Supplementary files of "Goodness of fit of logistic models for random graphs" : Appendix

Pierre Latouche
Laboratoire SAMM, EA 4543, Université Paris 1 Panthéon-Sorbonne, France
and
Stéphane Robin
AgroParisTech, UMR 518, MIA, Paris, France
INRA, UMR 518, MIA, Paris, France
and
Sarah Ouadah
AgroParisTech, UMR 518, MIA, Paris, France
INRA, UMR 518, MIA, Paris, France

November 23, 2015

A Appendix

A.1 Proof of Proposition 1

Let us start by showing that:

$$\log p(Y|Z,\alpha,\beta) \geq \log \sqrt{h(Z,\alpha,\beta,\xi)},$$

where ξ is an $n \times n$ positive real matrix. We use the bound on the log-logistic function introduced by [1] from Taylor expansions:

$$\log g(x) \geq \log g(\xi) + \frac{x-\xi}{2} - \lambda(\xi)(x^2-\xi^2), \forall (x,\xi) \in \mathbb{R} \times \mathbb{R}^+, \tag{1}$$

where $\lambda(\xi) = (g(\xi) - 1/2)/(2\xi)$. Note that (1) is an even function and therefore we can consider only positive values of x without loss of generality. Since

$$\log p(Y_{ij}|Z_i,Z_j,\alpha,\beta) = Y_{ij}(Z_i^\intercal \alpha Z_j + x_{ij}^\intercal \beta) + \log g(-Z_i^\intercal \alpha Z_j - x_{ij}^\intercal \beta),$$

then

$$\begin{aligned}
\log p(Y_{ij}|Z_i,Z_j,\alpha,\beta) \geq\ & Y_{ij}(Z_i^\intercal \alpha Z_j + x_{ij}^\intercal \beta) + \log g(\xi_{ij}) - \frac{Z_i^\intercal \alpha Z_j + x_{ij}^\intercal \beta + \xi_{ij}}{2} \\
& - \lambda(\xi_{ij})((Z_i^\intercal \alpha Z_j + x_{ij}^\intercal \beta)^2 - \xi_{ij}^2) \\
=\ & (Y_{ij} - \frac{1}{2})(Z_i^\intercal \alpha Z_j + x_{ij}^\intercal \beta) - \frac{\xi_{ij}}{2} + \log g(\xi_{ij}) \\
& - \lambda(\xi_{ij})((Z_i^\intercal \alpha Z_j + x_{ij}^\intercal \beta)^2 - \xi_{ij}^2).
\end{aligned}$$

Note that for undirected networks, the matrix ξ has to be symmetric, *i.e.* $\xi_{ij} = \xi_{ji}, \forall i \neq j$. We then have

$$\log p(Y|Z,\alpha,\beta) = \frac{1}{2}\sum_{i \neq j}^{n} \log p(Y_{ij}|Z_i,Z_j,\alpha,\beta).$$

Therefore

$$\log p(Y|Z,\alpha,\beta) \geq \log \sqrt{h(Z,\alpha,\beta,\xi)}.$$

Finally,

$$
\begin{aligned}
\mathcal{L}_K(q) &= \sum_Z \int q(Z,\pi,\alpha,\beta,\gamma,\eta) \log \frac{p(Y,Z,\pi,\alpha,\beta,\gamma,\eta)}{q(Z,\pi,\alpha,\beta,\gamma,\eta)} \mathrm{d}\pi \mathrm{d}\alpha \mathrm{d}\beta \mathrm{d}\gamma \mathrm{d}\eta \\
&= \sum_Z \int q(Z,\pi,\alpha,\beta,\gamma,\eta) \log \frac{p(Y|Z,\alpha,\beta) p(Z,\pi,\alpha,\beta,\gamma,\eta)}{q(Z,\pi,\alpha,\beta,\gamma,\eta)} \mathrm{d}\pi \mathrm{d}\alpha \mathrm{d}\beta \mathrm{d}\gamma \mathrm{d}\eta \\
&\geq \sum_Z \int q(Z,\pi,\alpha,\beta,\gamma,\eta) \log \frac{\sqrt{h(Z,\alpha,\beta,\xi)} p(Z,\pi,\alpha,\beta,\gamma,\eta)}{q(Z,\pi,\alpha,\beta,\gamma,\eta)} \mathrm{d}\pi \mathrm{d}\alpha \mathrm{d}\beta \mathrm{d}\gamma \mathrm{d}\eta \\
&= \mathcal{L}_K(q;\ \xi).
\end{aligned}
$$

A.2 Proof of Proposition 2

$$
\begin{aligned}
\log q(Z_i) &= \mathbb{E}_{Z^{\setminus i},\alpha,\beta,\pi} \left[\frac{1}{2} \log h(Z,\alpha,\beta,\xi) + \log p(Z|\pi) \right] + \mathrm{cst} \\
&= \mathbb{E}_{Z^{\setminus i},\alpha,\beta,\pi} \left[\frac{1}{2} \sum_{i \neq j}^{n} \left\{ (Y_{ij} - \frac{1}{2}) Z_i^\intercal \alpha Z_j - \lambda(\xi_{ij}) \left((Z_i^\intercal \alpha Z_j)^2 + 2 Z_i^\intercal \alpha Z_i x_{ij}^\intercal \beta \right) \right\} \right. \\
&\qquad \left. + \sum_{i=1}^{n} Z_{ik} \log \pi_k \right] + \mathrm{cst}.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\mathbb{E}_{Z_j,\alpha}[Z_i^\intercal \alpha Z_j] &= \mathbb{E}_{Z_j,\alpha} \left[\sum_{k,l}^{K} Z_{ik} \alpha_{kl} Z_{jl} \right] \\
&= \sum_{k=1}^{K} Z_{ik} \left\{ \sum_{l=1}^{K} (m_\alpha)_{kl} \tau_{jl} \right\}.
\end{aligned}
$$

Furthermore

$$
\begin{aligned}
\mathbb{E}_{Z_j,\alpha}[(Z_i^\intercal \alpha Z_j)^2] &= \mathbb{E}_{Z_j,\alpha} \left[\left(\sum_{k,l}^{K} Z_{ik} \alpha_{kl} Z_{kl} \right)^2 \right] \\
&= \mathbb{E}_{Z_j,\alpha} \left[\sum_{k,k',l,l'}^{K} Z_{ik} Z_{ik'} \alpha_{kl} \alpha_{k'l'} Z_{jl} Z_{jl'} \right].
\end{aligned} \tag{2}
$$

Because all vectors Z_i are sampled from a multinomial distribution with parameters $(1,\pi)$, all terms $Z_{ik} Z_{ik'} = 0$ for $k \neq k'$ and $Z_{ik}^2 = Z_{ik}$ in (2). Therefore

$$
(Z_i^\intercal \alpha Z_j)^2 = \sum_{k,l}^{K} Z_{ik} \alpha_{kl}^2 Z_{jl}. \tag{3}
$$

This leads to

$$
\mathbb{E}_{Z_j,\alpha}[(Z_i^\intercal \alpha Z_j)^2] = \sum_{k=1}^{K} Z_{ik} \left\{ \sum_{l=1}^{K} \mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] \tau_{jl} \right\}.
$$

Finally

$$\log q(Z_i) = \sum_{k=1}^{K} Z_{ik} \Big\{ \sum_{l=1}^{K} (m_\alpha)_{kl} \frac{1}{2} \sum_{j\neq i}^{n} \Big((Y_{ij} - \frac{1}{2}) - 2\lambda(\xi_{ij}) x_{ij}^\intercal m_\beta \Big) \tau_{jl} - \sum_{l=1}^{K} \mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] \frac{1}{2} \sum_{j\neq i}^{n} \lambda(\xi_{ij}) \tau_{jl}$$

$$+ \sum_{l=1}^{K} (m_\alpha)_{lk} \frac{1}{2} \sum_{j\neq i}^{n} \Big((Y_{ji} - \frac{1}{2}) - 2\lambda(\xi_{ji}) x_{ji}^\intercal m_\beta \Big) \tau_{jl} - \sum_{l=1}^{K} \mathbb{E}_{\alpha_{lk}}[\alpha_{lk}^2] \frac{1}{2} \sum_{j\neq i}^{n} \lambda(\xi_{ji}) \tau_{jl}$$

$$+ \psi(e_k^n) - \psi\Big(\sum_{l=1}^{K} e_l^n \Big) \Big\} + \text{cst}.$$

Since $(m_\alpha)_{kl} = (m_\alpha)_{lk}$, $\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] = \mathbb{E}_{\alpha_{lk}}[\alpha_{lk}^2]$, $Y_{ij} = Y_{ji}$, $x_{ij} = x_{ji}$, $\xi_{ij} = \xi_{ji}$, then

$$\log q(Z_i) = \sum_{k=1}^{K} Z_{ik} \Big\{ \sum_{l=1}^{K} (m_\alpha)_{kl} \sum_{j\neq i}^{n} \Big((Y_{ij} - \frac{1}{2}) - 2\lambda(\xi_{ij}) x_{ij}^\intercal m_\beta \Big) \tau_{jl} - \sum_{l=1}^{K} \mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] \sum_{j\neq i}^{n} \lambda(\xi_{ij}) \tau_{jl}$$

$$+ \psi(e_k^n) - \psi\Big(\sum_{l=1}^{K} e_l^n \Big) \Big\} + \text{cst}.$$

Therefore

$$q(Z_i) = \mathcal{M}(Z_i;\ 1, \tau_i),$$

where $\sum_{k=1}^{K} \tau_{ik} = 1$ and

$$\tau_{ik} \propto \exp \Big\{ \sum_{l=1}^{K} (m_\alpha)_{kl} \sum_{j\neq i}^{n} \Big((Y_{ij} - \frac{1}{2}) - 2\lambda(\xi_{ij}) x_{ij}^\intercal m_\beta \Big) \tau_{jl} - \sum_{l=1}^{K} \mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] \sum_{j\neq i}^{n} \lambda(\xi_{ij}) \tau_{jl}$$

$$+ \psi(e_k^n) - \psi\Big(\sum_{l=1}^{K} e_l^n \Big) \Big\}.$$

A.3 Proof of Proposition 3

$$\begin{aligned}
\log q(\pi) &= \mathbb{E}_Z \left[\log p(Z|\pi) + \log p(\pi) \right] + \text{cst} \\
&= \sum_{i=1}^{n} \sum_{k=1}^{K} \tau_{ik} \log \pi_k + \sum_{k=1}^{K} (e_0 - 1) \log \pi_k + \text{cst} \\
&= \sum_{k=1}^{K} \Big(e_0 + \sum_{i=1}^{n} \tau_{ik} - 1 \Big) \log \pi_k + \text{cst}.
\end{aligned}$$

Therefore

$$q(\pi) = \text{Dir}(\pi;\ e^n),$$

where $e_k^n = e_0 + \sum_{i=1}^{n} \tau_{ik}, \forall k \in \{1, \ldots, K\}$.

A.4 Proof of Proposition 4

$$\begin{aligned}
\log q(\beta) &= \mathbb{E}_{Z,\alpha,\eta}\left[\frac{1}{2}\log h(Z,\alpha,\beta,\xi) + \log p(\beta|\eta)\right] + \text{cst} \\
&= \mathbb{E}_{Z,\alpha,\eta}\left[\frac{1}{2}\sum_{i\neq j}^{n}\left\{(Y_{ij}-1/2)x_{ij}^\intercal\beta - \lambda(\xi_{ij})\Big((x_{ij}^\intercal\beta)^2 + 2Z_i^\intercal\alpha Z_j x_{ij}^\intercal\beta\Big)\right\} - \frac{\eta}{2}\beta^\intercal\beta\right] + \text{cst} \\
&= \beta^\intercal\left\{\frac{1}{2}\sum_{i\neq j}^{n}\Big(Y_{ij} - \frac{1}{2} - 2\lambda(\xi_{ij})\tau_i^\intercal m_\alpha\tau_j\Big)x_{ij}\right\} - \frac{1}{2}\beta^\intercal\left\{\frac{c_n}{d_n}I_d + \sum_{i\neq j}^{n}\lambda(\xi_{ij})x_{ij}x_{ij}^\intercal\right\}\beta + \text{cst}.
\end{aligned}$$

Therefore

$$q(\beta) = \mathcal{N}(\beta;\ m_\beta, S_\beta),$$

where

$$S_\beta^{-1} = \frac{c_n}{d_n}I_d + \sum_{i\neq j}^{n}\lambda(\xi_{ij})x_{ij}x_{ij}^\intercal,$$

and

$$m_\beta = S_\beta\frac{1}{2}\sum_{i\neq j}^{n}\left(Y_{ij} - \frac{1}{2} - 2\lambda(\xi_{ij})\tau_i^\intercal m_\alpha\tau_j\right)x_{ij}.$$

A.5 Proof of Proposition 5

$$\begin{aligned}
\log q(\gamma) &= \mathbb{E}_\alpha\left[\log p(\alpha|\gamma) + \log p(\gamma)\right] + \text{cst} \\
&= \mathbb{E}_\alpha\left[\sum_{k\leq l}^{K}\frac{1}{2}\log(\gamma) - \sum_{k\leq l}^{K}\frac{\gamma}{2}\alpha_{kl}^2\right] + (a_0-1)\log\gamma - b_0\gamma + \text{cst} \\
&= (a_0 + \frac{K(K+1)}{4} - 1)\log\gamma - \left(b_0 + \frac{1}{2}\sum_{k\leq l}^{K}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2]\right)\gamma + \text{cst}.
\end{aligned}$$

Therefore

$$q(\gamma) = \text{Gam}(\gamma;\ a_n, b_n),$$

where $a_n = a_0 + \frac{K(K+1)}{4}$ and $b_n = b_0 + \frac{1}{2}\sum_{k\leq l}^{K}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2]$.

A.6 Proof of Proposition 6

$$\begin{aligned}
\log q(\eta) &= \mathbb{E}_\beta\left[\log p(\beta|\eta) + \log p(\eta)\right] + \text{cst} \\
&= \mathbb{E}_\beta\left[\frac{d}{2}\log\eta - \frac{\eta}{2}\beta^\intercal\beta\right] + (c_0-1)\log\eta - d_0\eta + \text{cst} \\
&= (c_0 + \frac{d}{2} - 1)\log\eta - (d_0 + \frac{1}{2}\text{Tr}(S_\beta) + \frac{1}{2}m_\beta^\intercal m_\beta)\eta + \text{cst}.
\end{aligned}$$

Therefore

$$q(\eta) = \text{Gam}(\eta;\ c_n, d_n),$$

where $c_n = c_0 + \frac{d}{2}$ and $d_n = d_0 + \frac{1}{2}\text{Tr}(S_\beta) + \frac{1}{2}m_\beta^\intercal m_\beta$.

A.7 Proof of Proposition 7

$$\begin{aligned}\log q(\alpha) &= \mathbb{E}_{Z,\beta,\gamma}\left[\frac{1}{2}\log h(Z,\alpha,\beta,\xi) + \log p(\alpha|\gamma)\right] + \text{cst} \\ &= \mathbb{E}_{Z,\beta,\gamma}\left[\frac{1}{2}\sum_{i\neq j}^{n}\left\{(Y_{ij}-\frac{1}{2})Z_i^\intercal\alpha Z_j - \lambda(\xi_{ij})\Big((Z_i^\intercal\alpha Z_j)^2 + 2Z_i^\intercal\alpha Z_j x_{ij}^\intercal\beta\Big)\right\}\right. \\ &\qquad \left. - \sum_{k\leq l}^{K}\frac{\gamma}{2}\alpha_{kl}^2\right] + \text{cst}.\end{aligned} \tag{4}$$

We have $Z_i^\intercal\alpha Z_j = \sum_{k,l}^{K} Z_{ik}\alpha_{kl}Z_{kl}$ and $(Z_i^\intercal\alpha Z_j)^2 = \sum_{k,l}^{K} Z_{ik}\alpha_{kl}^2 Z_{kl} = Z_i^\intercal A Z_j$ (see Eq. 3) with A the $K\times K$ matrix such that $A_{kl} = \alpha_{kl}^2$. Moreover, any expression of the form $(1/2)\sum_{i\neq j}^{n} c_{ij}Z_i^\intercal B Z_j$ where B is a symmetric $K\times K$ matrix and $c_{ij} = c_{ji}$ can be written

$$\begin{aligned}\frac{1}{2}\sum_{i\neq j}^{n} c_{ij}Z_i^\intercal B Z_j &= \frac{1}{2}\sum_{i\neq j}^{n} c_{ij}\sum_{k,l}^{K} Z_{ik}B_{kl}Z_{jl} \\ &= \frac{1}{2}\sum_{i\neq j}^{n} c_{ij}\left(\sum_{k=1}^{K} Z_{ik}B_{kk}Z_{jk} + \sum_{k<l}^{K} Z_{ik}B_{kl}Z_{jl} + \sum_{k<l}^{K} Z_{jk}B_{lk}Z_{il}\right) \\ &= \sum_{k=1}^{K} B_{kk}\frac{1}{2}\sum_{i\neq j}^{n} c_{ij}Z_{ik}Z_{jk} + \sum_{k<l}^{K} B_{kl}\left(\frac{1}{2}\sum_{i\neq j}^{n} c_{ij}Z_{ik}Z_{jl} + \frac{1}{2}\sum_{i\neq j}^{n} c_{ij}Z_{jk}Z_{il}\right).\end{aligned}$$

By exchanging the role of i and j in the sum of the last term and since $c_{ij} = c_{ji}$, we obtain

$$\frac{1}{2}\sum_{i\neq j}^{n} c_{ij}Z_i^\intercal B Z_j = \sum_{k=1}^{K} B_{kk}\frac{1}{2}\sum_{i\neq j}^{n} c_{ij}Z_{ik}Z_{jk} + \sum_{k<l}^{K} B_{kl}\sum_{i\neq j}^{n} c_{ij}Z_{ik}Z_{jl}. \tag{5}$$

Using (5) in (4) leads to

$$\log q(\alpha) = \sum_{k=1}^{K} \alpha_{kk} \sum_{i\neq j}^{n} \left(\frac{1}{2}(Y_{ij} - \frac{1}{2}) - \lambda(\xi_{ij}) x_{ij}^{\intercal} m_\beta \right) \tau_{ik}\tau_{jk}$$
$$+ \sum_{k<l}^{K} \alpha_{kl} \sum_{i\neq j}^{n} \left((Y_{ij} - \frac{1}{2}) - 2\lambda(\xi_{ij}) x_{ij}^{\intercal} m_\beta \right) \tau_{ik}\tau_{jl}$$
$$- \sum_{k=1}^{K} \alpha_{kk}^2 \sum_{i\neq j}^{n} \frac{1}{2}\lambda(\xi_{ij})\tau_{ik}\tau_{jk} - \sum_{k<l}^{K} \alpha_{kl}^2 \sum_{i\neq j}^{n} \lambda(\xi_{ij})\tau_{ik}\tau_{jl}$$
$$- \sum_{k\leq l}^{K} \frac{a_n}{2b_n}\alpha_{kl}^2.$$

Therefore

$$q(\alpha) = \prod_{k\neq l}^{K} \mathcal{N}\left(\alpha_{kl}; (m_\alpha)_{kl}, (\sigma_\alpha^2)_{kl}\right),$$

where

$$(\sigma_\alpha^2)_{kk}^{-1} = \frac{a_n}{b_n} + \sum_{i\neq j}^{n} \lambda(\xi_{ij})\tau_{ik}\tau_{jk}, \forall k,$$
$$(\sigma_\alpha^2)_{kl}^{-1} = \frac{a_n}{b_n} + 2\sum_{i\neq j}^{n} \lambda(\xi_{ij})\tau_{ik}\tau_{jl}, \forall k\neq l,$$
$$(m_\alpha)_{kk} = (\sigma_\alpha^2)_{kk} \sum_{i\neq j}^{n} \left(\frac{1}{2}(Y_{ij} - \frac{1}{2}) - \lambda(\xi_{ij}) x_{ij}^{\intercal} m_\beta \right) \tau_{ik}\tau_{jk}$$
$$(m_\alpha)_{kl} = (\sigma_\alpha^2)_{kl} \sum_{i\neq j}^{n} \left((Y_{ij} - \frac{1}{2}) - 2\lambda(\xi_{ij}) x_{ij}^{\intercal} m_\beta \right) \tau_{ik}\tau_{jl}.$$

A.8 Proof of Proposition 8

Keeping only the terms that do depend on ξ_{ij} in (7), the lower bound is given by

$$\mathcal{L}_K(q;\, \xi) = \frac{1}{2}\sum_{i\neq j}^{n} \Bigg\{ \log g(\xi_{ij}) - \frac{\xi_{ij}}{2} - \lambda(\xi_{ij}) \Big(\mathbb{E}_{Z_i,Z_j,\alpha}[(Z_i^{\intercal}\alpha Z_j)^2] + 2\mathbb{E}_{Z_i,Z_j,\alpha}[Z_i^{\intercal}\alpha Z_j] x_{ij}^{\intercal}\mathbb{E}_\beta[\beta]$$
$$+ \mathbb{E}_\beta[(x_{ij}^{\intercal}\beta)^2] - \xi_{ij}^2 \Big) \Bigg\} + \text{cst}$$
$$= \frac{1}{2}\sum_{i\neq j}^{n} \Bigg\{ \log g(\xi_{ij}) - \frac{\xi_{ij}}{2} - \lambda(\xi_{ij}) \Big(\sum_{k,l}^{K} \tau_{ik}\tau_{jl}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] + 2\sum_{k,l}^{K} \tau_{ik}\tau_{jl}(m_\alpha)_{kl} x_{ij}^{\intercal} m_\beta$$
$$+ \text{Tr}(x_{ij}x_{ij}^{\intercal}(S_\beta + m_\beta m_\beta^{\intercal})) - \xi_{ij}^2 \Big) \Bigg\} + \text{cst}.$$

The partial derivative of the lower bound with respect to ξ_{ij} is

$$\frac{\partial \mathcal{L}_K}{\partial \xi_{ij}}(q;\ \xi) = g(-\xi_{ij}) - \frac{1}{2} - \lambda^{'}(\xi_{ij})\Big(\sum_{k,l}^{K} \tau_{ik}\tau_{jl}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] + 2\sum_{k,l}^{K} \tau_{ik}\tau_{jl}(m_\alpha)_{kl} x_{ij}^{\intercal} m_\beta \\ + \mathrm{Tr}(x_{ij}x_{ij}^{\intercal}(S_\beta + m_\beta m_\beta^{\intercal})) - \xi_{ij}^2\Big) + 2\lambda(\xi_{ij})\xi_{ij},$$

and

$$\frac{\partial \mathcal{L}_K}{\partial \xi_{ij}}(q;\ \xi) = -\lambda^{'}(\xi_{ij})\Big(\sum_{k,l}^{K} \tau_{ik}\tau_{jl}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] + 2\sum_{k,l}^{K} \tau_{ik}\tau_{jl}(m_\alpha)_{kl} x_{ij}^{\intercal} m_\beta \\ + \mathrm{Tr}(x_{ij}x_{ij}^{\intercal}(S_\beta + m_\beta m_\beta^{\intercal})) - \xi_{ij}^2\Big). \quad (6)$$

Finally, $\lambda(\xi_{ij})$ is a strictly decreasing function for positive values of ξ_{ij} . Thus, $\lambda^{'}(\xi_{ij}) \neq 0$ and therefore if we set (6) to zero, we obtain

$$\xi_{ij}^2 = \sum_{k,l}^{K} \tau_{ik}\tau_{jl}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] + 2\sum_{k,l}^{K} \tau_{ik}\tau_{jl}(m_\alpha)_{kl} x_{ij}^{\intercal} m_\beta + \mathrm{Tr}(x_{ij}x_{ij}^{\intercal}(S_\beta + m_\beta m_\beta^{\intercal})).$$

Note that $\xi_{ij} = \xi_{ji}$.

A.9 Proof of Proposition 9

$$\mathcal{L}_K(q;\xi) = \mathbb{E}_{Z,\pi,\alpha,\beta,\gamma,\eta}\Big[\log\sqrt{h(Z,\alpha,\beta,\xi)} + \log p(Z,\pi,\alpha,\beta,\gamma,\eta)\Big] - \mathbb{E}_{Z,\pi,\alpha,\beta,\gamma,\eta}\left[\log q(Z,\pi,\alpha,\beta,\gamma,\eta)\right]$$

$$
\begin{aligned}
\mathcal{L}_K(q;\ \xi) = \frac{1}{2}\sum_{i\neq j}^{n}\Bigg\{ &(Y_{ij} - \frac{1}{2})(\mathbb{E}_{Z_i,Z_j,\alpha}[Z_i^\intercal \alpha Z_j] + x_{ij}^\intercal \mathbb{E}_\beta[\beta]) + \log g(\xi_{ij}) - \frac{\xi_{ij}}{2} \\
&- \lambda(\xi_{ij})\left(\mathbb{E}_{Z_i,Z_j,\alpha}[(Z_i^\intercal \alpha Z_j)^2] + 2\mathbb{E}_{Z_i,Z_j,\alpha}[Z_i^\intercal \alpha Z_j] x_{ij}^\intercal \mathbb{E}_\beta[\beta] + \mathbb{E}_\beta[(x_{ij}^\intercal \beta)^2] - \xi_{ij}^2\right)\Bigg\} \\
&+ \sum_{i=1}^{n}\sum_{k=1}^{K}\mathbb{E}_{Z_i}[Z_{ik}]\mathbb{E}_\pi[\log \pi_k] - \log C(e) + \sum_{k=1}^{K}(e_0 - 1)\mathbb{E}_\pi[\log \pi_k] - \frac{K(K+1)}{4}\log(2\pi) \\
&+ \frac{K(K+1)}{4}\mathbb{E}_\gamma[\log \gamma] - \frac{\mathbb{E}_\gamma[\gamma]}{2}\sum_{k\leq l}^{K}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] - \frac{d}{2}\log(2\pi) + \frac{d}{2}\mathbb{E}_\eta[\log \eta] - \frac{\mathbb{E}_\eta[\eta]}{2}\mathbb{E}_\beta[\beta^\intercal \beta] \\
&- \log \Gamma(a_0) + a_0 \log b_0 + (a_0 - 1)\mathbb{E}_\gamma[\log \gamma] - b_0 \mathbb{E}_\gamma[\gamma] - \log \Gamma(c_0) \\
&+ c_0 \log d_0 + (c_0 - 1)\mathbb{E}_\eta[\log \eta] - d_0 \mathbb{E}_\eta[\eta] - \sum_{i=1}^{n}\sum_{k=1}^{K}\mathbb{E}_{Z_i}[Z_{ik}]\log \tau_{ik} + \log C(e^n) \\
&- \sum_{k=1}^{K}(e_k^n - 1)\mathbb{E}_\pi[\log \pi_k] + \frac{K(K+1)}{4}\log(2\pi) + \frac{1}{2}\sum_{k\leq l}^{K}\log(\sigma_\alpha^2)_{kl} + \frac{1}{2}\sum_{k\leq l}^{K}(\sigma_\alpha^2)_{kl}^{-1}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] \\
&- \sum_{k\leq l}^{K}(\sigma_\alpha^2)_{kl}^{-1}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}](m_\alpha)_{kl} + \frac{1}{2}\sum_{k\leq l}^{K}(\sigma_\alpha^2)_{kl}^{-1}(m_\alpha)_{kl}^2 + \frac{d}{2}\log(2\pi) + \frac{1}{2}\log|S_\beta| \\
&+ \frac{1}{2}\mathbb{E}_\beta[\beta^\intercal S_\beta^{-1}\beta] - \mathbb{E}_\beta[\beta]^\intercal S_\beta^{-1} m_\beta + \frac{1}{2} m_\beta S_\beta^{-1} m_\beta + \log \Gamma(a_n) - a_n \log b_n \\
&- (a_n - 1)\mathbb{E}_\gamma[\log \gamma] + b_n \mathbb{E}_\gamma[\gamma] + \log \Gamma(c_n) - c_n \log d_n - (c_n - 1)\mathbb{E}_\eta[\log \eta] \\
&+ d_n \mathbb{E}_\eta[\eta], \quad (7)
\end{aligned}
$$

where $C(x) = \frac{\prod_{k=1}^K \Gamma(x_k)}{\Gamma(\sum_{k=1}^K x_k)}$ and $\Gamma(\cdot)$ is the gamma function. The terms in $\mathbb{E}_\gamma[\log\gamma]$, $\mathbb{E}_\eta[\log\eta]$, $\mathbb{E}_\pi[\log\pi]$ and $\log(2\pi)$ do simplify in (7) after the VBEM update step. This leads to

$$
\begin{aligned}
\mathcal{L}_K(q;\,\xi) = \frac{1}{2}\sum_{i\neq j}^{n}\Bigg\{ & (Y_{ij} - \frac{1}{2})(\mathbb{E}_{Z_i,Z_j,\alpha}[Z_i^\intercal \alpha Z_j] + x_{ij}^\intercal \mathbb{E}_\beta[\beta]) + \log g(\xi_{ij}) - \frac{\xi_{ij}}{2} \\
& - \lambda(\xi_{ij})\left(\mathbb{E}_{Z_i,Z_j,\alpha}[(Z_i^\intercal \alpha Z_j)^2] + 2\mathbb{E}_{Z_i,Z_j,\alpha}[Z_i^\intercal \alpha Z_j] x_{ij}^\intercal \mathbb{E}_\beta[\beta] + \mathbb{E}_\beta[(x_{ij}^\intercal \beta)^2] - \xi_{ij}^2\right)\Bigg\} \\
& - \log C(e) - \frac{a_n}{2b_n}\sum_{k\leq l}^{K}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] - \frac{c_n}{2d_n}\mathrm{Tr}(S_\beta + m_\beta m_\beta^\intercal) \\
& - \log\Gamma(a_0) + a_0\log b_0 - b_0\frac{a_n}{b_n} - \log\Gamma(c_0) + c_0\log d_0 - d_0\frac{c_n}{d_n} \\
& - \sum_{i=1}^{n}\sum_{k=1}^{K}\tau_{ik}\log\tau_{ik} + \log C(e^n) + \frac{1}{2}\sum_{k\leq l}^{K}\log(\sigma_\alpha^2)_{kl} + \frac{1}{2}\sum_{k\leq l}^{K}(\sigma_\alpha^2)_{kl}^{-1}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2] \\
& - \frac{1}{2}\sum_{k\leq l}^{K}(\sigma_\alpha^2)_{kl}^{-1}(m_\alpha)_{kl}^2 + \frac{1}{2}\log|S_\beta| + \frac{1}{2}\mathrm{Tr}\big(S_\beta^{-1}(S_\beta + m_\beta(m_\beta)^\intercal)\big) \\
& - \frac{1}{2}m_\beta^\intercal S_\beta^{-1} m_\beta + \log\Gamma(a_n) - a_n\log b_n + b_n\frac{a_n}{b_n} + \log\Gamma(c_n) \\
& - c_n\log d_n + d_n\frac{c_n}{d_n}.
\end{aligned}
$$

Moreover, using (5), note that

$$
\begin{aligned}
\frac{1}{2}\sum_{i\neq j}^{n}(Y_{ij} - \frac{1}{2})\mathbb{E}_{Z_i,Z_j,\alpha}[Z_i^\intercal \alpha Z_j] = & \sum_{k=1}^{K}\mathbb{E}_{\alpha_{kk}}[\alpha_{kk}]\frac{1}{2}\sum_{i\neq j}^{n}(Y_{ij} - \frac{1}{2})\tau_{ik}\tau_{jk} \\
& + \sum_{k<l}^{K}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}]\sum_{i\neq j}^{n}(Y_{ij} - \frac{1}{2})\tau_{ik}\tau_{jl},
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{2}\sum_{i\neq j}^{n}2\lambda(\xi_{ij})\mathbb{E}_{Z_i,Z_j,\alpha}[Z_i^\intercal \alpha Z_j] x_{ij}^\intercal \mathbb{E}_\beta[\beta] = & \sum_{k=1}^{K}\mathbb{E}_{\alpha_{kk}}[\alpha_{kk}]\sum_{i\neq j}^{n}\lambda(\xi_{ij}) x_{ij}^\intercal \mathbb{E}_\beta[\beta]\tau_{ik}\tau_{jk} \\
& + \sum_{k<l}^{K}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}]2\sum_{i\neq j}^{n}\lambda(\xi_{ij}) x_{ij}^\intercal \mathbb{E}_\beta[\beta]\tau_{ik}\tau_{jl}.
\end{aligned}
$$

Using (3) and (5),

$$\frac{1}{2}\sum_{i\neq j}^{n}\lambda(\xi_{ij})\mathbb{E}_{Z_i,Z_j,\alpha}[(Z_i^\intercal \alpha Z_j)^2] = \sum_{k=1}^{K}\mathbb{E}_{\alpha_{kk}}[\alpha_{kk}^2]\frac{1}{2}\sum_{i\neq j}^{n}\lambda(\xi_{ij})\tau_{ik}\tau_{jk} + \sum_{k<l}^{K}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2]\sum_{i\neq j}^{n}\lambda(\xi_{ij})\tau_{ik}\tau_{jl}$$

Finally

$$\begin{aligned}
\mathbb{E}_\beta[(x_{ij}^\intercal\beta)^2] &= \mathbb{E}_\beta\left[x_{ij}^\intercal\beta x_{ij}^\intercal\beta\right] \\
&= \mathbb{E}_\beta\left[x_{ij}^\intercal\beta\beta^\intercal x_{ij}\right] \\
&= \mathrm{Tr}\left(x_{ij}x_{ij}^\intercal\mathbb{E}_\beta[\beta\beta^\intercal]\right) \\
&= \mathrm{Tr}\left(x_{ij}x_{ij}^\intercal(S_\beta + m_\beta m_\beta^\intercal)\right).
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathcal{L}_K(q;\,\xi) = &\frac{1}{2}\sum_{i\neq j}^{n}\left\{\log g(\xi_{ij}) - \frac{\xi_{ij}}{2} + \lambda(\xi_{ij})\xi_{ij}^2\right\} + \log\frac{C(e^n)}{C(e)} + \log\frac{\Gamma(a_n)}{\Gamma(a_0)} + \log\frac{\Gamma(c_n)}{\Gamma(c_0)} \\
&+ a_0\log b_0 + a_n(1-\frac{b_0}{b_n}-\log b_n) + c_0\log d_0 + c_n(1-\frac{d_0}{d_n}-\log d_n) \\
&+ \frac{1}{2}\sum_{k\leq l}^{K}\log(\sigma_\alpha^2)_{kl} + \frac{1}{2}\log|S_\beta| - \sum_{i=1}^{n}\sum_{k=1}^{K}\tau_{ik}\log\tau_{ik} - \frac{1}{2}\sum_{k\leq l}^{K}(\sigma_\alpha^2)_{kl}^{-1}(m_\alpha)_{kl}^2 - \frac{1}{2}m_\beta^\intercal S_\beta^{-1}m_\beta \\
&+ \sum_{k=1}^{K}(m_\alpha)_{kk}\sum_{i\neq j}^{n}\left(\frac{1}{2}(Y_{ij}-\frac{1}{2}) - \lambda(\xi_{ij})x_{ij}^\intercal m_\beta\right)\tau_{ik}\tau_{jk} \\
&+ \sum_{k<l}^{K}(m_\alpha)_{kl}\sum_{i\neq j}^{n}\left((Y_{ij}-\frac{1}{2}) - 2\lambda(\xi_{ij})x_{ij}^\intercal m_\beta\right)\tau_{ik}\tau_{jl} \\
&- \sum_{k=1}^{K}\mathbb{E}_{\alpha_{kk}}[\alpha_{kk}^2]\frac{1}{2}\left(\frac{a_n}{b_n} - (\sigma_\alpha^2)_{kk}^{-1} + \sum_{i\neq j}^{n}\lambda(\xi_{ij})\tau_{ik}\tau_{jk}\right) \\
&- \sum_{k<l}^{K}\mathbb{E}_{\alpha_{kl}}[\alpha_{kl}^2]\frac{1}{2}\left(\frac{a_n}{b_n} - (\sigma_\alpha^2)_{kl}^{-1} + 2\sum_{i\neq j}^{n}\lambda(\xi_{ij})\tau_{ik}\tau_{jl}\right) \\
&+ \frac{1}{2}m_\beta^\intercal\sum_{i\neq j}^{n}(Y_{ij}-\frac{1}{2})x_{ij} \\
&- \frac{1}{2}\mathrm{Tr}\left(\left(2\sum_{i\neq j}^{n}\lambda(\xi_{ij})x_{ij}x_{ij}^\intercal + \frac{c_n}{d_n}I_d - S_\beta^{-1}\right)(S_\beta + m_\beta m_\beta^\intercal)\right).
\end{aligned}$$

Finally, since the terms at the fourth and fifth line correspond exactly to $\sum_{k\leq l}^{K}(m_\alpha)_{kl}(\sigma_\alpha^2)_{kl}^{-1}(m_\alpha)_{kl}$, and after the VBEM update step

$$\begin{aligned}\mathcal{L}_K(q;\ \xi) = \frac{1}{2}\sum_{i\neq j}^{n}\left\{\log g(\xi_{ij}) - \frac{\xi_{ij}}{2} + \lambda(\xi_{ij})\xi_{ij}^2\right\} + \log\frac{C(e^n)}{C(e)} + \log\frac{\Gamma(a_n)}{\Gamma(a_0)} + \log\frac{\Gamma(c_n)}{\Gamma(c_0)} \\ + a_0\log b_0 + a_n(1 - \frac{b_0}{b_n} - \log b_n) + c_0\log d_0 + c_n(1 - \frac{d_0}{d_n} - \log d_n) \\ + \frac{1}{2}\sum_{k\leq l}^{K}\log(\sigma_\alpha^2)_{kl} + \frac{1}{2}\log|S_\beta| - \sum_{i=1}^{n}\sum_{k=1}^{K}\tau_{ik}\log\tau_{ik} + \frac{1}{2}\sum_{k\leq l}^{K}(\sigma_\alpha^2)_{kl}^{-1}(m_\alpha)_{kl}^2 - \frac{1}{2}m_\beta^\intercal S_\beta^{-1}m_\beta \\ + \frac{1}{2}m_\beta^\intercal\sum_{i\neq j}^{n}(Y_{ij} - \frac{1}{2})x_{ij}.\end{aligned}$$

References

[1] T.S. Jaakkola and M.I. Jordan. Bayesian parameter estimation via variational methods. *Statistics and Computing*, 10:25–37, 2000.