Solving Person Re-identification using Efficient Gibbs Sampling

May 1, 2013

1 The Model

Given a set of observations $\mathcal{X} = \{\mathbf{x}_i\}_{i=1}^N$ across multiple camera views, where each observation corresponds to a complete trajectory within a camera FOV, person re-identification can be defined as the problem of identifying the set of indicator variables associated with the observations $\mathbf{z} = \{z_i\}_{i=1}^N$. To identify the label $z_i \in [1, \dots, Z]$ associated with each trajectory, we utilise a combination of visual information, the appearance features, and the transitions between the cameras. To address the issues associated with appearance-based methods in our proposed person re-identification algorithm, we model each person's appearance using camera-specific illumination and camera gain. We identify the indicator labels by performing Bayesian inference using Gibbs sampling. Each observation $\mathbf{x}_i = \{\mathbf{a}_i, e_i, t_i, l_i\}$ consists of: $l_i \in [1, \dots, L]$ the camera that records the observation; the time of entry e_i in a camera's FOV; the time of leaving the camera's FOV t_i; and the observed appearance features $\mathbf{a}_i$. We define the likelihood as

$$p(\{\mathbf{x}_i\}_{i=1}^N | \{z_i\}_{i=1}^N) = \prod_{j=1}^N p(\mathbf{a}_j | z_j, l_j)\, p(l_j | \{l_i\}_{i=1}^{j-1}, \{z_i\}_{i=1}^{j})\, p(e_j | \{t_i\}_{i=1}^{j-1}, \{z_i\}_{i=1}^{j}) \tag{1}$$

where $p(\{\mathbf{a}_j\}_{i=1}^N | z_j, l_j)$ is modelled as $\mathbf{a}_i = g_l(\mathbf{r}_z + \mathbf{w}_l)$, where g_l is the multiplicative gain constant of camera l, $\mathbf{r}_z$ is the RGB color model, averaged over the entire trajectory, $\mathbf{w}_l$ is the illumination noise associated with camera l, and the terms are distributed as:

$$g_l \sim \text{Gamma}(\alpha_l^g, \beta_l^g), \text{ which we approximate as } \mathcal{N}(\mu_l^g, (\Lambda_l^g)^{-1}) \tag{2}$$
$$\mathbf{r}_z \sim \mathcal{N}(\mu_z, (\Lambda_z)^{-1}) \tag{3}$$
$$\mathbf{w}_l \sim \mathcal{N}(\mu_l^w, (\Lambda_l^w)^{-1}) \tag{4}$$

The transitions between cameras are modelled as

$$l_j \mid \{l_i\}_{i=1}^{j-1}, \{z_i\}_{i=1}^{j-1} \sim \text{Mult}(l_j; \theta_{l_i}), i : z_i = z_j \wedge z_k \neq z_j, i < k < j \tag{5}$$
$$e_j \mid \{t_i\}_{i=1}^{j-1}, \{z_i\}_{i=1}^{j} \sim \mathcal{N}(e_j - t_i; \mu_{l_i, l_j}, \Lambda_{l_i, l_j}^{-1}), i : z_i = z_j \wedge z_k \neq z_j, i < k < j \tag{6}$$

It is clear from its structure (Fig 1a) that this model does not allow for efficient inference, since the Markov blanket of any observation is the complete set of observations and indicator variables preceding it. Yet if the latent indicator variables are known, the observations of a person become independent of all other persons, and the model becomes much simpler (Fig 1b)

2 Gain per Camera: Derivation of the Conditional Distribution

The posterior distribution over the gain parameters for camera l in terms of the likelihood distribution and prior distribution is given as,

$$p(\alpha_l, \beta_l | A_l, \mu_l, \Lambda_l, \mu_z, \Lambda_z, l, z) = p(A_l | \alpha_l, \beta_l \mu_l, \Lambda_l, \mu_z, \Lambda_z, z, l) p(\alpha_l, \beta_l) \tag{7}$$

We define a gamma distribution for the gain parameters α_l, β_l. i.e. $G(\alpha_l, \beta_l)$. However, for easier analytical derivation we choose to represent $G(\alpha_l, \beta_l)$ using its Gaussian distribution approximation $N(\alpha\beta_l, (\alpha\beta_l^2)^{-1})$ or equivalently $N(\mu_l^g, \lambda_l^g)$. Thus, eqn 7 can now be wriiten as

$$p(\mu_l^g, \lambda_l^g | A_l, \mu_l^w, \Lambda_l^w, \mu_z, \Lambda_z, l, z) = p(A_l | \mu_l^g, \lambda_l^g, \mu_l^w, \Lambda_l^w, \mu_z, \Lambda_z, z, l) p(\mu_l^g, \lambda_l^g) \tag{8}$$

where, μ_l^w, Λ_l^w represent the mean and precision of the Gaussian distribution defined over the l camera's illumination value, while μ_l^g, λ_l^g represent the mean and precision of the Gaussian distribution defined over the l camera's gain value, and $A_l = \{a_i\}_{i=1}^{N_l}$ represents the set of observations, and N_l indicates the number of trajectories observed in camera l. The likelihood distribution is formulated as a Gaussian distribution given as,

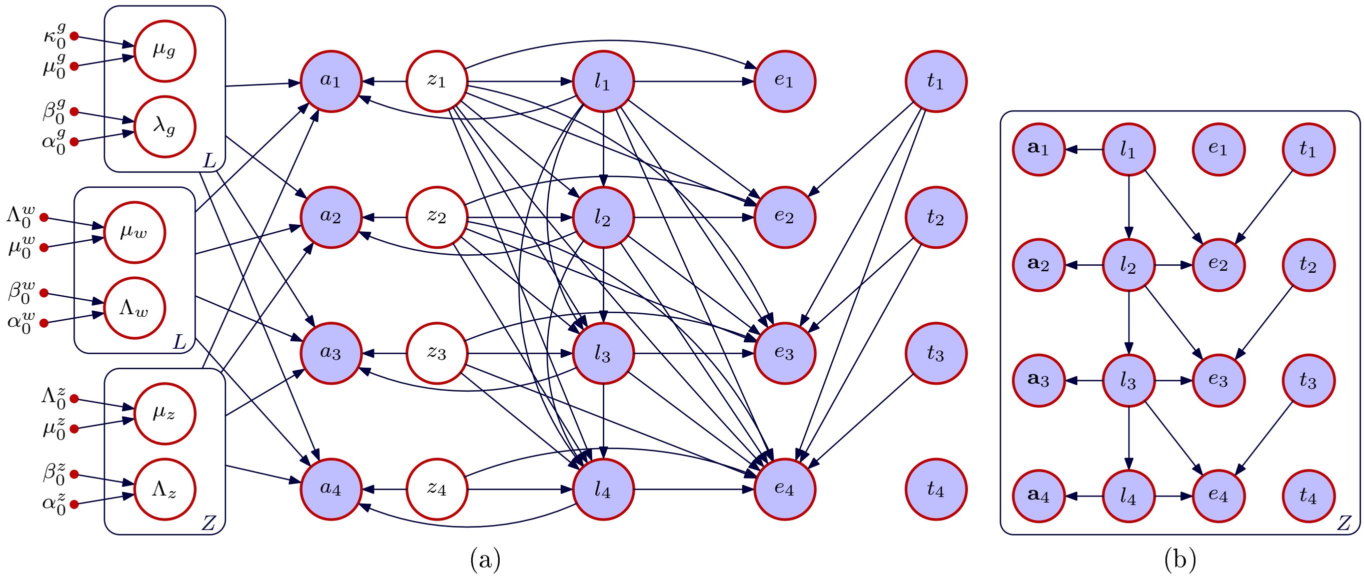

Figure 1: (a) Full graphical model of our probabilistic person re-identification algorithm and (b) Graphical model if the latent variables z_i are known.

$$p(A_l|\mu_l^g,\lambda_l^g,\mu_l^w,\Lambda_l^w,\mu_z,\Lambda_z,z,l) = \prod_{i=1}^{N_l} N(a_i|\mu_l^g(\mu_z+\mu_l^w),\lambda_l^g(\Lambda_z+\Lambda_l^w)) \tag{9}$$

Given, the Gaussian likelihood distribution, we choose a conjugate prior distribution to represent the multiplicative gain parameter. Thus, the prior distribution in 8 is given using a normal-gamma distribution as,

$$p(\mu_l^g,\lambda_l^g|\alpha_0^g,\kappa_0^g,\beta_0^g,\mu_0^g) \propto \lambda_l^{g(1/2)} exp(-\frac{\kappa_0^g\lambda_l^g}{2}(\mu_l^g-\mu_0^g))^2 \lambda_l^{g(\alpha_0^g-1)} exp^{-\lambda_l^g\beta_0^g} \tag{10}$$

where, $\alpha_0,\kappa_0,\beta_0,\mu_0$represent the fixed hyper-parameters for the prior distribution. Since, the normal-gamma distribution is a conjugate prior for the normal likelihood distribution, the resultant posterior is also a normal-gamma distribution given as,

$$p(\mu_l^g,\lambda_l^g|\alpha_n^g,\kappa_n^g,\beta_n^g,\mu_n^g) \propto \lambda_l^{g(1/2)} exp(-\frac{\kappa_n^g\lambda_l^g}{2}(\mu_l^g-\mu_n^g)^2 \lambda_l^{g(\alpha_n^g-1)} exp^{-\lambda_l^g\beta_n^g} \tag{11}$$

where, $\alpha_n^g,\kappa_n^g,\beta_n^g,\mu_n^g$represent the fixed hyper-parameters for the posterior distribution.

To solve for the posterior distribution, we first combine the likelihood distribution (eqn 9) with the prior distribution (10). Thus the resultant posterior distribution is given as,

$$\propto \prod_{i=1}^{N_l} N(a_i|\mu_l^g(\mu_z+\mu_l^w),\lambda_l^g(\Lambda_z+\Lambda_l^w))\lambda_l^{g(1/2)} exp(-\frac{\kappa_0^g\lambda_l^g}{2}(\mu_l^g-\mu_0^g)^2 \lambda_l^{g(\alpha_0^g-1)} exp^{-\lambda_l^g\beta_0^g} \tag{12}$$

$$\propto \prod_{z=1}^{Z} \left[|\lambda_l^g(\Lambda_z+\Lambda_l^w)|^{n_l^z/2}\right] exp\left(-\frac{1}{2}\left\{\sum_{z=1}^{Z}\sum_{i=1}^{n_l^z}(a_i-\mu_l^g(\mu_z+\mu_l^w))^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))(a_i-\mu_l^g(\mu_z+\mu_l^w)\right\}\right)$$

$$\dots\lambda_l^{g(1/2)} exp(-\frac{\kappa_0^g\lambda_l^g}{2}(\mu_l^g-\mu_0^g)^2 \lambda_l^{g(\alpha_0^g-1)} exp^{-\lambda_l^g\beta_0^g} \tag{13}$$

where, N_l represents the number of trajectories observed by each camera l and n_l^z represents the number of trajectories belonging to person z observed by camera l. Note that $N_l = \sum_z n_l^z$. Next, the likelihood distribution in 12 is factorised according to each person's appearance, resulting in the formulation in eqn 13. Before we proceed with our analytical derivation, we define certain terms. Firstly, we define the empirical mean and sum of square matrices, which can be given by,

$$\bar{a} = \frac{1}{n}\sum_i a_i$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(a_i-\bar{a})^T(a_i-\bar{a})$$

Additionally, for notational convenience, we represent the mean and precision within the first exponential term by the following,

$$\mu_z^* = \mu_l^g(\mu_z + \mu_l^w)$$
$$\Lambda_z^* = \lambda_l^g(\Lambda_z + \Lambda_l^w)$$

Using the above definitions, we can re-write the terms inside {} of the first exponential term in 13 as follows,

$$\sum_{z=1}^{Z}\sum_{i=1}^{n_l^z}(a_i - \mu_l^g(\mu_z + \mu_l^w))^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))(a_i - \mu_l^g(\mu_z + \mu_l^w) = \sum_{z=1}^{Z}\sum_{i=1}^{n_l^z}(a_i - \mu_z^*)^T(\Lambda_z^*)(a_i - \mu_z^*) \tag{14}$$

Now we can represent the RHS 14 in terms of the empirical mean and sum of square matrices, given as,

$$\begin{aligned}
&= \sum_{z=1}^{Z}\sum_{i=1}^{n_l^z}\left((a_i - \overline{a})^T(\Lambda_z^*)(a_i - \overline{a}) + (\overline{a} - \mu_z^*)^T(\Lambda_z^*)(\overline{a} - \mu_z^*)\right) \\
&= \sum_{z=1}^{Z}\left(S_z^2(n_l^z - 1)(\Lambda_z^*) + n_l^z(\overline{a} - \mu_z^*)^T(\Lambda_z^*)(\overline{a} - \mu_z^*)\right) \\
&= \sum_z S_z^2(n_l^z - 1)(\lambda_l^g(\Lambda_z + \Lambda_l^o)) \\
&\quad \ldots + \sum_z n_l^z(\overline{a} - (\mu_l^g(\mu_z + \mu_l^w)))^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))(\overline{a} - (\mu_l^g(\mu_z + \mu_l^w)))
\end{aligned} \tag{15}$$

Next, we replace Eqn 15 inside the {} of the first exponential term in Eqn 13 and re-write the equation, resulting in the posterior distribution

$$\begin{aligned}
&\propto \prod_{z=1}^{Z}\left[|\lambda_l^g(\Lambda_z + \Lambda_l^w)|^{n_l^z/2}\right] \\
&\ldots exp\left(-\frac{1}{2}\left\{\sum_z S_z^2(n_l^z - 1)(\lambda_l^g(\Lambda_z + \Lambda_l^w)) + \sum_z n_l^z(\overline{a} - (\mu_l^g(\mu_z + \mu_l^w)))^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))(\overline{a} - (\mu_l^g(\mu_z + \mu_l^w)))\right\}\right) \\
&\ldots \lambda_l^{g(1/2)} exp\left[-\frac{\kappa_0^g\lambda_l^g}{2}(\mu_l^g - \mu_0^g)^2\right]\lambda_l^{g(\alpha_0^g - 1)} exp^{-\lambda_l^g\beta_0^g} \\
&\propto \prod_{z=1}^{Z}\left[|\lambda_l^g(\Lambda_z + \Lambda_l^w)|^{n_l^z/2}\right]\lambda_l^{g(1/2)}\lambda_l^{g(\alpha_0^g - 1)} exp(-\lambda_l^g\beta_0^g) \\
&\ldots exp\left(-\frac{1}{2}\left\{\sum_z S_z^2(n_l^z - 1)(\lambda_l^g(\Lambda_z + \Lambda_l^w)) + \sum_z n_l^z(\overline{a} - (\mu_l^g(\mu_z + \mu_l^w)))^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))(\overline{a} - (\mu_l^g(\mu_z + \mu_l^w))) + \kappa_0^g\lambda_l^g(\mu_l^g - \mu_0^g)^2\right\}\right)
\end{aligned} \tag{16}$$

Next, we expanding the terms inside {} of the exp. of eqn 16. For clarity, we expand each individual term inside the {} of the exp separately. Firstly, the second term i.e. $\sum_z n_l^z(\overline{a} - (\mu_l^g(\mu_z + \mu_l^w)))^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))(\overline{a} - (\mu_l^g(\mu_z + \mu_l^w)))$ expands as follows,

$$
\begin{aligned}
=&\sum n_l^z \overline{a}^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\overline{a} - \sum_z n_l^z 2\overline{a}^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))(\mu_l^g(\mu_z+\mu_l^w)))\\
&\ldots+\sum_z n_l^z(\mu_l^g(\mu_z+\mu_l^w)))^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))(\mu_l^g(\mu_z+\mu_l^w)))\\
=&\sum_z n_l^z\overline{a}^T\lambda_l^g\Lambda_z\overline{a}+\sum_z n_l^z\overline{a}^T\lambda_l^g\Lambda_l^w\overline{a}-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g\Lambda_z)\mu_l^g\mu_z\\
&-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g\Lambda_z)\mu_l^g\mu_l^w-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g\Lambda_l^w)\mu_l^g\mu_z-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g\Lambda_l^w)\mu_l^g.\mu_l^w\ldots\\
&\ldots+\sum_z n_l^z\mu_l^{g2}\mu_z^T(\lambda_l^g\Lambda_z)\mu_z+\sum_z n_l^z\mu_l^{g2}\mu_z^T(\lambda_l^g\Lambda_l^w)\mu_z\\
&\ldots+\sum_z n_l^z\mu_l^{g2}\mu_l^{wT}(\lambda_l^g\Lambda_z)\mu_l^w+\sum_z n_l^z\mu_l^{g2}\mu_l^{wT}(\lambda_l^g\Lambda_l^w)\mu_l^w\\
&\ldots+\sum_z n_l^z 2\mu_l^{g2}\mu_z^T(\lambda_l^g\Lambda_z)\mu_l^w+\sum_z n_l^z 2\mu_l^{g2}\mu_z^T(\lambda_l^g\Lambda_l^w)\mu_l^w
\end{aligned}
\tag{17}
$$

Next, the third term i.e. $\kappa_0^g\lambda_l^g(\mu_l^g-\mu_0^g)^2$ expands as follows,

$$
=\kappa_0^g\lambda_l^g\mu_l^{g2}-2\kappa_0^g\lambda_l^g\mu_l^g\mu_0^g+\kappa_0^g\lambda_l^g\mu_0^{g2}
\tag{18}
$$

Using the expanded terms Eqn (1718) inside {} of the exp, we now get.

$$
\begin{aligned}
=&\sum_z S_z^2(n_l^z-1)\lambda_l^g\Lambda_z+\sum_z S_z^2(n_l^z-1)\Lambda_l^w\lambda_l^g+\sum_z n_l^z\overline{a}^T\lambda_l^g\Lambda_z\overline{a}+\sum_z n_l^z\overline{a}^T\lambda_l^g\Lambda_l^w\overline{a}-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g\Lambda_z)\mu_l^g\mu_z\ldots\\
&\ldots-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g\Lambda_z)\mu_l^g\mu_l^w-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g\Lambda_l^w)\mu_l^g\mu_z-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g\Lambda_l^w)\mu_l^g.\mu_l^w+\sum_z n_l^z\mu_l^{g2}\mu_z^T(\lambda_l^g\Lambda_z)\mu_z\ldots\\
&\ldots+\sum_z n_l^z\mu_l^{g2}\mu_z^T(\lambda_l^g\Lambda_l^w)\mu_z+\sum_z n_l^z\mu_l^{g2}\mu_l^{wT}(\lambda_l^g\Lambda_z)\mu_l^w+\sum_z n_l^z\mu_l^{g2}\mu_l^{wT}(\lambda_l^g\Lambda_l^w)\mu_l^w+\sum_z n_l^z 2\mu_l^{g2}\mu_z^T(\lambda_l^g\Lambda_z)\mu_l^w\ldots\\
&\ldots+\sum_z n_l^z 2\mu_l^{g2}\mu_z^T(\lambda_l^g\Lambda_l^w)\mu_l^w+\kappa_0^g\lambda_l^g\mu_l^{g2}-2\kappa_0^g\lambda_l^g\mu_l^g\mu_0^g+\kappa_0^g\lambda_l^g\mu_0^{g2}
\end{aligned}
$$

Rearranging the terms in the above equation in terms of $\mu_l^{g2}$and μ_l^gwe get

$$
\begin{aligned}
=&\,\mu_l^{g2}\left(\sum_z n_l^z\mu_z^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_z+\sum_z n_l^z\mu_l^{wT}(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_l^w+\sum_z n_l^z 2\mu_z^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_l^w+\kappa_0^g\lambda_l^g\right)\ldots\\
&\ldots+\mu_l^g\left(-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_z-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_l^w-2\kappa_0^g\lambda_l^g\mu_0^g\right)\ldots\\
&\ldots+\sum_z S_z^2(n_l^z-1)\lambda_l^g\Lambda_z+\sum_z S_z^2(n_l^z-1)\Lambda_l^w\lambda_l^g+\sum_z n_l^z\overline{a}^T\lambda_l^g\Lambda_z\overline{a}+\sum_z n_l^z\overline{a}^T\lambda_l^g\Lambda_l^w\overline{a}+\kappa_0^w\lambda_l^g\mu_0^{g2}
\end{aligned}
\tag{19}
$$

To obtain the solution for the posterior hyperparameters, we utilize completing the squares. We know that the result for completing the squares for a quadratic equation is given as

$$
\begin{aligned}
ax^2+bx+c&=a(x-h)^2+k\\
k=c-\frac{b^2}{4a};h&=-\frac{b}{2a}
\end{aligned}
$$

Comparing Eqn 19 with a quadratric equation ax^2+bx+c, we can write the coefficient and constant terms as
,

$$
\begin{aligned}
a&=\sum_z n_l^z\mu_z^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_z+\sum_z n_l^z\mu_l^{wT}(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_l^w+\sum_z n_l^z 2\mu_z^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_l^w+\kappa_0^g\lambda_l^g\\
b&=-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_z-\sum_z n_l^z 2\overline{a}^T(\lambda_l^g(\Lambda_z+\Lambda_l^w))\mu_l^o-2\kappa_0^g\lambda_l^g\mu_0^g\\
c&=+\sum_z S_z^2(n_l^z-1)\lambda_l^g\Lambda_z+\sum_z S_z^2(n_l^z-1)\Lambda_l^w\lambda_l^g+\sum_z n_l^z\overline{a}^T\lambda_l^g\Lambda_z\overline{a}+\sum_z n_l^z\overline{a}^T\lambda_l^g\Lambda_l^w\overline{a}+\kappa_0^g\lambda_l^g\mu_0^{g2}
\end{aligned}
\tag{20}
$$

Using the result for completing the squares and the expansion of the coefficients, we can complete the square for Eqn 19. This results in

$$a(\mu_l^g - h)^2 + k \tag{21}$$

$$a = \sum_z n_l^z \mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \mu_l^{wT} (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_l^w + \sum_z n_l^z 2\mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w)\mu_l^w + \kappa_0^g \lambda_l^g \tag{22}$$

$$h = \frac{\sum_z n_l^z \overline{a}^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \overline{a}^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_l^w + \kappa_0^g \lambda_l^g \mu_0^g}{\sum_z n_l^z \mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \mu_l^{wT} (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_l^w + \sum_z n_l^z 2\mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w)\mu_l^w + \kappa_0^g \lambda_l^g} \tag{23}$$

$$k = \sum_z S_z^2 (n_l^z - 1)\lambda_l^g \Lambda_z + \sum_z S_z^2 (n_l^z - 1)\lambda_l^g \Lambda_l^w + \sum_z n_l^z \overline{a}^T \lambda_l^g \Lambda_z \overline{a} + \sum_z n_l^z \overline{a}^T \lambda_l^g \Lambda_l^o \overline{a} + \kappa_0^g \lambda_l^g \mu_0^{g2} \dots \tag{24}$$

$$\dots \left(\frac{(\sum_z n_l^z \overline{a}^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \overline{a}^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_l^o + \kappa_0^g \lambda_l^g \mu_0^g)^2}{\sum_z n_l^z \mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \mu_l^{wT} (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_l^o + \sum_z n_l^z 2\mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_l^w + \kappa_0^g \lambda_l^g} \right) \tag{25}$$

Plugging the above results (17-21) into Eqn16, we get

$$\begin{aligned} p(\mu_l^g, \lambda_l^g | \alpha_n, \kappa_n, \beta_n, \mu_n) &\propto \prod_{z=1}^{Z} \left[|\lambda_l^g (\Lambda_z + \Lambda_l^w)|^{n_l^z/2} \right] \lambda_l^{g(1/2)} \lambda_l^{g(\alpha_0^g - 1)} exp \left(-\frac{1}{2} \left\{ a(\mu_l^g - h)^2 + k \right\} \right) exp(-\lambda_l^g \beta_0^g) \\ &\propto \prod_{z=1}^{Z} \left[|\lambda_l^g (\Lambda_z + \Lambda_l^w)|^{n_l^z/2} \right] \lambda_l^{g(1/2)} \lambda_l^{g(\alpha_0^g - 1)} exp \left(-\frac{a}{2} (\mu_l^g - h)^2 \right) exp \left(-\frac{k}{2} \right) exp(-\lambda_l^g \beta_0^g) \end{aligned} \tag{26}$$

Comparing Eqn 26 with the normal-gamma posterior distribution in Eqn 11, we can observe that $\mu_n^g = h$ or

$$\boxed{\mu_n^g = \frac{\sum_z n_l^z \overline{a}^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \overline{a}^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_l^w + \kappa_0^g \lambda_l^g \mu_0^g}{\sum_z n_l^z \mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \mu_l^{wT} (\lambda_l^g (\Lambda_z + \Lambda_l^w)\mu_l^w + \sum_z n_l^z 2\mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w)\mu_l^w + \kappa_0^g \lambda_l^g}} \tag{27}$$

Similarly by comparing the equations, we can also see that $\kappa_n^g \lambda_l^g = a$, which can be written as,

$$\boxed{\begin{aligned} \kappa_n^g \lambda_l^g &= \left(\sum_z n_l^z \mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \mu_l^{wT} (\Lambda_z + \Lambda_l^w)\mu_l^w + \sum_z n_l^z 2\mu_z^T (\Lambda_z + \Lambda_l^w)\mu_l^w + \kappa_0^g \right) \lambda_l^g \\ \kappa_n^g &= \sum_z n_l^z \mu_z^T (\lambda_l^g (\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \mu_l^{wT} (\Lambda_z + \Lambda_l^w)\mu_l^w + \sum_z n_l^z 2\mu_z^T (\Lambda_z + \Lambda_l^w)\mu_l^w + \kappa_0^g \end{aligned}} \tag{28}$$

Given solutions for $\mu_n^g$and κ_n^g, we next solve for α_n^g To solve for α_n^g, we consider the multiplicative terms before the exponential in Eqn 16 and expand them using the following identities

$$det(cA) = c^n det(A)$$
$$(x^a)^b = x^{ab}$$

;

where c is a scalar, $\mathbf{A}$ is a matrix and n denotes the size of the square matrix $\mathbf{A}$. Moreover, we know that $n_l = \sum_z n_l^z$.

$$\begin{aligned} \prod_{z=1}^{Z} \left[|\lambda_l^g (\Lambda_z + \Lambda_l^w)|^{n_l^z/2} \right] \lambda_l^{g(1/2)} \lambda_l^{g(\alpha_0^g - 1)} &= \prod_{z=1}^{Z} \left[|\lambda_l^g (\Lambda_z + \Lambda_l^w)|^{n_l^z/2} \right] \lambda_l^{g(\alpha_0^g - 1 + 1/2)} \\ &= \prod_{z=1}^{Z} \left[\left(\lambda_l^{g3} \right)^{n_l^z/2} |(\Lambda_z + \Lambda_l^w)|^{n_l^z/2} \right] \lambda_l^{g(\alpha_0^g - 1 + 1/2)} \\ &= \left(\lambda_l^{g3} \right)^{\frac{n_l}{2}} \prod_{z=1}^{Z} \left[|(\Lambda_z + \Lambda_l^w)|^{n_l^z/2} \right] \lambda_l^{g(\alpha_0^g - 1 + 1/2)} \\ &= (\lambda_l^g)^{\frac{3n_l}{2} + \alpha_0^g - 1 + \frac{1}{2}} \prod_{z=1}^{P} \left[|(\Lambda_z + \Lambda_l^w)|^{n_l^z/2} \right] \end{aligned} \tag{29}$$

Comparing the terms associated with α_0^gin Eqn 29 with the terms associated with α_n^gin Eqn 11 we get the following

$$\boxed{\begin{aligned}\alpha_n^g - 1 + \frac{1}{2} &= \frac{3n_l}{2} + \alpha_0^g - 1 + \frac{1}{2}\\ \alpha_n^g &= \alpha_0^g + \frac{3n_l}{2}\end{aligned} \quad (30)}$$

Finally, we solve for β_n^g To solve for the final hyper-parameter term β_n^g, we combine the second and third exponential term in Eqn 26 $exp\left(-\frac{k}{2}\right) exp(-\lambda_l^g \beta_0^g)$ and equate it to the second exp term in Eqn 11.

$$\lambda_l^g \beta_n^g = \lambda_l^g \beta_0^g + \frac{k}{2} \tag{31}$$

Next, we expand k

$$\begin{aligned}k &= \sum_z S_z^2(n_l^z - 1)\lambda_l^g \Lambda_z + \sum_z S_z^2(n_l^z - 1)\lambda_l^g \Lambda_l^w + \sum_z n_l^z \overline{a}^T \lambda_l^g \Lambda_z \overline{a} + \sum_z n_l^z \overline{a}^T \lambda_l^g \Lambda_l^w \overline{a} + \kappa_0^g \lambda_l^g \mu_0^{g2} \dots\\ &\dots - \left(\frac{(\sum_z n_l^z \overline{a}^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \overline{a}^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))\mu_l^w + \kappa_0^g \lambda_l^g \mu_0^g)^2}{\sum_z n_l^z \mu_z^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))\mu_z + \sum_z n_l^z \mu_l^{wT}(\lambda_l^g(\Lambda_z + \Lambda_l^w))\mu_l^w + \sum_z n_l^z 2\mu_z^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))\mu_l^w + \kappa_0^g \lambda_l^g}\right)\\ &= \sum_z S_z^2(n_l^z - 1)\lambda_l^g \Lambda_z + \sum_z S_z^2(n_l^z - 1)\lambda_l^g \Lambda_l^w + \sum_z n_l^z \overline{a}^T \lambda_l^g \Lambda_z \overline{a} + \sum_z n_l^z \overline{a}^T \lambda_l^g \Lambda_l^w \overline{a} + \kappa_0^g \lambda_l^g \mu_0^{g2} \dots\\ &\dots - \lambda_l^g \left(\frac{(\sum_z n_l^z \overline{a}^T(\Lambda_z + \Lambda_l^w)\mu_z + \sum_z n_l^z \overline{a}^T(\Lambda_z + \Lambda_l^w)\mu_l^w + \kappa_0^g \mu_0^g)^2}{\sum_z n_l^z \mu_z^T(\Lambda_z + \Lambda_l^w)\mu_z + \sum_z n_l^z \mu_l^{wT}(\Lambda_z + \Lambda_l^w)\mu_l^w + \sum_z n_l^z 2\mu_z^T(\Lambda_z + \Lambda_l^w)\mu_l^w + \kappa_0^g}\right)\end{aligned}$$

Continuing our expansion, we perform cross multiplication for simplification
Replacing k into Eqn 31, we get

$$\begin{aligned}\lambda_l^g \beta_n^g = \lambda_l^g \beta_0^g &+ 0.5\sum_z S_z^2(n_l^z - 1)(\lambda_l^g(\Lambda_z + \Lambda_l^w)) + 0.5\sum_z n_l^z \overline{a}^T(\lambda_l^g(\Lambda_z + \Lambda_l^w))\overline{a} + 0.5\kappa_0 \lambda_l^g \mu_0^{g2} - ...\\ &\lambda_l^g 0.5\left(\frac{(\sum_z n_l^z \overline{a}^T(\Lambda_z + \Lambda_l^w)\mu_z + \sum_z n_l^z \overline{a}^T(\Lambda_z + \Lambda_l^w)\mu_l^w + \kappa_0^g \mu_0^g)^2}{\sum_z n_l^z \mu_z^T(\Lambda_z + \Lambda_l^w)\mu_z + \sum_z n_l^z \mu_l^{oT}(\Lambda_z + \Lambda_l^o)\mu_l^o + \sum_z n_l^z 2\mu_z^T(\Lambda_z + \Lambda_l^o)\mu_l^o + \kappa_0}\right)\end{aligned}$$

Thus, β_n^g is given as,

$$\begin{aligned}\beta_n^g = \beta_0^g + \frac{\sum_z S_z^2(n_l^z - 1)(\Lambda_z + \Lambda_l^w)}{2} &+ \frac{\kappa_0^g \mu_0^{g2}\left(\sum_z n_l^z(\mu_l^w + \mu_z)^T(\Lambda_z + \Lambda_l^w)(\mu_l^w + \mu_z)\right)...}{2\sum_z n_l^z \mu_z^T(\Lambda_z + \Lambda_l^w)\mu_z + 2\sum_z n_l^z \mu_l^{wT}(\Lambda_z + \Lambda_l^w)\mu_l^w + 2\sum_z n_l^z 2\mu_z^T(\Lambda_z + \Lambda_l^w)\mu_l^w + 2\kappa_0^g}\\ &\frac{+\kappa_0^g\left(\sum_z n_l^z \overline{a}^T(\Lambda_z + \Lambda_l^w)\overline{a} - 2\mu_0^g \sum_z n_l^z \overline{a}^T(\Lambda_z + \Lambda_l^w)\mu_l^w - 2\mu_0^g \sum_z n_l^z \overline{a}^T(\Lambda_z + \Lambda_l^w)\mu_z\right)}{2\sum_z n_l^z \mu_z^T(\Lambda_z + \Lambda_l^w)\mu_z + 2\sum_z n_l^z \mu_l^{wT}(\Lambda_z + \Lambda_l^w)\mu_l^w + 2\sum_z n_l^z 2\mu_z^T(\Lambda_z + \Lambda_l^w)\mu_l^w + 2\kappa_0^g}\end{aligned}$$