# Formularius

## Contents



# 1 Binary SBM

## 1.1 Undirected/Directed

### 1.1.1 MAR

In the directed case, only $\mathcal{D}^{\text{o}}$ is changing in the following (i.e don't forget that the matrix is not symmetric: $Y_{ij}$ and $Y_{ji}$ are two different random variables).

**Proposition 1.** *The complete log-likelihood restricted to the observed variables is*

$$\log p_\theta(Y^{\text{o}}, Z) = \sum_{(i,j)\in\mathcal{D}^{\text{o}}} \sum_{(q,\ell)\in\mathcal{Q}^2} Z_{iq} Z_{j\ell} \log b(Y_{ij}, \pi_{q\ell}) + \sum_{i\in\mathcal{N}^{\text{o}}} \sum_{q\in\mathcal{Q}} Z_{iq} \log(\alpha_q),$$

*with $b(x,\pi) = \pi^x(1-\pi)^{1-x}$ the Bernoulli probability density function.*

**Proposition 2.** *The variationnal approximation is*

$$J_{\tau,\theta}(Y^{\text{o}}) = \sum_{(i,j)\in\mathcal{D}^{\text{o}}} \sum_{(q,\ell)\in\mathcal{Q}^2} \tau_{iq}\tau_{j\ell} \log b(Y_{ij}, \pi_{q\ell}) + \sum_{i\in\mathcal{N}^{\text{o}}} \sum_{q\in\mathcal{Q}} \tau_{iq} \log(\alpha_q/\tau_{iq}).$$

**Proposition 3.** *Consider the lower bound $J_{\tau,\theta}(Y^{\text{o}})$.*

1. *The parameters $\theta = (\alpha, \pi)$ maximizing $J_\theta(Y^{\text{o}})$ when $\tau$ is held fixed are*

$$\hat{\alpha}_q = \frac{\sum_{i\in\mathcal{N}^{\text{o}}} \hat{\tau}_{iq}}{\text{card}\,(\mathcal{N}^{\text{o}})}, \qquad \hat{\pi}_{q\ell} = \frac{\sum_{(i,j)\in\mathcal{D}^{\text{o}}} \hat{\tau}_{iq}\hat{\tau}_{j\ell} Y_{ij}}{\sum_{(i,j)\in\mathcal{D}^{\text{o}}} \hat{\tau}_{iq}\hat{\tau}_{j\ell}}.$$

2. *The variational parameters $\tau$ maximizing $J_\tau(Y^{\text{o}})$ when $\theta$ is held fixed are obtained thanks to the following fixed point relation:*

$$\hat{\tau}_{iq} \propto \alpha_q \left( \prod_{(i,j)\in\mathcal{D}^{\text{o}}} \prod_{\ell\in\mathcal{Q}} b(Y_{ij}; \pi_{q\ell})^{\hat{\tau}_{j\ell}} \right).$$

$$\hat{\tau}_{iq} \propto \alpha_q \left( \prod_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \prod_{\ell\in\mathcal{Q}} (b(Y_{ij};\pi_{q\ell}) b(Y_{ji};\pi_{\ell q}))^{\hat{\tau}_{j\ell}} \right).$$

**Proposition 4.** *For an SBM with $Q$ blocks and for $\hat{\theta} = \arg\max \log p_\theta(Y^{\mathrm{o}}, Z)$, the ICL criterion is given by*

$$\mathrm{ICL}(Q) = -2\mathbb{E}_{\tilde{p}_\tau}\left[\log p_{\hat{\theta}}(Y^{\mathrm{o}}, Z; Q)\right] + \frac{Q(Q+1)}{2} \log \operatorname{card}(\mathcal{D}^{\mathrm{o}}) + (Q-1) \log \operatorname{card}(\mathcal{N}^{\mathrm{o}}).$$

$$\mathrm{ICL}(Q) = -2\mathbb{E}_{\tilde{p}_\tau}\left[\log p_{\hat{\theta}}(Y^{\mathrm{o}}, Z; Q)\right] + Q^2 \log \operatorname{card}(\mathcal{D}^{\mathrm{o}}) + (Q-1) \log \operatorname{card}(\mathcal{N}^{\mathrm{o}}).$$

**Likelihood of Random-Pair sampling**

$$p_\rho(R) = \prod_{1\le i<j\le n} p_\rho(R_{ij}) = \left(\prod_{Y^{\mathrm{o}}} \rho\right)\left(\prod_{Y^{\mathrm{m}}} (1-\rho)\right) = \rho^{\operatorname{card}(Y^{\mathrm{o}})}(1-\rho)^{\operatorname{card}(Y^{\mathrm{m}})}$$

$$p_\rho(R) = \prod_{1\le i\neq j\le n} p_\rho(R_{ij}) = \left(\prod_{Y^{\mathrm{o}}} \rho\right)\left(\prod_{Y^{\mathrm{m}}} (1-\rho)\right) = \rho^{\operatorname{card}(Y^{\mathrm{o}})}(1-\rho)^{\operatorname{card}(Y^{\mathrm{m}})}$$

**Likelihood of star sampling** The likelihood is the same in symmetric and non-symmetric context.

$$p_\rho(S) = \prod_{1\le i\le n} p_\rho(S_i) = \left(\prod_{i,S_i=1} \rho\right)\left(\prod_{i,S_i=0} (1-\rho)\right) = \rho^{\sum_i S_i}(1-\rho)^{\sum_i (1-S_i)}$$

**ICL in random-pair sampling case**

**Proposition 5.** *For an SBM with $Q$ blocks and for $\hat{\theta} = \arg\max \log p_\theta(Y^{\mathrm{o}}, Z, R)$, the ICL criterion is given by*

$$\mathrm{ICL}(Q) = -2\mathbb{E}_{\tilde{p}_\tau}\left[\log p_{\hat{\theta}}(Y^{\mathrm{o}}, Z, R; Q)\right] + (1 + \frac{Q(Q+1)}{2}) \log \operatorname{card}(\mathcal{D}^{\mathrm{o}}) + (Q-1) \log \operatorname{card}(\mathcal{N}^{\mathrm{o}}).$$

$$\mathrm{ICL}(Q) = -2\mathbb{E}_{\tilde{p}_\tau}\left[\log p_{\hat{\theta}}(Y^{\mathrm{o}}, Z, R; Q)\right] + (1 + Q^2) \log \operatorname{card}(\mathcal{D}^{\mathrm{o}}) + (Q-1) \log \operatorname{card}(\mathcal{N}^{\mathrm{o}}).$$

**ICL in star sampling case**

**Proposition 6.** *For an SBM with $Q$ blocks and for $\hat{\theta} = \arg\max \log p_\theta(Y^{\mathrm{o}}, Z, R)$, the ICL criterion is given by*

$$\mathrm{ICL}(Q) = -2\mathbb{E}_{\tilde{p}_\tau}\left[\log p_{\hat{\theta}}(Y^{\mathrm{o}}, Z, R; Q)\right] + \frac{Q(Q+1)}{2} \log \operatorname{card}(\mathcal{D}^{\mathrm{o}}) + Q \log \operatorname{card}(\mathcal{N}^{\mathrm{o}}).$$

$$\mathrm{ICL}(Q) = -2\mathbb{E}_{\tilde{p}_\tau}\left[\log p_{\hat{\theta}}(Y^{\mathrm{o}}, Z, R; Q)\right] + Q^2 \log \operatorname{card}(\mathcal{D}^{\mathrm{o}}) + Q \log \operatorname{card}(\mathcal{N}^{\mathrm{o}}).$$

### 1.1.2 NMAR

Notice that $\mathcal{D}^{\mathrm{o}}$ and $\mathcal{D}^{\mathrm{m}}$ are changing in the following (i.e don't forget that the matrix is not symmetric: $Y_{ij}$ and $Y_{ji}$ are two different random variables).

**Proposition 7.** *The complete log-likelihood is*

$$\log p_{\theta,\psi}(Y^{\mathrm{o}}, R, Y^{\mathrm{m}}, Z) = \log p_{\psi}(R|Y^{\mathrm{o}}, Y^{\mathrm{m}}, Z) + \log p_{\theta}(Y^{\mathrm{o}}, Y^{\mathrm{m}}, Z),$$

*where*

$$\log p_{\theta}(Y, Z) = \sum_{1\le i<j\le n} \sum_{(q,\ell)\in\mathcal{Q}^2} Z_{iq} Z_{j\ell} \log b(Y_{ij}, \pi_{q\ell}) + \sum_{i\in\{1,\ldots n\}} \sum_{q\in\mathcal{Q}} Z_{iq} \log(\alpha_q),$$

$$\log p_{\theta}(Y, Z) = \sum_{1\le i\neq j\le n} \sum_{(q,\ell)\in\mathcal{Q}^2} Z_{iq} Z_{j\ell} \log b(Y_{ij}, \pi_{q\ell}) + \sum_{i\in\{1,\ldots n\}} \sum_{q\in\mathcal{Q}} Z_{iq} \log(\alpha_q),$$

*with $b(x, \pi) = \pi^x (1-\pi)^{1-x}$ the Bernoulli probability density function.*

**Proposition 8.** *The variationnal approximation is*

$$\begin{aligned} J_{\tau,\nu,\theta,\psi}(Y^{\mathrm{o}}, R) &= \mathbb{E}_{\tilde{p}_{\tau,\nu}}\left[\log p_{\theta,\psi}(Y^{\mathrm{o}}, R, Y^{\mathrm{m}}, Z)\right] - \mathbb{E}_{\tilde{p}_{\tau,\nu}}\left[\log \tilde{p}_{\tau,\nu}(Z, Y^{\mathrm{m}})\right] \quad (1)\\ &= \mathbb{E}_{\tilde{p}_{\tau,\nu}}\left[\log p_{\psi}(R|Y^{\mathrm{o}}, Y^{\mathrm{m}}, Z)\right] \\ &\quad + \sum_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \sum_{(q,\ell)\in\mathcal{Q}^2} \tau_{iq}\tau_{j\ell} \log b(Y_{ij}, \pi_{q\ell}) + \sum_{(i,j)\in\mathcal{D}^{\mathrm{m}}} \sum_{(q,\ell)\in\mathcal{Q}^2} \tau_{iq}\tau_{j\ell} \log b(\nu_{ij}, \pi_{q\ell}) \\ &\quad + \sum_{i\in\mathcal{N}} \sum_{q\in\mathcal{Q}} \tau_{iq} \log(\alpha_q/\tau_{iq}) - \sum_{(i,j)\in\mathcal{D}^{\mathrm{m}}} \nu_{ij} \log(\nu_{ij}) + (1-\nu_{ij}) \log(1-\nu_{ij}). \end{aligned}$$

**Proposition 9.** *Consider the lower bound $J_{\tau,\nu,\theta,\psi}(Y^{\mathrm{o}}, R)$ given by (1).*

1. *The parameters $\theta = (\alpha, \pi)$ maximizing (1) when all other parameters are held fixed are*

$$\hat{\alpha}_q = \frac{1}{n} \sum_{i\in\mathcal{N}} \hat{\tau}_{iq}, \qquad \hat{\pi}_{q\ell} = \frac{\sum_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \hat{\tau}_{iq}\hat{\tau}_{j\ell} Y_{ij} + \sum_{(i,j)\in\mathcal{D}^{\mathrm{m}}} \hat{\tau}_{iq}\hat{\tau}_{j\ell}\hat{\nu}_{ij}}{\sum_{(i,j)\in\mathcal{D}} \hat{\tau}_{iq}\hat{\tau}_{j\ell}}.$$

2. *The optimal $\tau$ in (1) when all other parameters are held fixed verifies*

$$\hat{\tau}_{iq} \propto \lambda_{iq}\alpha_q \left( \prod_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \prod_{\ell\in\mathcal{Q}} b(Y_{ij}; \pi_{q\ell})^{\hat{\tau}_{j\ell}} \right) \left( \prod_{(i,j)\in\mathcal{D}^{\mathrm{m}}} \prod_{\ell\in\mathcal{Q}} b(\nu_{ij}; \pi_{q\ell})^{\hat{\tau}_{j\ell}} \right).$$

$$\hat{\tau}_{iq} \propto \lambda_{iq}\alpha_q \left( \prod_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \prod_{\ell\in\mathcal{Q}} \left(b(Y_{ij}; \pi_{q\ell}) b(Y_{ji}; \pi_{\ell q})\right)^{\hat{\tau}_{j\ell}} \right) \left( \prod_{(i,j)\in\mathcal{D}^{\mathrm{m}}} \prod_{\ell\in\mathcal{Q}} \left(b(\nu_{ij}; \pi_{q\ell}) b(\nu_{ji}; \pi_{\ell q})\right)^{\hat{\tau}_{j\ell}} \right).$$

   *with $\lambda_{iq}$ a simple constant depending on the sampling design.*

**Double-standard sampling.** Recall that $S^{\mathrm{o}} = \sum_{(i,j)\in\mathcal{D}^{\mathrm{o}}} Y_{ij}$, $\bar{S}^{\mathrm{o}} = \sum_{(i,j)\in\mathcal{D}^{\mathrm{o}}} (1 - Y_{ij})$, and denote $S^{\mathrm{m}} = \sum_{(i,j)\in\mathcal{D}^{\mathrm{m}}} \nu_{ij}$, $\bar{S}^{\mathrm{m}} = \sum_{(i,j)\in\mathcal{D}^{\mathrm{m}}} (1 - \nu_{ij})$. Regarding the likelihood of the double standard sampling, we have

$$\mathbb{E}_{\tilde{p}} \log p_{\psi}(R|Y) = S^{\mathrm{o}} \log \rho_1 + \bar{S}^{\mathrm{o}} \log \rho_0 + S^{\mathrm{m}} \log(1-\rho_1) + \bar{S}^{\mathrm{m}} \log(1-\rho_0).$$

Based on this expression, we easily derive the following proposition:

**Proposition 10.** *Consider the maximization of the lower bound* (1) *in the double standard sampling.*

1. *The parameters* $\psi = (\rho_0, \rho_1)$ *maximizing* (1) *when all other parameters are held fixed are*

$$\hat{\rho}_0 = \frac{\bar{S}^{\mathrm{o}}}{\bar{S}^{\mathrm{o}} + \bar{S}^{\mathrm{m}}}, \qquad \hat{\rho}_1 = \frac{S^{\mathrm{o}}}{S^{\mathrm{o}} + S^{\mathrm{m}}}. \tag{1}$$

2. *The optimal* $\nu$ *in* (1) *when all other parameters are held fixed are*

$$\hat{\nu}_{ij} = \mathrm{logistic}\left(\log\left(\frac{1-\rho_1}{1-\rho_0}\right) + \sum_{(q,\ell)\in\mathcal{Q}^2} \tau_{iq}\tau_{j\ell}\log\left(\frac{\pi_{q\ell}}{1-\pi_{q\ell}}\right)\right).$$

*Moreover,* $\lambda_{iq} = 1\ \forall (i,q) \in \mathcal{N}\times\mathcal{Q}$ *for optimization of* $\tau$ *in Proposition 9.2).*

**Class sampling.** According to the likelihood of the class sampling, we derive the following expression of the conditional expectation under the variational approximation:

$$\mathbb{E}_{\tilde{p}} \log p_\psi(R|Y) = \sum_{i\in\mathcal{N}^{\mathrm{o}}}\sum_{q\in\mathcal{Q}} \tau_{iq}\log(\rho_q) + \sum_{i\in\mathcal{N}^{\mathrm{m}}}\sum_{q\in\mathcal{Q}} \tau_{iq}\log(1-\rho_q),$$

from which we derive the maximization of the remaining parameters for class sampling.

**Proposition 11.** *Consider the maximization of the lower bound* (1) *in the class sampling.*

1. *The parameters* $\psi = (\rho_1, \dots, \rho_Q)$ *maximizing* (1) *when all other parameters are held fixed are*

$$\hat{\rho}_q = \frac{\sum_{i\in\mathcal{N}^{\mathrm{o}}} \tau_{iq}}{\sum_{i\in\mathcal{N}} \tau_{iq}}. \tag{2}$$

2. *The optimal* $\nu$ *in* (1) *when all other parameters are held fixed verify*

$$\hat{\nu}_{ij} = \mathrm{logistic}\left(\sum_{(q,\ell)\in\mathcal{Q}^2} \tau_{iq}\tau_{j\ell}\log\left(\frac{\pi_{q\ell}}{1-\pi_{q\ell}}\right)\right).$$

*Moreover* $\lambda_{iq} = \rho_q^{\mathbb{1}_{\{i\in\mathcal{N}^{\mathrm{o}}\}}}(1-\rho_q)^{\mathbb{1}_{\{i\in\mathcal{N}^{\mathrm{m}}\}}}$ *for optimization of* $\tau$ *in Proposition 9.2).*

**Star degree sampling.** From Expression of the likelihood for star degree sampling, one has

$$\mathbb{E}_{\tilde{p}} \log p_\psi(R|Y) = -\sum_{i\in\mathcal{N}^{\mathrm{m}}}\left(a + b\tilde{D}_i\right) + \sum_{i\in\mathcal{N}} \mathbb{E}_{\tilde{p}}\left[-\log(1+e^{-(a+bD_i)})\right],$$

where $\tilde{D}_i = \mathbb{E}_{\tilde{p}}[D_i] = \sum_{i\in\mathcal{N}^{\mathrm{m}}} \nu_{ij} + \sum_{i\in\mathcal{N}^{\mathrm{o}}} Y_{ij}$ is the approximation of the degrees.

Because $\mathbb{E}_{\tilde{p}}\left[-\log(1+e^{-(a+bD_i)})\right]$ has no explicit form, we rely on an additional variational approximation. The principle is as follows: since $g(x) = -\log(1+e^{-x})$ is a convex function, we have from Taylor expansion

$$g(x) \geq g(\zeta) + \frac{x-\zeta}{2} + h(\zeta)(x^2-\zeta^2),\ \forall (x,\zeta)\in\mathbb{R}\times\mathbb{R}^+,$$

where $h(x) = \frac{-1}{2\zeta}\left[\text{logistic}(\zeta) - \frac{1}{2}\right]$ . This leads to a lower bound of the initial lower bound:

$$\log p_{\theta,\psi}(Y^{\mathrm{o}}, R) \geq J_{\tau,\nu,\theta,\psi}(Y^{\mathrm{o}}, R) \geq J_{\tau,\nu,\zeta,\theta,\psi}(Y^{\mathrm{o}}, R), \tag{3}$$

with $\zeta = (\zeta_i, i \in \mathcal{N})$ such that $\zeta_i > 0$ is an additional set of variational parameters used to approximate $-\log(1 + e^{-x})$.

By gathering all this, we have

$$\begin{aligned} J_\zeta &= C_{miss} + \sum_{i=1}^{n} \mathbb{E}_{\tilde{p}}\left[\text{logistic}(\zeta_i) + \frac{(a + bD_i) - \zeta_i}{2} + h(\zeta_i)((a + bD_i)^2 - \zeta_i^2)\right], \\ &= C_{miss} + \sum_{i=1}^{n}\left[\text{logistic}(\zeta_i) + \frac{(a + b\tilde{D}_i) - \zeta_i}{2} + h(\zeta_i)(a^2 + 2ab\tilde{D}_i + b^2\hat{D}_i - \zeta_i^2)\right], \end{aligned}$$

where $C_{miss} = \sum_{i\in\mathcal{N}^{\mathrm{m}}}\left(a + b\tilde{D}_i\right)$ and $\hat{D}_i = \mathbb{E}_{\tilde{p}}\left[{D_i}^2\right] = \mathbb{V}_{\tilde{p}}\left(D_i\right) + \mathbb{E}_{\tilde{p}}\left[D_i\right]^2 = \sum_{j\in\mathcal{N}^{\mathrm{m}}} \nu_{ij}(1 - \nu_{ij}) + \left(\sum_{i\in\mathcal{N}^{\mathrm{m}}} \nu_{ij} + \sum_{i\in\mathcal{N}^{\mathrm{o}}} Y_{ij}\right)^2$.

**Proposition 12.** *Consider the maximization of the lower bound* (3) *in the star degree sampling. Let us denote* $\hat{D}_i = \mathbb{E}_{\tilde{p}}\left[{D_i}^2\right]$ *and* $\tilde{D}_k^{-\ell} = \tilde{D}_k - \nu_{k\ell}$.

1. *The parameters* $\psi = (a, b)$ *maximizing* $J_{\tau,\nu,\zeta,\theta,\psi}(Y^{\mathrm{o}}, R)$ *when all other parameters are held fixed are*

$$\begin{aligned} \hat{b} &= \frac{2\left(\frac{n}{2} - \text{card}\left(\mathcal{N}^{\mathrm{m}}\right)\right)\sum_{i=1}^{n}(h(\zeta_i)\tilde{D}_i) - \left(\frac{1}{2}\sum_{i=1}^{n}\tilde{D}_i - \sum_{i\in\mathcal{N}^{\mathrm{m}}}\tilde{D}_i\right) \times \sum_{i=1}^{n} h(\zeta_i)}{2\sum_{i=1}^{n}(h(\zeta_i)\hat{D}_i) \times \sum_{i=1}^{n} h(\zeta_i) - \left(2\sum_{i=1}^{n} h(\zeta_i)\tilde{D}_i\right)^2}, \\ \hat{a} &= -\frac{\hat{b}\sum_{i=1}^{n}\left(h(\zeta_i)\tilde{D}_i\right) + \frac{n}{2} - \text{card}\left(\mathcal{N}^{\mathrm{m}}\right)}{\sum_{i=1}^{n} h(\zeta_i)}. \end{aligned} \tag{4}$$

2. *The parameters* $\zeta$ *maximizing* $J_{\tau,\nu,\zeta,\theta,\psi}(Y^{\mathrm{o}}, R)$ *when all other parameters are held fixed are*

$$\hat{\zeta}_i = \sqrt{a^2 + b^2\hat{D}_i + 2ab\tilde{D}_i}, \ \forall i \in \mathcal{N}.$$

3. *The optimal* $\nu$ *in* $J_{\tau,\nu,\zeta,\theta,\psi}(Y^{\mathrm{o}}, R)$ *when all other parameters are held fixed verify*

$$\begin{aligned} \hat{\nu}_{ij} = \text{logistic}\Bigg( &\sum_{(q,\ell)\in\mathcal{Q}^2} \tau_{iq}\tau_{j\ell} \log\left(\frac{\pi_{q\ell}}{1 - \pi_{q\ell}}\right) - b \\ &+ 2h(\zeta_i)\left(ab + b^2(1 + \tilde{D}_i^{-j})\right) + 2h(\zeta_j)\left(ab + b^2(1 + \tilde{D}_j^{-i})\right)\Bigg). \end{aligned} \tag{5}$$

*Moreover,* $\lambda_{iq} = 1\ \forall(i, q) \in \mathcal{N} \times \mathcal{Q}$ *for optimization of* $\tau$ *in Proposition 9.2).*

**Proposition 13.** *For a model with $Q$ blocks, a sampling design with a vector of parameters $\psi$ with dimension $K$ and $(\hat{\theta}, \hat{\psi}) = \arg\max_{(\theta,\psi)} \log p_{\theta,\psi}(Y^{\mathrm{o}}, Y^{\mathrm{m}}, R, Z)$, the ICL criterion is*

$$\mathrm{ICL} = -2\mathbb{E}_{\tilde{p}_{\tau,\nu};\hat{\theta},\hat{\psi}} \left[\log p_{\hat{\theta},\hat{\psi}}(Y^{\mathrm{o}}, Y^{\mathrm{m}}, R, Z|Q, K)\right] + \mathrm{pen}_{ICL},$$

*where*

$$\mathrm{pen}_{ICL} = \begin{cases} \left(K + \frac{Q(Q+1)}{2}\right) \log\left(\frac{n(n-1)}{2}\right) + (Q-1)\log(n) & \text{if the sampling design is dyad-centered,} \\ \frac{Q(Q+1)}{2} \log\left(\frac{n(n-1)}{2}\right) + (K+Q-1)\log(n) & \text{otherwise.} \end{cases}$$

$$\mathrm{pen}_{ICL} = \begin{cases} \left(K + Q^2\right) \log\left(n(n-1)\right) + (Q-1)\log(n) & \text{if the sampling design is dyad-centered,} \\ Q^2 \log\left(n(n-1)\right) + (K+Q-1)\log(n) & \text{otherwise.} \end{cases}$$

## 2 Poisson SBM

### 2.1 Undirected/Directed

#### 2.1.1 MAR

In the directed case, only $\mathcal{D}^{\mathrm{o}}$ is changing in the following (i.e don't forget that the matrix is not symmetric: $Y_{ij}$ and $Y_{ji}$ are two different random variables).

**Proposition 14.** *The complete log-likelihood restricted to the observed variables is*

$$\log p_\theta(Y^{\mathrm{o}}, Z) = \sum_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \sum_{(q,\ell)\in\mathcal{Q}^2} Z_{iq} Z_{j\ell} \log p(Y_{ij}, \lambda_{q\ell}) + \sum_{i\in\mathcal{N}^{\mathrm{o}}} \sum_{q\in\mathcal{Q}} Z_{iq} \log(\alpha_q),$$

*with $p(x, \lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$ the Poisson probability density function.*

**Proposition 15.** *The variationnal approximation is*

$$J_{\tau,\theta}(Y^{\mathrm{o}}) = \sum_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \sum_{(q,\ell)\in\mathcal{Q}^2} \tau_{iq}\tau_{j\ell} \log p(Y_{ij}, \lambda_{q\ell}) + \sum_{i\in\mathcal{N}^{\mathrm{o}}} \sum_{q\in\mathcal{Q}} \tau_{iq} \log(\alpha_q/\tau_{iq}).$$

**Proposition 16.** *Consider the lower bound $J_{\tau,\theta}(Y^{\mathrm{o}})$.*

1. *The parameters $\theta = (\alpha, \lambda)$ maximizing $J_\theta(Y^{\mathrm{o}})$ when $\tau$ is held fixed are*

$$\hat{\alpha}_q = \frac{\sum_{i\in\mathcal{N}^{\mathrm{o}}} \hat{\tau}_{iq}}{\mathrm{card}\,(\mathcal{N}^{\mathrm{o}})}, \qquad \hat{\lambda}_{q\ell} = \frac{\sum_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \hat{\tau}_{iq}\hat{\tau}_{j\ell} Y_{ij}}{\sum_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \hat{\tau}_{iq}\hat{\tau}_{j\ell}}.$$

2. *The variational parameters $\tau$ maximizing $J_\tau(Y^{\mathrm{o}})$ when $\theta$ is held fixed are obtained thanks to the following fixed point relation:*

$$\hat{\tau}_{iq} \propto \alpha_q \left( \prod_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \prod_{\ell\in\mathcal{Q}} p(Y_{ij}; \lambda_{q\ell})^{\hat{\tau}_{j\ell}} \right).$$

$$\hat{\tau}_{iq} \propto \alpha_q \left( \prod_{(i,j)\in\mathcal{D}^{\mathrm{o}}} \prod_{\ell\in\mathcal{Q}} \left(p(Y_{ij}; \lambda_{q\ell}) p(Y_{ji}; \lambda_{\ell q})\right)^{\hat{\tau}_{j\ell}} \right).$$

**Proposition 17.** *For an SBM with $Q$ blocks and for $\hat{\theta} = \arg\ \max \log p_\theta(Y^{\text{o}}, Z)$, the ICL criterion is given by*

$$\mathrm{ICL}(Q) = -2\mathbb{E}_{\tilde{p}_\tau}\left[\log p_{\hat{\theta}}(Y^{\text{o}}, Z; Q)\right] + \frac{Q(Q+1)}{2} \log \operatorname{card}\left(\mathcal{D}^{\text{o}}\right) + (Q-1) \log \operatorname{card}\left(\mathcal{N}^{\text{o}}\right).$$

$$\mathrm{ICL}(Q) = -2\mathbb{E}_{\tilde{p}_\tau}\left[\log p_{\hat{\theta}}(Y^{\text{o}}, Z; Q)\right] + Q^2 \log \operatorname{card}\left(\mathcal{D}^{\text{o}}\right) + (Q-1) \log \operatorname{card}\left(\mathcal{N}^{\text{o}}\right).$$