Random Mixture Cox Point Processes
Supplementary Material

Athanasios C. Micheas

September 25, 2024

1 Formulae from the main paper

The following formulae appear in the main paper and are referenced in the proofs below:

$$\varphi_j(\mathbf{s}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) = |2\pi\boldsymbol{\Sigma}_j|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\mathbf{s}-\boldsymbol{\mu}_j\right)^T \boldsymbol{\Sigma}_j^{-1}\left(\mathbf{s}-\boldsymbol{\mu}_j\right)\right\}, \tag{1}$$

$$f_j(\mathbf{s}|\boldsymbol{\theta}_j) = \exp\left\{\mathbf{u}\left(\mathbf{s}\right)^T \boldsymbol{\theta}_j - b(\boldsymbol{\theta}_j) - A(\mathbf{s})\right\}, \tag{2}$$

$$\pi(\boldsymbol{\theta}|\boldsymbol{\xi}) = \pi(\lambda|a_\lambda, b_\lambda)\pi(\mathbf{p}|\mathbf{d}) \prod_{j=1}^{m} \pi(\boldsymbol{\theta}_j|\boldsymbol{\eta}_j), \tag{3}$$

$$\pi(\boldsymbol{\theta}_j|\boldsymbol{\nu}_j, a_j) = \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j) + K(a_j, \boldsymbol{\nu}_j)\right\}, \tag{4}$$

$$I_1 = \int_0^{+\infty} \frac{e^{-\lambda}\lambda^n}{n!}\pi(\lambda|a_\lambda, b_\lambda)d\lambda, \tag{5}$$

$$I_2 = \int_{\mathcal{S}_m} \left[\prod_{j=1}^{m} p_j^{z_{\bullet j}}\right] \pi(\mathbf{p}|\mathbf{d})d\mathbf{p}, \tag{6}$$

$$I_{3j} = \int_{\Theta_j} \prod_{i=1}^{n} \left(f_j(\mathbf{s}_i|\boldsymbol{\theta}_j)\right)^{z_{ij}} \pi(\boldsymbol{\theta}_j|\boldsymbol{\eta}_j)d\boldsymbol{\theta}_j, \tag{7}$$

$$\pi(\boldsymbol{\theta}|\boldsymbol{\xi}) = \pi(\lambda|a_\lambda, b_\lambda)\pi(\mathbf{p}|\mathbf{d}) \prod_{j=1}^{m} \pi(\boldsymbol{\mu}_j|\boldsymbol{\mu}_0, \sigma_0^2, \boldsymbol{\Sigma}_j)\pi(\boldsymbol{\Sigma}_j|n_0, \mathbf{B}), \tag{8}$$

$$f_{N_C}(\varphi_n, n, \mathbf{z}_{1:n}|\boldsymbol{\xi}) = I_1 I_2 \prod_{j=1}^{m} I_{3j}, \tag{9}$$

$$\lambda_{\boldsymbol{\theta}}(\mathbf{s}) = \lambda \sum_{j=1}^{m} p_j f_j(\mathbf{s}|\boldsymbol{\theta}_j), \tag{10}$$

$$\begin{aligned} \mathcal{L}_{Q_M}(f) &= \int_{\mathbb{N}^f} \exp\left\{-\int_{\mathcal{R}^2} f(\mathbf{s})\Lambda(d\mathbf{s})\right\} Q_M(d\Lambda) \\ &= \int_{\boldsymbol{\Theta}} \exp\left\{-\int_{\mathcal{R}^2} f(\mathbf{s})\lambda_{\boldsymbol{\theta}}(\mathbf{s})\mu_2(d\mathbf{s})\right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi})\mu_2(d\boldsymbol{\theta}), \end{aligned} \tag{11}$$

2 Data Augmented RMCP Density

Lemma 1 (Data Augmented RMCP Density) *Assume that the mixture components of the RMCP are given by (2), and the driving density* $\pi(\boldsymbol{\theta}|\boldsymbol{\xi})$ *is of the form (3), with* $\lambda \sim Gamma(a_\lambda, b_\lambda)$, $\mathbf{p} \sim Dirichlet(\mathbf{d})$, *and the component parameters* $\boldsymbol{\theta}_j$ *follow the prior* $\pi(\boldsymbol{\theta}_j|\boldsymbol{\nu}_j, a_j)$ *of equation (4). Then, the distribution of the data augmented RMCP is given by*

$$f_{N_C}(\varphi_n, n, \mathbf{z}_{1:n}|\boldsymbol{\xi}) = \frac{\Gamma(a_\lambda+n)b_\lambda^n}{\Gamma(a_\lambda)(b_\lambda+1)^{a_\lambda+n}n!}\frac{\Gamma(d_0)}{\Gamma(d_0+n)}\prod_{j=1}^{m}\frac{\Gamma(d_j+z_{\bullet j})}{\Gamma(d_j)} \exp\left\{\sum_{j=1}^{m}K(a_j,\boldsymbol{\nu}_j)-\sum_{j=1}^{m}K\left(a_j^*,\boldsymbol{\nu}_j^*\right)-\sum_{i=1}^{n}A(\mathbf{s}_i)\right\}, \tag{12}$$

where $d_0 = \sum_{j=1}^{m} d_j$, $a_j^* = a_j + z_{\bullet j}$, $\boldsymbol{\nu}_j^* = \boldsymbol{\nu}_j + \sum_{i=1}^{n} z_{ij}\mathbf{u}(\mathbf{s}_i)$, $j = 1, 2, \ldots m$, *and* $\boldsymbol{\xi} = (a_\lambda, b_\lambda, \mathbf{d}, \mathbf{a}, \boldsymbol{\nu}_{1:m})$, $\mathbf{a} = (a_1, \ldots, a_m)$.

Proof. We require calculation of I_1, I_2 and I_{3j}, for all $j = 1, 2, \ldots m$. Using the $Gamma(a_\lambda, b_\lambda)$ prior for $\pi(\lambda|a_\lambda, b_\lambda)$, it is straightforward to show that (5) is given by

$$\begin{aligned} I_1 &= \int_0^{+\infty}\frac{e^{-\lambda}\lambda^n}{n!}\frac{\lambda^{a_\lambda-1}e^{-\lambda/b_\lambda}}{\Gamma(a_\lambda)b_\lambda^{a_\lambda}}d\lambda \\ &= \frac{\Gamma(a_\lambda+n)}{\Gamma(a_\lambda)b_\lambda^{a_\lambda}(1+1/b_\lambda)^{a_\lambda+n}n!}\int_0^{+\infty}\frac{\lambda^{a_\lambda+n-1}e^{-\lambda(1+1/b_\lambda)}}{\Gamma(a_\lambda+n)(1+1/b_\lambda)^{-a_\lambda-n}}d\lambda, \end{aligned}$$

and therefore

$$I_1 = \frac{\Gamma(a_\lambda+n)b_\lambda^n}{\Gamma(a_\lambda)(b_\lambda+1)^{a_\lambda+n}n!}. \tag{13}$$

Now let $d_0 = \sum_{j=1}^{m} d_j$, and use the Dirichlet prior

$$\pi(\mathbf{p}|\mathbf{d}) = \frac{\Gamma(d_0)}{\Gamma(d_1)\ldots\Gamma(d_m)}p_1^{d_1-1}\ldots p_m^{d_m-1},$$

in (6) to obtain

$$\begin{aligned} I_2 &= \int_{\mathcal{S}_m}\left[\prod_{j=1}^{m}p_j^{z_{\bullet j}}\right]\frac{\Gamma(d_0)}{\Gamma(d_1)\ldots\Gamma(d_m)}p_1^{d_1-1}\ldots p_m^{d_m-1}d\mathbf{p} \\ &= \frac{\Gamma(d_0)}{\Gamma(d_1)\ldots\Gamma(d_m)}\int_{\mathcal{S}_m}p_1^{d_1+z_{\bullet 1}-1}\ldots p_m^{d_m+z_{\bullet m}-1}d\mathbf{p} \\ &= \frac{\Gamma(d_0)}{\Gamma(d_0+n)}\prod_{j=1}^{m}\frac{\Gamma(d_j+z_{\bullet j})}{\Gamma(d_j)}\int_{\mathcal{S}_m}\frac{\Gamma(d_0+n)}{\prod_{j=1}^{m}\Gamma(d_j+z_{\bullet j})}p_1^{d_1+z_{\bullet 1}-1}\ldots p_m^{d_m+z_{\bullet m}-1}d\mathbf{p}, \end{aligned}$$

which reduces to

$$I_2 = \frac{\Gamma(d_0)}{\Gamma(d_0+n)}\prod_{j=1}^{m}\frac{\Gamma(d_j+z_{\bullet j})}{\Gamma(d_j)}, \tag{14}$$

since $\sum_{j=1}^{m} z_{\bullet j} = \sum_{i=1}^{n}\sum_{j=1}^{m} z_{ij} = n$, and the integral above is that of a $Dirichlet(d_1 + z_{\bullet 1}, \ldots, d_m + z_{\bullet m})$ density.

Finally, we consider (7) with $f_j(\mathbf{s}_i|\boldsymbol{\theta}_j)$ and $\pi(\boldsymbol{\theta}_j|\boldsymbol{\eta}_j)$ given by (2) and (4), respectively. We have

$$\begin{aligned} I_{3j} &= \int\limits_{\Theta_j} \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j) + K(a_j,\boldsymbol{\nu}_j)\right\} \\ &\quad \prod_{i=1}^n \exp\left\{z_{ij}\mathbf{u}\left(\mathbf{s}_i\right)^T\boldsymbol{\theta}_j - z_{ij}b(\boldsymbol{\theta}_j) - z_{ij}A(\mathbf{s}_i)\right\} d\boldsymbol{\theta}_j \\ &= \exp\left\{K(a_j,\boldsymbol{\nu}_j) - \sum_{i=1}^n z_{ij}A(\mathbf{s}_i)\right\} \\ &\quad \int\limits_{\Theta_j} \exp\left\{\boldsymbol{\theta}_j^T\left(\boldsymbol{\nu}_j + \sum_{i=1}^n z_{ij}\mathbf{u}\left(\mathbf{s}_i\right)\right) - \left(a_j + z_{\bullet j}\right) b(\boldsymbol{\theta}_j)\right\} d\boldsymbol{\theta}_j, \end{aligned}$$

so that we recognize a member of the exponential family priors (4), with parameters $a_j^* = a_j + z_{\bullet j}$ and $\boldsymbol{\nu}_j^* = \boldsymbol{\nu}_j + \sum\limits_{i=1}^n z_{ij}\mathbf{u}\left(\mathbf{s}_i\right)$. As a result, we have

$$\begin{aligned} I_{3j} &= \exp\left\{K(a_j,\boldsymbol{\nu}_j) - \sum_{i=1}^n z_{ij}A(\mathbf{s}_i) - K\left(a_j^*,\boldsymbol{\nu}_j^*\right)\right\} \\ &\quad \int\limits_{\Theta_j} \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j^* - a_j^* b(\boldsymbol{\theta}_j) + K\left(a_j^*,\boldsymbol{\nu}_j^*\right)\right\} d\boldsymbol{\theta}_j, \end{aligned}$$

and since the integral above is one, we have

$$I_{3j} = \exp\left\{K(a_j,\boldsymbol{\nu}_j) - \sum_{i=1}^n z_{ij}A(\mathbf{s}_i) - K\left(a_j^*,\boldsymbol{\nu}_j^*\right)\right\}. \tag{15}$$

Now putting together (13), (14) and (15), we have the result. ■

3 Exact RMCP Density

Lemma 2 (Exact RMCP Density) *Consider the setup of the previous lemma. The density of the RMCP is given by*

$$\begin{aligned} f_{N_C}(\varphi_n, n|\boldsymbol{\xi}) &= c\exp\left\{-\sum_{i=1}^n A(\mathbf{s}_i)\right\} \sum_{k_1=1}^m \dots \sum_{k_n=1}^m \prod_{j=1}^m \Gamma\left(d_j + \sum_{i=1}^n I(k_i = j)\right) \\ &\quad \exp\left\{-K\left(a_j + \sum_{i=1}^n I(k_i=j), \boldsymbol{\nu}_j + \sum_{i=1}^n I(k_i=j)\mathbf{u}\left(\mathbf{s}_i\right)\right)\right\}, \end{aligned} \tag{16}$$

where c is given by (17) in the proof below.

Proof. First we collect all quantities in equation (12) that do not involve the auxiliary variables $\mathbf{z}_{1:n}$, i.e., let

$$c = \frac{\Gamma(a_\lambda + n)b_\lambda^n}{\Gamma(a_\lambda)(b_\lambda+1)^{a_\lambda+n}n!} \frac{\Gamma\left(d_0\right)}{\Gamma(d_0+n)\prod\limits_{j=1}^m \Gamma(d_j)} \exp\left\{\sum_{j=1}^m K(a_j,\boldsymbol{\nu}_j)\right\}. \tag{17}$$

Therefore, we can rewrite (12) as

$$f_{N_C}(\varphi_n, n, \mathbf{z}_{1:n}|\boldsymbol{\xi}) = c\prod_{j=1}^m \Gamma(d_j + z_{\bullet j})\exp\left\{-K\left(a_j^*,\boldsymbol{\nu}_j^*\right)\right\},$$

and we are interested in the marginal density

$$\begin{aligned} f_{N_C}(\varphi_n, n|\boldsymbol{\xi}) &= \sum_{\mathbf{z}_1=(z_{1,1},\dots,z_{1,m})\in\mathcal{M}_z} \dots \sum_{\mathbf{z}_n=(z_{n,1},\dots,z_{n,m})\in\mathcal{M}_z} f_{N_C}(\varphi_n, n, \mathbf{z}_{1:n}|\boldsymbol{\xi}) \\ &= c\sum_{\mathbf{z}_1=(z_{1,1},\dots,z_{1,m})\in\mathcal{M}_z} \dots \sum_{\mathbf{z}_n=(z_{n,1},\dots,z_{n,m})\in\mathcal{M}_z} I(\mathbf{z}_1,\dots,\mathbf{z}_n), \end{aligned}$$

where

$$I(\mathbf{z}_1, ..., \mathbf{z}_n) = \prod_{j=1}^{m} \Gamma(d_j + z_{\bullet j}) \exp\left\{-K\left(a_j^*, \boldsymbol{\nu}_j^*\right)\right\}.$$

Let $\boldsymbol{\varepsilon}_{k_i} = (0, ..., 0, \underset{k_i^{th} \text{ position}}{1}, 0, ..., 0)_{1\times m}$, $k_i = 1, 2, \ldots, m$, $i = 1, 2, ..., n$, or $\boldsymbol{\varepsilon}_{k_i} = (\varepsilon_{k_i}^{(1)}, \ldots, \varepsilon_{k_i}^{(m)})$, with $\varepsilon_{k_i}^{(j)} = I(k_i = j)$, $j = 1, 2, ..., m$, so that all the elements of $\mathcal{M}_z$ are described by the vectors $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_m$. Consequently, we have

$$\sum_{\mathbf{z}_n=(z_{n,1},...,z_{n,m})\in\mathcal{M}_z} I(\mathbf{z}_1, ..., \mathbf{z}_n) = \sum_{k_n=1}^{m} I(\mathbf{z}_1, ..., \boldsymbol{\varepsilon}_{k_n}),$$

and using this n times we can write the density of the RMCP as

$$\begin{aligned} f_{N_C}(\varphi_n, n|\boldsymbol{\xi}) &= c\exp\left\{-\sum_{i=1}^{n} A(\mathbf{s}_i)\right\} \sum_{\mathbf{z}_1=(z_{1,1},...,z_{1,m})\in\mathcal{M}_z} \cdots \sum_{\mathbf{z}_{n-1}=(z_{n-1,1},...,z_{n-1,m})\in\mathcal{M}_z} \sum_{k_n=1}^{m} I(\mathbf{z}_1, ..., \boldsymbol{\varepsilon}_{k_n}) \\ &= \ldots = c\sum_{k_1=1}^{m} \ldots \sum_{k_n=1}^{m} I(\boldsymbol{\varepsilon}_{k_1}, ..., \boldsymbol{\varepsilon}_{k_n}), \end{aligned}$$

or in terms of $\varepsilon_{k_i}^{(j)}$, we have

$$\begin{aligned} f_{N_C}(\varphi_n, n|\boldsymbol{\xi}) \quad &= \quad c\exp\left\{-\sum_{i=1}^{n} A(\mathbf{s}_i)\right\} \sum_{k_1=1}^{m} \cdots \sum_{k_n=1}^{m} \prod_{j=1}^{m} \Gamma\left(d_j + \sum_{i=1}^{n} \varepsilon_{k_i}^{(j)}\right) \\ &\quad \exp\left\{-K\left(a_j + \sum_{i=1}^{n} \varepsilon_{k_i}^{(j)}, \boldsymbol{\nu}_j + \sum_{i=1}^{n} \varepsilon_{k_i}^{(j)} \mathbf{u}(\mathbf{s}_i)\right)\right\}, \end{aligned}$$

and we have the result. ■

4 Generating functional of the driving density

Theorem 3 (Generating functional of the driving density) *For the RMCP of equation (10) with driving density $\pi(\boldsymbol{\theta}|\boldsymbol{\xi})$, we have that the Laplace functional is given by*

$$\mathcal{L}_{Q_M}(f) = \int_{\boldsymbol{\Theta}} \exp\left\{-\lambda \sum_{j=1}^{m} p_j E_{\boldsymbol{\theta}_j}\right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta}, \tag{18}$$

where $\boldsymbol{\theta} = (\lambda, \mathbf{p}, \boldsymbol{\theta}_{1:m}) \in \boldsymbol{\Theta}$, and $E_{\boldsymbol{\theta}_j} = E^{f_j}(f(\mathbf{X}))$, where $\mathbf{X} \sim f_j(\mathbf{s})$, $j = 1, 2, \ldots m$.

Proof. Using (10) in (11) (and the Riemann versions of the integrals) we have

$$\mathcal{L}_{Q_M}(f) = \int_{\boldsymbol{\Theta}} \exp\left\{-\int_{\mathcal{R}^2} f(\mathbf{s})\lambda \sum_{j=1}^{m} p_j f_j(\mathbf{s}|\boldsymbol{\theta}_j) d\mathbf{s}\right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta},$$

where the inner integral is written as

$$\lambda \int_{\mathcal{R}^2} f(\mathbf{s}) \sum_{j=1}^{m} p_j f_j(\mathbf{s}|\boldsymbol{\theta}_j) d\mathbf{s} = \lambda \sum_{j=1}^{m} p_j \int_{\mathcal{R}^2} f(\mathbf{s}) f_j(\mathbf{s}|\boldsymbol{\theta}_j) d\mathbf{s},$$

and letting

$$E_{\boldsymbol{\theta}_j} = E^{f_j}(f(\mathbf{X})) = \int_{\mathcal{R}^2} f(\mathbf{s}) f_j(\mathbf{s}|\boldsymbol{\theta}_j) d\mathbf{s},$$

where $\mathbf{X} \sim f_j(\mathbf{s})$, we have the result. ■

5 RMCP functionals

Theorem 4 (RMCP functionals) *For the RMCP with driving density* $\pi(\boldsymbol{\theta}|\boldsymbol{\xi})$*, we have that*

$$\mathcal{L}_{N_C}(f) = \int_{\boldsymbol{\Theta}} \exp\left\{-\lambda \sum_{j=1}^{m} p_j E^{f_j}(1 - e^{-f(\mathbf{s})})\right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta}, \tag{19}$$

$$\mathcal{G}_{N_C}(g) = \int_{\boldsymbol{\Theta}} \exp\left\{-\lambda \sum_{j=1}^{m} p_j E^{f_j}(1 - g(\mathbf{s}))\right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta}, \tag{20}$$

and

$$v_{N_C}(K) = \int_{\boldsymbol{\Theta}} \exp\left\{-\lambda \sum_{j=1}^{m} p_j P(\mathbf{X} \in K|\mathbf{X} \sim f_j)\right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta}, \tag{21}$$

where f *is a non-negative* $\mathcal{B}(\mathcal{R}^2)$*-measurable function,* g *is a* $\mathcal{B}(\mathcal{R}^2)$*-measurable function with* $0 \le g(\mathbf{s}) \le 1$*,* $\forall \mathbf{s} \in \mathcal{R}^2$*, and* K *a compact subset of* $\mathcal{R}^2$*.*

Proof. A straightforward application of (18) gives the desired expressions. ■

6 RMCP moment measures

Theorem 5 (RMCP moment measures) *The first and second order moment measures of the driving measure* M *of the RMCP with driving density* $\pi(\boldsymbol{\theta}|\boldsymbol{\xi})$ *are given by*

$$\Lambda_M(B) = \Lambda_{N_C}(B) = \int_{\boldsymbol{\Theta}} \lambda \sum_{j=1}^{m} p_j P_{\boldsymbol{\theta}_j}(B) \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta},$$

and

$$\mu_M^{(2)}(B_1 \times B_2) = \int_{\boldsymbol{\Theta}} \lambda^2 \left[\sum_{j=1}^{m} p_j P_{\boldsymbol{\theta}_j}(B_1)\right]\left[\sum_{j=1}^{m} p_j P_{\boldsymbol{\theta}_j}(B_2)\right] \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta},$$

where $B, B_1, B_2 \in \mathcal{B}(\mathcal{R}^2)$*.*

Proof. For the proposed RMCP we have

$$\Lambda_{N_C}(B) = \Lambda_M(B) = \mu_{N_C}^{(1)}(B) = (-1) \lim_{a \downarrow 0} \tfrac{d}{da} \mathcal{L}_{N_C}(aI_B),$$

and since

$$\begin{aligned} E^{f_j}(1 - e^{-aI_B(\mathbf{X})}) &= \int_{\mathcal{R}^2} \left(1 - e^{-aI_B(\mathbf{s})}\right) f_j(\mathbf{s}|\boldsymbol{\theta}_j) d\mathbf{s} = (1 - e^{-a}) \int_B f_j(\mathbf{s}|\boldsymbol{\theta}_j) d\mathbf{s} \\ &= (1 - e^{-a}) P(\mathbf{X} \in B|\mathbf{X} \sim f_j), \end{aligned}$$

letting

$$P_{\boldsymbol{\theta}_j}(B) = P(\mathbf{X} \in B|\mathbf{X} \sim f_j),$$

it is easy to show that

$$\begin{aligned} \Lambda_{N_C}(B) &= (-1) \lim_{a \downarrow 0} \tfrac{d}{da} \int_{\boldsymbol{\Theta}} \exp\left\{-\lambda(1 - e^{-a}) \sum_{j=1}^{m} p_j P_{\boldsymbol{\theta}_j}(B)\right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta} \\ &= \lim_{a \downarrow 0} e^{-a} \int_{\boldsymbol{\Theta}} \lambda \sum_{j=1}^{m} p_j P_{\boldsymbol{\theta}_j}(B) \exp\left\{-\lambda(1 - e^{-a}) \sum_{j=1}^{m} p_j P_{\boldsymbol{\theta}_j}(B)\right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta} \\ &= \int_{\boldsymbol{\Theta}} \lambda \sum_{j=1}^{m} p_j P_{\boldsymbol{\theta}_j}(B) \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta}, \end{aligned}$$

provided that we can pass the derivative and the limit under the integral sign. Moreover, under the proposed RMCP we have

$$\mu_M^{(2)}(B_1 \times B_2) = \lim_{a_1,a_2 \downarrow 0} \frac{\partial^2}{\partial a_1 \partial a_2} \mathcal{G}_{N_C}(1 - a_1 I_{B_1} - a_2 I_{B_2}),$$

and since

$$\begin{aligned} E^{f_j}(1 - 1 + a_1 I_{B_1}(\mathbf{X}) + a_2 I_{B_2}(\mathbf{X})) &= \int_{\mathcal{R}^2} (a_1 I_{B_1}(\mathbf{s}) + a_2 I_{B_2}(\mathbf{s}))\, f_j(\mathbf{s}|\boldsymbol{\theta}_j) d\mathbf{s} \\ &= a_1 P_{\boldsymbol{\theta}_j}(B_1) + a_2 P_{\boldsymbol{\theta}_j}(B_2), \end{aligned}$$

it is straightforward to see that

$$\begin{aligned} \mu_M^{(2)}(B_1 \times B_2) &= \lim_{a_1,a_2 \downarrow 0} \frac{\partial^2}{\partial a_1 \partial a_2} \int_{\boldsymbol{\Theta}} \exp\left\{ -\lambda \sum_{j=1}^m p_j \left[a_1 P_{\boldsymbol{\theta}_j}(B_1) + a_2 P_{\boldsymbol{\theta}_j}(B_2) \right] \right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta} \\ &= \lim_{a_1,a_2 \downarrow 0} \frac{\partial}{\partial a_1} \int_{\boldsymbol{\Theta}} (-\lambda) \sum_{j=1}^m p_j P_{\boldsymbol{\theta}_j}(B_1) \\ &\quad \exp\left\{ -\lambda \sum_{j=1}^m p_j \left[a_1 P_{\boldsymbol{\theta}_j}(B_1) + a_2 P_{\boldsymbol{\theta}_j}(B_2) \right] \right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta} \\ &= \lim_{a_1,a_2 \downarrow 0} \int_{\boldsymbol{\Theta}} \lambda^2 \left[\sum_{j=1}^m p_j P_{\boldsymbol{\theta}_j}(B_1) \right] \left[\sum_{j=1}^m p_j P_{\boldsymbol{\theta}_j}(B_2) \right] \\ &\quad \exp\left\{ -\lambda \sum_{j=1}^m p_j \left[a_1 P_{\boldsymbol{\theta}_j}(B_1) + a_2 P_{\boldsymbol{\theta}_j}(B_2) \right] \right\} \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta}, \end{aligned}$$

provided that we can pass the partial derivatives and the limits under the integral sign, and therefore

$$\mu_M^{(2)}(B_1 \times B_2) = \int_{\boldsymbol{\Theta}} \lambda^2 \left[\sum_{j=1}^m p_j P_{\boldsymbol{\theta}_j}(B_1) \right] \left[\sum_{j=1}^m p_j P_{\boldsymbol{\theta}_j}(B_2) \right] \pi(\boldsymbol{\theta}|\boldsymbol{\xi}) d\boldsymbol{\theta}.$$

Note that the operations of passing the limits and derivatives under the integral signs are feasible since $\pi(\boldsymbol{\theta}|\boldsymbol{\xi})$ is well behaved (i.e., all integrals and their derivatives exist since the integrands are continuous and bounded with respect to $\boldsymbol{\theta}$, and $\pi(\boldsymbol{\theta}|\boldsymbol{\xi})$ is a proper density). ■

7 MLEs for the General RMCP

Theorem 6 (MLEs for the general RMCP) *Assume that* $\varphi_{n_1}^{(1)}, \ldots, \varphi_{n_K}^{(K)}$*, are point patterns from a RMCP, where* $\varphi_{n_k}^{(k)} = \left\{ \mathbf{s}_1^{(k)}, \ldots, \mathbf{s}_{n_k}^{(k)} \right\}$*,* $k = 1, 2, \ldots, K$*, and let* $\varphi = \{\varphi_{n_1}^{(1)}, \ldots, \varphi_{n_K}^{(K)}\}$*,* $\mathbf{n} = (n_1, \ldots, n_K)$*, denote all the numbers of events, and* $z = \{\mathbf{z}_{1:n_1}^{(1)}, \ldots, \mathbf{z}_{1:n_K}^{(K)}\}$*, the corresponding auxiliary variables, where* $\mathbf{z}_{1:n_k}^{(k)} = \{\mathbf{z}_1^{(k)}, \ldots, \mathbf{z}_{n_k}^{(k)}\}$ *and* $\mathbf{z}_i^{(k)} = \left(z_{i1}^{(k)}, \ldots, z_{im}^{(k)} \right)$*. The MLEs of* $\boldsymbol{\xi} = (a_\lambda, b_\lambda, \mathbf{d}, \mathbf{a}, \boldsymbol{\nu}_{1:m})$*,* $\mathbf{a} = (a_1, \ldots, a_m)$*, based on the data augmented RMCP of equation (12) are given by the solutions to the following equations*

$$\widehat{b}_\lambda = \frac{n_\bullet}{K \widehat{a}_\lambda}, \tag{22}$$

$$\sum_{k=1}^K \Psi(\widehat{a}_\lambda + n_k) - K\Psi(\widehat{a}_\lambda) = K \log\left(\frac{n_\bullet}{K \widehat{a}_\lambda} + 1 \right), \tag{23}$$

$$K\Psi(\widehat{d}_0) - \sum_{k=1}^K \Psi(\widehat{d}_0 + n_k) = K\Psi\left(\widehat{d}_j\right) - \sum_{k=1}^K \Psi(\widehat{d}_j + z_{\bullet j}^{(k)}), \tag{24}$$

$$K E^{\boldsymbol{\theta}_j | \widehat{\boldsymbol{\nu}}_j, \widehat{a}_j} \left[b(\boldsymbol{\theta}_j) \right] = \sum_{k=1}^K E^{\boldsymbol{\theta}_j | \widehat{\boldsymbol{\nu}}_j + \mathbf{u}_j^{(k)}, \widehat{a}_j + z_{\bullet j}^{(k)}} \left[b(\boldsymbol{\theta}_j) \right], \tag{25}$$

and

$$KE^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j,\widehat{a}_j}\left[\boldsymbol{\theta}_j\right]=\sum_{k=1}^{K}E^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j+\mathbf{u}_j^{(k)},\widehat{a}_j+z_{\bullet j}^{(k)}}\left[\boldsymbol{\theta}_j\right]. \tag{26}$$

for $j=1,2,\ldots m$*, where* $\mathbf{u}_j^{(k)}=\sum_{i=1}^{n}z_{ij}^{(k)}\mathbf{u}\left(\mathbf{s}_i^{(k)}\right)$, $n_\bullet=\sum_{k=1}^{K}n_k$, $\widehat{d}_0=\sum_{j=1}^{m}\widehat{d}_j$, *and* $\Psi(x)=\frac{d}{dx}\log\left(\Gamma(x)\right)$*, the digamma function.*

Proof. Using (12) we can write the log-likelihood as

$$l(\boldsymbol{\xi}|\varphi,\mathbf{n},z)=\sum_{k=1}^{K}\log f_{N_C}(\varphi_{n_k}^{(k)},n_k,\mathbf{z}_{1:n_k}^{(k)}|\boldsymbol{\xi})$$

and therefore

$$\begin{aligned}l(\boldsymbol{\xi}|\varphi,\mathbf{n},z) \;=\; & \sum_{k=1}^{K}\log\left(\Gamma(a_\lambda+n_k)\right)+\log\left(b_\lambda\right)\sum_{k=1}^{K}n_k-\sum_{k=1}^{K}\log(n_k!)-K\log(\Gamma(a_\lambda))\\ & -\left(Ka_\lambda+\sum_{k=1}^{K}n_k\right)\log(b_\lambda+1)+K\log(\Gamma\left(d_0\right))-\sum_{k=1}^{K}\log(\Gamma(d_0+n_k))\\ & +\sum_{k=1}^{K}\sum_{j=1}^{m}\log(\Gamma(d_j+z_{\bullet j}^{(k)}))-K\sum_{j=1}^{m}\log(\Gamma(d_j))+K\sum_{j=1}^{m}K(a_j,\boldsymbol{\nu}_j)\\ & -\sum_{k=1}^{K}\sum_{j=1}^{m}K\left(a_j+z_{\bullet j}^{(k)},\boldsymbol{\nu}_j+\sum_{i=1}^{n}z_{ij}^{(k)}\mathbf{u}\left(\mathbf{s}_i^{(k)}\right)\right)-\sum_{k=1}^{K}\sum_{i=1}^{n}A\left(\mathbf{s}_i^{(k)}\right),\end{aligned}$$

where $d_0=\sum_{j=1}^{m}d_j$, and we require maximization of the latter with respect to $\boldsymbol{\xi}=(a_\lambda,b_\lambda,\mathbf{d},\mathbf{a},\boldsymbol{\nu}_{1:m})$, $\mathbf{a}=(a_1,\ldots,a_m)$.

In order to maximize $l(\boldsymbol{\xi}|\varphi,n,z)$ with respect to a_λ and b_λ, we need to work with the equation

$$\begin{aligned}l(a_\lambda,b_\lambda|\mathbf{n}) \;=\; & \sum_{k=1}^{K}\log\left(\Gamma(a_\lambda+n)\right)+\log\left(b_\lambda\right)\sum_{k=1}^{K}n_k-\sum_{k=1}^{K}\log(n_k!)\\ & -K\log(\Gamma(a_\lambda))-\left(Ka_\lambda+\sum_{k=1}^{K}n_k\right)\log(b_\lambda+1)+\text{constant}.\end{aligned}$$

Let $n_\bullet=\sum_{k=1}^{K}n_k$, and recall that the digamma function is defined by $\Psi(x)=\frac{d}{dx}\log\left(\Gamma(x)\right)$. Then, taking derivatives with respect to a_λ and b_λ above yields

$$\frac{d}{da_\lambda}l(a_\lambda,b_\lambda|\mathbf{n})=\sum_{k=1}^{K}\Psi(a_\lambda+n_k)-K\Psi(a_\lambda)-K\log(b_\lambda+1), \tag{27}$$

and

$$\frac{d}{db_\lambda}l(a_\lambda,b_\lambda|n)=\frac{n_\bullet}{b_\lambda}-\frac{Ka_\lambda+n_\bullet}{b_\lambda+1}. \tag{28}$$

Now these equations are zero when evaluated at the MLEs $\widehat{a}_\lambda$, and $\widehat{b}_\lambda$, so that (28) yields

$$\widehat{b}_\lambda=\frac{n_\bullet}{K\widehat{a}_\lambda},$$

and therefore (27) reduces to

$$\sum_{k=1}^{K}\Psi(\widehat{a}_\lambda+n_k)-K\Psi(\widehat{a}_\lambda)=K\log\left(\frac{n_\bullet}{K\widehat{a}_\lambda}+1\right),$$

with the latter equation easily solved using numerical methods.

In order to maximize $l(\boldsymbol{\xi}|\varphi, n,z)$ with respect to $\mathbf{d}$, we need to work with the equation

$$\begin{aligned} l(\mathbf{d}|\mathbf{n}, z) &= K\log(\Gamma\left(d_0\right)) - \sum_{k=1}^{K} \log(\Gamma(d_0+n_k)) + \sum_{k=1}^{K}\sum_{j=1}^{m} \log(\Gamma(d_j + z_{\bullet j}^{(k)})) \\ &\quad -K\sum_{j=1}^{m} \log(\Gamma(d_j)) + \text{constant}. \end{aligned}$$

Now taking partial derivatives with respect to d_j, $j = 1, 2, \ldots m$, yields

$$\frac{\partial}{\partial d_j} l(\mathbf{d}|\mathbf{n}, \mathbf{z}_{1:n}) = K\Psi(d_0) - \sum_{k=1}^{K} \Psi(d_0+n_k) + \sum_{k=1}^{K} \Psi(d_j + z_{\bullet j}^{(k)}) - K\Psi(d_j),$$

and therefore, at the MLEs $\widehat{\mathbf{d}}$, we need to solve the equations

$$\begin{aligned} K\Psi(\widehat{d}_0) - \sum_{k=1}^{K} \Psi(\widehat{d}_0+n_k) &= K\Psi\left(\widehat{d}_j\right) - \sum_{k=1}^{K} \Psi(\widehat{d}_j + z_{\bullet j}^{(k)}), \\ \widehat{d}_j &> 0, \end{aligned}$$

for $j = 1, 2, \ldots m$, where $\widehat{d}_0 = \sum_{j=1}^{m} \widehat{d}_j$. Once again, we have m equations with m unknowns that can be easily solved numerically. In particular, let $\mathbf{f}(\mathbf{d}) = [f_1(\mathbf{d}), \ldots, f_m(\mathbf{d})]^T$, where

$$f_j(\mathbf{d}) = \Psi(d_0) - \Psi\left(d_j\right) + \frac{1}{K}\sum_{k=1}^{K} \left[\Psi(d_j + z_{\bullet j}^{(k)}) - \Psi(d_0+n_k)\right],$$

and note that

$$\frac{\partial f_j(\mathbf{d})}{\partial d_r} = \Psi_1(d_0) - \Psi_1\left(d_j\right)\delta_{rj} + \frac{1}{K}\sum_{k=1}^{K} \left[\Psi_1(d_j + z_{\bullet j}^{(k)})\delta_{rj} - \Psi_1(d_0+n_k)\right],$$

where $\Psi_1(x) = \frac{d}{dx}\Psi(x)$, the trigamma function, and δ_{jk} Kronecker's delta, with $\delta_{rj} = 1$, $j = r$, and 0, for $j \neq r$, $j, r = 1, 2, \ldots m$. Then, the MLE of $\mathbf{d}$ is given by the recursive formula

$$\mathbf{d}_{n+1} = \mathbf{d}_n - \mathbf{J}^{-1}(\mathbf{d}_n)\mathbf{f}(\mathbf{d}_n),$$

$n = 0, 1, \ldots$, where $\mathbf{J}(\mathbf{d}) = \left[\left(\frac{\partial f_j(\mathbf{d})}{\partial d_r}\right)\right]$ and $\mathbf{J}^{-1}$ is assumed to exist, for some starting value $\mathbf{d}_0$. Alternatively, the MLE of $\mathbf{d}$ is obtained using an optimization method such as gradient descent. In particular, for some starting value $\mathbf{d}_0$, we have

$$\mathbf{d}_{n+1} = \mathbf{d}_n - \zeta_n\mathbf{f}(\mathbf{d}_n),$$

where

$$\zeta_n = \frac{\left[\mathbf{d}_n - \mathbf{d}_{n-1}\right]^T\left[\mathbf{f}(\mathbf{d}_n) - \mathbf{f}(\mathbf{d}_{n-1})\right]}{\left\|\mathbf{f}(\mathbf{d}_n) - \mathbf{f}(\mathbf{d}_{n-1})\right\|^2}.$$

Finally, we turn to the maximization of $l(\boldsymbol{\xi}|\varphi, n,z)$ with respect to $\mathbf{a} = (a_1, \ldots, a_m)$ and $\boldsymbol{\nu}_{1:m}$, based on the log-likelihood

$$l(\mathbf{a}, \boldsymbol{\nu}_{1:m}|\varphi, \mathbf{n},z) = K\sum_{j=1}^{m} K(a_j, \boldsymbol{\nu}_j) - \sum_{k=1}^{K}\sum_{j=1}^{m} K\left(a_j + z_{\bullet j}^{(k)}, \boldsymbol{\nu}_j + \mathbf{u}_j^{(k)}\right) + \text{constant}, \tag{29}$$

where $\mathbf{u}_j^{(k)} = \sum_{i=1}^{n} z_{ij}^{(k)}\mathbf{u}\left(\mathbf{s}_i^{(k)}\right)$. From (4), since

$$\pi(\boldsymbol{\theta}_j|\boldsymbol{\nu}_j, a_j) = \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j) + K(a_j, \boldsymbol{\nu}_j)\right\},$$

is a density, we have

$$\exp\left\{-K(a_j,\boldsymbol{\nu}_j)\right\} = \int\limits_{\boldsymbol{\Theta}_j} \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j)\right\} d\boldsymbol{\theta}_j,$$

or

$$K(a_j,\boldsymbol{\nu}_j) = -\log\left(\int\limits_{\boldsymbol{\Theta}_j} \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j)\right\} d\boldsymbol{\theta}_j\right).$$

First we note that

$$\begin{aligned}
\frac{\partial}{\partial a_j}K(a_j,\boldsymbol{\nu}_j) &= -\frac{\int_{\boldsymbol{\Theta}_j} b(\boldsymbol{\theta}_j)\exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j)\right\} d\boldsymbol{\theta}_j}{\int_{\boldsymbol{\Theta}_j} \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j)\right\} d\boldsymbol{\theta}_j} \\
&= -\frac{\int_{\boldsymbol{\Theta}_j} b(\boldsymbol{\theta}_j)\exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j) + K(a_j,\boldsymbol{\nu}_j)\right\} d\boldsymbol{\theta}_j}{\int_{\boldsymbol{\Theta}_j} \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j) + K(a_j,\boldsymbol{\nu}_j)\right\} d\boldsymbol{\theta}_j} \\
&= -E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j,a_j}\left[b(\boldsymbol{\theta}_j)\right],
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{\nu}_j}K(a_j,\boldsymbol{\nu}_j) &= -\frac{\int_{\boldsymbol{\Theta}_j} \boldsymbol{\theta}_j\exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j)\right\} d\boldsymbol{\theta}_j}{\int_{\boldsymbol{\Theta}_j} \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j)\right\} d\boldsymbol{\theta}_j} \\
&= -\frac{\int_{\boldsymbol{\Theta}_j} \boldsymbol{\theta}_j\exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j) + K(a_j,\boldsymbol{\nu}_j)\right\} d\boldsymbol{\theta}_j}{\int_{\boldsymbol{\Theta}_j} \exp\left\{\boldsymbol{\theta}_j^T\boldsymbol{\nu}_j - a_j b(\boldsymbol{\theta}_j) + K(a_j,\boldsymbol{\nu}_j)\right\} d\boldsymbol{\theta}_j} \\
&= -E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j,a_j}\left[\boldsymbol{\theta}_j\right].
\end{aligned}$$

Now take partial derivatives of (29) with respect to a_j to obtain

$$\frac{\partial}{\partial a_j}l(\mathbf{a},\boldsymbol{\nu}_{1:m}|\varphi,\mathbf{n},z) = K\frac{\partial}{\partial a_j}K(a_j,\boldsymbol{\nu}_j) - \sum_{k=1}^{K}\frac{\partial}{\partial a_j}K\left(a_j + z_{\bullet j}^{(k)},\boldsymbol{\nu}_j + \mathbf{u}_j^{(k)}\right),$$

so that the MLE of a_j is the solution to the equation

$$KE^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j,\widehat{a}_j}\left[b(\boldsymbol{\theta}_j)\right] = \sum_{k=1}^{K} E^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j+\mathbf{u}_j^{(k)},\widehat{a}_j+z_{\bullet j}^{(k)}}\left[b(\boldsymbol{\theta}_j)\right].$$

Similarly, the partial derivative (gradient) of (29) with respect to $\boldsymbol{\nu}_j$ yields

$$\frac{\partial}{\partial \boldsymbol{\nu}_j}l(\mathbf{a},\boldsymbol{\nu}_{1:m}|\varphi,\mathbf{n},z) = K\frac{\partial}{\partial \boldsymbol{\nu}_j}K(a_j,\boldsymbol{\nu}_j) - \sum_{k=1}^{K}\frac{\partial}{\partial \boldsymbol{\nu}_j}K\left(a_j + z_{\bullet j}^{(k)},\boldsymbol{\nu}_j + \mathbf{u}_j^{(k)}\right),$$

and therefore, the MLE of $\boldsymbol{\nu}_j$ is given by the solution to the equations

$$KE^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j,\widehat{a}_j}\left[\boldsymbol{\theta}_j\right] = \sum_{k=1}^{K} E^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j+\mathbf{u}_j^{(k)},\widehat{a}_j+z_{\bullet j}^{(k)}}\left[\boldsymbol{\theta}_j\right].$$

■

8 RMCP Model with Normal Components

Lemma 7 (Conjugate Prior of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}^{-1}$) *The conjugate prior for the mixture component parameters $\boldsymbol{\mu}_j$ and $\boldsymbol{\Sigma}_j^{-1}$ (in non-canonical form), corresponding to the mixture component distribution of equation (1) is the product of a bivariate Normal and a Wishart distribution, i.e.,*

$$\boldsymbol{\mu}_j|\boldsymbol{\Sigma}_j^{-1}, a_j, \boldsymbol{\nu}_{j1} \sim N_2\left(\frac{1}{a_j}\boldsymbol{\nu}_{j1}, \frac{1}{a_j}\boldsymbol{\Sigma}_j\right),$$

and

$$\boldsymbol{\Sigma}_j^{-1}|a_j, \boldsymbol{\nu}_{j1}, \boldsymbol{\nu}_{j2} \sim W_2\left(a_j + 4, \left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right).$$

Proof. Without loss of generality and for better exposition, we drop the mixture component index. The mixture component density in canonical form is given by

$$\varphi(\mathbf{s}|\boldsymbol{\theta}) = \exp\{\boldsymbol{\theta}_1^T\mathbf{u}_1(\mathbf{s}) + \boldsymbol{\theta}_2^T\mathbf{u}_2(\mathbf{s}) - \frac{1}{2}\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \ln(2\pi) - \frac{1}{2}\ln(|\boldsymbol{\Sigma}|)\},$$

so that

$$b(\boldsymbol{\theta}) = \frac{1}{2}\ln(|\boldsymbol{\Sigma}|) + \frac{1}{2}\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu},$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2)^T$, $\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{21} & \sigma^{22} \end{bmatrix}$, $vec(\boldsymbol{\Sigma}^{-1})^T = (\sigma^{11}, \sigma^{21}, \sigma^{12}, \sigma^{22})^T$, so that

$$\begin{aligned}
\boldsymbol{\theta} &= (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)^T = \left(\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}, -\frac{1}{2}vec(\boldsymbol{\Sigma}^{-1})^T\right)^T \\
&= (\mu_1\sigma^{11} + \mu_2\sigma^{21}, \mu_1\sigma^{12} + \mu_2\sigma^{22}, -\frac{1}{2}\sigma^{11}, -\frac{1}{2}\sigma^{21}, -\frac{1}{2}\sigma^{12}, -\frac{1}{2}\sigma^{22})^T.
\end{aligned}$$

Based on the form of $b(\boldsymbol{\theta})$, the prior density under conjugacy (4) reduces to

$$\begin{aligned}
\pi(\boldsymbol{\theta}|\boldsymbol{\nu}, a) &= \exp\left\{\boldsymbol{\theta}^T\boldsymbol{\nu} - \frac{a}{2}\ln(|\boldsymbol{\Sigma}|) - \frac{a}{2}\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + K(a, \boldsymbol{\nu})\right\} \\
&= \exp\left\{\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\nu}_1 - \frac{1}{2}vec(\boldsymbol{\Sigma}^{-1})^T vec(\boldsymbol{\nu}_2) - \frac{a}{2}\ln(|\boldsymbol{\Sigma}|) - \frac{a}{2}\boldsymbol{\mu}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + K(a, \boldsymbol{\nu})\right\},
\end{aligned}$$

where $\boldsymbol{\nu} = (\boldsymbol{\nu}_1^T, vec(\boldsymbol{\nu}_2)^T)^T$, with $\boldsymbol{\nu}_1$ a 2×1 real vector and $\boldsymbol{\nu}_2$ is a symmetric, 2×2 real matrix. Now consider the transformation

$$\begin{aligned}
\boldsymbol{\theta} &\longmapsto \left(\boldsymbol{\mu}^T, vec(\boldsymbol{\Sigma}^{-1})^T\right)^T = (\mu_1, \mu_2, \sigma^{11}, \sigma^{21}, \sigma^{12}, \sigma^{22})^T \\
&= (u_1, u_2, u_3, u_4, u_5, u_6)^T = \mathbf{u}^T,
\end{aligned}$$

with Jacobian

$$\begin{aligned}
J &= \left|\frac{\partial\boldsymbol{\theta}}{\partial\mathbf{u}}\right| = \left[\left(\frac{\partial\theta_i}{\partial u_j}\right)\right] = \left\|\begin{bmatrix} \sigma^{11} & \sigma^{21} & \mu_1 & \mu_2 & 0 & 0 \\ \sigma^{12} & \sigma^{22} & 0 & 0 & \mu_1 & \mu_2 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \end{bmatrix}\right\| \\
&= \left(-\frac{1}{2}\right)^4 \left|\boldsymbol{\Sigma}^{-1}\right| = \frac{1}{16}|\boldsymbol{\Sigma}|^{-1},
\end{aligned}$$

since $\mathbf{\Sigma}$ is symmetric. As a result, the distribution of $\mathbf{u}$ is given by

$$\pi(\boldsymbol{\mu}, \mathbf{\Sigma}^{-1}|\boldsymbol{\nu}, a) = \frac{1}{16}|\mathbf{\Sigma}|^{-1}\exp\{\boldsymbol{\mu}^T\mathbf{\Sigma}^{-1}\boldsymbol{\nu}_1 - \frac{1}{2}vec(\mathbf{\Sigma}^{-1})^T vec(\boldsymbol{\nu}_2)$$
$$-\frac{a}{2}\ln(|\mathbf{\Sigma}|) - \frac{a}{2}\boldsymbol{\mu}^T\mathbf{\Sigma}^{-1}\boldsymbol{\mu} + K(a, \boldsymbol{\nu})\}$$
$$= \frac{1}{16}|\mathbf{\Sigma}|^{-1-\frac{a}{2}}\exp\{-\frac{1}{2}\left[\boldsymbol{\mu}^T\left(\frac{1}{a}\mathbf{\Sigma}\right)^{-1}\boldsymbol{\mu} - \boldsymbol{\mu}^T\left(\frac{1}{a}\mathbf{\Sigma}\right)^{-1}\left(\frac{1}{a}\boldsymbol{\nu}_1\right)\right]$$
$$-\frac{1}{2}tr(\mathbf{\Sigma}^{-1}\boldsymbol{\nu}_2) + K(a, \boldsymbol{\nu})\}$$
$$= \frac{1}{16}|\mathbf{\Sigma}|^{-\frac{1}{2}-\frac{a+1}{2}}\exp\left\{-\frac{1}{2}\left[\left(\boldsymbol{\mu} - \frac{1}{a}\boldsymbol{\nu}_1\right)^T\left(\frac{1}{a}\mathbf{\Sigma}\right)^{-1}\left(\boldsymbol{\mu} - \frac{1}{a}\boldsymbol{\nu}_1\right)\right]\right\}$$
$$\exp\left\{-\frac{1}{2}\left[\left(\frac{1}{a}\boldsymbol{\nu}_1\right)^T\left(\frac{1}{a}\mathbf{\Sigma}\right)^{-1}\left(\frac{1}{a}\boldsymbol{\nu}_1\right)\right] - \frac{1}{2}tr(\mathbf{\Sigma}^{-1}\boldsymbol{\nu}_2) + K(a, \boldsymbol{\nu})\right\},$$

and therefore

$$\pi(\boldsymbol{\mu}, \mathbf{\Sigma}^{-1}|a, \boldsymbol{\nu}) = \pi(\boldsymbol{\mu}|\mathbf{\Sigma}, a, \boldsymbol{\nu}_1)\pi(\mathbf{\Sigma}^{-1}|a, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2),$$

is a joint distribution where clearly we have

$$\boldsymbol{\mu}|\mathbf{\Sigma}, a, \boldsymbol{\nu}_1 \sim N_2\left(\frac{1}{a}\boldsymbol{\nu}_1, \frac{1}{a}\mathbf{\Sigma}\right),$$

while the distribution of $\mathbf{\Sigma}^{-1}$ (up to a constant) can be written as

$$\begin{aligned}\pi(\mathbf{\Sigma}^{-1}|a, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2) &\propto |\mathbf{\Sigma}|^{-\frac{a+1}{2}}\exp\left\{-\frac{1}{2a}\left[\boldsymbol{\nu}_1^T\mathbf{\Sigma}^{-1}\boldsymbol{\nu}_1\right] - \frac{1}{2}tr(\mathbf{\Sigma}^{-1}\boldsymbol{\nu}_2)\right\}\\ &\propto \left|\mathbf{\Sigma}^{-1}\right|^{\frac{a+1}{2}}\exp\left\{-\frac{1}{2}tr\left\{\mathbf{\Sigma}^{-1}\left(\boldsymbol{\nu}_2 + \frac{1}{a}\boldsymbol{\nu}_1\boldsymbol{\nu}_1^T\right)\right\}\right\}\\ &\propto \left|\mathbf{\Sigma}^{-1}\right|^{\frac{a+4-2-1}{2}}\exp\left\{-\frac{1}{2}tr\left\{\left(\boldsymbol{\nu}_2 + \frac{1}{a}\boldsymbol{\nu}_1\boldsymbol{\nu}_1^T\right)(\mathbf{\Sigma}^{-1})\right\}\right\},\end{aligned}$$

which is clearly the density of a Wishart distribution, i.e.,

$$\mathbf{\Sigma}^{-1}|a, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2 \sim W_2\left(a + 4, \left(\boldsymbol{\nu}_2 + \frac{1}{a}\boldsymbol{\nu}_1\boldsymbol{\nu}_1^T\right)^{-1}\right).$$

■

Theorem 8 *For the RMCP model with normal components and under the assumption of conjugacy for the driving density, the means required in equations (25) and (26) of Theorem 6, are given by*

$$E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j, a_j}[\boldsymbol{\theta}_j] = \left(\frac{a_j + 4}{a_j}\boldsymbol{\nu}_{j1}^T\left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}, -\frac{a_j + 4}{2}vec\left(\left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right)^T\right)^T,$$

and

$$E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j, a_j}[b(\boldsymbol{\theta}_j)] = -\log(2) - \Psi\left(\frac{a_j + 3}{2}\right) - \Psi\left(\frac{a_j + 4}{2}\right)$$
$$-0.5\log\left(\left|\left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right|\right) + \frac{1}{2a_j}\left(2 + \frac{a_j + 4}{a_j}\boldsymbol{\nu}_{j1}^T\left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\boldsymbol{\nu}_{j1}\right),$$

for $j = 1, 2, \ldots m$, where $\Psi(x) = \frac{d}{dx}\log(\Gamma(x))$, the digamma function.

Proof. Since $\boldsymbol{\theta}_j = \left(\boldsymbol{\mu}_j^T\boldsymbol{\Sigma}_j^{-1}, -\frac{1}{2}vec(\boldsymbol{\Sigma}_j^{-1})^T\right)^T$, we have

$$\begin{aligned}
E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j,a_j}\left[\boldsymbol{\theta}_j\right] &= E^{\boldsymbol{\mu}_j,\boldsymbol{\Sigma}_j^{-1}|\boldsymbol{\nu}_j,a_j}\left[\left(\boldsymbol{\mu}_j^T\boldsymbol{\Sigma}_j^{-1}, -\frac{1}{2}vec(\boldsymbol{\Sigma}_j^{-1})^T\right)^T\right] \\
&= E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left[E^{\boldsymbol{\mu}_j|\boldsymbol{\Sigma}_j^{-1},a_j,\boldsymbol{\nu}_{j1}}\left(\boldsymbol{\mu}_j^T\boldsymbol{\Sigma}_j^{-1}, -\frac{1}{2}vec(\boldsymbol{\Sigma}_j^{-1})^T\right)^T\right] \\
&= \left(\frac{1}{a_j}\boldsymbol{\nu}_{j1}^T E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left(\boldsymbol{\Sigma}_j^{-1}\right), -\frac{1}{2}vec(E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left(\boldsymbol{\Sigma}_j^{-1}\right))^T\right)^T \\
&= \left(\frac{a_j+4}{a_j}\boldsymbol{\nu}_{j1}^T\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}, -\frac{a_j+4}{2}vec\left(\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right)^T\right)^T
\end{aligned}$$

owing to the fact that

$$E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left(\boldsymbol{\Sigma}_j^{-1}\right) = (a_j+4)\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}.$$

Similarly,

$$\begin{aligned}
E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j,a_j}\left[b(\boldsymbol{\theta}_j)\right] &= E^{\boldsymbol{\mu}_j,\boldsymbol{\Sigma}_j^{-1}|\boldsymbol{\nu}_j,a_j}\left[\frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_j|\right)+\frac{1}{2}\boldsymbol{\mu}_j^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_j\right] \\
&= E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left[E^{\boldsymbol{\mu}_j|\boldsymbol{\Sigma}_j^{-1},a_j,\boldsymbol{\nu}_{j1}}\left(-\frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_j^{-1}|\right)+\frac{1}{2}\boldsymbol{\mu}_j^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_j\right)\right] \\
&= E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left[-\frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_j^{-1}|\right)+\frac{1}{2}E^{\boldsymbol{\mu}_j|\boldsymbol{\Sigma}_j^{-1},a_j,\boldsymbol{\nu}_{j1}}\left(\boldsymbol{\mu}_j^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_j\right)\right],
\end{aligned}$$

and since

$$\begin{aligned}
\boldsymbol{\mu}_j|\boldsymbol{\Sigma}_j^{-1},a_j,\boldsymbol{\nu}_{j1} &\sim N_2\left(\frac{1}{a_j}\boldsymbol{\nu}_{j1},\frac{1}{a_j}\boldsymbol{\Sigma}_j\right)\Rightarrow \\
\mathbf{x} &= \sqrt{a_j}\boldsymbol{\Sigma}_j^{-\frac{1}{2}}\boldsymbol{\mu}_j|\boldsymbol{\Sigma}_j^{-1},a_j,\boldsymbol{\nu}_{j1}\sim N_2\left(\frac{\sqrt{a_j}}{a_j}\boldsymbol{\Sigma}_j^{-\frac{1}{2}}\boldsymbol{\nu}_{j1},\mathbf{I}_2\right)\Rightarrow \\
\mathbf{x}^T\mathbf{x} &= a_j\boldsymbol{\mu}_j^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_j\sim\chi_2^2\left(\frac{1}{a_j}\boldsymbol{\nu}_{j1}^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\nu}_{j1}\right),
\end{aligned}$$

we have

$$E^{\boldsymbol{\mu}_j|\boldsymbol{\Sigma}_j^{-1},a_j,\boldsymbol{\nu}_{j1}}\left(\boldsymbol{\mu}_j^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_j\right) = \frac{1}{a_j}\left(2+\frac{1}{a_j}\boldsymbol{\nu}_{j1}^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\nu}_{j1}\right),$$

so that

$$\begin{aligned}
E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j,a_j}\left[b(\boldsymbol{\theta}_j)\right] &= E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left[-\frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_j^{-1}|\right)+\frac{1}{2a_j}\left(2+\frac{1}{a_j}\boldsymbol{\nu}_{j1}^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\nu}_{j1}\right)\right] \\
&= -\frac{1}{2}E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left(\ln\left(|\boldsymbol{\Sigma}_j^{-1}|\right)\right)+\frac{1}{2a_j}\left(2+\frac{1}{a_j}E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left[\boldsymbol{\nu}_{j1}^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\nu}_{j1}\right]\right).
\end{aligned}$$

Noting that if $\mathbf{A}\sim W_m\left(n,\boldsymbol{\Sigma}\right)$ then $\frac{\mathbf{y}^T\mathbf{A}\mathbf{y}}{\mathbf{y}^T\boldsymbol{\Sigma}\mathbf{y}}\sim\chi_n^2$, where $\mathbf{y}$ denotes any random vector independent of $\mathbf{A}$ (see Muirhead, 1982, Theorem 3.2.8), we have

$$E^{\boldsymbol{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left[\boldsymbol{\nu}_{j1}^T\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\nu}_{j1}\right] = (a_j+4)\,\boldsymbol{\nu}_{j1}^T\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\boldsymbol{\nu}_{j1}.$$

Moreover, if $\mathbf{A} \sim W_m(n, \mathbf{\Sigma})$, $n \geq m$, then $|\mathbf{A}| / |\mathbf{\Sigma}|$ has the same distribution as $\prod_{i=1}^{m} \chi^2_{n-i+1}$, where χ^2_{n-i+1} denote independent χ^2 random variable (see Muirhead, 1982, Theorem 3.2.15), and we can write

$$\begin{aligned} E[\log|\mathbf{A}|] &= E\left[\log\left(\frac{|\mathbf{A}|}{|\mathbf{\Sigma}|}|\mathbf{\Sigma}|\right)\right] = E\left[\log\left(\frac{|\mathbf{A}|}{|\mathbf{\Sigma}|}\right)\right] + \log(|\mathbf{\Sigma}|) \\ &= E\left[\log\left(\prod_{i=1}^{m}\chi^2_{n-i+1}\right)\right] + \log(|\mathbf{\Sigma}|) = \sum_{i=1}^{m} E\left[\log\left(\chi^2_{n-i+1}\right)\right] + \log(|\mathbf{\Sigma}|) \\ &= \sum_{i=1}^{m} E\left[\log\left(\chi^2_{n-i+1}\right)\right] + \log(|\mathbf{\Sigma}|), \end{aligned}$$

so that for $\mathbf{\Sigma}_j^{-1}|a_j, \boldsymbol{\nu}_{j1}, \boldsymbol{\nu}_{j2} \sim W_2\left(a_j+4, \left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right)$, we have

$$\begin{aligned} E^{\mathbf{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left(\ln\left(|\mathbf{\Sigma}_j^{-1}|\right)\right) &= E\left[\log\left(\chi^2_{a_j+4-1+1}\right)\right] + E\left[\log\left(\chi^2_{a_j+4-2+1}\right)\right] \\ &\quad + \log\left(\left|\left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right|\right) \\ &= E\left[\log\left(\chi^2_{a_j+4}\right)\right] + E\left[\log\left(\chi^2_{a_j+3}\right)\right] \\ &\quad + \log\left(\left|\left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right|\right). \end{aligned}$$

Now let $Y = log(X) \Rightarrow X = e^Y$, where $X \sim \chi^2_a$, so that

$$\begin{aligned} f_Y(y) &= f_X(e^y)e^y = \frac{e^{\left(\frac{a}{2}-1\right)y - e^y/2}}{\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}}e^y \\ &= \frac{e^{\frac{a}{2}y - e^y/2}}{\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}},\ y \in \mathcal{R}, \end{aligned}$$

where

$$\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}} = \int_{-\infty}^{+\infty} e^{\frac{a}{2}y - e^y/2}dy.$$

As a result, we have

$$\begin{aligned} E[\log(X)] &= E(Y) = \frac{1}{\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}}\int_{-\infty}^{+\infty} y e^{\frac{a}{2}y - e^y/2}dy = \frac{1}{\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}}\int_{-\infty}^{+\infty} 2\frac{d}{da}\left[e^{\frac{a}{2}y - e^y/2}\right]dy \\ &= \frac{2}{\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}}\frac{d}{da}\left[\int_{-\infty}^{+\infty} e^{\frac{a}{2}y - e^y/2}dy\right] = \frac{2}{\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}}\frac{d}{da}\left[\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}\right] \\ &= \frac{2}{\Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}}\left[2^{\frac{a}{2}}\frac{1}{2}\frac{d}{da}\Gamma\left(\frac{a}{2}\right) + \Gamma\left(\frac{a}{2}\right)2^{\frac{a}{2}}\frac{1}{2}\log(2)\right] \\ &= \frac{d}{da}\log\Gamma\left(\frac{a}{2}\right) + \log(2), \end{aligned}$$

so that

$$\begin{aligned} E^{\mathbf{\Sigma}_j^{-1}|a_j,\boldsymbol{\nu}_{j1},\boldsymbol{\nu}_{j2}}\left(\ln\left(|\mathbf{\Sigma}_j^{-1}|\right)\right) &= \frac{d}{da_j}\log\Gamma\left(\frac{a_j+3}{2}\right) + 2\log(2) \\ \frac{d}{da_j}\log\Gamma\left(\frac{a_j+4}{2}\right) &+ \log\left(\left|\left(\boldsymbol{\nu}_{j2} + \frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right|\right). \end{aligned}$$

Finally, we can write

$$
\begin{aligned}
E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j,a_j}\left[b(\boldsymbol{\theta}_j)\right] &= -\log(2)-0.5\frac{d}{da_j}\log\Gamma\left(\frac{a_j+3}{2}\right)-0.5\frac{d}{da_j}\log\Gamma\left(\frac{a_j+4}{2}\right)\\
&\quad -0.5\log\left(\left|\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right|\right)\\
&\quad +\frac{1}{2a_j}\left(2+\frac{a_j+4}{a_j}\boldsymbol{\nu}_{j1}^T\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\boldsymbol{\nu}_{j1}\right),
\end{aligned}
$$

which completes the proof. ■

Theorem 9 *For the RMCP model with normal components and under the assumption of conjugacy for the driving density, the solutions to equations (25) and (26) of Theorem 6, yield the MLEs of* $\boldsymbol{\nu}_j = (\boldsymbol{\nu}_{j1}^T, vec(\boldsymbol{\nu}_{j2})^T)^T$ *and* a_j*, are given by*

$$
E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j,a_j}\left[\boldsymbol{\theta}_j\right]=\left(\frac{a_j+4}{a_j}\boldsymbol{\nu}_{j1}^T\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1},-\frac{a_j+4}{2}vec\left(\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right)^T\right)^T,
$$

and

$$
E^{\boldsymbol{\theta}_j|\boldsymbol{\nu}_j,a_j}\left[b(\boldsymbol{\theta}_j)\right]=-\log(2)-\Psi\left(\frac{a_j+3}{2}\right)-\Psi\left(\frac{a_j+4}{2}\right)
$$

$$
-0.5\log\left(\left|\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\right|\right)+\frac{1}{2a_j}\left(2+\frac{a_j+4}{a_j}\boldsymbol{\nu}_{j1}^T\left(\boldsymbol{\nu}_{j2}+\frac{1}{a_j}\boldsymbol{\nu}_{j1}\boldsymbol{\nu}_{j1}^T\right)^{-1}\boldsymbol{\nu}_{j1}\right),
$$

Proof. Without loss of generality and for better exposition, we drop the mixture component index j.

$$
E^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j,\widehat{a}_j}\left[b(\boldsymbol{\theta}_j)\right]=\frac{1}{K}\sum_{k=1}^{K}E^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j+\mathbf{u}_j^{(k)},\widehat{a}_j+z_{\bullet j}^{(k)}}\left[b(\boldsymbol{\theta}_j)\right], \tag{30}
$$

and

$$
E^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j,\widehat{a}_j}\left[\boldsymbol{\theta}_j\right]=\frac{1}{K}\sum_{k=1}^{K}E^{\boldsymbol{\theta}_j|\widehat{\boldsymbol{\nu}}_j+\mathbf{u}_j^{(k)},\widehat{a}_j+z_{\bullet j}^{(k)}}\left[\boldsymbol{\theta}_j\right]. \tag{31}
$$

Recall that $\mathbf{u}^{(k)}=\sum_{i=1}^{n}z_i^{(k)}\mathbf{u}\left(\mathbf{s}_i^{(k)}\right)$, and using equation (26) we have two equations

$$
\begin{aligned}
K\frac{a+4}{a}\boldsymbol{\nu}_1^T\left(\boldsymbol{\nu}_2+\frac{1}{a}\boldsymbol{\nu}_1\boldsymbol{\nu}_1^T\right)^{-1} &= \sum_{k=1}^{K}\frac{a+z_\bullet^{(k)}+4}{a+z_\bullet^{(k)}}\left(\boldsymbol{\nu}_1+\mathbf{u}_1^{(k)}\right)^T\\
&\quad \left(\boldsymbol{\nu}_2+\mathbf{u}_2^{(k)}+\frac{1}{a+z_\bullet^{(k)}}\left(\boldsymbol{\nu}_1+\mathbf{u}_1^{(k)}\right)\left(\boldsymbol{\nu}_1+\mathbf{u}_1^{(k)}\right)^T\right)^{-1},
\end{aligned} \tag{32}
$$

and

$$
\begin{aligned}
-K\frac{a+4}{2}vec\left(\left(\boldsymbol{\nu}_2+\frac{1}{a}\boldsymbol{\nu}_1\boldsymbol{\nu}_1^T\right)^{-1}\right)^T &= -\sum_{k=1}^{K}\frac{a+z_\bullet^{(k)}+4}{2}\\
&\quad vec\left(\left(\boldsymbol{\nu}_2+\mathbf{u}_2^{(k)}+\frac{1}{a+z_\bullet^{(k)}}\left(\boldsymbol{\nu}_1+\mathbf{u}_1^{(k)}\right)\left(\boldsymbol{\nu}_1+\mathbf{u}_1^{(k)}\right)^T\right)^{-1}\right)\Rightarrow\\
K\left(a+4\right)\left(\boldsymbol{\nu}_2+\frac{1}{a}\boldsymbol{\nu}_1\boldsymbol{\nu}_1^T\right)^{-1} &= \sum_{k=1}^{K}\left(a+z_\bullet^{(k)}+4\right)\\
&\quad \left(\boldsymbol{\nu}_2+\mathbf{u}_2^{(k)}+\frac{1}{a+z_\bullet^{(k)}}\left(\boldsymbol{\nu}_1+\mathbf{u}_1^{(k)}\right)\left(\boldsymbol{\nu}_1+\mathbf{u}_1^{(k)}\right)^T\right)^{-1},
\end{aligned}
$$

so that (32) yields

$$\frac{1}{a}\boldsymbol{\nu}_1^T \sum_{k=1}^{K} \left(a + z_{\bullet}^{(k)} + 4\right) \left(\boldsymbol{\nu}_2 + \mathbf{u}_2^{(k)} + \frac{1}{a + z_{\bullet}^{(k)}} \left(\boldsymbol{\nu}_1 + \mathbf{u}_1^{(k)}\right) \left(\boldsymbol{\nu}_1 + \mathbf{u}_1^{(k)}\right)^T\right)^{-1} =$$

$$\sum_{k=1}^{K} \frac{a + z_{\bullet}^{(k)} + 4}{a + z_{\bullet}^{(k)}} \left(\boldsymbol{\nu}_1 + \mathbf{u}_1^{(k)}\right)^T \left(\boldsymbol{\nu}_2 + \mathbf{u}_2^{(k)} + \frac{1}{a + z_{\bullet}^{(k)}} \left(\boldsymbol{\nu}_1 + \mathbf{u}_1^{(k)}\right) \left(\boldsymbol{\nu}_1 + \mathbf{u}_1^{(k)}\right)^T\right)^{-1} \Rightarrow$$

$$\sum_{k=1}^{K} \left(\frac{a + z_{\bullet}^{(k)} + 4}{a}\boldsymbol{\nu}_1 - \frac{a + z_{\bullet}^{(k)} + 4}{a + z_{\bullet}^{(k)}} \left(\boldsymbol{\nu}_1 + \mathbf{u}_1^{(k)}\right)\right)^T$$

$$\left(\boldsymbol{\nu}_2 + \mathbf{u}_2^{(k)} + \frac{1}{a + z_{\bullet}^{(k)}} \left(\boldsymbol{\nu}_1 + \mathbf{u}_1^{(k)}\right) \left(\boldsymbol{\nu}_1 + \mathbf{u}_1^{(k)}\right)^T\right)^{-1} = 0$$

■