1 Lower bound

$$\begin{aligned}
\mathcal{L}(q) &= \int Q(\theta) \ln \frac{p(\boldsymbol{x},\theta)}{Q(\theta)}\, d\theta \\
&= \left\langle \ln \frac{p(\boldsymbol{x},\theta)}{Q(\theta)} \right\rangle_Q \\
&= \underbrace{\langle \ln p(\boldsymbol{x},\theta)\rangle_Q}_{(2)} - \underbrace{\langle \ln q(\theta)\rangle_Q}_{(3)} \qquad (1)
\end{aligned}$$

$$\begin{aligned}
\langle \ln p(\boldsymbol{x},\theta)\rangle_Q &= \langle \ln p(\boldsymbol{x},\boldsymbol{Z},\boldsymbol{\mu},\boldsymbol{W},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\tau})\rangle_Q \\
&= \langle \ln p(\boldsymbol{x}|\boldsymbol{Z},\boldsymbol{\mu},\boldsymbol{W},\boldsymbol{\alpha},\boldsymbol{\delta},\boldsymbol{\tau})p(\boldsymbol{Z})p(\boldsymbol{\mu})p(\boldsymbol{W}|\boldsymbol{\alpha})p(\boldsymbol{\alpha})p(\boldsymbol{\delta})p(\boldsymbol{\tau})\rangle_Q \\
&= \underbrace{\langle \ln p(\boldsymbol{x}|\boldsymbol{Z},\boldsymbol{\mu},\boldsymbol{W},\boldsymbol{\delta},\boldsymbol{\tau})\rangle_Q}_{(4)} + \underbrace{\langle \ln p(\boldsymbol{Z})\rangle_Q}_{(5)} + \underbrace{\langle \ln p(\boldsymbol{\mu})\rangle_Q}_{(6)} + \\
&\quad \underbrace{\langle \ln p(\boldsymbol{W}|\boldsymbol{\alpha})\rangle_Q}_{(7)} + \underbrace{\langle \ln \boldsymbol{\alpha}\rangle_Q}_{(8)} + \underbrace{\langle \ln p(\boldsymbol{\delta})\rangle_Q}_{(9)} + \underbrace{\langle \ln p(\boldsymbol{\tau})\rangle_Q}_{(10)} \qquad (2)
\end{aligned}$$

$$\begin{aligned}
\langle \ln Q(\theta)\rangle_Q &= \langle \ln q(\boldsymbol{W})q(\boldsymbol{Z})q(\boldsymbol{\mu})q(\boldsymbol{\alpha})q(\boldsymbol{\delta})q(\boldsymbol{\tau})\rangle_Q \\
&= \underbrace{\langle \ln q(\boldsymbol{W})\rangle_Q}_{11} + \underbrace{\langle \ln q(\boldsymbol{Z})\rangle_Q}_{12} + \underbrace{\langle \ln q(\boldsymbol{\mu})\rangle_Q}_{13} + \underbrace{\langle \ln q(\boldsymbol{\alpha})\rangle_Q}_{16} + \underbrace{\langle \ln q(\boldsymbol{\delta})\rangle_Q}_{17} + \underbrace{\langle \ln q(\boldsymbol{\tau})\rangle_Q}_{18} \qquad (3)
\end{aligned}$$

$$
\begin{aligned}
\langle \ln p(\boldsymbol{X}|\boldsymbol{Z},\boldsymbol{\mu},\boldsymbol{W},\boldsymbol{\delta},\tau)\rangle_Q &= \left\langle \ln \prod_b^B \prod_t^T \mathcal{N}\left(\boldsymbol{x_t^{(b)}}|\boldsymbol{W}diag(\boldsymbol{\delta}^{(b)})\boldsymbol{z}_t+\boldsymbol{\mu}^{(b)},\boldsymbol{\tau}^{(b)}I_V\right)\right\rangle_Q \\
&= \sum_b^B\sum_t^T \left\langle \ln \mathcal{N}\left(\boldsymbol{x_t^{(b)}}|\boldsymbol{W}diag(\boldsymbol{\delta}^{(b)})\boldsymbol{z}_t+\boldsymbol{\mu}^{(b)},\boldsymbol{\tau}^{(b)}I_V\right)\right\rangle_Q \\
&= \sum_b^B\sum_t^T \left\langle -\frac{V}{2}\ln 2\pi\right\rangle_Q + \left\langle \frac{\ln|\boldsymbol{\tau}^{(b)}I_V|}{2}\right\rangle_Q \\
&- \frac{1}{2}\left\langle (\boldsymbol{x_t^{(b)}}-\boldsymbol{W}diag(\boldsymbol{\delta}^{(b)})\boldsymbol{z}_t+\boldsymbol{\mu}^{(b)})^T(\boldsymbol{\tau}^{(b)}I_V)^{-1}(\boldsymbol{x}_t^{(b)}-\boldsymbol{W}diag(\boldsymbol{\delta}^{(b)})\boldsymbol{z}_t+\boldsymbol{\mu}^{(b)})\right\rangle_Q \\
&= \sum_b^B\sum_t^T \frac{1}{2}\left\langle ln|\boldsymbol{\tau}^{(b)}I_V|\right\rangle_Q \\
&- \frac{1}{2}\left\langle (\boldsymbol{x}_t^{(b)}-\boldsymbol{W}diag(\boldsymbol{\delta}^{(b)})\boldsymbol{z}_t+\mu^{(b)})^T(\boldsymbol{\tau}^{(b)}I_V)^{-1}(\boldsymbol{x}_t^{(b)}-\boldsymbol{W}diag(\boldsymbol{\delta}^{(b)})\boldsymbol{z}_t+\boldsymbol{\mu}^{(b)})\right\rangle_Q \\
&+ const \\
&= \sum_b^B\sum_t^T \frac{1}{2}\left\langle ln|\boldsymbol{\tau}^{(b)}I_V|\right\rangle_Q \\
&- \frac{1}{2}\left\langle \boldsymbol{x}_t^{(b)T}(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{x}_t^{(b)}\right\rangle_Q + \frac{1}{2}\left\langle \boldsymbol{x}_t^{(b)T}(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{W}diag(\boldsymbol{\delta^{(b)}})\boldsymbol{z_t}\right\rangle_Q \\
&- \frac{1}{2}\left\langle \boldsymbol{x}_t^{(b)T}(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{\mu}^{(b)}\right\rangle_Q + \frac{1}{2}\left\langle \boldsymbol{z}_t^T diag(\boldsymbol{\delta}^{(b)})\boldsymbol{W}^T(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{x}_t^{(b)}\right\rangle_Q \\
&- \frac{1}{2}\left\langle \boldsymbol{z}_t^T diag(\boldsymbol{\delta}^{(b)})\boldsymbol{W}^T(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{W}diag(\boldsymbol{\delta^{(b)}})\boldsymbol{z_t}\right\rangle_Q \\
&+ \frac{1}{2}\left\langle \boldsymbol{z}_t^T diag(\boldsymbol{\delta}^{(b)})\boldsymbol{W}^T(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{\mu}^{(b)}\right\rangle_Q - \frac{1}{2}\left\langle \boldsymbol{\mu}^{(b)T}(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{x}_t^{(b)}\right\rangle_Q \\
&+ \frac{1}{2}\left\langle \boldsymbol{\mu}^{(b)T}(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{W}diag(\boldsymbol{\delta}^{(b)})\boldsymbol{z}_t\right\rangle_Q - \frac{1}{2}\left\langle \boldsymbol{\mu}^{(b)T}(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{\mu}^{(b)}\right\rangle_Q \\
&+ const \\
&= \sum_b^B V\ln\boldsymbol{\tau}^{(b)} \\
&+ \frac{1}{2}\sum_b^B\sum_t^T -Tr((\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{W}diag(\boldsymbol{\delta^{(b)}})) + \boldsymbol{\mu^{(b)T}}(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{\mu^{(b)}} \\
&+ \left\langle \boldsymbol{x}_t^{(b)T}(\boldsymbol{\tau}^{(b)}I_V)^{-1}\boldsymbol{W}diag(\boldsymbol{\delta^{(b)}})\boldsymbol{z_t}\right\rangle_Q
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
\langle \ln p(\boldsymbol{Z}) \rangle_Q &= \left\langle \ln \prod_{n=1}^{N} N(\boldsymbol{z}_n|\boldsymbol{0}, \boldsymbol{I}_q) \right\rangle_Q \\
&= \left\langle \sum_{n=1}^{N} \ln N(\boldsymbol{z}_n|\boldsymbol{0}, \boldsymbol{I}_q) \right\rangle_Q \\
&= -\frac{Nq}{2} \ln 2\pi - \frac{1}{2} \sum_{n=1}^{N} \left\langle \|\boldsymbol{z}_n\|^2 \right\rangle_Q \\
&= -\frac{1}{2} \sum_{n=1}^{N} \left\langle \|\boldsymbol{z}_n\|^2 \right\rangle_Q + const
\end{aligned} \tag{5}
$$

$$
\begin{aligned}
\langle \ln p(\boldsymbol{\mu}) \rangle_Q &= \left\langle \ln N(\mu|\boldsymbol{0}, \beta^{-1}\boldsymbol{I}_d) \right\rangle_Q \\
&= \left\langle -\frac{d}{2} \ln 2\pi - \frac{1}{2} \ln \left| \beta^{-1} \boldsymbol{I}_d \right| - \frac{1}{2} \boldsymbol{\mu}^T \beta I_d \boldsymbol{\mu} \right\rangle_Q \\
&= -\frac{\beta}{2} \left\langle \|\boldsymbol{\mu}\|^2 \right\rangle_Q + const
\end{aligned} \tag{6}
$$

$$
\begin{aligned}
\langle \ln p(\boldsymbol{W}|\boldsymbol{\alpha}) \rangle_Q &= \left\langle \ln \prod_{i=1}^{q} \left(\frac{\alpha_i}{2\pi} \right)^{T/2} \exp\left(-\frac{1}{2} \alpha_i \|\boldsymbol{w}_i\|^2 \right) \right\rangle_Q \\
&= \frac{1}{2} \sum_{i=1}^{q} T \langle \ln \alpha_i \rangle_Q - T \ln 2\pi - \langle \alpha_i \rangle_Q \left\langle \|\boldsymbol{w}_i\|^2 \right\rangle_Q \\
&= \frac{1}{2} \sum_{i=1}^{q} T(\psi(\tilde{a}_\alpha) - \ln \tilde{b}_{\alpha i}) - \langle \alpha_i \rangle_Q \left\langle \|\boldsymbol{w}_i\|^2 \right\rangle_Q + const \\
&= -\frac{1}{2} \sum_{i=1}^{q} T \ln \tilde{b}_{\alpha i} + \langle \alpha_i \rangle_Q \left\langle \|\boldsymbol{w}_i\|^2 \right\rangle_Q + const
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
\langle \ln p(\boldsymbol{\alpha}) \rangle_Q &= \left\langle \prod_{i=1}^{q} \ln \Gamma(\alpha_i|a_\alpha, b_{\alpha i}) \right\rangle_Q \\
&= \sum_{i=1}^{q} \langle a_\alpha \ln b_{\alpha i} - \ln \Gamma(a_\alpha) + (a_\alpha - 1) \ln \alpha_i - b_{\alpha i} \alpha_i \rangle_Q \\
&= \sum_{i=1}^{q} \tilde{a}_\alpha \ln \tilde{b}_{\alpha i} + (\tilde{a}_\alpha - 1) \langle \ln \alpha_i \rangle_Q - \tilde{b}_{\alpha i} \langle \alpha_i \rangle_Q + const \\
&= \sum_{i=1}^{q} \tilde{a}_\alpha \ln \tilde{b}_{\alpha i} + (\tilde{a}_\alpha - 1)(\psi(\tilde{a}_\alpha) - \ln \tilde{b}_{\alpha i}) - \tilde{b}_{\alpha i} \frac{\tilde{a}_\alpha}{\tilde{b}_{\alpha i}} + const \\
&= \sum_{i=1}^{q} \ln \tilde{b}_{\alpha i} + const
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
\langle \ln p(\boldsymbol{\delta}) \rangle_Q &= \left\langle \ln \prod_b^B N(\boldsymbol{\delta}^{(b)}|0, I_q) \right\rangle_Q \\
&= \sum_b^B \left\langle \ln N(\boldsymbol{\delta}^{(b)}|0, I_q) \right\rangle_Q \\
&= \sum_b^B \left\langle -\frac{q}{2} \ln 2\pi - \frac{1}{2} \ln |I_q| - \frac{1}{2} \boldsymbol{\delta}^{(b)T} \boldsymbol{\delta}^{(b)} \right\rangle_Q \\
&= \sum_b^B \left\langle \|\boldsymbol{\delta}^{(b)}\|^2 \right\rangle_Q + const
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
\langle \ln p(\boldsymbol{\tau}) \rangle_Q &= \left\langle \ln \prod_b^B \prod_t^T \Gamma(\boldsymbol{\tau}_t^{(b)}|\alpha_{\tau^{(b)}}, b_{\tau^{(b)}}) \right\rangle_Q \\
&= \sum_b^B \sum_t^T \left\langle \ln \Gamma(\boldsymbol{\tau}_t^{(b)}|\alpha_{\tau^{(b)}}, b_{\tau^{(b)}}) \right\rangle_Q \\
&= \sum_b^B \sum_t^T \psi(\alpha_{\tau^{(b)}}) - \ln b_{\tau^{(b)}} \\
&= T \sum_b^B \psi(\alpha_{\tau^{(b)}}) - \ln b_{\tau^{(b)}}
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
\langle \ln Q(\boldsymbol{W}) \rangle_Q &= \left\langle \ln \prod_{i=1}^d \mathcal{N}\left(\boldsymbol{W}_i|\mu_{\boldsymbol{W}_i}, \Sigma_{\boldsymbol{W}}\right) \right\rangle_Q \\
&= \sum_{i=1}^d \langle \ln \mathcal{N}\left(\boldsymbol{W}_i|\mu_{\boldsymbol{W}_i}, \Sigma_{\boldsymbol{W}}\right) \rangle_Q \\
&= -\frac{dq}{2}(1 + \ln(2\pi)) - \sum_{i=1}^d \frac{1}{2} \ln |\Sigma_{\boldsymbol{W}}| \\
&= -\frac{d}{2} \ln |\Sigma_{\boldsymbol{W}}| + const
\end{aligned} \tag{11}
$$

$$
\begin{aligned}
\langle \ln Q(\boldsymbol{Z}) \rangle_Q &= \left\langle \ln \prod_{n=1}^{N} \mathcal{N}\left(\boldsymbol{z}_n | \mu_{\boldsymbol{z}}, \Sigma_{\boldsymbol{z}}\right) \right\rangle_Q \\
&= \sum_{n=1}^{N} \left\langle \ln \mathcal{N}\left(\boldsymbol{z}_n | \mu_{\boldsymbol{z}}, \Sigma_{\boldsymbol{z}}\right) \right\rangle_Q \\
&= \sum_{n=1}^{N} -\frac{q}{2}(1 + \ln(2\pi)) - \frac{1}{2} \ln |\Sigma_{\boldsymbol{z}}| \\
&= -\frac{Nq}{2}(1 + \ln(2\pi)) - \frac{N}{2} \ln |\Sigma_{\boldsymbol{z}}| \\
&= -\frac{N}{2} \ln |\Sigma_{\boldsymbol{z}}| + const
\end{aligned} \tag{12}
$$

$$
\begin{aligned}
\langle \ln Q(\boldsymbol{\mu}) \rangle_Q &= \left\langle \ln \mathcal{N}\left(\boldsymbol{\mu} | \mu_{\boldsymbol{\mu}}, \Sigma_{\boldsymbol{\mu}}\right) \right\rangle_Q \\
&= -\frac{q}{2}(1 + \ln(2\pi)) - \frac{1}{2} \ln |\Sigma_{\boldsymbol{\mu}}| \\
&= -\frac{1}{2} \ln |\Sigma_{\boldsymbol{\mu}}| + const
\end{aligned} \tag{13}
$$

$$
\langle \ln Q(\boldsymbol{\alpha}) \rangle_Q = \left\langle \ln \prod_{i=1}^{q} \Gamma(\alpha_i | a_\alpha, b_{\alpha i}) \right\rangle_Q \tag{14}
$$

$$
= \sum_{i=1}^{q} -\tilde{a}_\alpha + \ln \tilde{b}_{\alpha i} - \ln \Gamma(\tilde{a}_\alpha) - (1 - \tilde{a}_\alpha)\psi(\tilde{a}_\alpha) \tag{15}
$$

$$
= \sum_{i=1}^{q} \ln \tilde{b}_{\alpha i} + const \tag{16}
$$

$$
\begin{aligned}
\langle \ln Q(\boldsymbol{\delta}) \rangle_Q &= \left\langle \ln \prod_{b}^{B} \mathcal{N}(\boldsymbol{\delta}^{(b)} | \mu_{\boldsymbol{\delta}^{(b)}}, \Sigma_{\boldsymbol{\delta}}) \right\rangle_Q \\
&= \sum_{b}^{B} \left\langle \ln \mathcal{N}(\boldsymbol{\delta}^{(b)} | \mu_{\boldsymbol{\delta}^{(b)}}, \Sigma_{\boldsymbol{\delta}}) \right\rangle_Q \\
&= \sum_{b}^{B} -\frac{q}{2}(1 + \ln 2\pi) - \frac{1}{2} \ln |\Sigma_{\boldsymbol{\delta}}| \\
&= -\frac{B}{2} \ln |\Sigma_{\boldsymbol{\delta}}| + const
\end{aligned} \tag{17}
$$

$$
\begin{aligned}
\langle \ln Q(\boldsymbol{\tau}) \rangle_Q &= \left\langle \ln \prod_b^B \prod_t^T \Gamma(\boldsymbol{\tau}_t^{(b)} | \tilde{\alpha}_{\boldsymbol{\tau}_t^{(b)}}, \tilde{\beta}_{\boldsymbol{\tau}_t^{(b)}}) \right\rangle_Q \\
&= \sum_b^B \sum_t^T \left\langle \ln \Gamma(\boldsymbol{\tau}_t^{(b)} | \tilde{\alpha}_{\boldsymbol{\tau}_t^{(b)}}, \tilde{\beta}_{\boldsymbol{\tau}_t^{(b)}}) \right\rangle_Q \\
&= \sum_b^B \sum_t^T -\tilde{\alpha}_{\boldsymbol{\tau}_t^{(b)}} + \ln \tilde{\beta}_{\boldsymbol{\tau}_t^{(b)}} - \ln \Gamma(\tilde{\alpha}_{\boldsymbol{\tau}_t^{(b)}}) - (1 - \tilde{\alpha}_{\boldsymbol{\tau}_t^{(b)}}) \psi(\tilde{\alpha}_{\boldsymbol{\tau}_t^{(b)}})
\end{aligned}
\tag{18}
$$

2 Update rules

$$
\Sigma_{\boldsymbol{\delta}} = I + \sum_b^B \left\langle \boldsymbol{\delta}^{(b)T} \boldsymbol{\delta}^{(b)} \right\rangle_Q \sum_t^T \left\langle \boldsymbol{\tau}_t^{(b)} \right\rangle_Q \left\langle \alpha_{\boldsymbol{w}_t^T \boldsymbol{w}_t} \right\rangle_Q \tag{19}
$$

$$
\mu_{\boldsymbol{\delta_i}} = \Sigma_{\boldsymbol{\delta}} \sum_b^B \left\langle \boldsymbol{\delta}^{(b)T} \right\rangle_Q \langle W^T \rangle_Q \left\langle \boldsymbol{\tau}^{(b)} \right\rangle_Q x_i^{(b)} \tag{20}
$$

$$
\Sigma_{w_t} = (\langle \alpha \rangle_Q I + \sum_b^B \boldsymbol{\tau^{(b)}} \left\langle \boldsymbol{\delta}^{(b)T} \boldsymbol{\delta}^{(b)} \right\rangle_Q \sum_v^V \left\langle \boldsymbol{z}_v \boldsymbol{z}_v^T \right\rangle_Q)^{-1} \tag{21}
$$

$$
\mu_{w_t} = \Sigma_{w_t} \left\langle (\sum_b^B \left\langle \boldsymbol{\tau}_t^{(b)} \right\rangle_Q \left\langle diag(\boldsymbol{\delta}^{(b)} \right\rangle_Q) \sum_v^V \boldsymbol{x}_v^{(b)T} \langle \boldsymbol{z}_v \rangle_Q \right\rangle_Q \tag{22}
$$

$$
\Sigma_{\boldsymbol{\delta}^{(b)}} = (I + \sum_t^T \left\langle \boldsymbol{\tau}_t^{(b)} \right\rangle_Q \langle \boldsymbol{w}_{:,t}^T \boldsymbol{w}_{:,t} \rangle_Q \sum_v^V \left\langle \boldsymbol{z}_v^T \boldsymbol{z}_v \right\rangle_Q)^{-1} \tag{23}
$$

$$
\mu_{\boldsymbol{\delta}^{(b)}} = \Sigma_{\boldsymbol{\delta}^{(b)}} \sum_v^V \sum_t^T \left\langle \boldsymbol{x_{v,t}^{(b)}} \right\rangle_Q \boldsymbol{\tau}_t^{(b)} \langle \boldsymbol{z}_v \rangle_Q \langle \boldsymbol{w}_{:,t} \rangle_Q \tag{24}
$$

$$
\tilde{\alpha_\alpha} = \alpha_\alpha + \frac{T}{2} \tag{25}
$$

$$
\tilde{b_t^\alpha} = b_t^\alpha + \frac{\langle \boldsymbol{w}_t^T \boldsymbol{w}_t \rangle_Q}{2} \tag{26}
$$

$$
\tilde{\alpha}_{\boldsymbol{\tau}^{(b)}} = \alpha_{\boldsymbol{\tau}^{(b)}} + \frac{V}{2} \tag{27}
$$

$$
\tilde{b}_t^{\boldsymbol{\tau}^{(b)}} = b_t^{\boldsymbol{\tau}^{(b)}} + \frac{1}{2} \sum_v^V \left\langle (\boldsymbol{x}_{v,t} - \boldsymbol{w}_t diag(\boldsymbol{\delta}^{(b)}) \boldsymbol{z}_v)^2 \right\rangle_Q \tag{28}
$$