Appendix 5.1: Conditional Distributions and Implementation of MH Algorithm Related to SEMs with Continuous and Ordered Categorical Variables

We first consider the conditional distribution in Step (a) of the Gibbs sampler. We note that as the underlying continuous measurements in $\mathbf{Y}$ are given, $\mathbf{Z}$ gives no additional information to this conditional distribution. Moreover, as $\mathbf{v}_i$ are conditionally independent, and $\boldsymbol{\omega}_i$ are also conditionally independent among themselves and independent of $\mathbf{Z}$, we have

$$p(\boldsymbol{\Omega}|\boldsymbol{\alpha}, \boldsymbol{\theta}, \mathbf{Y}, \mathbf{X}, \mathbf{Z}) = \prod_{i=1}^{n} p(\boldsymbol{\omega}_i|\mathbf{v}_i, \boldsymbol{\theta}).$$

It can be shown that

$$[\boldsymbol{\omega}_i|\mathbf{v}_i, \boldsymbol{\theta}] \stackrel{D}{=} N[\boldsymbol{\Sigma}^* \boldsymbol{\Lambda}^T \boldsymbol{\Psi}_\epsilon^{-1}(\mathbf{v}_i - \boldsymbol{\mu}), \boldsymbol{\Sigma}^*], \tag{5.A1}$$

in which $\boldsymbol{\Sigma}^* = (\boldsymbol{\Sigma}_\omega^{-1} + \boldsymbol{\Lambda}^T \boldsymbol{\Psi}_\epsilon^{-1} \boldsymbol{\Lambda})^{-1}$, where $\boldsymbol{\Pi}_0 = \mathbf{I} - \boldsymbol{\Pi}$, and

$$\boldsymbol{\Sigma}_\omega = \begin{bmatrix} \boldsymbol{\Pi}_0^{-1}(\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}^T + \boldsymbol{\Psi}_\delta)\boldsymbol{\Pi}_0^{-T} & \boldsymbol{\Pi}_0^{-1}\boldsymbol{\Gamma}\boldsymbol{\Phi} \\ \boldsymbol{\Phi}\boldsymbol{\Gamma}^T\boldsymbol{\Pi}_0^{-T} & \boldsymbol{\Phi} \end{bmatrix},$$

is the covariance matrix of $\boldsymbol{\omega}_i$. An alternative expression for this conditional distribution can be obtained by the following result, $p(\boldsymbol{\omega}_i|\mathbf{v}_i, \boldsymbol{\theta}) \propto p(\mathbf{v}_i|\boldsymbol{\omega}_i, \boldsymbol{\theta})p(\boldsymbol{\eta}_i|\boldsymbol{\xi}_i, \boldsymbol{\theta})p(\boldsymbol{\xi}_i|\boldsymbol{\theta})$. Based on the definition of the model and assumptions, $p(\boldsymbol{\omega}_i|\mathbf{v}_i, \boldsymbol{\theta})$ is proportional to

$$\begin{aligned} \exp\Big\{ -\frac{1}{2}\big[(\mathbf{v}_i - \boldsymbol{\mu} - \boldsymbol{\Lambda}\boldsymbol{\omega}_i)^T \boldsymbol{\Psi}_\epsilon^{-1}(\mathbf{v}_i - \boldsymbol{\mu} - \boldsymbol{\Lambda}\boldsymbol{\omega}_i) \\ + (\boldsymbol{\eta}_i - \boldsymbol{\Pi}\boldsymbol{\eta}_i - \boldsymbol{\Gamma}\boldsymbol{\xi}_i)^T \boldsymbol{\Psi}_\delta^{-1}(\boldsymbol{\eta}_i - \boldsymbol{\Pi}\boldsymbol{\eta}_i - \boldsymbol{\Gamma}\boldsymbol{\xi}_i) + \boldsymbol{\xi}_i^T \boldsymbol{\Phi}^{-1}\boldsymbol{\xi}_i\big]\Big\}. \end{aligned} \tag{5.A2}$$

Based on the practical experience available so far, simulating observations on the basis of (5.A1) or (5.A2) give similar and acceptable results for statistical inference.

To derive the conditional distributions with respect to the structural parameters in Step (b), let $\boldsymbol{\theta}_v$ be the unknown parameters in $\boldsymbol{\mu}, \boldsymbol{\Lambda}$, and $\boldsymbol{\Psi}_\epsilon$ associated with (5.1), and let $\boldsymbol{\theta}_\omega$ be the unknown parameters in $\boldsymbol{\Lambda}_\omega, \boldsymbol{\Phi}$, and $\boldsymbol{\Psi}_\delta$ associated with (5.2). It is natural to take prior distributions such that $p(\boldsymbol{\theta}) = p(\boldsymbol{\theta}_v)p(\boldsymbol{\theta}_\omega)$.

We first consider the conditional distributions corresponding to $\boldsymbol{\theta}_v$. Similar as before, the following commonly used conjugate type prior distributions are used:

$$\boldsymbol{\mu} \stackrel{D}{=} N[\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0], \quad \psi_{\epsilon k}^{-1} \stackrel{D}{=} Gamma[\alpha_{0\epsilon k}, \beta_{0\epsilon k}],$$
$$[\boldsymbol{\Lambda}_k|\psi_{\epsilon k}] \stackrel{D}{=} N[\boldsymbol{\Lambda}_{0k}, \psi_{\epsilon k}\mathbf{H}_{0vk}], \quad k = 1, \cdots, p,$$

where $\psi_{\epsilon k}$ is the kth diagonal element of $\boldsymbol{\Psi}_\epsilon$, $\boldsymbol{\Lambda}_k^T$ is a $1 \times l_k$ row vector that only contains the unknown parameters in the kth row of $\boldsymbol{\Lambda}$; $\alpha_{0\epsilon k}$, $\beta_{0\epsilon k}$, $\boldsymbol{\mu}_0$, $\boldsymbol{\Lambda}_{0k}$, $\mathbf{H}_{0vk}$, and $\boldsymbol{\Sigma}_0$ are hyperparameters whose values are assumed to be given. For $k \neq h$, it is assumed that $(\psi_{\epsilon k}, \boldsymbol{\Lambda}_k)$ and $(\psi_{\epsilon h}, \boldsymbol{\Lambda}_h)$ are independent. To cope with the case with fixed known elements in $\boldsymbol{\Lambda}$, let $\mathbf{L} = (l_{kj})_{p\times q}$ be the index matrix such that $l_{kj} = 0$ if λ_{kj} is known and $l_{kj} = 1$ if λ_{kj} is unknown, and $l_k = \sum_{j=1}^q l_{kj}$. Let $\boldsymbol{\Omega}_k$ be a submatrix of $\boldsymbol{\Omega}$ such that the jth row with $l_{kj} = 0$ deleted, and let $\mathbf{v}_k^* = (v_{1k}^*, \cdots, v_{nk}^*)^T$ with

$$v_{ik}^* = v_{ik} - \mu_k - \sum_{j=1}^q \lambda_{kj}\omega_{ij}(1 - l_{kj}),$$

where v_{ik} is the kth element of $\mathbf{v}_i$, and μ_k is the kth element of $\boldsymbol{\mu}$. Let $\boldsymbol{\Sigma}_{vk} = (\mathbf{H}_{0vk}^{-1} + \boldsymbol{\Omega}_k\boldsymbol{\Omega}_k^T)^{-1}, \boldsymbol{\mu}_{vk} = \boldsymbol{\Sigma}_{vk}[\mathbf{H}_{0vk}^{-1}\boldsymbol{\Lambda}_{0k} + \boldsymbol{\Omega}_k\mathbf{v}_k^*]$, and $\beta_{\epsilon k} = \beta_{0\epsilon k} + 2^{-1}(\mathbf{v}_k^{*T}\mathbf{v}_k^* - \boldsymbol{\mu}_{vk}^T\boldsymbol{\Sigma}_{vk}^{-1}\boldsymbol{\mu}_{vk} + \boldsymbol{\Lambda}_{0k}^T\mathbf{H}_{0vk}^{-1}\boldsymbol{\Lambda}_{0k})$. Then, it can be shown that for $k = 1, \cdots, p$,

$$\begin{aligned}
&[\psi_{\epsilon k}^{-1}|\boldsymbol{\mu}, \mathbf{V}, \boldsymbol{\Omega}] \stackrel{D}{=} Gamma[n/2 + \alpha_{0\epsilon k}, \beta_{\epsilon k}],\\
&[\boldsymbol{\Lambda}_k|\psi_{\epsilon k}, \boldsymbol{\mu}, \mathbf{V}, \boldsymbol{\Omega}] \stackrel{D}{=} N[\boldsymbol{\mu}_{vk}, \psi_{\epsilon k}\boldsymbol{\Sigma}_{vk}], \qquad (5.A3)\\
&[\boldsymbol{\mu}|\boldsymbol{\Lambda}, \boldsymbol{\Psi}_\epsilon, \mathbf{V}, \boldsymbol{\Omega}] \stackrel{D}{=} N[(\boldsymbol{\Sigma}_0^{-1} + n\boldsymbol{\Psi}_\epsilon^{-1})^{-1}(n\boldsymbol{\Psi}_\epsilon^{-1}\widetilde{\mathbf{V}} + \boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0),\ (\boldsymbol{\Sigma}_0^{-1} + n\boldsymbol{\Psi}_\epsilon^{-1})^{-1}],
\end{aligned}$$

where $\widetilde{\mathbf{V}} = \sum_{i=1}^n(\mathbf{v}_i - \boldsymbol{\Lambda}\boldsymbol{\omega}_i)/n$.

Now, consider the conditional distribution of $\boldsymbol{\theta}_\omega$. As the parameters in $\boldsymbol{\theta}_\omega$ are only involved in the structural equation, this conditional distribution is proportional to $p(\boldsymbol{\Omega}|\boldsymbol{\theta}_\omega)$

$p(\boldsymbol{\theta}_\omega)$, which is independent of $\mathbf{V}$ and $\mathbf{Z}$. Let $\boldsymbol{\Omega}_1 = (\boldsymbol{\eta}_1, \cdots, \boldsymbol{\eta}_n)$ and $\boldsymbol{\Omega}_2 = (\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_n)$. Since the distribution of $\boldsymbol{\xi}_i$ only involves $\boldsymbol{\Phi}$, $p(\boldsymbol{\Omega}_2|\boldsymbol{\theta}_\omega) = p(\boldsymbol{\Omega}_2|\boldsymbol{\Phi})$. Moreover, we take the prior distribution of $\boldsymbol{\Phi}$ such that it is independent of the prior distributions of $\boldsymbol{\Lambda}_\omega$ and $\boldsymbol{\Psi}_\delta$. It follows that $p(\boldsymbol{\Omega}|\boldsymbol{\theta}_\omega)p(\boldsymbol{\theta}_\omega) \propto [p(\boldsymbol{\Omega}_1|\boldsymbol{\Omega}_2, \boldsymbol{\Lambda}_\omega, \boldsymbol{\Psi}_\delta)p(\boldsymbol{\Lambda}_\omega, \boldsymbol{\Psi}_\delta)][p(\boldsymbol{\Omega}_2|\boldsymbol{\Phi})p(\boldsymbol{\Phi})]$. Hence, the marginal conditional densities of $(\boldsymbol{\Lambda}_\omega, \boldsymbol{\Psi}_\delta)$ and $\boldsymbol{\Phi}$ can be treated separately.

Consider a conjugate type prior distribution for $\boldsymbol{\Phi}$ with $\boldsymbol{\Phi}^{-1} \stackrel{D}{=} W_{q_2}[\mathbf{R}_0, \rho_0]$, where ρ_0 and the positive definite matrix $\mathbf{R}_0$ are the given hyperparameters. It can be shown that

$$[\boldsymbol{\Phi}|\boldsymbol{\Omega}_2] \stackrel{D}{=} IW_{q_2}[(\boldsymbol{\Omega}_2\boldsymbol{\Omega}_2^T + \mathbf{R}_0^{-1}), n + \rho_0]. \tag{5.A4}$$

Similar as before, the prior distributions of elements in $(\boldsymbol{\Psi}_\delta, \boldsymbol{\Lambda}_\omega)$ are taken as

$$\psi_{\delta k}^{-1} \stackrel{D}{=} Gamma[\alpha_{0\delta k}, \beta_{0\delta k}], \quad [\boldsymbol{\Lambda}_{\omega k}|\psi_{\delta k}] \stackrel{D}{=} N[\boldsymbol{\Lambda}_{0\omega k}, \psi_{\delta k}\mathbf{H}_{0\omega k}],$$

where $k = 1, \cdots, q_1, \boldsymbol{\Lambda}_{\omega k}^T$ is a $1 \times l_{\omega k}$ row vector that contains the unknown parameters in the kth row of $\boldsymbol{\Lambda}_\omega$; $\alpha_{0\delta k}, \beta_{0\delta k}, \boldsymbol{\Lambda}_{0\omega k}$, and $\mathbf{H}_{0\omega k}$ are given hyperparameters. For $h \neq k$, $(\psi_{\delta k}, \boldsymbol{\Lambda}_{\omega k})$ and $(\psi_{\delta h}, \boldsymbol{\Lambda}_{\omega h})$ are assumed to be independent. Let $\mathbf{L}_\omega = (l_{\omega kj})_{q_1 \times q}$ be the index matrix associated with $\boldsymbol{\Lambda}_\omega$, and $l_{\omega k} = \sum_{j=1}^q l_{\omega kj}$. Let $\boldsymbol{\Omega}_k^*$ be the submatrix of $\boldsymbol{\Omega}$ such that all the jth row corresponding to $l_{\omega kj} = 0$ are deleted; and $\boldsymbol{\Omega}_{\eta k}^* = (\eta_{1k}^*, \cdots, \eta_{nk}^*)^T$ with

$$\eta_{ik}^* = \eta_{ik} - \sum_{j=1}^q \lambda_{\omega kj}\omega_{ij}(1 - l_{\omega kj}),$$

where ω_{ij} is the jth element of $\boldsymbol{\omega}_i$. Then, it can be shown that

$$[\psi_{\delta k}^{-1}|\boldsymbol{\Omega}] \stackrel{D}{=} Gamma[n/2 + \alpha_{0\delta k}, \beta_{\delta k}], \quad [\boldsymbol{\Lambda}_{\omega k}|\boldsymbol{\Omega}, \psi_{\delta k}] \stackrel{D}{=} N[\boldsymbol{\mu}_{\omega k}, \psi_{\delta k}\boldsymbol{\Sigma}_{\omega k}], \tag{5.A5}$$

where $\boldsymbol{\Sigma}_{\omega k} = (\mathbf{H}_{0\omega k}^{-1} + \boldsymbol{\Omega}_k^*\boldsymbol{\Omega}_k^{*T})^{-1}$, $\boldsymbol{\mu}_{\omega k} = \boldsymbol{\Sigma}_{\omega k}[\mathbf{H}_{0\omega k}^{-1}\boldsymbol{\Lambda}_{0\omega k} + \boldsymbol{\Omega}_k^*\boldsymbol{\Omega}_{\eta k}^*]$, and $\beta_{\delta k} = \beta_{0\delta k} + 2^{-1}(\boldsymbol{\Omega}_{\eta k}^{*T}\boldsymbol{\Omega}_{\eta k}^* - \boldsymbol{\mu}_{\omega k}^T\boldsymbol{\Sigma}_{\omega k}^{-1}\boldsymbol{\mu}_{\omega k} + \boldsymbol{\Lambda}_{0\omega k}^T\mathbf{H}_{0\omega k}^{-1}\boldsymbol{\Lambda}_{0\omega k})$.

Finally, we consider the joint conditional distribution of $(\boldsymbol{\alpha}, \mathbf{Y})$ given $\boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{X}$, and $\mathbf{Z}$. Suppose that the model in relation to the subvector $\mathbf{y}_i = (y_{i1}, \cdots, y_{is})^T$ of $\mathbf{v}_i$ is given by:

$$\mathbf{y}_i = \boldsymbol{\mu}_y + \boldsymbol{\Lambda}_y\boldsymbol{\omega}_i + \boldsymbol{\epsilon}_{yi},$$

where $\boldsymbol{\mu}_y$ $(s \times 1)$ is a subvector of $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}_y$ $(s \times q)$ is a submatrix of $\boldsymbol{\Lambda}$, $\boldsymbol{\epsilon}_{yi}$ $(s \times 1)$ is a subvector of $\boldsymbol{\epsilon}_i$ with diagonal covariance submatrix $\boldsymbol{\Psi}_y$ of $\boldsymbol{\Psi}_\epsilon$. Let $\mathbf{z}_i = (z_{i1}, \cdots, z_{is})^T$ be the ordered categorical observation corresponding to $\mathbf{y}_i$, $i = 1, \cdots, n$. We use the following non-informative prior distribution for the unknown thresholds in $\boldsymbol{\alpha}_k$:

$$p(\alpha_{k,2}, \cdots, \alpha_{k,b_k-1}) \propto C, \quad \text{for} \quad \alpha_{k,2} < \cdots < \alpha_{k,b_k-1}, \quad k = 1, \cdots, s,$$

where C is a constant. Given $\boldsymbol{\Omega}$ and the fact that the covariance matrix $\boldsymbol{\Psi}_y$ is diagonal, the ordered categorical data $\mathbf{Z}$ and the thresholds corresponding to different rows are also conditionally independent. For $k = 1, \cdots, s$, let $\mathbf{Y}_k^T$ and $\mathbf{Z}_k^T$ be the kth rows of $\mathbf{Y}$ and $\mathbf{Z}$, respectively, it can be shown that

$$p(\boldsymbol{\alpha}_k, \mathbf{Y}_k | \mathbf{Z}_k, \boldsymbol{\theta}, \boldsymbol{\Omega}) = p(\boldsymbol{\alpha}_k | \mathbf{Z}_k, \boldsymbol{\theta}, \boldsymbol{\Omega}) p(\mathbf{Y}_k | \boldsymbol{\alpha}_k, \mathbf{Z}_k, \boldsymbol{\theta}, \boldsymbol{\Omega}), \tag{5.A6}$$

$$\propto p(Z_k | \alpha_k, \theta, \Omega) p(\alpha_k) \propto p(Z_k | \alpha_k, \theta, \Omega)$$

$$= p(\alpha_{k,Z_k} < y_k < \alpha_{k,Z_k+1})$$

with

$$p(\boldsymbol{\alpha}_k | \mathbf{Z}_k, \boldsymbol{\theta}, \boldsymbol{\Omega}) \propto \prod_{i=1}^{n} \Big[\Phi^* \Big\{ \psi_{yk}^{-1/2} (\alpha_{k,z_{ik}+1} - \mu_{yk} - \boldsymbol{\Lambda}_{yk}^T \boldsymbol{\omega}_i) \Big\} - \Phi^* \Big\{ \psi_{yk}^{-1/2} (\alpha_{k,z_{ik}} - \mu_{yk} - \boldsymbol{\Lambda}_{yk}^T \boldsymbol{\omega}_i) \Big\} \Big], \tag{5.A7}$$

and $p(\mathbf{Y}_k | \boldsymbol{\alpha}_k, \mathbf{Z}_k, \boldsymbol{\theta}, \boldsymbol{\Omega})$ is the product of $p(y_{ik} | \boldsymbol{\alpha}_k, \mathbf{Z}_k, \boldsymbol{\theta}, \boldsymbol{\Omega})$, where

$$[y_{ik} | \boldsymbol{\alpha}_k, \mathbf{Z}_k, \boldsymbol{\theta}, \boldsymbol{\Omega}] \stackrel{D}{=} N[\mu_{yk} + \boldsymbol{\Lambda}_{yk}^T \boldsymbol{\omega}_i, \psi_{yk}] I_{[\alpha_{k,z_{ik}}, \alpha_{k,z_{ik}+1})}(y_{ik}), \tag{5.A8}$$

in which ψ_{yk} is the kth diagonal element of $\boldsymbol{\Psi}_y$, μ_{yk} is the kth element of $\boldsymbol{\mu}_y$, $\boldsymbol{\Lambda}_{yk}^T$ is the kth row of $\boldsymbol{\Lambda}_y$, $I_A(y)$ is an index function which takes 1 if $y \in A$ and 0 otherwise, and $\Phi^*(\cdot)$ denotes the distribution function of $N[0, 1]$. As a result,

$$p(\boldsymbol{\alpha}_k, \mathbf{Y}_k | \mathbf{Z}_k, \boldsymbol{\theta}, \boldsymbol{\Omega}) \propto \prod_{i=1}^{n} \phi \Big\{ \psi_{yk}^{-1/2} (y_{ik} - \mu_{yk} - \boldsymbol{\Lambda}_{yk}^T \boldsymbol{\omega}_i) \Big\} I_{[\alpha_{k,z_{ik}}, \alpha_{k,z_{ik}+1})}(y_{ik}), \tag{5.A9}$$

where $\phi(\cdot)$ is the standard normal density.

To sample from the conditional distributions (5.A2) and (5.A9), the MH algorithm is implemented as follows.

For $p(\boldsymbol{\omega}_i|\mathbf{v}_i,\boldsymbol{\theta})$, we choose $N[\mathbf{0},\sigma^2\boldsymbol{\Sigma}^*]$ as the proposal distribution, where $\boldsymbol{\Sigma}^{*-1} = \boldsymbol{\Sigma}_\omega^{-1} + \boldsymbol{\Lambda}^T\boldsymbol{\Psi}_\epsilon^{-1}\boldsymbol{\Lambda}$, with

$$\boldsymbol{\Sigma}_\omega^{-1} = \begin{bmatrix} \boldsymbol{\Pi}_0^T\boldsymbol{\Psi}_\delta^{-1}\boldsymbol{\Pi}_0 & -\boldsymbol{\Pi}_0^T\boldsymbol{\Psi}_\delta^{-1}\boldsymbol{\Gamma} \\ -\boldsymbol{\Gamma}^T\boldsymbol{\Psi}_\delta^{-1}\boldsymbol{\Pi}_0 & \boldsymbol{\Phi}^{-1} + \boldsymbol{\Gamma}^T\boldsymbol{\Psi}_\delta^{-1}\boldsymbol{\Gamma} \end{bmatrix}.$$

Let $p(\cdot|\mathbf{0},\sigma^2,\boldsymbol{\Sigma}^*)$ be the proposal density corresponding to $N[\mathbf{0},\sigma^2\boldsymbol{\Sigma}^*]$, where σ^2 is an appropriate preassigned constant. The MH algorithm is implemented as follows: At the jth MH iteration with a current value $\boldsymbol{\omega}_i^{(j)}$, a new candidate $\boldsymbol{\omega}_i$ is generated from $p(\cdot|\boldsymbol{\omega}_i^{(j)},\sigma^2,\boldsymbol{\Sigma}^*)$, and accepting this new candidate with the probability

$$\min\left\{1, \frac{p(\boldsymbol{\omega}_i|\mathbf{v}_i,\boldsymbol{\theta})}{p(\boldsymbol{\omega}_i^{(j)}|\mathbf{v}_i,\boldsymbol{\theta})}\right\},$$

where $p(\boldsymbol{\omega}_i|\mathbf{v}_i,\boldsymbol{\theta})$ is given by (5.A2).

For $p(\boldsymbol{\alpha}_k,\mathbf{Y}_k|\mathbf{Z}_k,\boldsymbol{\theta},\boldsymbol{\Omega})$, we use the equality (5.A6) from Cowles (1996) to construct a joint proposal density for $\boldsymbol{\alpha}_k$, and $\mathbf{Y}_k$ in the MH algorithm for generating observations from it. At the jth MH iteration, we generate a vector of thresholds $(\alpha_{k,2},\cdots,\alpha_{k,b_k-1})$ from the following univariate truncated normal distribution:

$$\alpha_{k,z} \stackrel{D}{=} N[\alpha_{k,z}^{(j)},\sigma_{\alpha_k}^2]I_{(\alpha_{k,z-1},\alpha_{k,z+1}^{(j)}]}(\alpha_{k,z}) \quad \text{for} \quad z = 2,\cdots,b_k-1,$$

where $\alpha_{k,z}^{(j)}$ is the current value of $\alpha_{k,z}$ at the jth iteration of the Gibbs sampler, and $\sigma_{\alpha_k}^2$ is an appropriate preassigned constant. Random observations from the above univariate truncated normal are simulated via the algorithm of Roberts (1995). Then, the acceptance probability for $(\boldsymbol{\alpha}_k,\mathbf{Y}_k)$ as a new observation is $\min\{1,R_k\}$, where

$$R_k = \prod_{z=2}^{b_k-1} \frac{\Phi^*\{(\alpha_{k,z+1}^{(j)} - \alpha_{k,z}^{(j)})/\sigma_{\alpha_k}\} - \Phi^*\{(\alpha_{k,z-1} - \alpha_{k,z}^{(j)})/\sigma_{\alpha_k}\}}{\Phi^*\{(\alpha_{k,z+1} - \alpha_{k,z})/\sigma_{\alpha_k}\} - \Phi^*\{(\alpha_{k,z-1}^{(j)} - \alpha_{k,z})/\sigma_{\alpha_k}\}} \times$$
$$\prod_{i=1}^{n} \frac{\Phi^*\left\{\psi_{yk}^{-1/2}(\alpha_{k,z_{ik}+1} - \mu_{yk} - \boldsymbol{\Lambda}_{yk}^T\boldsymbol{\omega}_i)\right\} - \Phi^*\left\{\psi_{yk}^{-1/2}(\alpha_{k,z_{ik}} - \mu_{yk} - \boldsymbol{\Lambda}_{yk}^T\boldsymbol{\omega}_i)\right\}}{\Phi^*\left\{\psi_{yk}^{-1/2}(\alpha_{k,z_{ik}+1}^{(j)} - \mu_{yk} - \boldsymbol{\Lambda}_{yk}^T\boldsymbol{\omega}_i)\right\} - \Phi^*\left\{\psi_{yk}^{-1/2}(\alpha_{k,z_{ik}}^{(j)} - \mu_{yk} - \boldsymbol{\Lambda}_{yk}^T\boldsymbol{\omega}_i)\right\}}.$$

As R_k only depends on the old and new values of $\boldsymbol{\alpha}_k$ and not on $\mathbf{Y}_k$, it does not require to generate a new $\mathbf{Y}_k$ in any iteration in which the new value of $\boldsymbol{\alpha}_k$ is not accepted (see Cowles, 1996). For an accepted $\boldsymbol{\alpha}_k$, a new $\mathbf{Y}_k$ is simulated from (5.A8).