Computational Appendix to "Count Roy Model with Finite Mixtures" by Murat K. Munkin.

A1. MCMC Algorithm.

The set of parameters is divided into blocks $\mathbf{R}_i^1, \mathbf{R}_i^2$, $\mathbf{z}_i^1$, $\mathbf{z}_i^2$, $\boldsymbol{\gamma}^1$, $\boldsymbol{\gamma}^2$, $\boldsymbol{\mu}_i^1$, $\boldsymbol{\mu}_i^2$, D_i, $[\boldsymbol{\psi}, \boldsymbol{\alpha}]$, $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, $\boldsymbol{\sigma}$ and η.

1. Draw R_{ij}^1 $(j = 2, ..., k_1)$ for each observation i in the treated state independently from $N\left(\mathbf{V}_i\boldsymbol{\gamma}_{1j}, 1\right)$.
2. Draw R_{ij}^2 $(j = 2, ..., k_2)$ for each observation i in the untreated state independently from $N\left(\mathbf{V}_i\boldsymbol{\gamma}_{2j}, 1\right)$.
3. Draw $\mathbf{z}_i^1 = (z_{i1}^1, z_{i2}^1, ...z_{ik_1}^1)$ for all observations in the treated state to assign them to a component. Assume that the current component for observation i is j $\left(z_{ij}^1 = 1\right)$. Identify that component j^* for which random variable $R_{ij^*}^1$ takes the largest value
$$R_{ij^*}^1 \geqslant R_{il}^1 \text{ for } \forall l,\ l = 1, ..., k_1.$$
Assign observation i to component j^* with probability
$$\min\left\{\frac{\exp\left(-\exp\left(\mu_{ij^*}^1\right)\right)\exp\left(\mu_{ij^*}^1 Y_i\right)}{\exp\left(-\exp\left(\mu_{ij}^1\right)\right)\exp\left(\mu_{ij}^1 Y_i\right)}, 1\right\}.$$
If the proposal component j^* is accepted then $z_{ij^*}^1 = 1$. Otherwise $z_{ij}^1 = 1$.
4. Draw $\mathbf{z}_i^2 = (z_{i1}^2, z_{i2}^2, ...z_{ik_2}^2)$ for all observations in the untreated state to assign them to a component. Assume that the current component for observation i is j $\left(z_{ij}^2 = 1\right)$. Identify that component j^* for which random variable $R_{ij^*}^2$ takes the largest value
$$R_{ij^*}^2 \geqslant R_{il}^2 \text{ for } \forall l,\ l = 1, ..., k_2.$$
Assign observation i to component j^* with probability
$$\min\left\{\frac{\exp\left(-\exp\left(\mu_{ij^*}^2\right)\right)\exp\left(\mu_{ij^*}^2 Y_i\right)}{\exp\left(-\exp\left(\mu_{ij}^2\right)\right)\exp\left(\mu_{ij}^2 Y_i\right)}, 1\right\}.$$
If the proposal component j^* is accepted then $z_{ij^*}^2 = 1$. Otherwise $z_{ij}^2 = 1$.
5. Draw $\boldsymbol{\gamma}_{1j}$ $(j = 2, ..., k_1)$ from the full conditional density $\boldsymbol{\gamma}_{1j} \sim N\left(\overline{\boldsymbol{\gamma}}_{1j}, \overline{\mathbf{H}}_{\gamma_{1j}}^{-1}\right)$ where
$$\begin{aligned}\overline{\mathbf{H}}_{\gamma_{1j}} &= \underline{\mathbf{H}}_{\gamma} + \sum_{\substack{i=1\\ s.t.\ d_i=1}}^{N} \mathbf{V}_i'\mathbf{V}_i \\ \overline{\boldsymbol{\gamma}}_{1j} &= \overline{\mathbf{H}}_{\gamma_{1j}}^{-1}\left[\underline{\mathbf{H}}_{\gamma}\underline{\boldsymbol{\gamma}} + \sum_{\substack{i=1\\ s.t.\ d_i=1}}^{N} \mathbf{V}_i' R_{ij}^1\right]\end{aligned}$$

6. Draw $\boldsymbol{\gamma}_{2j}$ $(j = 2, ..., k_2)$ from the full conditional density $\boldsymbol{\gamma}_{2j} \sim N\left(\overline{\boldsymbol{\gamma}}_{2j}, \overline{\mathbf{H}}_{\gamma_{2j}}^{-1}\right)$ where

$$\begin{aligned}
\overline{\mathbf{H}}_{\gamma_{2j}} &= \underline{\mathbf{H}}_{\gamma} + \sum_{\substack{i=1 \\ s.t.\ d_i=0}}^{N} \mathbf{V}_i' \mathbf{V}_i \\
\overline{\boldsymbol{\gamma}}_{2j} &= \overline{\mathbf{H}}_{\gamma_{2j}}^{-1} \left[\underline{\mathbf{H}}_{\gamma} \underline{\boldsymbol{\gamma}} + \sum_{\substack{i=1 \\ s.t.\ d_i=0}}^{N} \mathbf{V}_i' R_{ij}^2 \right]
\end{aligned}$$

7. Draw μ_{ij}^1 for observation i in the treated state. If $z_{ij}^1 = 0$, then μ_{ij}^1 is drawn from

$$N\left(\mathbf{X}_i \boldsymbol{\beta}_{1j} + \delta_{1j}\left(D_i - \mathbf{P}_i \mathbf{R}^{-1} \boldsymbol{\psi} - \mathbf{W}_i \boldsymbol{\alpha}\right), \sigma_{1j}^2\right).$$

However, if observation i belongs to component j $\left(z_{ij}^1 = 1\right)$ then the full conditional density of μ_{ij}^1 is proportional to

$$\begin{aligned}
\pi\left(\mu_{ij}^1 | Y_i, D_i, \boldsymbol{\Delta}_i\right) &= \exp\left(-\exp\left(\mu_{ij}^1\right)\right) \exp\left(\mu_{ij}^1 Y_i\right) \\
&\quad \times \exp\left(-0.5\sigma_{1j}^{-2}\left(\mu_{ij}^1 - \mathbf{X}_i \boldsymbol{\beta}_{1j} - \delta_{1j}\left(D_i - \mathbf{P}_i \mathbf{R}^{-1} \boldsymbol{\psi} - \mathbf{W}_i \boldsymbol{\alpha}\right)\right)^2\right)
\end{aligned}$$

To sample μ_{ij}^1 utilize the Metropolis-Hasting algorithm using a normal distribution centered at the modal value of the full conditional density as the proposal density. Let

$$\widehat{\mu}_{ij}^1 = \arg\max \log \pi\left(\mu_{ij}^1 | Y_i, D_i, \boldsymbol{\Delta}_i\right)$$

and $V_{\widehat{\mu}_{ij}^1} = -(H_{\widehat{\mu}_{ij}^1})^{-1}$ be the negative inverse of the Hessian of $\log \pi\left(\mu_{ij}^1 | Y_i, D_i, \boldsymbol{\Delta}_i\right)$ evaluated at the mode $\widehat{\mu}_{ij}^1$. Choose the normal proposal distribution $q\left(\mu_{ij}^1\right) = f(\mu_{ij}^1 | \widehat{\mu}_{ij}^1, \tau V_{\widehat{\mu}_{ij}^1})$, where τ is a tuning parameter selected to have desired acceptance rates. When μ_{ij}^{1*} is drawn the chain moves from the current value μ_{ij}^1 to the proposal value with probability

$$\alpha(\mu_{ij}^1, \mu_{ij}^{1*}) = \min\left\{\frac{\pi\left(\mu_{ij}^{1*} | Y_i, D_i, \boldsymbol{\Delta}_i\right) q\left(\mu_{ij}^1\right)}{\pi\left(\mu_{ij}^1 | Y_i, D_i, \boldsymbol{\Delta}_i\right) q\left(\mu_{ij}^{1*}\right)}, 1\right\}.$$

If the proposal value is rejected then the next state of the chain is the current value μ_{ij}^1.

8. Draw μ_{ij}^2 for observation i in the untreated state. If $z_{ij}^2 = 0$, then μ_{ij}^2 is drawn from

$$N\left(\mathbf{X}_i \boldsymbol{\beta}_{2j} + \delta_{2j}\left(D_i - \mathbf{P}_i \mathbf{R}^{-1} \boldsymbol{\psi} - \mathbf{W}_i \boldsymbol{\alpha}\right), \sigma_{2j}^2\right).$$

However, if observation i belongs to component j $\left(z_{ij}^2 = 1\right)$ then the full conditional density of μ_{ij}^2 is proportional to

$$\begin{aligned}
\pi\left(\mu_{ij}^2 | Y_i, D_i, \boldsymbol{\Delta}_i\right) &= \exp\left(-\exp\left(\mu_{ij}^2\right)\right) \exp\left(\mu_{ij}^2 Y_i\right) \\
&\quad \times \exp\left(-0.5\sigma_{2j}^{-2}\left(\mu_{ij}^2 - \mathbf{X}_i \boldsymbol{\beta}_{2j} - \delta_{2j}\left(D_i - \mathbf{P}_i \mathbf{R}^{-1} \boldsymbol{\psi} - \mathbf{W}_i \boldsymbol{\alpha}\right)\right)^2\right)
\end{aligned}$$

To sample μ_{ij}^2 use a normal distribution centered at the modal value of the full conditional density as the proposal density. Let

$$\widehat{\mu}_{ij}^2 = \arg\max \log \pi\left(\mu_{ij}^2|Y_i, D_i, \mathbf{\Delta}_i\right)$$

and $V_{\widehat{\mu}_{ij}^2} = -(H_{\widehat{\mu}_{ij}^2})^{-1}$ be the negative inverse of the Hessian of $\log \pi\left(\mu_{ij}^2|Y_i, D_i, \mathbf{\Delta}_i\right)$ evaluated at the mode $\widehat{\mu}_{ij}^2$. Choose the normal proposal distribution $q\left(\mu_{ij}^2\right) = f(\mu_{ij}^2|\widehat{\mu}_{ij}^2, \tau V_{\widehat{\mu}_{ij}^2})$ with tuning parameter τ. When μ_{ij}^{2*} is drawn the chain moves from the current value μ_{ij}^2 to the proposal value with probability

$$\alpha(\mu_{ij}^2, \mu_{ij}^{2*}) = \min\left\{\frac{\pi\left(\mu_{ij}^{2*}|Y_i, D_i, \mathbf{\Delta}_i\right) q\left(\mu_{ij}^2\right)}{\pi\left(\mu_{ij}^2|Y_i, D_i, \mathbf{\Delta}_i\right) q\left(\mu_{ij}^{2*}\right)}, 1\right\}.$$

If the proposal value is rejected then the chain stays at the current value μ_{ij}^2.

9. Draw latent vectors D_i $(i = 1, ...N)$ from independent truncated normal distributions $D_i \overset{iid}{\sim} N\left(\overline{D}_i, \overline{H}_i^{-1}\right)$ where

$$\begin{aligned}
\overline{H}_i &= 1 + d_i \sum_{j=1}^{k_1} \delta_{1j}^2 \sigma_{1j}^{-2} + (1-d_i) \sum_{j=1}^{k_2} \delta_{2j}^2 \sigma_{2j}^{-2}, \\
\overline{D}_i &= \mathbf{P}_i \mathbf{R}^{-1} \boldsymbol{\psi} + \mathbf{W}_i \boldsymbol{\alpha} + \overline{H}_i^{-1} \left[d_i \sum_{j=1}^{k_1} \delta_{1j} \sigma_{1j}^{-2} \left(\mu_{ij}^1 - \mathbf{X}_i \boldsymbol{\beta}_{1j}\right) + (1-d_i) \sum_{j=1}^{k_2} \delta_{2j} \sigma_{2j}^{-2} \left(\mu_{ij}^2 - \mathbf{X}_i \boldsymbol{\beta}_{2j}\right) \right],
\end{aligned}$$

and each variable is truncated such that

$$D_i \geqslant 0 \text{ if } d_i = 1 \text{ and } D_i < 0 \text{ if } d_i = 0.$$

10. Define $\boldsymbol{\varphi}' = \left(\boldsymbol{\psi}', \boldsymbol{\alpha}\right)$ and denote $\mathbf{G}_i = \left(\mathbf{P}_i \mathbf{R}^{-1}, \mathbf{W}_i\right)$. The prior distribution $N\left(\underline{\boldsymbol{\varphi}}, \underline{\mathbf{H}}_{\boldsymbol{\varphi}}^{-1}\right)$ is formed from the prior distributions $N\left(\underline{\boldsymbol{\psi}}, \underline{\mathbf{H}}_{\psi}^{-1}\right)$ and $N\left(\underline{\boldsymbol{\alpha}}, \underline{\mathbf{H}}_{\alpha}^{-1}\right)$. Then draw $\boldsymbol{\varphi}$ from $N\left(\overline{\boldsymbol{\varphi}}, \overline{\mathbf{H}}_{\boldsymbol{\varphi}}^{-1}\right)$ where

$$\begin{aligned}
\overline{\mathbf{H}}_{\boldsymbol{\varphi}} &= \underline{\mathbf{H}}_{\boldsymbol{\varphi}} + \sum_{i=1}^{N} \mathbf{G}_i' \left(1 + d_i \sum_{j=1}^{k_1} \delta_{1j}^2 \sigma_{1j}^{-2} + (1-d_i) \sum_{j=1}^{k_2} \delta_{2j}^2 \sigma_{2j}^{-2}\right) \mathbf{G}_i \\
\overline{\boldsymbol{\varphi}} &= \overline{\mathbf{H}}_{\boldsymbol{\varphi}}^{-1} \left[\underline{\mathbf{H}}_{\boldsymbol{\varphi}} \underline{\boldsymbol{\varphi}} + \sum_{i=1}^{N} \mathbf{G}_i' \left(1 + d_i \sum_{j=1}^{k_1} \delta_{1j}^2 \sigma_{1j}^{-2} + (1-d_i) \sum_{j=1}^{k_2} \delta_{2j}^2 \sigma_{2j}^{-2}\right) D_i \right. \\
&\quad \left. - \sum_{i=1}^{N} \mathbf{G}_i' d_i \sum_{j=1}^{k_1} \delta_{1j} \sigma_{1j}^{-2} \left(\mu_{ij}^1 - \mathbf{X}_i \boldsymbol{\beta}_{1j}\right) - \sum_{i=1}^{N} \mathbf{G}_i' (1-d_i) \sum_{j=1}^{k_2} \delta_{2j} \sigma_{2j}^{-2} \left(\mu_{ij}^2 - \mathbf{X}_i \boldsymbol{\beta}_{2j}\right) \right].
\end{aligned}$$

11. Draw $\boldsymbol{\phi}_{1j}$ $(j = 1, ..., k_1)$ from the full conditional distribution $\boldsymbol{\phi}_{1j} \sim N\left(\overline{\boldsymbol{\phi}}_{1j}, \overline{\mathbf{H}}_{\phi_{1j}}^{-1}\right)$ in which the prior $\pi(\boldsymbol{\phi}_j)\tilde{\ } N\left(\underline{\boldsymbol{\phi}}, \underline{\mathbf{H}}_{\phi}^{-1}\right)$ is formed from the prior distributions $N\left(\underline{\boldsymbol{\beta}}, \underline{\mathbf{H}}_{\beta}^{-1}\right)$ and $N\left(\underline{\delta}, \underline{H}_{\delta}^{-1}\right)$ and $\mathbf{C}_i = (\mathbf{X}_i, D_i - \mathbf{P}_i\mathbf{R}^{-1}\boldsymbol{\psi} - \mathbf{W}_i\boldsymbol{\alpha})$ such that

$$\begin{aligned} \overline{\mathbf{H}}_{\phi_{1j}} &= \underline{\mathbf{H}}_{\phi} + \sum_{\substack{i=1 \\ s.t.\ d_i=1}}^{N} \mathbf{C}_i' \sigma_{1j}^{-2} \mathbf{C}_i \\ \overline{\boldsymbol{\phi}}_{1j} &= \overline{\mathbf{H}}_{\phi_{1j}}^{-1} \left[\underline{\mathbf{H}}_{\phi}\underline{\boldsymbol{\phi}} + \sum_{\substack{i=1 \\ s.t.\ d_i=1}}^{N} \mathbf{C}_i' \sigma_{1j}^{-2} \mu_{ij}^1 \right] \end{aligned}$$

12. Draw $\boldsymbol{\phi}_{2j}$ $(j = 1, ..., k_2)$ from the full conditional distribution $\boldsymbol{\phi}_{2j} \sim N\left(\overline{\boldsymbol{\phi}}_{2j}, \overline{\mathbf{H}}_{\phi_{2j}}^{-1}\right)$ where

$$\begin{aligned} \overline{\mathbf{H}}_{\phi_{2j}} &= \underline{\mathbf{H}}_{\phi} + \sum_{\substack{i=1 \\ s.t.\ d_i=0}}^{N} \mathbf{C}_i' \sigma_{2j}^{-2} \mathbf{C}_i \\ \overline{\boldsymbol{\phi}}_{2j} &= \overline{\mathbf{H}}_{\phi_{2j}}^{-1} \left[\underline{\mathbf{H}}_{\phi}\underline{\boldsymbol{\phi}} + \sum_{\substack{i=1 \\ s.t.\ d_i=0}}^{N} \mathbf{C}_i' \sigma_{2j}^{-2} \mu_{ij}^2 \right] \end{aligned}$$

13. The full conditional of σ_{1j}^{-2} $(j = 1, ..., k_1)$, given the prior $\sigma_{1j}^{-2} \sim G\left(n/2, (g/2)^{-1}\right)$, is

$$G\left(\frac{n + \sum_{i=1}^{N} d_i}{2}, \left[\frac{g}{2} + \sum_{i=1}^{N} d_i \frac{\left(\mu_{ij}^1 - \mathbf{X}_i\boldsymbol{\beta}_{1j} - \delta_{1j}\left(D_i - \mathbf{P}_i\mathbf{R}^{-1}\boldsymbol{\psi} - \mathbf{W}_i\boldsymbol{\alpha}\right)\right)^2}{2} \right]^{-1} \right).$$

14. The full conditional of σ_{2j}^{-2} $(j = 1, ..., k_2)$, given the prior $\sigma_{2j}^{-2} \sim G\left(n/2, (g/2)^{-1}\right)$, is

$$G\left(\frac{n + \sum_{i=1}^{N} (1 - d_i)}{2}, \left[\frac{g}{2} + \sum_{i=1}^{N} (1 - d_i) \frac{\left(\mu_{ij}^2 - \mathbf{X}_i\boldsymbol{\beta}_{2j} - \delta_{2j}\left(D_i - \mathbf{P}_i\mathbf{R}^{-1}\boldsymbol{\psi} - \mathbf{W}_i\boldsymbol{\alpha}\right)\right)^2}{2} \right]^{-1} \right).$$

15. Finally, given the prior $\eta \sim IG(a, b)$, the full conditional for η is

$$\eta \sim IG\left(\frac{k_\gamma - 2}{2} + a, \left(b^{-1} + \frac{1}{2}\sum_{i=3}^{k_\gamma} \psi_j^2 \right)^{-1} \right).$$

A2. Marginal Likelihood.

Estimate the logarithm of the marginal likelihood

$$\log m(y) = \log l(y|\boldsymbol{\theta}^*) + \log \pi(\boldsymbol{\theta}^*) - \log \pi(\boldsymbol{\theta}^*|y),$$

evaluated at the posterior mean $\boldsymbol{\theta}^*$. The prior term $\log \pi(\boldsymbol{\theta}^*)$ is simply evaluated at $\boldsymbol{\theta}^*$, however, the log-likelihood $\log l(y|\boldsymbol{\theta}^*)$ and the logarithm of the posterior ordinate $\log \widehat{\pi}(\boldsymbol{\theta}^*|y)$ need to be estimated.

Log-likelihood

The log-likelihood function $\log l(y|\boldsymbol{\theta})$ is defined as

$$\log l(y|\boldsymbol{\theta}) = \sum_{i=1}^{N} \log l(Y_i, d_i|\boldsymbol{\theta}),$$

however, $l(Y_i, d_i|\boldsymbol{\theta})$ does not have a closed form. Given the latent variables the conditional likelihood

$$l(Y_i, d_i|D_i, \boldsymbol{\mu}_i^1, \boldsymbol{\mu}_i^2, \mathbf{R}_i^1, \mathbf{R}_i^2, \mathbf{z}_i^1, \mathbf{z}_i^2, \boldsymbol{\theta})$$

has a closed form. Integrating out latent variables R_{ij}^1 and R_{ij}^2 leads to $\Pr\left(z_{ij}^1\right)$ and $\Pr\left(z_{ij}^2\right)$ expressed as $(k_1 - 1)$ and $(k_2 - 1)$ variate integrals respectively. Since the models under consideration satisfy $k_1 \leqslant 3$ and $k_2 \leqslant 3$ conditions these integrals can be calculated numerically, for example, in Matlab with function mvncdf. Latent variable D_i can be integrated out analytically given the full conditional density $D_i \overset{iid}{\sim} N\left(\overline{D}_i, \overline{H}_i^{-1}\right)$ derived in step 9 of the MCMC algorithm. However, μ_{ij}^1 and μ_{ij}^2 can not be integrated out analytically. Then $l(Y_i, d_i|\boldsymbol{\theta})$ could be approximated numerically based on

$$l(Y_i, d_i|\boldsymbol{\theta}) = \int l(Y_i, d_i|\boldsymbol{\mu}_i^1, \boldsymbol{\mu}_i^2, \boldsymbol{\theta}) f(\boldsymbol{\mu}_i^1, \boldsymbol{\mu}_i^2|\boldsymbol{\theta}) d\boldsymbol{\mu}_i^1 d\boldsymbol{\mu}_i^2,$$

where $f(\boldsymbol{\mu}_i^1, \boldsymbol{\mu}_i^2|\boldsymbol{\theta}) = f(\boldsymbol{\mu}_i^1|\boldsymbol{\theta}) f(\boldsymbol{\mu}_i^2|\boldsymbol{\theta})$ and it can be shown that $f(\mu_{ij}^1) \sim N\left(\mathbf{X}_i\boldsymbol{\beta}_{1j}, \delta_{1j}^2 + \sigma_{1j}^2\right)$ $(j = 1, ..., k_1)$ and $f(\mu_{ij}^2) \sim N\left(\mathbf{X}_i\boldsymbol{\beta}_{2j}, \delta_{2j}^2 + \sigma_{2j}^2\right)$ $(j = 1, ..., k_2)$ respectively. These densities can be used as importance sampling functions, to approximate

$$\widehat{l}(Y_i, d_i|\boldsymbol{\theta}) = \frac{1}{S}\sum_{s=1}^{S} l(Y_i, d_i|\boldsymbol{\mu}_i^{1s}, \boldsymbol{\mu}_i^{2s}, \boldsymbol{\theta}),$$

where $\boldsymbol{\mu}_i^{1s}$and $\boldsymbol{\mu}_i^{2s}$ $(s = 1, ..., S)$ are drawn from $f(\boldsymbol{\mu}_i^1|\boldsymbol{\theta})$ and $f(\boldsymbol{\mu}_i^2|\boldsymbol{\theta})$ respectively.

Posterior Ordinates

The posterior ordinate term $\log \widehat{\pi}(\boldsymbol{\theta}^*|y)$ is estimated as

$$\log \widehat{\pi}(\boldsymbol{\sigma}^*|y) + \log \widehat{\pi}(\boldsymbol{\phi}^*|y, \boldsymbol{\sigma}^*) + \log \widehat{\pi}(\boldsymbol{\gamma}^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*) + \log \widehat{\pi}(\eta^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*, \boldsymbol{\gamma}^*) + \log \widehat{\pi}(\boldsymbol{\psi}^*, \boldsymbol{\alpha}^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*, \boldsymbol{\gamma}^*, \eta^*).$$

Estimate $\widehat{\pi}(\boldsymbol{\sigma}^*|y)$, $\widehat{\pi}(\boldsymbol{\phi}^*|y, \boldsymbol{\sigma}^*)$, $\widehat{\pi}(\boldsymbol{\gamma}^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*)$, $\widehat{\pi}(\eta^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*, \boldsymbol{\gamma}^*)$ and $\widehat{\pi}(\boldsymbol{\psi}^*, \boldsymbol{\alpha}^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*, \boldsymbol{\gamma}^*, \eta^*)$ through numerical approximation.

1. First, use output from the MCMC algorithm for the latent data and all parameters. Given the full conditional density of $\boldsymbol{\sigma}$, the posterior ordinate $\widehat{\pi}(\boldsymbol{\sigma}^*|y)$ is estimated as

$$\widehat{\pi}(\boldsymbol{\sigma}^*|y) = \frac{1}{S}\sum_{s=1}^{S}\prod_{t=1}^{2}\prod_{j=1}^{k_t} \pi(\sigma_{tj}^{*-2}|\boldsymbol{\mu}_j^{ts}, \mathbf{D}^s, \boldsymbol{\Delta}^s)$$

where the full conditionals of σ_{tj}^{-2} $(t = 1, 2;\ j = 1, ..., k_t)$ are given in steps 13 and 14 of the MCMC algorithm.

2. Continue sampling using the MCMC algorithm but with the reduced set of parameters in which σ_{tj}^{-2} $(t = 1, 2;\ j = 1, ..., k_t)$ are fixed at σ_{tj}^{*-2}. Then

$$\widehat{\pi}(\boldsymbol{\phi}^*|y, \boldsymbol{\sigma}^*) = \frac{1}{S}\sum_{s=1}^{S}\prod_{t=1}^{2}\prod_{j=1}^{k_t}\phi(\boldsymbol{\phi}_{tj}^*|\overline{\boldsymbol{\phi}}_{tj}^s, \overline{\mathbf{H}}_{\phi_{tj}^s}^{-1}),$$

where $\overline{\boldsymbol{\phi}}_{tj}^s$ and $\overline{\mathbf{H}}_{\phi_{tj}^s}^{-1s}$ $(t = 1, 2;\ j = 1, ..., k_t)$ are given in steps 11 and 12 of the MCMC algorithm subject to $\sigma_{tj}^{-2} = \sigma_{tj}^{*-2}$ restriction.

3. Continue sampling using the MCMC algorithm but with $\sigma_{tj}^{-2} = \sigma_{tj}^{*-2}$ and $\boldsymbol{\phi}_{tj} = \boldsymbol{\phi}_{tj}^*$ $(t = 1, 2;\ j = 1, ..., k_t)$ restrictions. Then

$$\widehat{\pi}(\boldsymbol{\gamma}^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*) = \frac{1}{S}\sum_{s=1}^{S}\prod_{t=1}^{2}\prod_{j=1}^{k_t}\phi(\boldsymbol{\gamma}_{tj}^*|\overline{\boldsymbol{\gamma}}_{tj}^s, \overline{\mathbf{H}}_{\gamma_{tj}^s}^{-1}),$$

where $\overline{\boldsymbol{\gamma}}_{tj}^s$ and $\overline{\mathbf{H}}_{\gamma_{tj}^s}^{-1}$ are defined in steps 5 and 6 of the MCMC algorithm subject to $\sigma_{tj}^{-2} = \sigma_{tj}^{*-2}$ and $\boldsymbol{\phi}_{tj} = \boldsymbol{\phi}_{tj}^*$ $(t = 1, 2;\ j = 1, ..., k_t)$ restrictions.

4. Continue sampling with $\sigma_{tj}^{-2} = \sigma_{tj}^{*-2}$, $\boldsymbol{\phi}_{tj} = \boldsymbol{\phi}_{tj}^*$ and $\boldsymbol{\gamma}_{tj} = \boldsymbol{\gamma}_{tj}^*$ $(t = 1, 2;\ j = 1, ..., k_t)$ restrictions. Then

$$\widehat{\pi}(\eta^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*, \boldsymbol{\gamma}^*) = \frac{1}{S}\sum_{s=1}^{S}\pi(\eta^*|\boldsymbol{\mu}_j^{ts}, \mathbf{D}^s, \boldsymbol{\Delta}^s)$$

where the full conditional $\pi(\eta|\boldsymbol{\mu}_j^{ts}, \mathbf{D}^s, \boldsymbol{\Delta}^s)$ is specified in step 15 of the MCMC algorithm.

5. Finally, continue sampling fixing $\eta = \eta^*$, $\sigma_{tj}^{-2} = \sigma_{tj}^{*-2}$, $\boldsymbol{\phi}_{tj} = \boldsymbol{\phi}_{tj}^*$, $\boldsymbol{\gamma}_{tj} = \boldsymbol{\gamma}_{tj}^*$ $(t = 1, 2;\ j = 1, ..., k_t)$. Then

$$\widehat{\pi}(\boldsymbol{\psi}^*, \boldsymbol{\alpha}^*|y, \boldsymbol{\sigma}^*, \boldsymbol{\phi}^*, \boldsymbol{\gamma}^*, \eta^*) = \frac{1}{S}\sum_{s=1}^{S}\phi(\boldsymbol{\psi}^*, \boldsymbol{\alpha}^*|\overline{\boldsymbol{\varphi}}^s, \overline{\mathbf{H}}_{\boldsymbol{\varphi}^s}^{-1}),$$

where $\overline{\boldsymbol{\varphi}}^s$ and $\overline{\mathbf{H}}_{\boldsymbol{\varphi}^s}^{-1}$ are specified in step 10 of the MCMC algorithm.