1 The model

The model is:

- $Z_g \sim \pi$ $\qquad g \in [G]$.
- $M_{g,j,\mathscr{T}} | \{Z_g = j\} \sim \text{Normal}(\mu_0, \tau_0^2)$ $\qquad g \in [G], \mathscr{T} \in \mathbb{T}_j$.
- $X_{i,g,t} | \{Z_g = j, M_{g,j,\mathscr{T}} = \mu_{g,j,\mathscr{T}}\} \sim \text{Normal}(\mu_{g,j,\mathscr{T}}, \tau^2)$ $\qquad i \in [n], g \in [G], \mathscr{T} \in \mathbb{T}_j, t \in \mathscr{T}$.
- $S_{i,g} = \sum_{t=1}^T p_{\sigma_i,t} X_{i,g,t}$ $\qquad i \in [n], g \in [G]$.

$S_{i,g}$ is observed. The parameters $\theta = (p_{\sigma,t}, \pi_j, \tau, \tau_0, \mu_0)$. We require that $\sigma \mapsto p_{\sigma,t}$ is monotone.

1.1 The distribution of $S_g | \{Z_g = j\}$.

Claim 1. *Let $S_g = (S_{1,g}, S_{2,g}, \ldots, S_{n,g})$. Then $S_g | \{Z_g = j\} \sim \text{Normal}(\mu_0 \mathbb{1}, \text{cov}_j)$ where, letting $p_{\sigma,\mathscr{T}} = \sum_{t \in \mathscr{T}} p_{\sigma,t}$,*

$$\text{cov}_j = \tau^2 \begin{bmatrix} \sum_{t=1}^T p_{\sigma_1,t}^2 & 0 & \cdots & 0 \\ 0 & \sum_{t=1}^T p_{\sigma_2,t}^2 & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \sum_{t=1}^T p_{\sigma_n,t}^2 \end{bmatrix} + \tau_0^2 \begin{bmatrix} \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_1,\mathscr{T}}^2 & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_1,\mathscr{T}} p_{\sigma_2,\mathscr{T}} & \cdots & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_1,\mathscr{T}} p_{\sigma_n,\mathscr{T}} \\ \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_2,\mathscr{T}} p_{\sigma_1,\mathscr{T}} & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_2,\mathscr{T}}^2 & \cdots & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_2,\mathscr{T}} p_{\sigma_n,\mathscr{T}} \\ & & \ddots & \\ \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_n,\mathscr{T}} p_{\sigma_1,\mathscr{T}} & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_n,\mathscr{T}} p_{\sigma_2,\mathscr{T}} & \cdots & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_n,\mathscr{T}}^2 \end{bmatrix}.$$

Proof. Step 1. Note that

$$S_g = (\sum_{t=1}^T p_{\sigma_i,t} X_{i,g,t} : i = 1, \ldots, n)'$$

$$= \begin{bmatrix} p_{\sigma_1,1} & p_{\sigma_1,2} & p_{\sigma_1,T} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{\sigma_2,1} & p_{\sigma_2,2} & p_{\sigma_2,T} & \cdots & 0 & 0 & 0 \\ \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & p_{\sigma_n,1} & p_{\sigma_n,2} & p_{\sigma_n,T} \end{bmatrix} \begin{bmatrix} X_{1,g,1} \\ X_{1,g,2} \\ X_{1,g,T} \\ X_{2,g,1} \\ X_{2,g,2} \\ X_{2,g,T} \\ \vdots \\ X_{n,g,1} \\ X_{n,g,2} \\ X_{n,g,T} \end{bmatrix} = QY_g,$$

where, letting $t(k) = ((k-1) \bmod T) + 1$,

$$Q = (q_{i,k})_{i=1}^{n}{}_{k=1}^{nT}, \qquad q_{i,k} = \begin{cases} p_{\sigma_i,t(k)} & \text{if } (i-1)T < k \le iT \\ 0 & \text{otherwise} \end{cases}$$

$$Y_g = (Y_{g,k})_{k=1}^{nT}, \qquad Y_{g,k} = X_{g,t(k)}.$$

Step 2. Let $\mathscr{F}_{g,j} = \{\omega : Z_g(\omega) = j, M_{g,j}(\omega) = \mu_{g,j}\}$. Let $\mathscr{T}(j,t)$ be the element of $\mathbb{T}_j$ that contains t. Note

$$Y_g | \mathscr{F}_{g,j} = \begin{bmatrix} X_{1,g,1} \\ X_{1,g,2} \\ X_{1,g,T} \\ \vdots \\ X_{n,g,1} \\ X_{n,g,2} \\ X_{n,g,T} \end{bmatrix} | \mathscr{F}_{g,j} \sim \text{Normal}(\begin{bmatrix} \mu_{g,j,\mathscr{T}(j,1)} \\ \mu_{g,j,\mathscr{T}(j,2)} \\ \mu_{g,j,\mathscr{T}(j,T)} \\ \vdots \\ \mu_{g,j,\mathscr{T}(j,1)} \\ \mu_{g,j,\mathscr{T}(j,2)} \\ \mu_{g,j,\mathscr{T}(j,T)} \end{bmatrix}, \tau^2 I) = \text{Normal}(\nu_{g,j}, \tau^2 I)$$

where $\nu_{g,j} = (\nu_{g,j,k})_{k=1}^{nT}, \nu_{g,j,k} = \mu_{g,j,\mathscr{T}(j,t(k))}$. This holds because, for fixed g, the $X_{i,g,t}$ are conditionally independent given $\mathscr{F}_{g,j}$.

Note that $\mathscr{F}_{g,j} = \{\omega : Z_g(\omega) = j, N_{g,j}(\omega) = \nu_{g,j}\}$.

Step 3. Next,

$$N_{g,j} = \begin{bmatrix} M_{g,j,\mathscr{T}(j,1)} \\ M_{g,j,\mathscr{T}(j,2)} \\ M_{g,j,\mathscr{T}(j,T)} \\ \vdots \\ M_{g,j,\mathscr{T}(j,1)} \\ M_{g,j,\mathscr{T}(j,2)} \\ M_{g,j,\mathscr{T}(j,T)} \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \quad\quad \Delta_j \quad\quad \\ \\ \\ \\ \end{bmatrix} \begin{bmatrix} M_{g,j,\mathscr{T}_{j,1}} \\ M_{g,j,\mathscr{T}_{j,2}} \\ \vdots \\ M_{g,j,\mathscr{T}_{j,|\mathbb{T}_j|}} \end{bmatrix} = \Delta_j M_{g,j}$$

where

$$\Delta_j = (\delta_{j,k,l})_{k=1}^{nT}{}_{l=1}^{|\mathbb{T}_j|}, \qquad \delta_{j,k,l} = \begin{cases} 1 & \text{if } \mathscr{T}(j,t(k)) = \mathscr{T}_{j,l} \\ 0 & \text{otherwise} \end{cases}$$

and $\mathscr{T}_{j,1}, \mathscr{T}_{j,2}, \ldots, \mathscr{T}_{j,|\mathbb{T}_j|}$ is an enumeration of $\mathbb{T}_j$. Note that each row of Δ_j has exactly one nonzero element.

Step 4. Altogether, we have:

- $S_g = QY_g$.
- $Y_g | \mathscr{F}_{g,j} \sim \text{Normal}(\nu_{g,j}, \tau^2 I)$.
- $N_{g,j} = \Delta_j M_{g,j}$.
- $M_{g,j} | \{Z_g = j\} \sim \text{Normal}(\mu_0 \mathbb{1}, \tau_0^2 I)$.

Step 5. In general, if $X \sim \text{Normal}(\mu, \Sigma)$ and $Y = AX + b$, then $Y \sim \text{Normal}(A\mu + b, A\Sigma A')$.

In our case, $M_g | \{Z_g = j\} \sim \text{Normal}(\mu_0 \mathbb{1}, \tau_0^2 I)$ and $N_{g,j} = \Delta_j M_{g,j}$, so

$$\begin{aligned} N_{g,j} | \{Z_g = j\} &\sim \text{Normal}(\Delta_j \mu_0 \mathbb{1}, \Delta_j \tau_0^2 I \Delta_j') \\ &= \text{Normal}(\mu_0 \mathbb{1}, \tau_0^2 \Delta_j \Delta_j'). \end{aligned}$$

The equality holds because each row of Δ_j has exactly one nonzero element, equal to 1, so $\Delta_j \mathbb{1} = \mathbb{1}$.

Next, $Y_g | \mathscr{F}_{g,j} \sim \text{Normal}(\nu_{g,j}, \tau^2 I)$ and $S_g = QY_g$, so

$$\begin{aligned} S_g | \mathscr{F}_{g,j} &\sim \text{Normal}(Q\nu_{g,j}, Q\tau^2 I Q') \\ &= \text{Normal}(Q\nu_{g,j}, \tau^2 QQ') \end{aligned}$$

Step 6. In general, if $X \sim \text{Normal}(\mu, \Sigma)$ and $Y | \{X = x\} \sim \text{Normal}(AX + b, S)$, then $Y \sim \text{Normal}(A\mu + b, S + A\Sigma A')$.

In our case, $N_{g,j} | \{Z_g = j\} \sim \text{Normal}(\mu_0 \mathbb{1}, \tau_0^2 \Delta_j \Delta_j')$ and $S_g | \{Z_g = j, N_{g,j} = \nu_{g,j}\} \sim \text{Normal}(Q\nu_{g,j}, \tau^2 QQ')$, so

$$\begin{aligned} S_g | \{Z_g = j\} &\sim \text{Normal}(Q\mu_0 \mathbb{1}, \tau^2 QQ' + Q\tau_0^2 \Delta_j \Delta_j' Q') \\ &= \text{Normal}(\mu_0 \mathbb{1}, \tau^2 QQ' + \tau_0^2 Q\Delta_j \Delta_j' Q') \end{aligned}$$

The equality holds because each row of Q sums to one, so $Q\mathbb{1} = \mathbb{1}$.

Note that QQ' is an $n \times n$ matrix with (i,i') entry

$$\begin{aligned}
&\sum_{k=1}^{nT} p_{\sigma_i,t(k)} p_{\sigma_{i'},t(k)} \underbrace{1((i-1)T < k \le iT)1((i'-1)T < k < i'T)}_{(*)} \\
&= \sum_{k=1}^{nT} p^2_{\sigma_i,t(k)} 1((i-1)T < k \le iT)1(i = i') \quad \text{since (*) is nonzero only if } i = i' \\
&= 1(i = i') \sum_{t=1}^{T} p^2_{\sigma_i,t}
\end{aligned}$$

Note that $Q\Delta_j$ is an $n \times |\mathbb{T}_j|$ matrix with (i,l) entry

$$\begin{aligned}
\sum_{k=1}^{nT} q_{i,k}\delta_{j,k,l} &= \sum_{k=1}^{nT} p_{\sigma_i,t(k)} 1((i-1)T < k \le iT)1(\mathscr{T}(j,t(k)) = \mathscr{T}_{j,l}) \\
&= \sum_{k=1}^{nT} p_{\sigma_i,t(k)} 1((i-1)T < k \le iT)1(t(k) \in \mathscr{T}_{j,l}) \\
&= \sum_{t \in \mathscr{T}_l} p_{\sigma_i,t}
\end{aligned}$$

Note that $Q\Delta_j\Delta_j'Q' = (Q\Delta_j)(Q\Delta_j)'$ is an $n \times n$ matrix with (i,i') entry

$$\begin{aligned}
&\sum_{l=1}^{|\mathbb{T}_j|} (Q\Delta_j)_{i,l}(Q\Delta_j)_{i',l} \\
&= \sum_{l=1}^{|\mathbb{T}_j|} \Big(\sum_{t \in \mathscr{T}_l} p_{\sigma_i,t}\Big)\Big(\sum_{t \in \mathscr{T}_l} p_{\sigma_{i'},t}\Big) \\
&= \sum_{\mathscr{T} \in \mathbb{T}_j} \Big(\sum_{t \in \mathscr{T}} p_{\sigma_i,t}\Big)\Big(\sum_{t \in \mathscr{T}} p_{\sigma_{i'},t}\Big)
\end{aligned}$$

Thus,

$$S_g | \{Z_g = j\} \sim \text{Normal}(\mu_0 \mathbb{1}, \text{cov}_j)$$

where, letting $p_{\sigma,\mathscr{T}} = \sum_{t \in \mathscr{T}} p_{\sigma,t}$,

$$\begin{aligned}
\text{cov}_j &= \tau^2 QQ' + \tau_0^2 Q\Delta_j\Delta_j'Q' \\
&= \tau^2 \begin{bmatrix} \sum_{t=1}^{T} p^2_{\sigma_1,t} & 0 & \cdots & 0 \\ 0 & \sum_{t=1}^{T} p^2_{\sigma_2,t} & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \sum_{t=1}^{T} p^2_{\sigma_n,t} \end{bmatrix} + \tau_0^2 \begin{bmatrix} \sum_{\mathscr{T} \in \mathbb{T}_j} p^2_{\sigma_1,\mathscr{T}} & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_1,\mathscr{T}} p_{\sigma_2,\mathscr{T}} & \cdots & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_1,\mathscr{T}} p_{\sigma_n,\mathscr{T}} \\ \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_2,\mathscr{T}} p_{\sigma_1,\mathscr{T}} & \sum_{\mathscr{T} \in \mathbb{T}_j} p^2_{\sigma_2,\mathscr{T}} & \cdots & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_2,\mathscr{T}} p_{\sigma_n,\mathscr{T}} \\ & & \ddots & \\ \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_n,\mathscr{T}} p_{\sigma_1,\mathscr{T}} & \sum_{\mathscr{T} \in \mathbb{T}_j} p_{\sigma_n,\mathscr{T}} p_{\sigma_2,\mathscr{T}} & \cdots & \sum_{\mathscr{T} \in \mathbb{T}_j} p^2_{\sigma_n,\mathscr{T}} \end{bmatrix}.
\end{aligned}$$

as claimed. ■

1.2 Block Cholesky decomposition

For each $k \in [\Sigma]$, $g \in [G]$, let

- $m_k = |\{i : \sigma_i = k\}|$.
- $\tilde{S}_{k,g} = (S_{i,g} : \sigma_i = k)$, a block column vector with m_k entries.
- $\tilde{S}_g = (\tilde{S}_{k,g} : k \in [\Sigma])$, a block column vector with n entries.
- $\tilde{\sigma}_i$ the stage of the tissue sample $(\tilde{S}_g)_i$; in other words, $\tilde{\sigma}_i = \sigma_{f(i)}$ where f is such that $(\tilde{S}_g)_i = S_{f(i),g}$.

By Claim 1, $\tilde{S}_g|\{Z_g = j\} \sim \text{Normal}(\mu_0 \mathbb{1}, \text{cov}_j)$, where cov_j is as given in the claim.

We call the $n \times n$ identity matrix I_n and the $n \times m$ all-ones matrix $J_{n,m}$.

Claim 2. *For each $j \in [J]$, the covariance matrix $C = \text{cov}_j$ has the block Cholesky decomposition $C = LDL'$, where*

- $L = \begin{bmatrix} I_{m_1} & 0 & 0 & 0 & \cdots & 0 \\ L_{2,1} & I_{m_2} & 0 & 0 & \cdots & 0 \\ L_{3,1} & L_{3,2} & I_{m_2} & 0 & \cdots & 0 \\ \cdots \\ L_{\Sigma,1} & L_{\Sigma,2} & L_{\Sigma,3} & \cdots & L_{\Sigma,\Sigma-1} & I_{m_\Sigma} \end{bmatrix}$, $L_{k,l} = \lambda_{k,l} J_{m_k,m_l}$ $\qquad k \in [\Sigma], l < k.$
- $D = \begin{bmatrix} D_1 & & & & \\ & D_2 & & & \\ & & D_3 & & \\ & & & \ddots & \\ & & & & D_\Sigma \end{bmatrix}$, $D_k = \alpha_k I_{m_k} + \delta_{k,k} J_{m_k,m_k}$ $\qquad k \in [\Sigma].$
- $\alpha_k = \tau^2 \sum_{t=1}^{T} p_{k,t}^2$ $\qquad k \in [\Sigma].$
- $\beta_{k,l} = \tau_0^2 \sum_{\mathscr{T} \in \mathbb{T}_j} (\sum_{t \in \mathscr{T}} p_{k,t})(\sum_{t \in \mathscr{T}} p_{l,t})$ $\qquad k \in [\Sigma], l \leq k.$
- $\gamma_k = \alpha_k + m_k \delta_{k,k}$ $\qquad k \in [\Sigma].$
- $\delta_{k,l} = \beta_{k,l} - \sum_{h<l} \delta_{k,h} \delta_{l,h} m_h / \gamma_h$ $\qquad k \in [\Sigma], l \leq k.$
- $\lambda_{k,l} = \delta_{k,l} / \gamma_l$ $\qquad k \in [\Sigma], l \leq k.$

Claim 3. *Let $y = L^{-1}x$ where $C = LDL'$. Partition x and y into Σ blocks $x_k = (x_{k,i}) \in \mathbb{R}^{m_k}$ and $y_k = (y_{k,i}) \in \mathbb{R}^{m_k}$, $k \in [\Sigma]$. Then for all $k \in [\Sigma]$,*

$$y_k = x_k - \sum_{l<k} \lambda_{k,l} \Big(\sum_{i=1}^{m_l} y_{l,i} \Big) \mathbb{1}_{m_k}.$$

Claim 4. *Let D be such that $C = LDL'$. Then*

$$D^{-1} = \begin{bmatrix} D_{1,1}^{-1} & & & \\ & D_{2,2}^{-1} & & \\ & & \ddots & \\ & & & D_{\Sigma,\Sigma}^{-1} \end{bmatrix}$$

where

$$D_{k,k}^{-1} = \frac{1}{\alpha_k} I_{m_k} - \frac{\delta_{k,k}}{\alpha_k \gamma_k} J_{m_k,m_k}$$

Claim 5. *Let $y = L^{-1}x$ where $C = LDL'$. Partition x and y into Σ blocks $x_k = (x_{k,i}) \in \mathbb{R}^{m_k}$ and $y_k = (y_{k,i}) \in \mathbb{R}^{m_k}$, $k \in [\Sigma]$. Then*

$$
\begin{aligned}
y' D^{-1} y &= \sum_{k=1}^{\Sigma} \frac{y_k' y_k}{\alpha_k} - \frac{\delta_{k,k}}{\alpha_k \gamma_k} \Big(\sum_{i=1}^{m_k} y_{k,i} \Big)^2 \\
&= \sum_{k=1}^{\Sigma} \frac{y_k' y_k}{\alpha_k} - \frac{\lambda_{k,k} (\mathbb{1}' y_k)^2}{\alpha_k}.
\end{aligned}
$$

Claim 6. *Let $y = L^{-1}x$ and $\tilde{y} = L^{-1}\tilde{x}$, where $C = LDL'$. Partition x, y, and $\tilde{y}$ into Σ blocks $x_k = (x_{k,i}) \in \mathbb{R}^{m_k}$, $y_k = (y_{k,i}) \in \mathbb{R}^{m_k}$, and $\tilde{y}_k = (\tilde{y}_{k,i}) \in \mathbb{R}^{m_k}$, $k \in [\Sigma]$. Then*

$$
\begin{aligned}
\tilde{y}' D^{-1} y &= \sum_{k=1}^{\Sigma} \frac{\tilde{y}_k' y_k}{\alpha_k} - \frac{\delta_{k,k}}{\alpha_k \gamma_k} \Big(\sum_{i=1}^{m_k} \tilde{y}_{k,i} \Big) \Big(\sum_{i=1}^{m_k} y_{k,i} \Big) \\
&= \sum_{k=1}^{\Sigma} \frac{\tilde{y}_k' y_k}{\alpha_k} - \frac{\lambda_{k,k} (\mathbb{1}' \tilde{y}_k)(\mathbb{1}' y_k)}{\alpha_k}.
\end{aligned}
$$

Claim 7. *Let $y = L^{-1}\tilde{\mathbb{1}}$, where $C = LDL'$. Partition y into Σ blocks $y_k = (y_{k,i}) \in \mathbb{R}^{m_k}$, $k \in [\Sigma]$. Then*

$$
y_k = \upsilon_k \mathbb{1}_{m_k} \quad \text{where } \upsilon_k = 1 - \sum_{l<k} \lambda_{k,l} \upsilon_l m_l
$$

Proof. We use the definition of $L^{-1}x$ from Claim 3. First, $y_1 = x_1 = \mathbb{1}_{m_1}$. For induction, assume that the claim holds for all $l < k$. Then

$$
\begin{aligned}
y_k &= \mathbb{1}_{m_k} - \sum_{l<k} \lambda_{k,l} (\mathbb{1}_{m_l}' y_l) \mathbb{1}_{m_k} \\
&= \mathbb{1}_{m_k} - \sum_{l<k} \lambda_{k,l} (\mathbb{1}_{m_l}' (\upsilon_l \mathbb{1}_{m_l})) \mathbb{1}_{m_k} \\
&= \mathbb{1}_{m_k} - \sum_{l<k} \lambda_{k,l} \upsilon_l m_l \mathbb{1}_{m_k} \\
&= (1 - \sum_{l<k} \lambda_{k,l} \upsilon_l m_l) \mathbb{1}_{m_k} \\
&= \upsilon_k \mathbb{1}_{m_k}
\end{aligned}
$$

as claimed. ■

Claim 8. *Let $y = L^{-1}x$ and $\tilde{y} = L^{-1}\mathbb{1}$, where $C = LDL'$. Partition x, y, and $\tilde{y}$ into Σ blocks $x_k = (x_{k,i}) \in \mathbb{R}^{m_k}$, $y_k = (y_{k,i}) \in \mathbb{R}^{m_k}$, and $\tilde{y}_k = (\tilde{y}_{k,i}) \in \mathbb{R}^{m_k}$, $k \in [\Sigma]$. Then*

$$
y' D^{-1} \tilde{y} = \sum_{k=1}^{\Sigma} \frac{\upsilon_k}{\alpha_k} (1 - m_k \lambda_{k,k}) (\mathbb{1}' y_k)
$$

where υ_k is from Claim 7.

Proof. Note that, from Claim 7,

$$
\begin{aligned}
y_k' \tilde{y}_k &= y_k' (\upsilon_k \mathbb{1}) = \upsilon_k (\mathbb{1}' y_k) \\
\mathbb{1}' \tilde{y}_k &= \mathbb{1}' (\upsilon_k \mathbb{1}) = m_k \upsilon_k
\end{aligned}
$$

Plugging into Claim 6,

$$\begin{aligned}\tilde{y}'D^{-1}y &= \sum_{k=1}^{\Sigma} \frac{\tilde{y}_k'y_k}{\alpha_k} - \frac{\lambda_{k,k}(\mathbb{1}'\tilde{y}_k)(\mathbb{1}'y_k)}{\alpha_k} \\ &= \sum_{k=1}^{\Sigma} \frac{\upsilon_k(\mathbb{1}'y_k)}{\alpha_k} - \frac{\lambda_{k,k}m_k\upsilon_k(\mathbb{1}'y_k)}{\alpha_k} \\ &= \sum_{k=1}^{\Sigma} \frac{\upsilon_k}{\alpha_k}(1-m_k\lambda_{k,k})(\mathbb{1}'y_k)\end{aligned}$$

as claimed ■

Claim 9. *The determinant of $C = LDL'$ is*

$$\begin{aligned}\det C &= \prod_{k=1}^{\Sigma} \alpha_k^{m_k-1}(\alpha_k + m_k\delta_{k,k}) \\ &= \prod_{k=1}^{\Sigma} \alpha_k^{m_k-1}\gamma_k\end{aligned}$$

and

$$\log\det C = \sum_{k=1}^{\Sigma}(m_k-1)\log(\alpha_k) + \log(\alpha_k + m_k\delta_{k,k})$$

Corollary 10. *The multivariate normal density* $\text{Normal}(x|\mu_0\mathbb{1}_n, C)$ *can be evaluated as follows:*

- *Compute α_k, $\beta_{k,l}$, γ_k, $\delta_{k,l}$, and $\lambda_{k,l}$ for all k and l via Claim 2.*
- *Compute $y = L^{-1}(x-\mu_0\mathbb{1}_n)$ via Claim 3.*
- *Compute $y'D^{-1}y$ via Claim 5.*
- *Compute* $\log\det C$ *via Claim 9.*
- *Compute* $\log\text{Normal}(x|\mu_0\mathbb{1}_n, C) = -\frac{1}{2}(n\log(2\pi) + \log\det C + y'D^{-1}y)$.

Claim 11. $(\mathbb{1}'x_k)_{k=1}^{\Sigma}$ *and* $(x_k'x_k)_{k=1}^{\Sigma}$ *are sufficient statistics for* $\text{Normal}(x|\mu_0\mathbb{1}, C)$.

1.3 Gradient and Hessian

We want the gradient and Hessian of the log likelihood

$$\begin{aligned}\ell(\theta) &= \sum_{g\in[G]} \text{logsumexp}_{j\in[J]}(\log\pi_j + \log\text{Normal}(s_g|\mu_0\mathbb{1}, C_j)) \\ &= \sum_{g\in[G]} \log\sum_{j\in[J]} \pi_j\text{Normal}(s_g|\mu_0\mathbb{1}, C_j)\end{aligned}$$

Note that

$$\begin{aligned}\frac{\partial\ell}{\partial\pi_{j_0}} &= \sum_{g\in[G]} \frac{\partial}{\partial\pi_{j_0}} \log\sum_{j\in[J]} \pi_j\text{Normal}(s_g|\mu_0\mathbb{1}, C_j) \\ &= \sum_{g\in[G]} \frac{1}{\sum_{j\in[J]}\pi_j\text{Normal}(s_g|\mu_0\mathbb{1}, C_j)} \frac{\partial}{\partial\pi_{j_0}} \sum_{j\in[J]} \pi_j\text{Normal}(s_g|\mu_0\mathbb{1}, C_j) \\ &= \sum_{g\in[G]} \frac{\text{Normal}(s_g|\mu_0\mathbb{1}, C_{j_0})}{\sum_{j\in[J]}\pi_j\text{Normal}(s_g|\mu_0\mathbb{1}, C_j)}\end{aligned}$$

For any other parameter $\eta \in \{p_{k,t}, \mu_0, \tau_0^2, \tau^2\}$,

$$\begin{aligned}\frac{\partial \ell}{\partial \eta} &= \sum_{g\in[G]} \frac{\partial}{\partial \eta} \log \sum_{j\in[J]} \pi_j \text{Normal}(s_g|\mu_0\mathbb{1}, C_j) \\ &= \sum_{g\in[G]} \frac{1}{\sum_{j\in[J]} \pi_j \text{Normal}(s_g|\mu_0\mathbb{1}, C_j)} \sum_{j\in[J]} \pi_j \frac{\partial}{\partial \eta} \text{Normal}(s_g|\mu_0\mathbb{1}, C_j)\end{aligned}$$

The gradient of the normal is as follows, letting $\tilde{x}_g = x_g - \mu_0\mathbb{1}$:

$$\begin{aligned}\frac{\partial \text{Normal}(s_g|\mu_0\mathbb{1}, C_j)}{\partial \eta} &= (2\pi)^{-n/2}\Big[\Big(\frac{\partial}{\partial \eta}|C_j|^{-1/2}\Big)\exp\Big(-\frac{1}{2}\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big) + |C_j|^{-1/2}\frac{\partial}{\partial \eta}\exp\Big(-\frac{1}{2}\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big)\Big] \\ &= (2\pi)^{-n/2}\Big[\Big(-\frac{1}{2}|C_j|^{-3/2}\frac{\partial}{\partial \eta}|C_j|\Big)\exp\Big(-\frac{1}{2}\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big) + |C_j|^{-1/2}\exp\Big(-\frac{1}{2}\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big)\Big(-\frac{1}{2}\Big)\frac{\partial}{\partial \eta}\Big(\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big)\Big] \\ &= (2\pi)^{-n/2}|C_j|^{-1/2}\exp\Big(-\frac{1}{2}\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big)\Big[-\frac{1}{2}|C_j|^{-1}\frac{\partial}{\partial \eta}|C_j| - \frac{1}{2}\frac{\partial}{\partial \eta}\Big(\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big)\Big] \\ &= -\frac{1}{2}\Big[|C_j|^{-1}\frac{\partial |C_j|}{\partial \eta} + \frac{\partial}{\partial \eta}\Big(\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big)\Big]\text{Normal}(s_g|\mu_0\mathbb{1}, C_j)\end{aligned}$$

The gradient of the normal is as follows, letting $\langle x_g, x_g\rangle_j = (x_g - \mu_0\mathbb{1})' C_j^{-1}(x_g - \mu_0\mathbb{1})$:

$$\begin{aligned}\frac{\partial \text{Normal}(s_g|\mu_0\mathbb{1}, C_j)}{\partial \eta} &= (2\pi)^{-n/2}\Big[\Big(\frac{\partial}{\partial \eta}|C_j|^{-1/2}\Big)\exp\Big(-\frac{1}{2}\langle x_g, x_g\rangle_j\Big) + |C_j|^{-1/2}\frac{\partial}{\partial \eta}\exp\Big(-\frac{1}{2}\langle x_g, x_g\rangle_j\Big)\Big] \\ &= (2\pi)^{-n/2}\Big[\Big(-\frac{1}{2}|C_j|^{-3/2}\frac{\partial}{\partial \eta}|C_j|\Big)\exp\Big(-\frac{1}{2}\langle x_g, x_g\rangle_j\Big) + |C_j|^{-1/2}\exp\Big(-\frac{1}{2}\langle x_g, x_g\rangle_j\Big)\Big(-\frac{1}{2}\Big)\frac{\partial}{\partial \eta}\Big(\langle x_g, x_g\rangle_j\Big)\Big] \\ &= (2\pi)^{-n/2}|C_j|^{-1/2}\exp\Big(-\frac{1}{2}\langle x_g, x_g\rangle_j\Big)\Big[-\frac{1}{2}|C_j|^{-1}\frac{\partial}{\partial \eta}|C_j| - \frac{1}{2}\frac{\partial}{\partial \eta}\langle x_g, x_g\rangle_j\Big] \\ &= -\frac{1}{2}\Big(|C_j|^{-1}\frac{\partial |C_j|}{\partial \eta} + \frac{\partial \langle x_g, x_g\rangle_j}{\partial \eta}\Big)\text{Normal}(s_g|\mu_0\mathbb{1}, C_j)\end{aligned}$$

For $\eta = \mu_0$, $\frac{\partial}{\partial \mu_0}|C_j| = 0$ and

$$\begin{aligned}\frac{\partial}{\partial \mu_0}\Big(\tilde{x}_g' C_j^{-1}\tilde{x}_g\Big) &= \frac{\partial}{\partial \mu_0}\Big((x_g - \mu_0\mathbb{1})' C_j^{-1}(x_g - \mu_0\mathbb{1})\Big) \\ &= \frac{\partial}{\partial \mu_0}\Big(x_g' C_j^{-1}x_g - 2\mu_0\mathbb{1}' C_j^{-1}x_g + \mu_0^2\mathbb{1}' C_j^{-1}\mathbb{1}\Big) \\ &= 2(\mu_0\mathbb{1}' C_j^{-1}\mathbb{1} - \mathbb{1}' C_j^{-1}x_g)\end{aligned}$$

For $\eta \in \{p_{k,t}, \tau_0^2, \tau^2\}$, we use the block LDL' decomposition as follows. First, by Claim 9

$$\begin{aligned}\frac{\partial |C_j|}{\partial \eta} &= \frac{\partial}{\partial \eta}\prod_{k=1}^{\Sigma}\alpha_k^{m_k-1}\gamma_k \\ &= \sum_{k=1}^{\Sigma}\Big(\prod_{k'\neq k}\alpha_{k'}^{m_{k'}-1}\gamma_{k'}\Big)\Big((m_k-1)\alpha_k^{m_k-2}\frac{\partial \alpha_k}{\partial \eta}\gamma_k + \alpha_k^{m_k-1}\frac{\partial \gamma_k}{\partial \eta}\Big)\end{aligned}$$

Next, by Corollary 10, $\tilde{x}_g' C_j^{-1}\tilde{x}_g = y' D^{-1}y$, where $y = L^{-1}\tilde{x}_g$ and $C_j = LDL'$ is the block LDL' decomposition. Plugging in the formula for $y'D^{-1}y$ from Claim 5, we get

$$\begin{aligned}
\frac{\partial(\tilde{x}_g' C_j^{-1}\tilde{x}_g)}{\partial\eta} &= \frac{\partial(y'D^{-1}y)}{\partial\eta} \\
&= \frac{\partial}{\partial\eta}\Big(\sum_{k=1}^{\Sigma}\frac{y_k'y_k}{\alpha_k} - \frac{\lambda_{k,k}(\mathbb{1}'y_k)^2}{\alpha_k}\Big) \\
&= \sum_{k=1}^{\Sigma}\Big[\frac{\partial}{\partial\eta}\Big(y_k'y_k - \lambda_{k,k}(\mathbb{1}'y_k)^2\Big)\Big](1/\alpha_k) + \Big(y_k'y_k - \lambda_{k,k}(\mathbb{1}'y_k)^2\Big)\frac{\partial(1/\alpha_k)}{\partial\eta} \\
&= \sum_{k=1}^{\Sigma}\Big(\frac{\partial y_k'y_k}{\partial\eta} - \frac{\partial\lambda_{k,k}}{\partial\eta}(\mathbb{1}'y_k)^2 - 2\lambda_{k,k}(\mathbb{1}'y_k)\frac{\partial\mathbb{1}'y_k}{\partial\eta}\Big)(1/\alpha_k) - \frac{y_k'y_k - \lambda_{k,k}(\mathbb{1}'y_k)^2}{\alpha_k^2}\frac{\partial\alpha_k}{\partial\eta}
\end{aligned}$$

Next, the inner products involving y_k have partial derivatives as follows. From Eq. (2),

$$\begin{aligned}
\frac{\partial\mathbb{1}'y_k}{\partial\eta} &= \frac{\partial}{\partial\eta}\Big(\mathbb{1}'x_k - m_k\mu_0 - \sum_{l<k} m_k\lambda_{k,l}\mathbb{1}'y_l\Big) \\
&= -\sum_{l<k} m_k\frac{\partial}{\partial\eta}\Big(\lambda_{k,l}\mathbb{1}'y_l\Big) \\
&= -\sum_{l<k} m_k\Big(\frac{\partial\lambda_{k,l}}{\partial\eta}\mathbb{1}'y_l + \lambda_{k,l}\frac{\partial\mathbb{1}'y_l}{\partial\eta}\Big)
\end{aligned}$$

and from Eq. (3) and the discussion around it

$$\begin{aligned}
\frac{\partial y_k'y_k}{\partial\eta} &= \frac{\partial}{\partial\eta}\Big(z_k'z_k - 2(\mathbb{1}'z_k)\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l) + m_k\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big)^2\Big) \\
&= -\frac{\partial}{\partial\eta}\Big(2(\mathbb{1}'z_k)\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big) + \frac{\partial}{\partial\eta}\Big(m_k\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big)^2\Big) \\
&= -2\frac{\partial\mathbb{1}'z_k}{\partial\eta}\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big) - 2(\mathbb{1}'z_k)\frac{\partial}{\partial\eta}\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big) + 2m_k\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big)\frac{\partial}{\partial\eta}\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big) \\
&= -2(\mathbb{1}'z_k)\frac{\partial}{\partial\eta}\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big) + 2m_k\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big)\frac{\partial}{\partial\eta}\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big) \\
&= \Big(-2(\mathbb{1}'z_k) + 2m_k\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big)\frac{\partial}{\partial\eta}\Big(\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big) \\
&= \Big(-2(\mathbb{1}'z_k) + 2m_k\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big)\sum_{l<k}\Big(\frac{\partial\lambda_{k,l}}{\partial\eta}\mathbb{1}'y_l + \lambda_{k,l}\frac{\partial\mathbb{1}'y_l}{\partial\eta}\Big) \\
&= \Big(2(\mathbb{1}'z_k)m_k^{-1} - 2\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\Big)\frac{\partial\mathbb{1}'y_k}{\partial\eta}
\end{aligned}$$

where we have used the fact that $z_k := x_k - \mu_0\mathbb{1}$ has partials $\frac{\partial z_k'z_k}{\partial\eta} = \frac{\partial\mathbb{1}'z_k}{\partial\eta} = 0$.

Next, we find the partial derivatives of γ_k, $\delta_{k,l}$, and $\lambda_{k,l}$ from the LDL' decomposition:

- $\frac{\partial\gamma_k}{\partial\eta} = \frac{\partial\alpha_k}{\partial\eta} + m_k\frac{\partial\delta_{k,k}}{\partial\eta}$ $\quad k\in[\Sigma]$.
- $\frac{\partial\delta_{k,l}}{\partial\eta} = \frac{\partial\beta_{k,l}}{\partial\eta} - \sum_{h<l}\big(\frac{\partial\delta_{k,h}}{\partial\eta}\delta_{l,h}m_h\gamma_h^{-1} + \delta_{k,h}\frac{\partial\delta_{l,h}}{\partial\eta}m_h\gamma_h^{-1} - \delta_{k,h}\delta_{l,h}m_h\gamma_h^{-2}\frac{\partial\gamma_h}{\partial\eta}\big)$ $\quad k\in[\Sigma], l\le k$.
- $\frac{\partial\lambda_{k,l}}{\partial\eta} = \frac{\partial\delta_{k,l}}{\partial\eta}\gamma_l^{-1} - \delta_{k,l}\gamma_l^{-2}\frac{\partial\gamma_l}{\partial\eta}$ $\quad k\in[\Sigma], l\le k$.

Finally, the quantities α_k and $\beta_{k,l}$ have partial derivatives with respect to η as follows:

- $\frac{\partial \alpha_k}{\partial \tau^2} = \sum_{t=1}^{T} p_{k,t}^2$.
- $\frac{\partial \alpha_k}{\partial \tau_0^2} = 0$.
- $\frac{\partial \alpha_k}{\partial p_{k,t}} = 2\tau^2 p_{k,t}$.
- $\frac{\partial \alpha_k}{\partial p_{k',t}} = 0$ if $k \neq k'$.

and

- $\frac{\partial \beta_{k,l}}{\partial \tau^2} = 0$.
- $\frac{\partial \beta_{k,l}}{\partial \tau_0^2} = \sum_{\mathscr{T} \in \mathbb{T}_j} (\sum_{t \in \mathscr{T}} p_{k,t})(\sum_{t \in \mathscr{T}} p_{l,t})$.
- $\frac{\partial \beta_{k,l}}{\partial p_{k',t}} = 0$ if $k \neq k'$ and $l \neq k'$.
- $\frac{\partial \beta_{k,l}}{\partial p_{k,t}} = \tau_0^2 \sum_{t' \in \mathscr{T}} p_{l,t'}$ where $\mathscr{T}$ is the element of $\mathbb{T}_j$ that contains t.
- $\frac{\partial \beta_{k,l}}{\partial p_{l,t}} = \tau_0^2 \sum_{t' \in \mathscr{T}} p_{k,t'}$ where $\mathscr{T}$ is the element of $\mathbb{T}_j$ that contains t.

1.3.1 Note on the slope constraints

We want to reformulate the slope constraints

$$p_{1,t} \leq p_{2,t} \leq \cdots \leq p_{\Sigma,t} \tag{1}$$

as a set of constraints

$$\text{lower}_i \leq \text{cons}_i(p) \leq \text{upper}_i$$

Indeed, note that $p_{k,t} \leq p_{k+1,t}$ iff $p_{k+1,t} - p_{k,t} \geq 0$, so Eq. (1) is equivalent to

$$\begin{aligned} 0 &\leq p_{2,t} - p_{1,t} \leq \infty \\ 0 &\leq p_{3,t} - p_{2,t} \leq \infty \\ &\vdots \\ 0 &\leq p_{\Sigma,t} - p_{\Sigma-1,t} \leq \infty \end{aligned}$$

as desired.

1.4 Proofs from Section 1.2

Proof of Claim 2. From Claim 1, one can see that the covariance matrix C has the following form:

$$C = \tau^2 \begin{bmatrix} \sum_{t=1}^{T} p_{\tilde{\sigma}_1,t}^2 & 0 & \cdots & 0 \\ 0 & \sum_{t=1}^{T} p_{\tilde{\sigma}_2,t}^2 & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \sum_{t=1}^{T} p_{\tilde{\sigma}_n,t}^2 \end{bmatrix} + \tau_0^2 \begin{bmatrix} \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_1,\mathscr{T}}^2 & \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_1,\mathscr{T}} p_{\tilde{\sigma}_2,\mathscr{T}} & \cdots & \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_1,\mathscr{T}} p_{\tilde{\sigma}_n,\mathscr{T}} \\ \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_2,\mathscr{T}} p_{\tilde{\sigma}_1,\mathscr{T}} & \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_2,\mathscr{T}}^2 & \cdots & \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_2,\mathscr{T}} p_{\tilde{\sigma}_n,\mathscr{T}} \\ & & \ddots & \\ \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_n,\mathscr{T}} p_{\tilde{\sigma}_1,\mathscr{T}} & \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_n,\mathscr{T}} p_{\tilde{\sigma}_2,\mathscr{T}} & \cdots & \sum_{\mathscr{T}\in\mathbb{T}_j} p_{\tilde{\sigma}_n,\mathscr{T}}^2 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{\tilde{\sigma}_1} & 0 & \cdots & 0 \\ 0 & \alpha_{\tilde{\sigma}_2} & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \alpha_{\tilde{\sigma}_n} \end{bmatrix} + \begin{bmatrix} \beta_{\tilde{\sigma}_1,\tilde{\sigma}_1} & \beta_{\tilde{\sigma}_1,\tilde{\sigma}_2} & \cdots & \beta_{\tilde{\sigma}_1,\tilde{\sigma}_n} \\ \beta_{\tilde{\sigma}_2,\tilde{\sigma}_1} & \beta_{\tilde{\sigma}_2,\tilde{\sigma}_2} & \cdots & \beta_{\tilde{\sigma}_2,\tilde{\sigma}_n} \\ & & \ddots & \\ \beta_{\tilde{\sigma}_n,\tilde{\sigma}_1} & \beta_{\tilde{\sigma}_n,\tilde{\sigma}_2} & \cdots & \beta_{\tilde{\sigma}_n,\tilde{\sigma}_n} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1 I_{m_1} & 0 & \cdots & 0 \\ 0 & \alpha_2 I_{m_2} & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \alpha_\Sigma I_{m_\Sigma} \end{bmatrix} + \begin{bmatrix} \beta_{1,1} J_{m_1,m_1} & \beta_{1,2} J_{m_1,m_2} & \cdots & \beta_{1,\Sigma} J_{m_1,m_\Sigma} \\ \beta_{2,1} J_{m_2,m_1} & \beta_{2,2} J_{m_2,m_2} & \cdots & \beta_{2,\Sigma} J_{m_2,m_\Sigma} \\ & & \ddots & \\ \beta_{\Sigma,1} J_{m_\Sigma,m_1} & \beta_{\Sigma,2} J_{m_\Sigma,m_2} & \cdots & \beta_{\Sigma,\Sigma} J_{m_\Sigma,m_\Sigma} \end{bmatrix}$$

Claim 2 now follows from Lemmas 12 and 13 below. ■

Lemma 12. *The block off-diagonal entries of LDL' equal the corresponding block off-diagonal entries of C. In other words, for all $k > k'$,*

$$(LDL')_{k,k'} = \beta_{k,k'} J_{m_k,m_{k'}}$$

Proof.

$$
\begin{aligned}
(LDL')_{k,k'} &= \sum_{l=1}^{\Sigma} L_{k,l} \sum_{h=1}^{\Sigma} D_{l,h}(L')_{h,k'} \quad \text{(matrix multiplication by blocks)} \\
&= \sum_{l=1}^{\Sigma} L_{k,l} D_{l,l}(L')_{l,k'} \quad (\text{since } D_{l,h} = 0 \text{ for all } h \neq l) \\
&= \sum_{l=1}^{\Sigma} L_{k,l} D_{l,l} L'_{k',l} \quad (\text{since } (L')_{l,k'} = (L_{k',l})') \\
&= \sum_{l<k'} L_{k,l} D_{l,l} L'_{k',l} + L_{k,k'} D_{k',k'} \quad (\text{since } L_{k',l} = 0 \text{ for all } l > k' \text{ and } L_{k',k'} = I_{m_{k'}}) \\
&= \Big(\sum_{l<k'} (\lambda_{k,l} J_{m_k,m_l})(\alpha_l I_{m_l} + \delta_{l,l} J_{m_l,m_l})(\lambda_{k',l} J_{m_l,m'_k})\Big) + (\lambda_{k,k'} J_{m_k,m_{k'}})(\alpha_{k'} I_{m_{k'}} + \delta_{k',k'} J_{m_{k'},m_{k'}}) \quad \text{(subs. dfns)} \\
&= \Big[\Big(\sum_{l<k'} \alpha_l \lambda_{k,l} \lambda_{k',l} m_l + \delta_{l,l} \lambda_{k,l} \lambda_{k',l} m_l^2\Big) + \Big(\alpha_{k'} \lambda_{k,k'} + \delta_{k',k'} \lambda_{k,k'} m_{k'}\Big)\Big] J_{m_k,m_{k'}} \\
&\qquad (\text{since } J_{n,m} J_{m,p} = m J_{n,p} \text{ and } J_{n,m} J_{m,p} J_{p,q} = mp J_{n,q}) \\
&= \Big[\Big(\sum_{l<k'} \lambda_{k,l} \lambda_{k',l} m_l (\alpha_l + \delta_{l,l} m_l)\Big) + \Big(\lambda_{k,k'} (\alpha_{k'} + \delta_{k',k'} m_{k'})\Big)\Big] J_{m_k,m_{k'}} \quad \text{(factor)} \\
&= \Big[\Big(\sum_{l<k'} \delta_{k,l} \delta_{k',l} m_l (\alpha_l + \delta_{l,l} m_l)/\gamma_l^2\Big) + \Big(\delta_{k,k'} (\alpha_{k'} + \delta_{k',k'} m_{k'})/\gamma_{k'}\Big)\Big] J_{m_k,m_{k'}} \quad (\text{subs. dfn. of } \lambda_{k,l}) \\
&= \Big[\Big(\sum_{l<k'} \delta_{k,l} \delta_{k',l} m_l/\gamma_l\Big) + \delta_{k,k'}\Big] J_{m_k,m_{k'}} \quad (\text{simplify via dfn. of } \gamma_{k,l}) \\
&= \Big[\Big(\sum_{l<k'} \delta_{k,l} \delta_{k',l} m_l/\gamma_l\Big) + \Big(\beta_{k,k'} - \sum_{l<k'} \delta_{k,l} \delta_{k',l} m_l/\gamma_l\Big)\Big] J_{m_k,m_{k'}} \quad (\text{subs. dfn. of } \delta_{k,k'}) \\
&= \beta_{k,k'} J_{m_k,m_{k'}} \quad \text{(cancel)}
\end{aligned}
$$

as claimed. ■

Lemma 13. *The block diagonal entries of LDL' equal the corresponding block diagonal entries of C. In other words,*

$$
(LDL')_{k,k} = \alpha_k I_{m_k} + \beta_{k,k} J_{m_k,m_k}
$$

Proof.

$$
\begin{aligned}
(LDL')_{k,k} &= \sum_{l=1}^{\Sigma} L_{k,l} \sum_{h=1}^{\Sigma} D_{l,h} (L')_{h,k} \quad \text{(matrix multiplication by blocks)} \\
&= \sum_{l=1}^{\Sigma} L_{k,l} D_{l,l} (L')_{l,k} \quad (\text{since } D_{l,h} = 0 \text{ for all } h \neq l) \\
&= \sum_{l=1}^{\Sigma} L_{k,l} D_{l,l} L'_{k,l} \quad (\text{since } (L')_{l,k} = (L_{k,l})') \\
&= \sum_{l<k} L_{k,l} D_{l,l} L'_{k,l} + D_{k,k} \quad (\text{since } L_{k,l} = 0 \text{ for all } l > k \text{ and } L_{k,k} = I_{m_k}) \\
&= \Big(\sum_{l<k} (\lambda_{k,l} J_{m_k,m_l})(\alpha_l I_{m_l} + \delta_{l,l} J_{m_l,m_l})(\lambda_{k,l} J_{m_l,m_k}) \Big) + (\alpha_k I_{m_k} + \delta_{k,k} J_{m_k,m_k}) \quad \text{(subs. dfns)} \\
&= \Big[\Big(\sum_{l<k} \alpha_l \lambda_{k,l}^2 m_l + \delta_{l,l} \lambda_{k,l}^2 m_l^2 \Big) + \delta_{k,k} \Big] J_{m_k,m_k} + \alpha_k I_{m_k} \quad (\text{since } J_{n,m} J_{m,p} = m J_{n,p} \text{ and } J_{n,m} J_{m,p} J_{p,q} = mp J_{n,q}) \\
&= \Big[\Big(\sum_{l<k} \lambda_{k,l}^2 m_l (\alpha_l + \delta_{l,l} m_l) \Big) + \delta_{k,k} \Big] J_{m_k,m_k} + \alpha_k I_{m_k} \quad \text{(factor)} \\
&= \Big[\Big(\sum_{l<k} \delta_{k,l}^2 m_l (\alpha_l + \delta_{l,l} m_l) / \gamma_l^2 \Big) + \delta_{k,k} \Big] J_{m_k,m_k} + \alpha_k I_{m_k} \quad (\text{subs. dfn. of } \lambda_{k,l}) \\
&= \Big[\Big(\sum_{l<k} \delta_{k,l}^2 m_l / \gamma_l \Big) + \delta_{k,k} \Big] J_{m_k,m_k} + \alpha_k I_{m_k} \quad (\text{simplify via dfn. of } \gamma_{k,l}) \\
&= \Big[\Big(\sum_{l<k} \delta_{k,l}^2 m_l / \gamma_l \Big) + \Big(\beta_{k,k} - \sum_{l<k} \delta_{k,l}^2 m_l / \gamma_l \Big) \Big] J_{m_k,m_k} + \alpha_k I_{m_k} \quad (\text{subs. dfn. of } \delta_{k,k'}) \\
&= \beta_{k,k} J_{m_k,m_k} + \alpha_k I_{m_k} \quad \text{(cancel)}
\end{aligned}
$$

as claimed. ∎

Proof of Claim 3. Note that y_k is such that $(Ly)_k = x_k$. In other words, since L is block lower triangular and has identity matrices on the diagonal,

$$L_{k,1} y_1 + L_{k,2} y_2 + \cdots + L_{k,k-1} y_{k-1} + y_k = x_k.$$

So, substituting the definition $L_{k,l} = \lambda_{k,l} J_{m_k,m_l}$,

$$
\begin{aligned}
y_k &= x_k - \sum_{l<k} L_{k,l} y_l \\
&= x_k - \sum_{l<k} \lambda_{k,l} J_{m_k,m_l} y_l \\
&= x_k - \sum_{l<k} \lambda_{k,l} \Big(\sum_{i=1}^{m_l} y_l \Big) \mathbb{1}_{m_k}
\end{aligned}
$$

as claimed. ∎

Proof of Claim 4. Note that DD^{-1} is a diagonal matrix with blocks $D_k D_k^{-1}$, where

$$
\begin{aligned}
D_k D_k^{-1} &= \Big(\alpha_k I_{m_k} + \delta_{k,k} J_{m_k,m_k} \Big) \Big(\frac{1}{\alpha_k} I_{m_k} - \frac{\delta_{k,k}}{\alpha_k \gamma_k} J_{m_k,m_k} \Big) \\
&= \frac{\alpha_k}{\alpha_k} I_{m_k} - \frac{\alpha_k \delta_{k,k}}{\alpha_k \gamma_k} J_{m_k,m_k} + \frac{\delta_{k,k}}{\alpha_k} J_{m_k,m_k} - \frac{m_k \delta_{k_k}^2}{\alpha_k \gamma_k} J_{m_k,m_k}
\end{aligned}
$$

We just need to check that the coefficients of the last three terms sum to 0. Indeed,

$$-\frac{\alpha_k\delta_{k,k}}{\alpha_k\gamma_k}+\frac{\delta_{k,k}}{\alpha_k}-\frac{m_k\delta_{k_k}^2}{\alpha_k\gamma_k}=\frac{\delta_{k,k}}{\alpha_k}-\frac{(\alpha_k+m_k\delta_{k,k})\delta_{k_k}}{\alpha_k\gamma_k}=0,$$

since $\gamma_k=\alpha_k+m_k\delta_{k,k}$. ■

Proof of Claim 5. Note that $y'D^{-1}y=\sum_{k=1}^{\Sigma}y_k'(D^{-1})_{k,k}y_k$, since $(D^{-1})_{k,l}=0$ if $k\neq l$. Substituting $(D^{-1})_{k,k}$ from Claim 4,

$$\begin{aligned}y'D^{-1}y&=\sum_{k=1}^{\Sigma}y_k'\Big(\frac{1}{\alpha_k}I_{m_k}-\frac{\delta_{k,k}}{\alpha_k\gamma_k}J_{m_k,m_k}\Big)y_k\\&=\sum_{k=1}^{\Sigma}\frac{y_k'y_k}{\alpha_k}-\frac{\delta_{k,k}}{\alpha_k\gamma_k}\Big(\sum_{i=1}^{m_k}y_{k,i}\Big)^2\end{aligned}$$

since

$$y_k'J_{m_k,m_k}y_k=\sum_{i=1}^{m_k}y_{k,i}(J_{m_k,m_k}y_k)_i=\sum_{i=1}^{m_k}y_{k,i}\sum_{i'=1}^{m_k}(J_{m_k,m_k})_{i,i'}y_{k,i'}=\Big(\sum_{i=1}^{m_k}y_{k,i}\Big)^2.$$

■

Proof of Claim 9. Note that $\det C=\det(LDL')=(\det L)^2\det D=\det D$, since L is block lower triangular with identity matrices on the diagonal. The determinant of a block diagonal matrix is the product of the determinant of the blocks. In general,

$$\det(aI_n+bJ_n)=a^{n-1}(a+nb),$$

so, in our case,

$$\det D_{k,k}=\det(\alpha_kI_{m_k}+\delta_{k,k}J_{m_k,m_k})=\alpha_k^{m_k-1}(\alpha_k+m_k\delta_{k,k})$$

which gives the claimed formula. ■

Proof of Claim 11. To establish the claim, it is sufficient to write $y'D^{-1}y$, where $y=L^{-1}(x-\mu_0\mathbb{1})$, in terms of the sufficient statistics. To do this, it is sufficient (according to the formula in Claim 5) to show that $\mathbb{1}'y_k$ and $y_k'y_k$ can be written in terms of the sufficient statistics.

First, we show that $(\mathbb{1}'y_k)_{k=1}^{\Sigma}$ can be written in terms of the sufficient statistics. Since, according to Claim 3, $y_1=x_1-\mu_0\mathbb{1}$, we can write $\mathbb{1}'y_1=\mathbb{1}'(x_1-\mu_0\mathbb{1})=\mathbb{1}'x_1-m_1\mu_0$, which is in terms of the sufficient statistics. For induction, assume that $\mathbb{1}'y_1,\dots,\mathbb{1}'y_{k-1}$ can be written in terms of the sufficient statistics. Then, again according to Claim 3,

$$y_k=x_k-\mu_0\mathbb{1}-\sum_{l<k}\lambda_{k,l}\Big(\sum_{i=1}^{m_l}y_{l,i}\Big)\mathbb{1}_{m_k}=x_k-\mu_0\mathbb{1}-\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\mathbb{1}$$

so

$$\begin{aligned}\mathbb{1}'y_k&=\mathbb{1}'\Big(x_k-\mu_0\mathbb{1}-\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)\mathbb{1}\Big)\\&=\mathbb{1}'x_k-m_k\mu_0-\sum_{l<k}\lambda_{k,l}(\mathbb{1}'y_l)(\mathbb{1}'\mathbb{1})\\&=\mathbb{1}'x_k-m_k\mu_0-\sum_{l<k}m_k\lambda_{k,l}\mathbb{1}'y_l\end{aligned}\qquad(2)$$

which can be written in terms of the sufficient statistics using the inductive assumption.

Next, we show that $(y_k' y_k)_{k=1}^{\Sigma}$ can be written in terms of the sufficient statistics. Note that $y_1' y_1 = (x_1 - \mu_0 \mathbb{1})'(x_1 - \mu_0 \mathbb{1}) = x_1' x_1 - 2\mu_0 \mathbb{1}' x_1 + m_1 \mu_0^2$, which is in terms of the sufficient statistics. For induction, assume that $y_1' y_1, \ldots, y_{k-1}' y_{k-1}$ can be written in terms of the sufficient statistics. Then, letting $z_k = x_k - \mu_0 \mathbb{1}$,

$$\begin{aligned}
y_k' y_k &= \Big(z_k - \sum_{l<k} \lambda_{k,l} (\mathbb{1}' y_l) \mathbb{1}\Big)' \Big(z_k - \sum_{l<k} \lambda_{k,l} (\mathbb{1}' y_l) \mathbb{1}\Big) \\
&= z_k' z_k - 2 z_k' \Big(\sum_{l<k} \lambda_{k,l} (\mathbb{1}' y_l) \mathbb{1}\Big) + \Big(\sum_{l<k} \lambda_{k,l} (\mathbb{1}' y_l) \mathbb{1}\Big)' \Big(\sum_{l<k} \lambda_{k,l} (\mathbb{1}' y_l) \mathbb{1}\Big) \\
&= z_k' z_k - 2(\mathbb{1}' z_k) \sum_{l<k} \lambda_{k,l} (\mathbb{1}' y_l) + m_k \Big(\sum_{l<k} \lambda_{k,l} (\mathbb{1}' y_l)\Big)^2. \qquad (3)
\end{aligned}$$

Note that $z_k' z_k = (x_k - \mu_0 \mathbb{1})'(x_k - \mu_0 \mathbb{1}) = x_k' x_k - 2\mu_0 \mathbb{1}' x_k + m_k \mu_0^2$, and $\mathbb{1}' z_k = \mathbb{1}'(x_k - \mu_0 \mathbb{1}) = \mathbb{1}' x_k - m_k \mu_0$. Thus, using the inductive assumption, $y_k' y_k$ can be written in terms of the sufficient statistics. ■