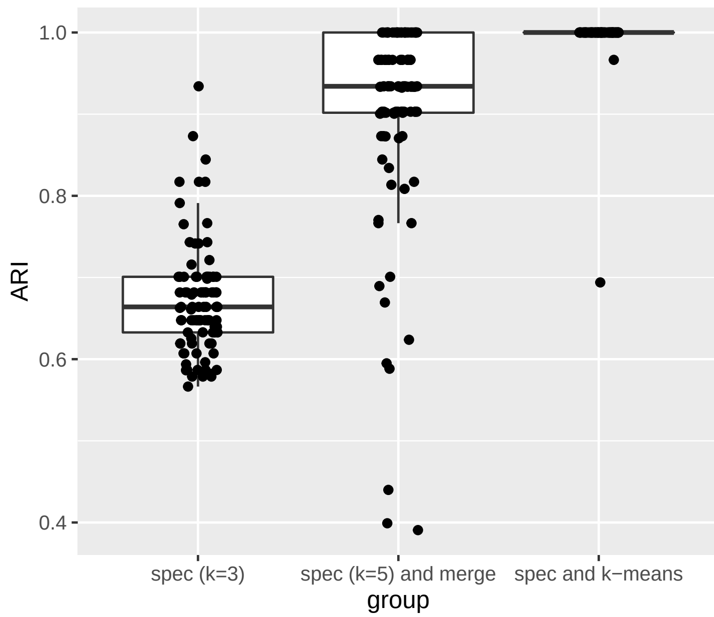
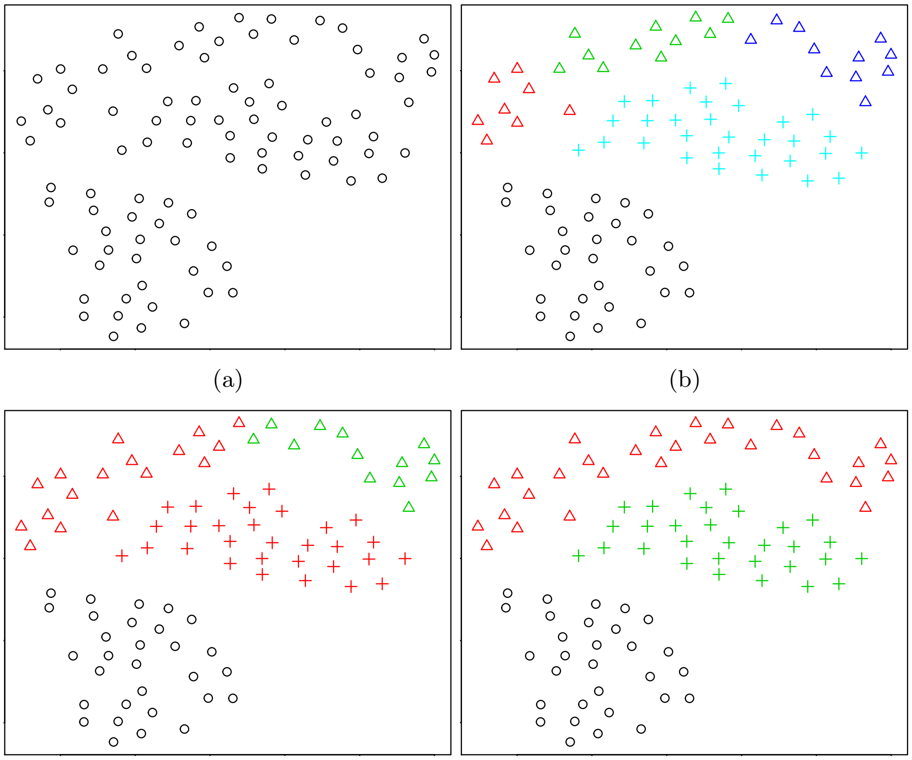


Fig 2: Clustering result.



(a) (b) (c) (d)

Fig 3: (a): 2 dimensional representation of aggregated point processes. (b): Result of spectral clustering with three clusters. (c): Result of spectral clustering with five clusters. (d): Result of merging the five clusters into three.

6. Conclusion

Appendix A: Appendix section

TO DO:

1. Re-write Method section. [by 4/21]
2. Reading: basic knowledge about point process, and shift operation. [by 4/21]
3. Literature review: SBM, best permutation algorithm, SIM, etc. [by 4/21]
4. Proposal. [draft, by 04/28]
5. Slides. [draft, by 04/28]
6. Theory. [05/1-28, if time permits]

7. Apply algorithm of Matias' paper.
8. Real data.

A.1. Plots

Acknowledgements

See Supplement A for the supplementary material example.

Supplementary Material

Supplement A: Title of the Supplement A
(http://www.e-publications.org/ims/support/dowload/imsart-ims.zip). Dum esset rex in accubitu suo, nardus mea dedit odorem suavitatis. Quoniam confortavit seras portarum tuarum, benedixit filiis tuis in te. Qui posuit fines tuos

References

[1] Bigot, J. and Gendre, X. (2013). Minimax properties of Fréchet means of discretely sampled curves. *Ann. Stat.* **41** 923–956.
[2] Gao, C., Lu, Y. and Zhou, H. H. (2015). Rate-optimal graphon estimation. *Ann. Stat.* **43** 2624–2652.
[3] Holland, P. W., Blackmond, K. and Leinhardt, S. (1983). STOCHASTIC BLOCKMODELS: FIRST STEPS * Educational Testing Seroice ** Technical Report.
[4] Matias, C. and Miele, V. (2017). Statistical clustering of temporal networks through a dynamic stochastic block model. *J. R. Stat. Soc. Ser. B (Statistical Methodol.* **79** 1119–1141.
[5] Matias, C., Rebafka, T. and Villers, F. (2018). A semiparametric extension of the stochastic block model for longitudinal networks. *Biometrika* **105** 665–680.
[6] Pensky, M. (2019). Dynamic network models and graphon estimation. *Ann. Stat.* **47** 2378–2403.
[7] Wan, Y., Wei, Z., Looger, L. L., Koyama, M., Druckmann, S., Keller Correspondence, P. J. and Keller, P. J. (2019). Single-Cell Reconstruction of Emerging Population Activity in an Entire Developing Circuit. *Cell* **179**.
[8] Xu, K. S. (2015). Stochastic block transition models for dynamic networks. *J. Mach. Learn. Res.* **38** 1079–1087.
[9] Xu, K. S. and Hero, A. O. (2014). Dynamic stochastic blockmodels for time-evolving social networks. *IEEE J. Sel. Top. Signal Process.* **8** 552–562.
[10] Yang, T., Chi, Y., Zhu, S., Gong, Y. and Jin, R. (2011). Detecting communities and their evolutions in dynamic social networks - A Bayesian approach. *Mach. Learn.* **82** 157–189.