We update the CAVI algorithm with

$$q_j^*(\nu_j) \propto \exp\left(\mathbb{E}_{q_{-\nu_j}}\left[\ln p(\nu_j, \boldsymbol{\nu}_{-j}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r}, \mathbf{y})\right]\right) = \exp\left(\mathbb{E}_{q_{-\nu_j}}\left[\ln p(\mathbf{y}|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r}) p(\boldsymbol{\nu}, \boldsymbol{\gamma})\right]\right),$$

hence we would like to calculate

$$\mathbb{E}_{q_{-\nu_j}}[\ln p(\mathbf{y}|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r})] \quad \text{and} \quad \mathbb{E}_{q_{-\nu_j}}[\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma})],$$

where

- $q_{-\nu_j}$ is the distribution over $(\boldsymbol{\nu}_{-j}, \boldsymbol{\gamma})$ that assumes the variables are independent Gaussians with mean vector $(\mathbf{m}_{\boldsymbol{\nu}_{-j}}, \mathbf{m}_{\boldsymbol{\gamma}})$ and variance vector $(\mathbf{s}_{\boldsymbol{\nu}_{-j}}, \mathbf{s}_{\boldsymbol{\gamma}})$,
- $p(\mathbf{y}|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r}) = \prod_{i=1}^n p(y_i|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}_i, \mathbf{r}_i)$ with $p(y_i|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r}) \sim \mathcal{N}(\boldsymbol{\nu}^T \mathbf{t}_i + \boldsymbol{\gamma}^T \mathbf{r}_i, \sigma_\epsilon^2)$, and
- $p(\boldsymbol{\nu}, \boldsymbol{\gamma}) = p(\boldsymbol{\nu}) p(\boldsymbol{\gamma})$ with $p(\boldsymbol{\nu}) \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\nu}}, \boldsymbol{\Sigma}_{\boldsymbol{\nu}})$ and $p(\boldsymbol{\gamma}) \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\gamma}}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}})$.

1 Calculating $\mathbb{E}_{q_{-\nu_j}}[\ln p(\mathbf{y}|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r})]$

First notice that

$$\begin{aligned}
\ln p(\mathbf{y}|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r}) &= \ln \prod_{i=1}^n p(y_i|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}_i, \mathbf{r}_i) \\
&= \sum_{i=1}^n \ln p(y_i|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}_i, \mathbf{r}_i) \qquad (1) \\
&= -\frac{n}{2} \ln\left(2\pi\sigma_\epsilon^2\right) - \frac{1}{2\sigma_\epsilon^2} \sum_{i=1}^n \left(y_i - \boldsymbol{\nu}^T \mathbf{t}_i - \boldsymbol{\gamma}^T \mathbf{r}_i\right)^2,
\end{aligned}$$

Therefore, we have

$$\mathbb{E}_{q_{-\nu_j}}[\ln p(\mathbf{y}|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r})] = -\frac{n}{2} \ln\left(2\pi\sigma_\epsilon^2\right) - \frac{1}{2\sigma_\epsilon^2} \sum_{i=1}^n \mathbb{E}_{q_{-\nu_j}}\left[\left(y_i - \boldsymbol{\nu}^T \mathbf{t}_i - \boldsymbol{\gamma}^T \mathbf{r}_i\right)^2\right], \qquad (2)$$

and

$$\begin{aligned}
&\mathbb{E}_{q_{-\nu_j}}\left[\left(y - \boldsymbol{\nu}^T \mathbf{t} - \boldsymbol{\gamma}^T \mathbf{r}\right)^2\right] = \mathbb{E}_{q_{-\nu_j}}\left[y^2 + (\boldsymbol{\nu}^T \mathbf{t})^2 + (\boldsymbol{\gamma}^T \mathbf{r})^2 - 2y\boldsymbol{\nu}^T \mathbf{t} - 2y\boldsymbol{\gamma}^T \mathbf{r} + 2(\boldsymbol{\nu}^T \mathbf{t})(\boldsymbol{\gamma}^T \mathbf{r})\right] \\
&= y^2 + \mathbb{E}_{q_{-\nu_j}}\left[(\boldsymbol{\nu}^T \mathbf{t})^2\right] + \mathbb{E}_{q_{-\nu_j}}\left[(\boldsymbol{\gamma}^T \mathbf{r})^2\right] - 2y\nu_j t_j - 2y\mathbf{m}_{\boldsymbol{\nu}_{-j}}^T \mathbf{t}_{-j} - 2\mathbf{m}_{\boldsymbol{\gamma}}^T \mathbf{r} + 2(\nu_j t_j + \mathbf{m}_{\boldsymbol{\nu}_{-j}}^T \mathbf{t}_{-j})(\mathbf{m}_{\boldsymbol{\gamma}}^T \mathbf{r}) \qquad (3)
\end{aligned}$$

where we simplify further using that

$$
\begin{aligned}
\mathbb{E}_{q_{-\nu_j}}\left[(\boldsymbol{\nu}^T\mathbf{t})^2\right] &= \mathbb{E}_{q_{-\nu_j}}\left[\sum_{i=1}^{n_{\boldsymbol{\nu}}}\sum_{i'=1}^{n_{\boldsymbol{\nu}}}\nu_i\nu_{i'}t_it_{i'}\right] \\
&= \sum_{i=1}^{n_{\boldsymbol{\nu}}}\sum_{i'=1}^{n_{\boldsymbol{\nu}}}t_it_{i'}\mathbb{E}_{q_{-\nu_j}}\left[\nu_i\nu_{i'}\right] \\
&= t_j^2\nu_j^2 + \sum_{i'\neq j}t_jt_{i'}\nu_j[m_{\boldsymbol{\nu}}]_{i'} + \sum_{i\neq j}t_it_j[m_{\boldsymbol{\nu}}]_i\nu_j + \sum_{i\neq j}\sum_{i'\neq j}t_it_{i'}\mathbb{E}_{q_{-\nu_j}}\left[\nu_i\nu_{i'}\right] \\
&= t_j^2\nu_j^2 + 2\sum_{i'\neq j}t_jt_{i'}\nu_j[m_{\boldsymbol{\nu}}]_{i'} + \sum_{i\neq j}\sum_{i'\neq j;i'\neq i}t_it_{i'}[m_{\boldsymbol{\nu}}]_i[m_{\boldsymbol{\nu}}]_{i'} + \sum_{i\neq j}t_i^2\mathbb{E}_{q_{-\nu_j}}\left[\nu_i^2\right] \\
&= t_j^2\nu_j^2 + 2\sum_{i'\neq j}t_jt_{i'}\nu_j[m_{\boldsymbol{\nu}}]_{i'} + \sum_{i\neq j}\sum_{i'\neq j;i'\neq i}t_it_{i'}[m_{\boldsymbol{\nu}}]_i[m_{\boldsymbol{\nu}}]_{i'} + \sum_{i\neq j}t_i^2\left([s_{\boldsymbol{\nu}}]_i + [m_{\boldsymbol{\nu}}]_i^2\right) \\
&= t_j^2\nu_j^2 + 2\sum_{i'\neq j}t_jt_{i'}\nu_j[m_{\boldsymbol{\nu}}]_{i'} + \sum_{i\neq j}\sum_{i'\neq j}t_it_{i'}[m_{\boldsymbol{\nu}}]_i[m_{\boldsymbol{\nu}}]_{i'} + \sum_{i\neq j}t_i^2[s_{\boldsymbol{\nu}}]_i
\end{aligned}
\tag{4}
$$

and

$$
\begin{aligned}
\mathbb{E}_{q_{-\nu_j}}\left[(\boldsymbol{\gamma}^T\mathbf{r})^2\right] &= \sum_{i=1}^{n_{\boldsymbol{\gamma}}}\sum_{i'=1}^{n_{\boldsymbol{\gamma}}}r_ir_{i'}\mathbb{E}_{q_{-\nu_j}}\left[\gamma_i\gamma_{i'}\right] \\
&= \sum_{i=1}^{n_{\boldsymbol{\gamma}}}r_i^2\mathbb{E}_{q_{-\nu_j}}\left[\gamma_i^2\right] + \sum_{i=1}^{n_{\boldsymbol{\gamma}}}\sum_{i'\neq i}r_ir_{i'}\mathbb{E}_{q_{-\nu_j}}\left[\gamma_i\gamma_{i'}\right] \\
&= \sum_{i=1}^{n_{\boldsymbol{\gamma}}}r_i^2\left([s_{\boldsymbol{\gamma}}]_i + [m_{\boldsymbol{\gamma}}]_i^2\right) + \sum_{i=1}^{n_{\boldsymbol{\gamma}}}\sum_{i'\neq i}r_ir_{i'}[m_{\boldsymbol{\gamma}}]_i[m_{\boldsymbol{\gamma}}]_{i'} \\
&= \sum_{i=1}^{n_{\boldsymbol{\gamma}}}r_i^2[s_{\boldsymbol{\gamma}}]_i + \sum_{i=1}^{n_{\boldsymbol{\gamma}}}\sum_{i'=1}^{n_{\boldsymbol{\gamma}}}r_ir_{i'}[m_{\boldsymbol{\gamma}}]_i[m_{\boldsymbol{\gamma}}]_{i'} \\
&= \mathbf{s}_{\boldsymbol{\gamma}}^T\mathbf{r}^2 + (\mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r})^2.
\end{aligned}
\tag{5}
$$

Excluding terms that don't contain a ν_j we have

$$
\begin{aligned}
&\mathbb{E}_{q_{-\nu_j}}[\ln p(\mathbf{y}|\boldsymbol{\nu},\boldsymbol{\gamma},\mathbf{t},\mathbf{r})] = -\frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{n}\left(t_{ij}^2\nu_j^2 + 2t_{ij}\nu_j\sum_{k\neq j}t_{ik}[m_{\boldsymbol{\nu}}]_k - 2\nu_jt_{ij}(y_i - \mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i)\right) \\
&= -\frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{n}\left(t_{ij}^2\nu_j^2 + 2\nu_jt_{ij}(\mathbf{m}_{\boldsymbol{\nu}}^T\mathbf{t}_i - t_{ij}[m_{\boldsymbol{\nu}}]_j) - 2\nu_jt_{ij}(y_i - \mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i)\right) \\
&= -\frac{\nu_j^2}{2\sigma_\epsilon^2}\sum_{i=1}^{n}t_{ij}^2 + \frac{\nu_j}{\sigma_\epsilon^2}\sum_{i=1}^{n}t_{ij}\left(y_i - \mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i - \mathbf{m}_{\boldsymbol{\nu}}^T\mathbf{t}_i + [m_{\boldsymbol{\nu}}]_jt_{ij}\right).
\end{aligned}
\tag{6}
$$

2 Calculating $\mathbb{E}_{q_{-\nu_j}}[\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma})]$

First, we simplify:

$$\begin{aligned}
&\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma}) = \ln p(\boldsymbol{\nu}) + \ln p(\boldsymbol{\gamma}) \\
&= -\left(\frac{n_{\boldsymbol{\nu}}}{2} + \frac{n_{\boldsymbol{\gamma}}}{2}\right) \ln(2\pi) - \frac{1}{2}(\ln|\boldsymbol{\Sigma}_{\boldsymbol{\nu}}| + \ln|\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}|) - \frac{1}{2}(\boldsymbol{\nu} - \boldsymbol{\mu}_{\boldsymbol{\nu}})^T \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\mu}_{\boldsymbol{\nu}}) - \frac{1}{2}(\boldsymbol{\gamma} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})^T \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1} (\boldsymbol{\gamma} - \boldsymbol{\mu}_{\boldsymbol{\gamma}}).
\end{aligned} \tag{7}$$

Then, we find

$$\begin{aligned}
\mathbb{E}_{q_{-\nu_j}}[\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma})] = &-\left(\frac{n_{\boldsymbol{\nu}}}{2} + \frac{n_{\boldsymbol{\gamma}}}{2}\right) \ln(2\pi) - \frac{1}{2}(\ln|\boldsymbol{\Sigma}_{\boldsymbol{\nu}}| + \ln|\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}|) \\
&- \frac{1}{2}\mathbb{E}_{q_{-\nu_j}}[(\boldsymbol{\nu} - \boldsymbol{\mu}_{\boldsymbol{\nu}})^T \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\mu}_{\boldsymbol{\nu}})] - \frac{1}{2}\mathbb{E}_{q_{-\nu_j}}[(\boldsymbol{\gamma} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})^T \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1} (\boldsymbol{\gamma} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})],
\end{aligned} \tag{8}$$

where we find the final result using

$$\begin{aligned}
&\mathbb{E}_{q_{-\nu_j}}[(\boldsymbol{\nu} - \boldsymbol{\mu}_{\boldsymbol{\nu}})^T \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1} (\boldsymbol{\nu} - \boldsymbol{\mu}_{\boldsymbol{\nu}})] = \sum_{i=1}^{n_{\boldsymbol{\nu}}} \sum_{i'=1}^{n_{\boldsymbol{\nu}}} [\Sigma_{\boldsymbol{\nu}}^{-1}]_{ii'} \mathbb{E}_{q_{-\nu_j}}[(\nu_i - [\mu_{\boldsymbol{\nu}}]_i)(\nu_{i'} - [\mu_{\boldsymbol{\nu}}]_{i'})] \\
&= \sum_{i=1}^{n_{\boldsymbol{\nu}}} \sum_{i'=1}^{n_{\boldsymbol{\nu}}} [\Sigma_{\boldsymbol{\nu}}^{-1}]_{ii'} \mathbb{E}_{q_{-\nu_j}}\left[\nu_i \nu_{i'} - [\mu_{\boldsymbol{\nu}}]_i \nu_{i'} - [\mu_{\boldsymbol{\nu}}]_{i'} \nu_i + [\mu_{\boldsymbol{\nu}}]_i [\mu_{\boldsymbol{\nu}}]_{i'}\right] \\
&= \sum_{i=1}^{n_{\boldsymbol{\nu}}} \sum_{i'=1}^{n_{\boldsymbol{\nu}}} [\Sigma_{\boldsymbol{\nu}}^{-1}]_{ii'} \mathbb{E}_{q_{-\nu_j}}\left[\nu_i \nu_{i'} - [\mu_{\boldsymbol{\nu}}]_i \nu_{i'} - [\mu_{\boldsymbol{\nu}}]_{i'} \nu_i + [\mu_{\boldsymbol{\nu}}]_i [\mu_{\boldsymbol{\nu}}]_{i'}\right] \\
&= [\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj} \left(\nu_j^2 - 2[\mu_{\boldsymbol{\nu}}]_j \nu_j + [\mu_{\boldsymbol{\nu}}]_j^2\right) + 2\sum_{i \neq j} [\Sigma_{\boldsymbol{\nu}}^{-1}]_{ij} \left([m_{\boldsymbol{\nu}}]_i \nu_j - [\mu_{\boldsymbol{\nu}}]_i \nu_j - [\mu_{\boldsymbol{\nu}}]_j [m_{\boldsymbol{\nu}}]_i + [\mu_{\boldsymbol{\nu}}]_i [\mu_{\boldsymbol{\nu}}]_j\right) \\
&\quad + \sum_{i \neq j} \sum_{i' \neq j} [\Sigma_{\boldsymbol{\nu}}^{-1}]_{ii'} \left([m_{\boldsymbol{\nu}}]_i [m_{\boldsymbol{\nu}}]_{i'} - [\mu_{\boldsymbol{\nu}}]_i [m_{\boldsymbol{\nu}}]_{i'} - [\mu_{\boldsymbol{\nu}}]_{i'} [m_{\boldsymbol{\nu}}]_i + [\mu_{\boldsymbol{\nu}}]_i [\mu_{\boldsymbol{\nu}}]_{i'}\right),
\end{aligned} \tag{9}$$

and

$$\mathbb{E}_{q_{-\nu_j}}[(\boldsymbol{\gamma} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})^T \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1} (\boldsymbol{\gamma} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})] = (\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})^T \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1} (\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}}). \tag{10}$$

Putting this together and excluding terms that don't contain a ν_j we have

$$\begin{aligned}
\mathbb{E}_{q_{-\nu_j}}[\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma})] &= -\frac{1}{2}\left([\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj} \left(\nu_j^2 - 2[\mu_{\boldsymbol{\nu}}]_j \nu_j\right) + 2\nu_j \sum_{i \neq j} [\Sigma_{\boldsymbol{\nu}}^{-1}]_{ij} \left([m_{\boldsymbol{\nu}}]_i - [\mu_{\boldsymbol{\nu}}]_i\right)\right) \\
&= -\frac{\nu_j^2}{2} [\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj} + \nu_j [\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj} [\mu_{\boldsymbol{\nu}}]_j.
\end{aligned} \tag{11}$$

3 CAVI Update

Therefore, we have

$$\ln q_j^*(\nu_j) \propto -\frac{\nu_j^2}{2}\left([\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj} + \frac{1}{\sigma_\epsilon^2}\sum_{i=1}^n t_{ij}^2\right) + \nu_j\left([\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj}[\mu_{\boldsymbol{\nu}}]_j + \frac{1}{\sigma_\epsilon^2}\sum_{i=1}^n t_{ij}\left(y_i - \mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i - \mathbf{m}_{\boldsymbol{\nu}}^T\mathbf{t}_i + [m_{\boldsymbol{\nu}}]_j t_{ij}\right)\right),$$

hence $q_j^*(\nu_j)$ is Gaussian with mean

$$\frac{[\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj}[\mu_{\boldsymbol{\nu}}]_j + \frac{1}{\sigma_\epsilon^2}\sum_{i=1}^n t_{ij}\left(y_i - \mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i - \mathbf{m}_{\boldsymbol{\nu}}^T\mathbf{t}_i + [m_{\boldsymbol{\nu}}]_j t_{ij}\right)}{[\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj} + \frac{1}{\sigma_\epsilon^2}\sum_{i=1}^n t_{ij}^2},$$

and variance

$$\left([\Sigma_{\boldsymbol{\nu}}^{-1}]_{jj} + \frac{1}{\sigma_\epsilon^2}\sum_{i=1}^n t_{ij}^2\right)^{-1}.$$

The $\boldsymbol{\gamma}$ updates are similar.

4 Calculating the ELBO

Recall that the ELBO equals

$$ELBO(q) = \mathbb{E}_q[\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r}, \mathbf{y})] - \mathbb{E}_q[\ln q(\boldsymbol{\nu}, \boldsymbol{\gamma})]$$

where in any iteration of the CAVI algorithm, q is the distribution over $(\boldsymbol{\nu}, \boldsymbol{\gamma})$ that assumes the variables are independent Gaussians with mean vector $(\mathbf{m}_{\boldsymbol{\nu}}, \mathbf{m}_{\boldsymbol{\gamma}})$ and variance vector $(\mathbf{s}_{\boldsymbol{\nu}}, \mathbf{s}_{\boldsymbol{\gamma}})$. We know that

$$\mathbb{E}_q\left[\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r}, \mathbf{y})\right] = \mathbb{E}_q\left[\ln p(\mathbf{y}|\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r})\right] + \mathbb{E}_q\left[\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma})\right],$$

hence, from work similar to what was previously computed,

$$\begin{aligned}
&\mathbb{E}_q\left[\ln p(\boldsymbol{\nu}, \boldsymbol{\gamma}, \mathbf{t}, \mathbf{r}, \mathbf{y})\right] \\
&= -\frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^n \left[y_i^2 + \mathbf{s}_{\boldsymbol{\nu}}^T\mathbf{t}_i^2 + (\mathbf{m}_{\boldsymbol{\nu}}^T\mathbf{t}_i)^2 + \mathbf{s}_{\boldsymbol{\gamma}}^T\mathbf{r}_i^2 + (\mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i)^2 - 2y_i\mathbf{m}_{\boldsymbol{\nu}}^T\mathbf{t}_i - 2\mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i + 2(\mathbf{m}_{\boldsymbol{\nu}}^T\mathbf{t}_i)(\mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i)\right] \\
&\quad - \frac{1}{2}(\mathbf{m}_{\boldsymbol{\nu}} - \boldsymbol{\mu}_{\boldsymbol{\nu}})^T\boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1}(\mathbf{m}_{\boldsymbol{\nu}} - \boldsymbol{\mu}_{\boldsymbol{\nu}}) - \frac{1}{2}(\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})^T\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1}(\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}}) \\
&= -\frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^n \left(y_i - \mathbf{m}_{\boldsymbol{\nu}}^T\mathbf{t}_i - \mathbf{m}_{\boldsymbol{\gamma}}^T\mathbf{r}_i\right)^2 - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^n \left(\mathbf{s}_{\boldsymbol{\nu}}^T\mathbf{t}_i^2 + \mathbf{s}_{\boldsymbol{\gamma}}^T\mathbf{r}_i^2\right) \\
&\quad - \frac{1}{2}(\mathbf{m}_{\boldsymbol{\nu}} - \boldsymbol{\mu}_{\boldsymbol{\nu}})^T\boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1}(\mathbf{m}_{\boldsymbol{\nu}} - \boldsymbol{\mu}_{\boldsymbol{\nu}}) - \frac{1}{2}(\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})^T\boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1}(\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}}).
\end{aligned} \tag{12}$$

In the above, we have dropped constants that will not change as the algorithm updates.

We finally note that

$$q(\boldsymbol{\nu}, \boldsymbol{\gamma}) = \prod_{k=1}^{n_{\boldsymbol{\nu}}} \prod_{k'=1}^{n_{\boldsymbol{\gamma}}} q(\nu_k) q(\gamma_{k'}) = \prod_{k=1}^{n_{\boldsymbol{\nu}}} \prod_{k'=1}^{n_{\boldsymbol{\gamma}}} \frac{1}{2\pi\sqrt{[s_{\boldsymbol{\nu}}]_k [s_{\boldsymbol{\gamma}}]_{k'}}} \exp\left(-\frac{(\nu_k - [m_{\boldsymbol{\nu}}]_k)^2}{2[s_{\boldsymbol{\nu}}]_k} - \frac{(\gamma_{k'} - [m_{\boldsymbol{\gamma}}]_{k'})^2}{2[s_{\boldsymbol{\gamma}}]_{k'}} \right).$$

Hence,

$$\ln q(\boldsymbol{\nu}, \boldsymbol{\gamma}) = -\left(\sum_{k=1}^{n_{\boldsymbol{\nu}}} \sum_{k'=1}^{n_{\boldsymbol{\gamma}}} \ln\left(2\pi \sqrt{[s_{\boldsymbol{\nu}}]_k [s_{\boldsymbol{\gamma}}]_{k'}} \right) + \frac{(\nu_k - [m_{\boldsymbol{\nu}}]_k)^2}{2[s_{\boldsymbol{\nu}}]_k} + \frac{(\gamma_{k'} + [m_{\boldsymbol{\gamma}}]_{k'})^2}{2[s_{\boldsymbol{\gamma}}]_{k'}} \right),$$

and so it follows

$$\begin{aligned} \mathbb{E}_q \left[\ln q(\boldsymbol{\nu}, \boldsymbol{\gamma}) \right] &= -\left(\sum_{k=1}^{n_{\boldsymbol{\nu}}} \sum_{k'=1}^{n_{\boldsymbol{\gamma}}} \ln\left(2\pi \sqrt{[s_{\boldsymbol{\nu}}]_k [s_{\boldsymbol{\gamma}}]_{k'}} \right) + \frac{\mathbb{E}_q \left[(\nu_k - [m_{\boldsymbol{\nu}}]_k)^2 \right]}{2[s_{\boldsymbol{\nu}}]_k} + \frac{\mathbb{E}_q \left[(\gamma_{k'} + [m_{\boldsymbol{\gamma}}]_{k'})^2 \right]}{2[s_{\boldsymbol{\gamma}}]_{k'}} \right) \\ &= -\sum_{k=1}^{n_{\boldsymbol{\nu}}} \sum_{k'=1}^{n_{\boldsymbol{\gamma}}} \left(\ln\left(2\pi \sqrt{[s_{\boldsymbol{\nu}}]_k [s_{\boldsymbol{\gamma}}]_{k'}} \right) + 1 \right). \end{aligned} \tag{13}$$

Next, combining (12) and (13) we can compute the ELBO as

$$\begin{aligned} ELBO(q) = \frac{1}{\sigma_\epsilon^2} \sum_{i=1}^{n} \left(y_i - \mathbf{m}_{\boldsymbol{\nu}}^T \mathbf{t}_i - \mathbf{m}_{\boldsymbol{\gamma}}^T \mathbf{r}_i \right)^2 + \frac{1}{\sigma_\epsilon^2} \sum_{i=1}^{n} \left(\mathbf{s}_{\boldsymbol{\nu}}^T \mathbf{t}_i^2 + \mathbf{s}_{\boldsymbol{\gamma}}^T \mathbf{r}_i^2 \right) \\ + (\mathbf{m}_{\boldsymbol{\nu}} - \boldsymbol{\mu}_{\boldsymbol{\nu}})^T \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1} (\mathbf{m}_{\boldsymbol{\nu}} - \boldsymbol{\mu}_{\boldsymbol{\nu}}) + (\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})^T \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1} (\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}}) - \sum_{k=1}^{n_{\boldsymbol{\nu}}} \sum_{k'=1}^{n_{\boldsymbol{\gamma}}} \ln\left([s_{\boldsymbol{\nu}}]_k [s_{\boldsymbol{\gamma}}]_{k'} \right). \end{aligned} \tag{14}$$

Noticing further that $\mathbf{s}_{\boldsymbol{\nu}}$ and $\mathbf{s}_{\boldsymbol{\gamma}}$ do not change as the algorithm iterates, we have

$$ELBO(q) = \frac{1}{\sigma_\epsilon^2} \sum_{i=1}^{n} \left(y_i - \mathbf{m}_{\boldsymbol{\nu}}^T \mathbf{t}_i - \mathbf{m}_{\boldsymbol{\gamma}}^T \mathbf{r}_i \right)^2 + (\mathbf{m}_{\boldsymbol{\nu}} - \boldsymbol{\mu}_{\boldsymbol{\nu}})^T \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1} (\mathbf{m}_{\boldsymbol{\nu}} - \boldsymbol{\mu}_{\boldsymbol{\nu}}) + (\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}})^T \boldsymbol{\Sigma}_{\boldsymbol{\gamma}}^{-1} (\mathbf{m}_{\boldsymbol{\gamma}} - \boldsymbol{\mu}_{\boldsymbol{\gamma}}). \tag{15}$$