5.35

$$E_n = -\ln\left\{\sum_{k=1}^{K}\pi_k(\mathbf{x}_n,\mathbf{w})\mathcal{N}\left(\mathbf{t}_n|\boldsymbol{\mu}_k(\mathbf{x}_n,\mathbf{w}),\left(\sigma_k(\mathbf{x}_n,\mathbf{w})^2\mathbf{I}\right)\right)\right\}$$

Since $a^{\mu}_{kl} = \mu_{kln}$,

$$\frac{\partial E_n}{\partial a^{\mu}_{kl}} = -\frac{1}{\left\{\sum_{k=1}^{K}\pi_{kn}\mathcal{N}_{kn}\right\}}\left(\pi_{kn}\frac{\partial \mathcal{N}_{kn}}{\partial a^{\mu}_{kl}}\right)$$

$$\frac{\partial \mathcal{N}_{kn}}{\partial a^{\mu}_{kl}} = \frac{\partial \mathcal{N}_{kn}}{\partial \mu_{kln}}$$

$$= \frac{\partial}{\partial \mu_{kln}}\left(\frac{1}{(2\pi)^{L/2}|\,(\sigma^2_{kn}\mathbf{I})\,|^{(1/2)}}\exp\left\{-\frac{1}{2}(\mathbf{t}_n-\boldsymbol{\mu}_{kn})^T\left(\sigma^2_{kn}\mathbf{I}\right)^{-1}(\mathbf{t}_n-\boldsymbol{\mu}_{kn})\right\}\right)$$

$$= \frac{1}{(2\pi)^{L/2}|\,(\sigma^2_{kn}\mathbf{I})\,|^{(1/2)}}\left(\frac{\partial}{\partial \mu_{kln}}\exp\left\{-\frac{1}{2\sigma^2_{kn}}(\mathbf{t}_n-\boldsymbol{\mu}_{kn})^T(\mathbf{t}_n-\boldsymbol{\mu}_{kn})\right\}\right)$$

$$= \frac{1}{(2\pi)^{L/2}|\,(\sigma^2_{kn}\mathbf{I})\,|^{(1/2)}}\exp\left\{-\frac{1}{2\sigma^2_{kn}}(\mathbf{t}_n-\boldsymbol{\mu}_{kn})^T\left(\sigma^2_{kn}\mathbf{I}\right)^{-1}(\mathbf{t}_n-\boldsymbol{\mu}_{kn})\right\}$$
$$\left(\frac{\partial}{\partial \mu_{kln}}\left(-\frac{1}{2\sigma^2_{kn}}(\mathbf{t}_n-\boldsymbol{\mu}_{kn})^T(\mathbf{t}_n-\boldsymbol{\mu}_{kn})\right)\right)$$

$$= \frac{1}{(2\pi)^{L/2}|\,(\sigma^2_{kn}\mathbf{I})\,|^{(1/2)}}\exp\left\{-\frac{1}{2}(\mathbf{t}_n-\boldsymbol{\mu}_{kn})^T\left(\sigma^2_{kn}\mathbf{I}\right)^{-1}(\mathbf{t}_n-\boldsymbol{\mu}_{kn})\right\}$$
$$\left(-\frac{1}{2\sigma^2_{kn}}\left(\frac{\partial}{\partial \mu_{kln}}\left(\mathbf{t}_n^T\mathbf{t}_n - 2\boldsymbol{\mu}_{kn}^T\mathbf{t}_n + \boldsymbol{\mu}_{kn}^T\boldsymbol{\mu}_{kn}\right)\right)\right)$$

Calculating the partial derivatives inside the braces:

$$\frac{\partial \mathbf{t}_n^T\mathbf{t}_n}{\partial \mu_{kln}} = 0$$

$$\frac{\partial(-2\boldsymbol{\mu}_{kn}^T\mathbf{t}_n)}{\partial\mu_{kln}} = -2t_l$$

$$\frac{\partial\boldsymbol{\mu}_{kn}^T\boldsymbol{\mu}_{kn}}{\partial\mu_{kln}} = (2\boldsymbol{\mu}_{kn})\frac{\partial\boldsymbol{\mu}_{kn}}{\partial\mu_{kln}} = 2\mu_{kln}$$

Substituting these results back, we get:

$$\begin{aligned}
\frac{\partial E_n}{\partial a_{kl}^{\mu}} &= -\frac{1}{\left\{\sum_{k=1}^{K}\pi_{kn}\mathcal{N}_{kn}\right\}}\left(\pi_{kn}\mathcal{N}_{kn}\left(-\frac{1}{2\sigma_{kn}^2}\left(-2t_l + 2\mu_{kln}\right)\right)\right) \\
&= \frac{\pi_{kn}\mathcal{N}_{kn}}{\left\{\sum_{k=1}^{K}\pi_{kn}\mathcal{N}_{kn}\right\}}\left(\frac{\mu_{kln} - t_l}{\sigma_{kn}^2}\right) \\
&= \gamma_{kn}\left(\frac{\mu_{kln} - t_l}{\sigma_{kn}^2}\right)
\end{aligned}$$

which is the same as the result in 5.156.