Hence

$$\begin{split} \ln\left(p(\boldsymbol{X},\boldsymbol{Z}|\boldsymbol{\theta})\right) &= \sum_{n=1}^{N} a_n \ln p\left(\boldsymbol{x_n}|\boldsymbol{z_n},\boldsymbol{\Sigma},\boldsymbol{\mu}\right) + \sum_{n=1}^{N} a_n \ln p\left(\boldsymbol{z_n}|\boldsymbol{\pi}\right) \\ &= \sum_{n=1}^{N} a_n \sum_{k=1}^{K} (\ln \pi_k) z_{kn} + \sum_{n=1}^{N} a_n \sum_{k=1}^{K} (\ln \phi_k(\boldsymbol{x_n})) z_{kn} \\ &= \ln \boldsymbol{\pi}^T \cdot \boldsymbol{Z} \cdot \boldsymbol{a} + \\ &\sum_{n=1}^{N} a_n \sum_{k=1}^{K} \left[-\frac{l}{2} \ln(2\pi) - \frac{1}{2} \ln |\boldsymbol{\Sigma}_k| - \frac{1}{2} (\boldsymbol{x_n} - \mu_k)^T \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{x_n} - \mu_k) \right] z_{kn} \\ &= (\ln \boldsymbol{\pi}^T - \frac{1}{2} |\boldsymbol{\Sigma}|^T) \cdot \boldsymbol{Z} \cdot \boldsymbol{a} - \frac{l}{2} \ln(2\pi) ones(1, K) \cdot \boldsymbol{Z} \cdot \boldsymbol{a} \\ &- \frac{1}{2} \sum_{k=1}^{K} \left[\operatorname{tr} \left((\boldsymbol{X} - \mu_k \cdot \boldsymbol{1}^T)^T \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{X} - \mu_k \cdot \boldsymbol{1}^T) \cdot \operatorname{diag}(\boldsymbol{a} \odot \boldsymbol{z_k}) \right) \right] \end{split}$$

Now find close form of Q

 $\sum_{m{Z}} p\left(m{Z}|m{X}, m{ heta}^{
m old}
ight.)$ is [probability distribution] random variable **of** $m{Z}$ [probability distribution] random variable

Here
$$q^k = p\left(\mathbf{Z}|\mathbf{X}, \boldsymbol{\theta}^k\right) = p\left(\mathbf{Z}|\mathbf{X}, \boldsymbol{\theta}^{ ext{ old }}\right)$$

$$Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{ ext{ old }}\right) \equiv \sum_{\mathbf{Z}} p\left(\mathbf{Z}|\mathbf{X}, \boldsymbol{\theta}^{ ext{ old }}\right) \ln p(\mathbf{X}, \mathbf{Z}|\boldsymbol{\theta})$$

$$= \left\{ \left[\sum_{n=1}^{N} a_n \sum_{\mathbf{z}_n} p\left(\mathbf{z}_n | \mathbf{x}_n, \boldsymbol{\theta}^{ ext{ old }}\right) \mathbf{z}_n^T \right] \left(\ln \boldsymbol{\pi} - \frac{1}{2} |\boldsymbol{\Sigma}| - \frac{l}{2} \ln(2\pi) \cdot 1.\right) \right\}$$

$$- \frac{1}{2} \left\{ \sum_{n=1}^{N} a_n \sum_{k=1}^{K} \left[\sum_{\mathbf{z}_n} p\left(\mathbf{z}_n | \mathbf{x}_n, \boldsymbol{\theta}^{ ext{ old }}\right) \left(\mathbf{x}_n - \mu_k\right)^T \Sigma_k^{-1} (\mathbf{x}_n - \mu_k) z_{kn} \right] \right\}$$

$$= \left[\boldsymbol{\gamma} \cdot \boldsymbol{a} \right]^T \left(\ln \boldsymbol{\pi} - \frac{1}{2} |\boldsymbol{\Sigma}| - \frac{l}{2} \ln(2\pi) \cdot 1.\right)$$

$$- \frac{1}{2} \sum_{k=1}^{K} \operatorname{tr} \left(\operatorname{diag}(\boldsymbol{a} \odot \boldsymbol{\gamma}_k) \cdot (\boldsymbol{X} - \mu_k \cdot 1^T)^T \Sigma_k^{-1} (\boldsymbol{X} - \mu_k \cdot 1^T) \right)$$

Here

$$p(oldsymbol{z_n} | oldsymbol{x_n}, oldsymbol{ heta}^{ ext{old}}) = rac{p(oldsymbol{x_n} | oldsymbol{z_n}, oldsymbol{ heta}^{ ext{old}}) p(oldsymbol{z_n} | oldsymbol{ heta}^{ ext{old}})}{\sum\limits_{oldsymbol{z_n}} p(oldsymbol{x_n} | oldsymbol{z_n}, oldsymbol{ heta}^{ ext{old}}) p(oldsymbol{z_n} | oldsymbol{ heta}^{ ext{old}})}$$

So, have $\mathbf{1}.^T oldsymbol{\gamma} = ones(\mathbf{1}, N)$

$$egin{aligned} p(oldsymbol{z_n} = k | oldsymbol{x_n}, oldsymbol{ heta}^{ ext{old}}) &= rac{\phi_k^{ ext{old}}(oldsymbol{x_n}) \pi_k^{ ext{old}}}{\sum\limits_{k=1}^K \phi_k^{ ext{old}}(oldsymbol{x_n}) \pi_k^{ ext{old}}} \equiv \gamma_{kn} \ oldsymbol{\gamma} &= oldsymbol{\phi}^{ ext{old}}(X) \odot (\pi_k^{ ext{old}} \cdot 1.^T) \odot (rac{1.}{1. \cdot (oldsymbol{\pi}^{ ext{old}})^T oldsymbol{\phi}^{ ext{old}}(X)}) \end{aligned}$$

maximize Q

To sum up

$$egin{aligned} oldsymbol{\pi} &= rac{\left[oldsymbol{\gamma} \cdot oldsymbol{a}
ight]}{1.^T \left[oldsymbol{\gamma} \cdot oldsymbol{a}
ight]} = rac{\left[oldsymbol{\gamma} \cdot oldsymbol{a}
ight]}{1.^T oldsymbol{a}} \ \mu_k &= rac{X(oldsymbol{a} \odot oldsymbol{\gamma_k})}{1.^T (oldsymbol{a} \odot oldsymbol{\gamma_k})} \ \Sigma_k &= rac{(X - \mu_k \cdot 1^T) diag(oldsymbol{a} \odot oldsymbol{\gamma_k}) \cdot (X - \mu_k \cdot 1^T)^T}{1.^T (oldsymbol{a} \odot oldsymbol{\gamma_k})} \end{aligned}$$

where, have $\mathbf{1.}^{T} \boldsymbol{\gamma} = ones(\mathbf{1}, N)$

$$egin{aligned} p(oldsymbol{z_n} = k | oldsymbol{x_n}, oldsymbol{ heta}^{ ext{old}}) &= rac{oldsymbol{\phi}_k^{ ext{old}}(oldsymbol{x_n}) \pi_k^{ ext{old}}}{\sum\limits_{k=1}^K oldsymbol{\phi}_k^{ ext{old}}(oldsymbol{x_n}) \pi_k^{ ext{old}}} &\equiv \gamma_{kn} \ oldsymbol{\gamma} &= oldsymbol{\phi}^{ ext{old}}(X) \odot (\pi_k^{ ext{old}} \cdot 1.^T) \odot (rac{1.}{1. \cdot (oldsymbol{\pi}^{ ext{old}})^T oldsymbol{\phi}^{ ext{old}}(X)}) \end{aligned}$$

here γ_k is (N x 1) matrix

$$oldsymbol{\gamma} = \left[egin{array}{c} dots \ oldsymbol{\gamma}_k^T \ dots \end{array}
ight]$$

Actually, Hinton calculate the π

$$\pi_k = \operatorname{sigmoid}\left(\lambda\left(eta_a - \left(eta_u + \ln|\Sigma_k|
ight)\left(oldsymbol{a^T}oldsymbol{\gamma_k}
ight)
ight)
ight)$$