Supplemental material for "Dynamic degree-corrected blockmodels for social networks: a nonparametric approach"

Linda S. L. Tan
statsll@nus.edu.sg
National University of Singapore

Maria De Iorio
m.deiorio@ucl.ac.uk
University College London

S1 Simulated data

In this example, we investigate the ability of the static model to recover the membership of the communities and popularity clusters. We let $n = 30$ and generate a network consisting of three communities, each with $\beta_k^* = 1.5$. We create two popularity clusters, by selecting ten nodes from two of the communities (five from each community) and assign these nodes a higher θ_l^* of 0.5 as compared to -1 for the rest of the nodes. In particular, communities 1, 2, 3 consists of the nodes 1–10, 11–20 and 21–30 respectively. Nodes 1–5 and 26–30 are the nodes with much higher activity level. Fitting the static model using Algorithm 1 took 72 s. We use 20000 iterations, discard the first half and then apply a thinning factor of 5. We set $a_\nu = b_\nu = a_\alpha = b_\alpha = 5$ and $\sigma_\theta^2 = \sigma_\beta^2 = 1$. Applying Binder's loss function to the posterior similarity matrices, we obtain the results shown in Figure S1. For the communities, the memberships of all the nodes were recovered accurately except for nodes 8 and 16. For the popularities, nodes 1 to 5 were identified correctly as the nodes with higher activity levels. However, nodes 26 to 30 were grouped together with the large group of lower activity nodes. Instead, the community that they belonged to was interpreted as having a higher within-group interaction rate. As the communities and popularities are "competing" to observe the network, we note that it may be challenging for the model to distinguish whether interactions should be attributed to communities or popularities at times.

S2 Dolphins social network

Lusseau et al. (2003) constructed an undirected social network describing the associations among a community of 62 bottlenose dolphins living off Doubtful Sound, New Zealand

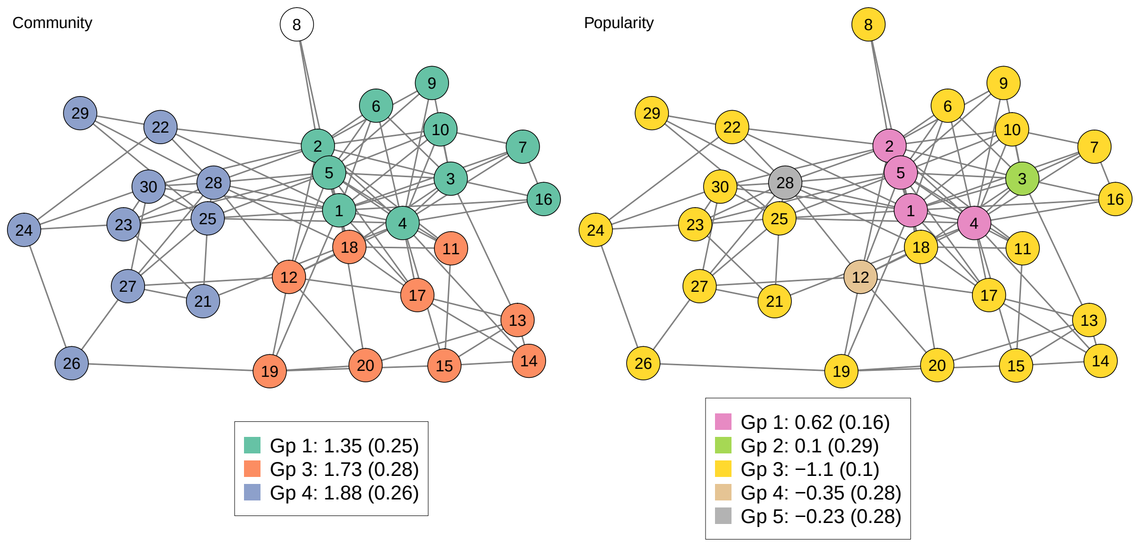

Figure S1: Fit of static model to simulated data.

after observing them for seven years from 1994–2001. This dataset has been widely studied in community detection, see for instance, Lusseau and Newman (2004) and Cao et al. (2015). In this network, the nodes represent dolphins and the ties represent higher than expected frequency of being sighted together. Of the 62 dolphins, 33 are males, 25 are females and the gender of the remaining 4 are unknown.

We apply Algorithm 1 to this network, using 15,000 iterations with a burn-in of 5000 iterations and a thinning factor of 5 in each chain. Three chains were run in parallel and the total runtime is 250 seconds. We set $a_\nu = b_\nu = a_\alpha = b_\alpha = 10$ and $\sigma_\theta^2 = \sigma_\beta^2 = 1$. The marginal posterior distributions of K, L, ν and α are shown in Figure S2. The posterior of K is concentrated on larger values as compared to L and K has a mode of 7 while the mode of L is 2. The posterior similarity matrices in Figure S3 show the community and

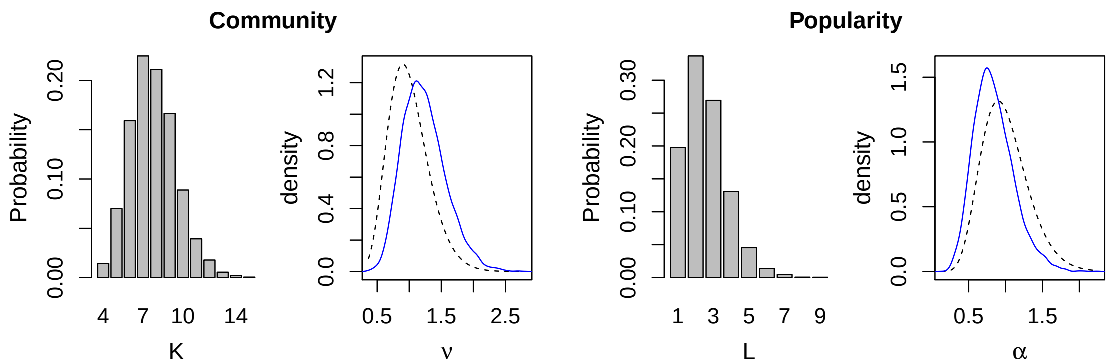

Figure S2: Posterior distributions of K, ν, L and α. For ν and α, prior distributions are shown in dotted lines and posterior distributions in solid (blue) lines.

popularity clustering structure in this network. Around five communities can be seen in the matrix on the left while the right matrix shows faint outlines of two clusters.

We use Binder's loss function to obtain clustering estimates for the community struc-

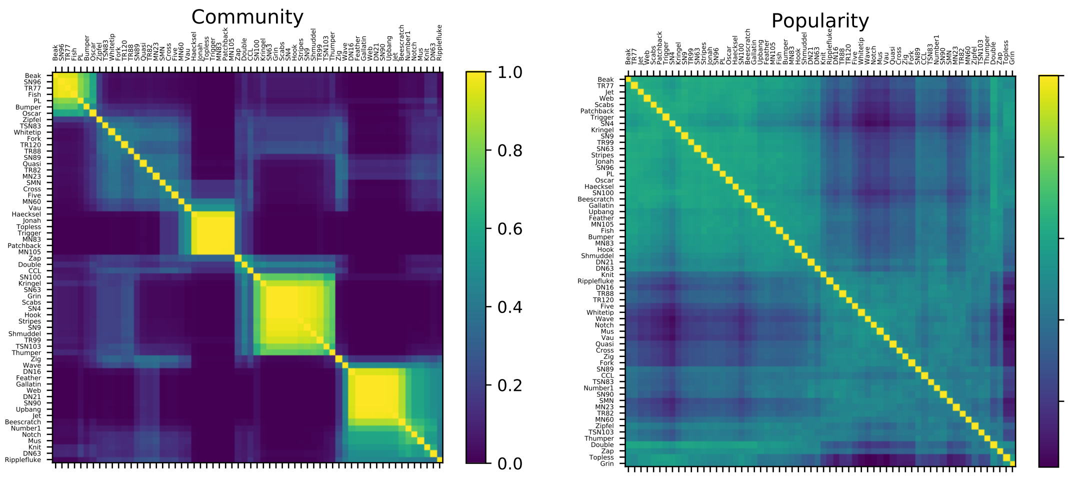

Figure S3: Posterior similarity matrices for community (left) and popularity (right).

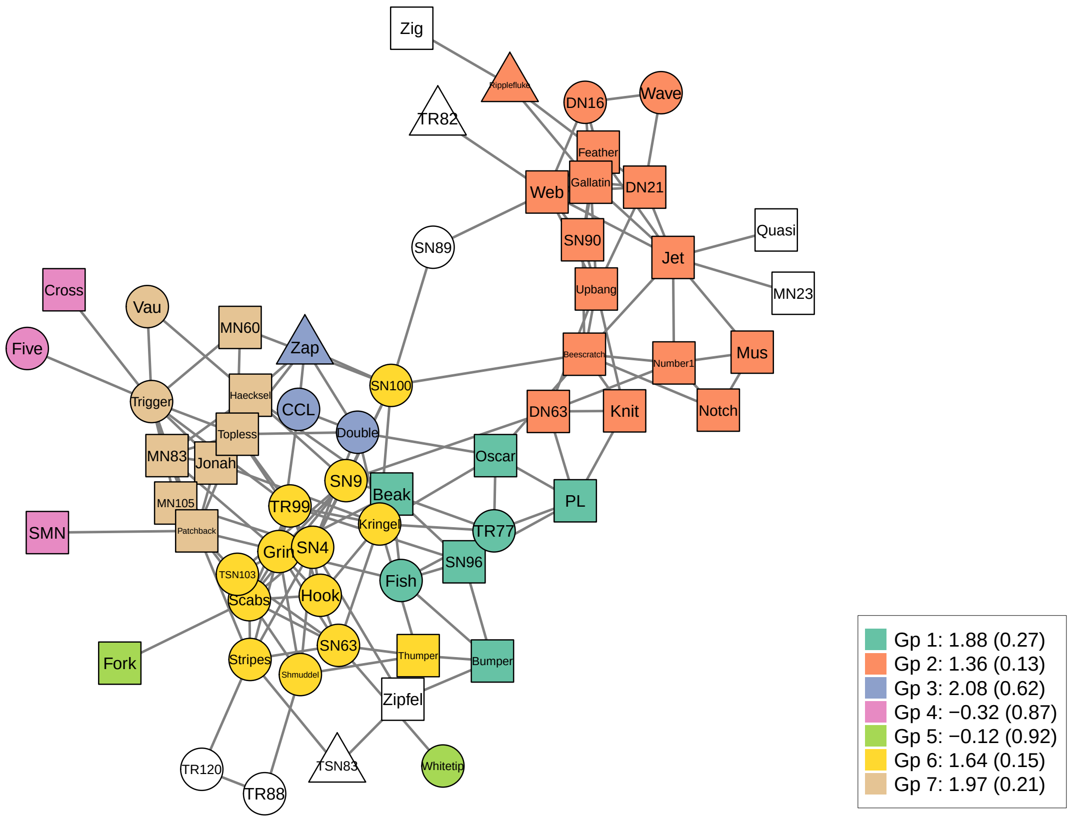

Figure S4: Dolphins social network. Males are represented by squares, females by circles and unknown gender by triangles. Nodes of the same color belong to the same community. Singletons are not colored.

ture and popularity clusterings based on the MCMC samples. This yields 16 communities and a single popularity cluster. Of the 16 communities, 9 are singletons so there are essentially 7 communities. We run Algorithm 1 again, fixing c and z to obtain estimates of β^* and θ^* for these clusterings. The estimate of θ^* is -0.92 ± 0.03. Figure S4 shows the observed dolphins social network where the nodes are labeled with the names of the dolphins. From Figure S4, most of the singletons can be regarded as peripherals (e.g. Zig, TR82, Quasi, MN23); they have few links and lie at the margins of the network. While some of them can clearly be pushed into certain clusters, others such as SN89 lie at the edge of different groups. The estimate of β_k^* is indicative of the rate of interaction for each group k and this is shown in the legend along with the standard deviation in brackets. Groups 1–3 and 6–7 represent close-knit communities while groups 4–5 have low within-group interaction rates.

Recall that in the static model, the communities and popularities are “competing" to explain the network. For this data, application of Binder's function results in a single popularity cluster, which implies that the estimated popularities of all the nodes are close to some common value. Hence the network structure is explained almost wholly using communities, which results in the emergence of weakly connected “communities" with negative β_k^*'s. From the posterior similarity matrix, we observe some trace of a “peripheral group", {Zipfel, TSN83, Whitetip, Fork, TR120, TR 88, SN89, Quasi, TR82, MN23, SMN, Cross, Five}. From the posterior similarity matrix, group 4 arises from this “peripheral group" due to their greater similarity in behavior of being in the same group as actors in group 7 while group 5 arises likely due to their common association with group 6. The model is likely unable to match the nodes in the “peripheral" group with their “most likely" community because they are interacting with members in that community at a rate that is much lower than others. While the introduction of popularities is intended to handle this issue, it appears that the static model explains this network better with the emergence of a “peripheral group".

Previously, Lusseau and Newman (2004) studied the community structure of this network by using a clustering algorithm based on removing links with high “betweeness" measures to extract the groupings (Girvan and Newman, 2002). They also investigated the role that gender and age homophily played in the formation of communities. They concluded that there are 2 main communities and 4 sub-communities; the first sub-community matches group 1 exactly, the second matches group 2 together with the singletons (Zig, TR82, Quasi, MN23), the third matches groups 7 and 4 combined and the fourth matches groups 3 and 6 combined plus the singletons TR120, TR88, TSN83, Zipfel and SN89. We note that the posterior similarity matrix does suggests some of these combinations. Thus, the communities detected by Algorithm 1 agree largely with the results of Lusseau and Newman (2004) and also that of Cao et al. (2015). In addition, Figure S4 also provides some evidence of assortative mixing by sex. For example, group 6 consists almost entirely of females, while groups 2 and 7 are composed of mainly males.

S3 Updates of static model

Let $\mathcal{H}_s = \{(i, j) \in \mathcal{S} | i = s \text{ or } j = s\}$.

- For $i < j$,

$$
\begin{aligned}
p(\zeta_{ij}|\text{rest}) &\propto p(y_{ij}|\zeta_{ij})p(\zeta_{ij}|c_i, c_j, \theta^*, z_i, z_j, \beta^*) \\
&\propto \mathrm{P}(\zeta_{ij} > 0)^{y_{ij}}\mathrm{P}(\zeta_{ij} \le 0)^{1-y_{ij}} \exp\Big\{ -\frac{1}{2}\big[\zeta_{ij}^2 - 2\zeta_{ij}(\theta^*_{c_i} + \theta^*_{c_j} + Z_{ij}^T\beta^*)\big]\Big\}.
\end{aligned}
$$

- For $s = 1, \dots, n$, $p(z_s|\text{rest}) \propto p(z_s|z_{-s}, \nu) \prod_{(i,j)\in\mathcal{H}_s} p(\zeta_{ij}|\theta^*_{c_i}, \theta^*_{c_j}, \beta^*_{z_s})$. Therefore

$$
\mathrm{P}(z_s = k|\text{rest}) = a' m_{-s,k} \exp\Big\{-\frac{1}{2}\sum_{(i,j)\in\mathcal{H}_s}(\zeta_{ij}-\theta^*_{c_i}-\theta^*_{c_j}-\beta^*_k\mathbb{1}\{z_i = z_j = k\})^2\Big\} \text{ for } k \in z_{-s},
$$

$$
\begin{aligned}
&\mathrm{P}(z_s \ne z_j \text{ for all } j \ne s|\text{rest}) \\
&= a'\nu \int \exp\Big\{ -\frac{1}{2}\sum_{(i,j)\in\mathcal{H}_s}(\zeta_{ij} - \theta^*_{c_i} - \theta^*_{c_j} - \beta^*_k\mathbb{1}\{z_i = z_j = k\})^2\Big\} \frac{1}{\sqrt{2\pi}\sigma_\beta} \exp\Big\{ -\frac{\beta_k^{*2}}{2\sigma_\beta^2}\Big\}\, d\beta^*_k \\
&= a'\nu \exp\Big\{ -\frac{1}{2}\sum_{(i,j)\in\mathcal{H}_s}(\zeta_{ij} - \theta^*_{c_i} - \theta^*_{c_j})^2\Big\},
\end{aligned}
$$

where a' is a normalizing constant that ensures the probabilities sum to one. Hence we can simplify the expressions to that in (3.2).

- $$
\begin{aligned}
p(\beta^*|\text{rest}) &\propto \exp\Big\{ -\frac{1}{2}\sum_{i<j}(\zeta_{ij} - \theta^*_{c_i} - \theta^*_{c_j} - Z_{ij}^T\beta^*)^2\Big\} \exp\Big(-\frac{\beta^{*T}\beta^*}{2\sigma_\beta^2}\Big) \\
&\propto \exp\Big\{ -\frac{1}{2}\Big(\beta^{*T}\Big(Z^TZ + \frac{1}{\sigma_\beta^2}\Big)\beta^* - 2\beta^{*T}\sum_{i<j} Z_{ij}(\zeta_{ij} - \theta^*_{c_i} - \theta^*_{c_j})\Big\}.
\end{aligned}
$$

Note that $Z^TZ = \sum_{i<j} Z_{ij}Z_{ij}^T$ is a $K \times K$ diagonal matrix where the kth diagonal element counts the number of pairs of (z_i, z_j) that assume a common value k.

- For $i = 1, \dots, n$,

$$
\begin{aligned}
\mathrm{P}(c_i|\text{rest}) &\propto \exp\Big\{ -\frac{1}{2}\sum_{i<j}(\zeta_{ij} - \theta^*_{c_i} - \theta^*_{c_j} - Z_{ij}^T\beta^*)^2\Big\}p(c|\alpha) \\
&\propto \exp\Big\{\theta^*_{c_i}\sum_{j:j\ne i}(\zeta_{ij} - \theta^*_{c_j} - Z_{ij}^T\beta^*) - \frac{n-1}{2}{\theta^*_{c_i}}^2\Big\}p(c|\alpha).
\end{aligned}
$$

$$
\begin{aligned}
&\therefore P(c_i \ne c_j \text{ for all } j \ne i|\text{rest}) \\
&\propto \alpha \int \exp\Big\{\theta^*_{c_i}\sum_{j:j\ne i}(\zeta_{ij} - \theta^*_{c_j} - Z_{ij}^T\beta^*) - \frac{n-1}{2}{\theta^*_{c_i}}^2\Big\} \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp\{-\frac{{\theta^*_{c_i}}^2}{2\sigma_\theta^2}\}\, d\theta^*_{c_i} \\
&= \frac{\alpha}{\sigma_\theta\sqrt{2\pi}} \int \exp\Big\{\theta^*_{c_i}\sum_{j:j\ne i}(\zeta_{ij} - \theta^*_{c_j} - Z_{ij}^T\beta^*) - \frac{1}{2}\Big(n-1+\frac{1}{\sigma_\theta^2}\Big){\theta^*_{c_i}}^2\Big\}\, d\theta^*_{c_i} = \frac{\alpha\sigma_c}{\sigma_\theta} \exp\Big\{\frac{\mu_{c_i}^2}{2\sigma_c^2}\Big\}.
\end{aligned}
$$

- For $m = 1, \dots, L$,

$$p(\theta^*_m|\text{rest}) \propto \exp\Big\{ -\frac{1}{2}\sum_{i<j}(\zeta_{ij} - \theta^*_{c_i} - \theta^*_{c_j} - Z_{ij}^T\beta^*)^2 \Big\} \exp\Big\{ -\frac{\theta^{*\,2}_m}{2\sigma^2_\theta} \Big\}$$

$$\propto \exp\Big\{ \theta^*_m \Big(2\sum_{\mathcal{S}_m}(\zeta_{ij} - Z_{ij}^T\beta^*) + \sum_{\mathcal{P}_m}(\zeta_{ij} - \theta^*_{c_j} - Z_{ij}^T\beta^*) \Big) - \frac{\theta^{*\,2}_m}{2}\Big(\frac{1}{\sigma^2_\theta} + \sum_{\mathcal{S}_m} 4 + \sum_{\mathcal{P}_m} 1 \Big) \Big\}.$$

S4 Updates of dynamic model I

- For $t = 1, \dots, T$, $i < j$,

$$p(\zeta_{t,ij}|\text{rest}) \propto P(\zeta_{t,ij} > 0)^{y_{t,ij}} \mathrm{P}(\zeta_{t,ij} \le 0)^{1-y_{t,ij}} \exp\Big\{ -\frac{1}{2}\Big(\zeta_{t,ij}^2 - 2\zeta_{t,ij}(\theta^*_{c_{it}} + \theta^*_{c_{jt}} + Z_{ij}^T\beta^*) \Big) \Big\}.$$

- For $s = 1, \dots, n$, $p(z_s|\text{rest}) \propto p(z_s|z_{-s}, \nu) \prod_t \prod_{(i,j)\in\mathcal{H}_s} p(\zeta_{t,ij}|c_{it}, c_{jt}, \theta^*, \beta^*_k)$.

$$\therefore \mathrm{P}(z_s = k|\text{rest}) = a' m_{-s,k} \exp\Big\{ -\frac{1}{2}\sum_t \sum_{(i,j)\in\mathcal{H}_s} (\zeta_{t,ij} - \theta^*_{c_{it}} - \theta^*_{c_{jt}} - \beta^*_k \mathbb{1}\{z_i = z_j = k\})^2 \Big\}$$

for $k \in z_{-s}$ and

$$\mathrm{P}(z_s \ne z_j \text{ for all } j \ne s|\text{rest}) = a'\nu \exp\Big\{ -\frac{1}{2}\sum_t \sum_{(i,j)\in\mathcal{H}_s} (\zeta_{t,ij} - \theta^*_{c_{it}} - \theta^*_{c_{jt}})^2 \Big\}$$

where a' is a normalizing constant to ensure probabilities sum to one. Hence we can simplify the expressions to (3.4).

- $$p(\beta^*|\text{rest}) \propto \exp\Big\{ -\frac{1}{2}\sum_t \sum_{i<j}(\zeta_{t,ij} - \theta^*_{c_{it}} - \theta^*_{c_{jt}} - Z_{ij}^T\beta^*)^2 \Big\} \exp\Big(-\frac{\beta^{*T}\beta^*}{2\sigma^2_\beta} \Big)$$

$$\propto \exp\Big\{ -\frac{1}{2}\Big(\beta^{*T} T Z^T Z \beta^* - 2\beta^{*T} \sum_{i<j} Z_{ij} \sum_t (\zeta_{t,ij} - \theta^*_{c_{it}} - \theta^*_{c_{jt}}) + \frac{\beta^{*T}\beta^*}{\sigma^2_\beta} \Big) \Big\}.$$

- For $i = 1, \dots, n$, $t = 1, \dots, T$,

$$\mathrm{P}(c_{it}|\text{rest}) \propto \exp\Big\{ -\frac{1}{2}\sum_{i<j}(\zeta_{t,ij} - \theta^*_{c_{it}} - \theta^*_{c_{jt}} - Z_{ij}^T\beta^*)^2 \Big\} p(c|\alpha).$$

$$\propto \exp\Big\{ -\frac{n-1}{2}\theta^{*\,2}_{c_{it}} + \theta^*_{c_{it}} \sum_{j:j\ne i}(\zeta_{t,ij} - \theta^*_{c_{jt}} - Z_{ij}^T\beta^*) \Big\} p(c|\alpha).$$

$\therefore \mathrm{P}(c_{it} \neq c_{jt'} \text{ for all } j \neq i \text{ or } t' \neq t|\text{rest})$

$$= b\alpha \int \exp\Big\{ -\frac{n-1}{2}{\theta^*_{c_{it}}}^2 + \theta^*_{c_{it}} \sum_{j:j\neq i}(\zeta_{t,ij} - \theta^*_{c_{jt}} - Z_{ij}^T\beta^*)\Big\} \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp\{-\frac{{\theta^*_{c_{it}}}^2}{2\sigma_\theta^2}\}\, d\theta^*_{c_{it}}$$

$$= \frac{b\alpha}{\sigma_\theta\sqrt{2\pi}} \int \exp\Big\{ -\frac{1}{2}\Big(n-1+\frac{1}{\sigma_\theta^2}\Big){\theta^*_{c_{it}}}^2 + \theta^*_{c_{it}} \sum_{j:j\neq i}(\zeta_{t,ij} - \theta^*_{c_{jt}} - Z_{ij}^T\beta^*)\Big\}\, d\theta^*_{c_{it}}.$$

- For $m = 1, \dots, L$,

$$p(\theta^*_m|\text{rest}) \propto \exp\Big\{ -\frac{1}{2}\sum_t\sum_{i<j}(\zeta_{t,ij} - \theta^*_{c_{it}} - \theta^*_{c_{jt}} - Z_{ij}^T\beta^*)^2\Big\} \exp\Big\{ -\frac{{\theta^*_m}^2}{2\sigma_\theta^2}\Big\}$$

$$\propto \exp\Big\{ -\frac{{\theta^*_m}^2}{2}\Big(\frac{1}{\sigma_\theta^2} + \sum_t\sum_{\mathcal{S}_{t,m}} 4 + \sum_t\sum_{\mathcal{P}_{t,m}} 1\Big) + \theta^*_m\Big(2\sum_t\sum_{\mathcal{S}_{t,m}}(\zeta_{t,ij} - Z_{ij}^T\beta^*) + \sum_t\sum_{\mathcal{P}_{t,m}}(\zeta_{t,ij} - \theta^*_{c_{jt}} - Z_{ij}^T\beta^*)\Big)\Big\}.$$

S5 Updates of dynamic model II

- For $t = 1, \dots, T$, $(i,j) \in \mathcal{S}$,

$$p(\zeta_{t,ij}|\text{rest}) \propto \mathrm{P}(\zeta_{t,ij} > 0)^{y_{t,ij}}\mathrm{P}(\zeta_{t,ij} \leq 0)^{1-y_{t,ij}} \exp\Big\{ -\frac{1}{2}\Big({\zeta_{t,ij}}^2 - 2\zeta_{t,ij}(\eta y_{t-1,ij}\mathbb{1}\{t>1\} + \theta^*_{c_i} + \theta^*_{c_j} + Z_{ij}^T\beta^*)\Big)\Big\}.$$

- $p(z_s|\text{rest}) \propto p(z_s|z_{-s},\nu)\prod_t\prod_{(i,j)\in\mathcal{H}_s} p(\zeta_{t,ij}|\theta^*_{c_i},\theta^*_{c_j},\eta,y,\beta^*_{z_s})$

$$\propto p(z_s|z_{-s},\nu)\exp\Big\{ -\frac{1}{2}\sum_t\sum_{(i,j)\in\mathcal{H}_s}(\tilde{\zeta}_{t,ij} - \theta^*_{c_i} - \theta^*_{c_j} - \beta^*_k\mathbb{1}\{z_i = z_j = k\})^2\Big\}.$$

  For $k \in z_{-s}$,

$$\mathrm{P}(z_s = k|\text{rest}) = a' m_{-s,k}\exp\Big\{ -\frac{1}{2}\sum_t\sum_{(i,j)\in\mathcal{H}_s}(\tilde{\zeta}_{t,ij} - \theta^*_{c_i} - \theta^*_{c_j} - \beta^*_k\mathbb{1}\{z_i = z_j = k\})^2\Big\},$$

  and

$$\mathrm{P}(z_s \neq z_j \text{ for all } j \neq s|\text{rest}) = a'\nu\exp\Big\{ -\frac{1}{2}\sum_t\sum_{(i,j)\in\mathcal{H}_s}(\tilde{\zeta}_{t,ij} - \theta^*_{c_i} - \theta^*_{c_j})^2\Big\},$$

  where a' is a normalizing constant to ensure probabilities sum to one. Hence we can simplify the expressions to (3.6).

- $p(\beta^*|\text{rest}) \propto \exp\Big\{-\dfrac{1}{2}\sum_t\sum_{i<j}(\tilde{\zeta}_{t,ij} - \theta^*_{c_{it}} - \theta^*_{c_{jt}} - Z_{ij}^T\beta^*)^2\Big\}\exp\Big(-\dfrac{\beta^{*T}\beta^*}{2\sigma_\beta^2}\Big)$

$$\propto \exp\Big\{-\frac{1}{2}\Big(\beta^{*T}TZ^TZ\beta^* - 2\beta^{*T}\sum_{i<j}Z_{ij}\sum_t(\tilde{\zeta}_{t,ij} - \theta^*_{c_{it}} - \theta^*_{c_{jt}}) + \frac{\beta^{*T}\beta^*}{\sigma_\beta^2}\Big)\Big\}.$$

- For $i = 1, \ldots, n$,

$$p(c_i|\text{rest}) \propto p(c_i|c_{-i},\alpha)\exp\Big\{-\frac{1}{2}\sum_t\sum_{i<j}\big(\tilde{\zeta}_{t,ij} - \theta^*_{c_i} - \theta^*_{c_j} - Z_{ij}^T\beta^*\big)^2\Big\}.$$

$$\propto p(c_i|c_{-i},\alpha)\exp\Big\{\theta^*_{c_i}\sum_t\sum_{j:j\neq i}(\tilde{\zeta}_{t,ij} - \theta^*_{c_j} - Z_{ij}^T\beta^*) - \frac{T(n-1)}{2}{\theta^*_{c_i}}^2\Big\}.$$

$\therefore \mathrm{P}(c_i \neq c_j \text{ for all } j \neq i|\text{rest})$

$$\propto \alpha\int\exp\Big\{\theta^*_{c_i}\sum_t\sum_{j:j\neq i}(\tilde{\zeta}_{t,ij} - \theta^*_{c_j} - Z_{ij}^T\beta^*) - \frac{T(n-1)}{2}{\theta^*_{c_i}}^2\Big\}\frac{1}{\sqrt{2\pi}\sigma_\theta}\exp\{-\frac{{\theta^*_{c_i}}^2}{2\sigma_\theta^2}\}\,d\theta^*_{c_i}$$

$$\propto \frac{\alpha}{\sigma_\theta\sqrt{2\pi}}\int\exp\Big\{\theta^*_{c_i}\sum_t\sum_{j:j\neq i}(\tilde{\zeta}_{t,ij} - \theta^*_{c_j} - Z_{ij}^T\beta^*) - \frac{1}{2}\Big(T(n-1) + \frac{1}{\sigma_\theta^2}\Big){\theta^*_{c_i}}^2\Big\}\,d\theta^*_{c_i}.$$

- For $m = 1, \ldots, L$,

$$p(\theta^*_m|\text{rest}) \propto \exp\Big\{-\frac{1}{2}\sum_t\sum_{i<j}(\tilde{\zeta}_{t,ij} - \theta^*_{c_i} - \theta^*_{c_j} - Z_{ij}^T\beta^*)^2\Big\}\exp\Big\{-\frac{{\theta^*_m}^2}{2\sigma_\theta^2}\Big\}$$

$$\propto \exp\Big\{-\frac{1}{2}{\theta^*_m}^2\Big(\frac{1}{\sigma_\theta^2} + \sum_{\mathcal{S}_m}4T + \sum_{\mathcal{P}_m}T\Big) + \theta^*_c\Big(2\sum_t\sum_{\mathcal{S}_m}(\tilde{\zeta}_{t,ij} - Z_{ij}^T\beta^*)$$
$$+\sum_t\sum_{\mathcal{P}_m}(\tilde{\zeta}_{t,ij} - \theta^*_{c_j} - Z_{ij}^T\beta^*)\Big)\Big\}.$$

- $p(\eta|\text{rest}) \propto \exp\Big\{-\dfrac{1}{2}\sum_{t=2}^{T}\sum_{i<j}(\zeta_{t,ij} - \eta y_{t-1,ij} - \theta^*_{c_i} - \theta^*_{c_j} - \beta^TZ_{ij})^2 - \dfrac{\eta^2}{2\sigma_\eta^2}\Big\}$

$$\propto \exp\Big\{-\frac{\eta^2}{2}\Big(\frac{1}{\sigma_\eta^2} + \sum_{t=2}^{T}\sum_{i<j}y_{t-1,ij}^2\Big) + \eta\sum_{t=2}^{T}\sum_{i<j}y_{t-1,ij}(\zeta_{t,ij} - \theta^*_{c_i} - \theta^*_{c_j} - \beta^TZ_{ij})\Big\}.$$

S6 OpenBUGS code for static model

```
model{
for (i in 1:(n-1)) {
 for (j in (i+1):n) {
   y[i,j] ~ dbern(p[i,j])
   p[i,j] <- phi(theta[i] + theta[j] + I[i,j]*betastar[z[i]])
   I[i,j] <- equals(z[i],z[j])
```

```
}}
# DP for popularity parameters (L: upper bound on number of components)
for (i in 1:n){
 theta[i] <- x[s[i]]
 s[i] ~ dcat(u[])
}
for (l in 1:L){ x[l] ~ dnorm(0, 1) }
for (l in 1:(L-1)){ r[l] ~ dbeta(1, alpha) }
r[L] <- 1
u[1] <- r[1]
for (l in 2:L){ u[l] <- r[l]*(1-r[l-1])*u[l-1]/r[l-1] }

# DP for interaction parameters (K: upper bound on number of components)
for (i in 1:n){
 z[i] ~ dcat(b[])
 beta[i] <- betastar[z[i]]
}
for (k in 1:K){ betastar[k] ~ dnorm(0, 1) }
for (k in 1:(K-1)){ a[k] ~ dbeta(1, nu) }
a[K] <- 1
b[1] <- a[1]
for (k in 2:K){ b[k] <- a[k]*(1-a[k-1])*b[k-1]/a[k-1] }

# hyperparameters
alpha ~ dgamma(10, 10)
nu ~ dgamma(10, 10)
}
```

References

Cao, X., X. Wang, D. Jin, X. Guo, and X. Tang (2015). A stochastic model for detecting overlapping and hierarchical community structure. *PLoS ONE 10*, e0119171.

Girvan, M. and M. E. J. Newman (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences 99*, 7821–7826.

Lusseau, D. and M. E. J. Newman (2004). Identifying the role that animals play in their social networks. *Proceedings of the Royal Society of London B: Biological Sciences 271*, S477–S481.

Lusseau, D., K. Schneider, O. Boisseau, P. Haase, E. Slooten, and S. Dawson (2003). The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology 54*, 396–405.