Let $\boldsymbol{\beta}_k = (\beta_{k1}, ..., \beta_{kP})^T$ be the kth component's coefficient vector. We write the log-likelihood as

$$l_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log \sum_{k=1}^{K} \pi_k \phi\left(y_i; \boldsymbol{x}_i^T \boldsymbol{\beta}_k, \sigma_k^2\right).$$

Define the penalty function as

$$\text{pen}_n(\boldsymbol{\theta}) = \sum_{k=1}^{K} \pi_k \left[\sum_{j=1}^{P} p_{nk}(\beta_{kj})\right]$$

where $p_{nk}(\beta)$ is defined via its derivative

$$p'_{nk}(\beta) = \gamma_{nk}\sqrt{n}\mathbb{I}\left(\sqrt{n}|\beta| \le \gamma_{nk}\right) + \frac{\sqrt{n}\left(a\gamma_{nk} - \sqrt{n}|\beta|\right)_+}{a-1}\mathbb{I}\left(\sqrt{n}|\beta| > \gamma_{nk}\right).$$

Luckily, we can approximate $p_{nk}(\beta)$ by

$$\tilde{p}_{nk}(\beta) = p_{nk}\left(\tilde{\beta}\right) + \frac{p'_{nk}\left(\tilde{\beta}\right)}{2\tilde{\beta}}\left(\beta^2 - \tilde{\beta}^2\right)$$

arround the neighbourhood of $\tilde{\beta}$.

We wish to maximize the penalized objective

$$h_n(\boldsymbol{\theta}) = l_n(\boldsymbol{\theta}) - \text{pen}_n(\boldsymbol{\theta}).$$

We can approximate $l_n(\boldsymbol{\theta})$ by

$$\begin{aligned}
Q\left(\boldsymbol{\theta}; \boldsymbol{\theta}^{(m)}\right) &= \sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\left[\log \pi_k + \log \phi\left(y_i; \boldsymbol{x}_i^T \boldsymbol{\beta}_k, \sigma_k^2\right)\right] \\
&= \sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\log \pi_k \\
&- \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\log\left(\sigma_k^2\right) \\
&- \frac{1}{2\sigma_k^2}\sum_{i=1}^{n}\sum_{k=1}^{K} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\left(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_k\right)^2 \\
&+ C
\end{aligned}$$

where C is a constant.

Here

$$\tau_{ik}(\boldsymbol{\theta}) = \frac{\pi_k \phi\left(y_i; \boldsymbol{x}_i^T \boldsymbol{\beta}_k, \sigma_k^2\right)}{\sum_{k'=1}^{K} \pi_{k'} \phi\left(y_i; \boldsymbol{x}_i^T \boldsymbol{\beta}_{k'}, \sigma_{k'}^2\right)}.$$

Similarly, we approximate $\text{pen}_n(\boldsymbol{\theta})$ by

$$\begin{aligned}
q\left(\boldsymbol{\theta}; \boldsymbol{\theta}^{(m)}\right) &= \sum_{k=1}^{K} \pi_k \left[\sum_{j=1}^{P} \tilde{p}_{nk}(\beta_{kj})\right] \\
&= \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} \left[p_{nk}\left(\beta_{kj}^{(m)}\right) + \frac{p'_{nk}\left(\beta_{kj}^{(m)}\right)}{2\beta_{kj}^{(m)}}\left(\beta_{kj}^2 - \beta_{kj}^{(m)2}\right)\right] \\
&= \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} p_{nk}\left(\beta_{jk}^{(m)}\right) + \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} \frac{p'_{nk}\left(\beta_{kj}^{(m)}\right)}{2\beta_{kj}^{(m)}}\left(\beta_{kj}^2 - \beta_{kj}^{(m)2}\right) \\
&= D_1 + \sum_{k=1}^{K} \pi_k \left(\boldsymbol{\beta}_k - \boldsymbol{\beta}_k^{(m)}\right)^T \boldsymbol{W}_k \left(\boldsymbol{\beta}_k - \boldsymbol{\beta}_k^{(m)}\right),
\end{aligned}$$

where

$$D_1 = \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} p_{nk}\left(\beta_{jk}^{(m)}\right)$$

and

$$\boldsymbol{W}_k = \operatorname{diag}\left(\frac{p'_{nk}\left(\beta_{k1}^{(m)}\right)}{2\beta_{k1}^{(m)}}, \ldots, \frac{p'_{nk}\left(\beta_{kP}^{(m)}\right)}{2\beta_{kP}^{(m)}}\right).$$

Thus, we wish to maximize the approximate function

$$\eta_n\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(m)}\right) = Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(m)}\right) - q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(m)}\right).$$

Differentiating $Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(m)}\right)$ with respect to $\boldsymbol{\beta}_k$ yields

$$\frac{\partial Q}{\partial \boldsymbol{\beta}_k} = \frac{1}{\sigma_k^2}\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\boldsymbol{x}_i\left(y_i - \boldsymbol{x}_i^T\boldsymbol{\beta}_k\right)$$

and $q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(m)}\right)$ with respects to $\boldsymbol{\beta}_k$ yields

$$\frac{\partial q}{\partial \boldsymbol{\beta}_k} = 2\pi_k \boldsymbol{W}_k\left(\boldsymbol{\beta}_k - \boldsymbol{\beta}_k^{(m)}\right).$$

Together, we wish to solve

$$\begin{aligned}
\frac{\partial Q}{\partial \boldsymbol{\beta}_k} - \frac{\partial q}{\partial \boldsymbol{\beta}_k} &= \mathbf{0} \\
\frac{1}{\sigma_k^2}\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\boldsymbol{x}_i\left(y_i - \boldsymbol{x}_i^T\boldsymbol{\beta}_k\right) &= 2\pi_k \boldsymbol{W}_k\left(\boldsymbol{\beta}_k - \boldsymbol{\beta}_k^{(m)}\right) \\
\frac{1}{\sigma_k^2}\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\boldsymbol{x}_i y_i - \frac{1}{\sigma_k^2}\left[\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\boldsymbol{x}_i\boldsymbol{x}_i^T\right]\boldsymbol{\beta}_k &= 2\pi_k \boldsymbol{W}_k\boldsymbol{\beta}_k - 2\pi_k \boldsymbol{W}_k\boldsymbol{\beta}_k^{(m)} \\
\frac{1}{\sigma_k^2}\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\boldsymbol{x}_i y_i + 2\pi_k \boldsymbol{W}_k\boldsymbol{\beta}_k^{(m)} &= \left(2\pi_k \boldsymbol{W}_k + \frac{1}{\sigma_k^2}\left[\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\boldsymbol{x}_i\boldsymbol{x}_i^T\right]\right)\boldsymbol{\beta}_k
\end{aligned}$$

so

$$\boldsymbol{\beta}_k = \left(2\pi_k \boldsymbol{W}_k + \frac{1}{\sigma_k^2}\left[\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\boldsymbol{x}_i\boldsymbol{x}_i^T\right]\right)^{-1}\left[\frac{1}{\sigma_k^2}\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right)\boldsymbol{x}_i y_i + 2\pi_k \boldsymbol{W}_k\boldsymbol{\beta}_k^{(m)}\right].$$

Correction

Similarly, we approximate $\operatorname{pen}_n(\boldsymbol{\theta})$ by

$$\begin{aligned}
q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(m)}\right) &= \sum_{k=1}^{K} \pi_k\left[\sum_{j=1}^{P} \tilde{p}_{nk}\left(\beta_{kj}\right)\right] \\
&= \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P}\left[p_{nk}\left(\beta_{kj}^{(m)}\right) + \frac{p'_{nk}\left(\beta_{kj}^{(m)}\right)}{2\beta_{kj}^{(m)}}\left(\beta_{kj}^2 - \beta_{kj}^{(m)2}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} p_{nk}\left(\beta_{jk}^{(m)}\right) + \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} \frac{p'_{nk}\left(\beta_{kj}^{(m)}\right)}{2\beta_{kj}^{(m)}} \left(\beta_{kj}^{2} - \beta_{kj}^{(m)2}\right) \\
&= \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} p_{nk}\left(\beta_{jk}^{(m)}\right) + \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} \frac{p'_{nk}\left(\beta_{kj}^{(m)}\right)}{2\beta_{kj}^{(m)}} \beta_{kj}^{2} - \sum_{k=1}^{K} \frac{\pi_k}{2} \sum_{j=1}^{P} \frac{p'_{nk}\left(\beta_{kj}^{(m)}\right)}{\beta_{kj}^{(m)}} \beta_{kj}^{(m)2} \\
&= D_1 + \sum_{k=1}^{K} \pi_k \boldsymbol{\beta}_k^T \boldsymbol{W}_k \boldsymbol{\beta}_k,
\end{aligned}
$$

where

$$
D_1 = \sum_{k=1}^{K} \pi_k \sum_{j=1}^{P} p_{nk}\left(\beta_{jk}^{(m)}\right) - \sum_{k=1}^{K} \frac{\pi_k}{2} \sum_{j=1}^{P} \frac{p'_{nk}\left(\beta_{kj}^{(m)}\right)}{\beta_{kj}^{(m)}} \beta_{kj}^{(m)2}
$$

and

$$
\boldsymbol{W}_k = \operatorname{diag}\left(\frac{p'_{nk}\left(\beta_{k1}^{(m)}\right)}{\beta_{k1}^{(m)}}, ..., \frac{p'_{nk}\left(\beta_{kP}^{(m)}\right)}{\beta_{kP}^{(m)}} \right).
$$

Thus, we wish to maximize the approximate function

$$
\eta_n\left(\boldsymbol{\theta}; \boldsymbol{\theta}^{(m)}\right) = Q\left(\boldsymbol{\theta}; \boldsymbol{\theta}^{(m)}\right) - q\left(\boldsymbol{\theta}; \boldsymbol{\theta}^{(m)}\right).
$$

Differentiating $Q\left(\boldsymbol{\theta}; \boldsymbol{\theta}^{(m)}\right)$ with respect to $\boldsymbol{\beta}_k$ yields

$$
\frac{\partial Q}{\partial \boldsymbol{\beta}_k} = \frac{1}{\sigma_k^2} \sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right) \boldsymbol{x}_i \left(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_k\right)
$$

and $q\left(\boldsymbol{\theta}; \boldsymbol{\theta}^{(m)}\right)$ with respects to $\boldsymbol{\beta}_k$ yields

$$
\frac{\partial q}{\partial \boldsymbol{\beta}_k} = \pi_k \boldsymbol{W}_k \boldsymbol{\beta}_k.
$$

Together, we wish to solve

$$
\begin{aligned}
\frac{\partial Q}{\partial \boldsymbol{\beta}_k} - \frac{\partial q}{\partial \boldsymbol{\beta}_k} &= \mathbf{0} \\
\frac{1}{\sigma_k^2} \sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right) \boldsymbol{x}_i \left(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_k\right) &= \pi_k \boldsymbol{W}_k \boldsymbol{\beta}_k \\
\frac{1}{\sigma_k^2} \sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right) \boldsymbol{x}_i y_i - \frac{1}{\sigma_k^2} \left[\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right) \boldsymbol{x}_i \boldsymbol{x}_i^T\right] \boldsymbol{\beta}_k &= \pi_k \boldsymbol{W}_k \boldsymbol{\beta}_k \\
\frac{1}{\sigma_k^2} \sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right) \boldsymbol{x}_i y_i &= \left(\pi_k \boldsymbol{W}_k + \frac{1}{\sigma_k^2} \left[\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right) \boldsymbol{x}_i \boldsymbol{x}_i^T\right]\right) \boldsymbol{\beta}_k
\end{aligned}
$$

so

$$
\boldsymbol{\beta}_k = \left(\pi_k \boldsymbol{W}_k + \frac{1}{\sigma_k^2} \left[\sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right) \boldsymbol{x}_i \boldsymbol{x}_i^T\right]\right)^{-1} \left[\frac{1}{\sigma_k^2} \sum_{i=1}^{n} \tau_{ik}\left(\boldsymbol{\theta}^{(m)}\right) \boldsymbol{x}_i y_i\right].
$$