1 
$$x_t^{(j)}$$
 and  $\mu_t^{(j)}$ 

Let the  $t^{\text{th}}$  column observation and firing rate of the  $j^{\text{th}}$  cluster be  $\widetilde{\boldsymbol{y}}_t = vec\left(\{y_{it}^{(j)}|z_i=j\}\right)$  and  $\widetilde{\boldsymbol{\lambda}}_t^{(j)} = vec\left(\{\lambda_{it}|z_i=j\}\right)$ . The number of neurons in cluster j is  $n_j = |\{i: z_i=j\}|$ . The corresponding loading for these  $n_j$  neurons is  $\boldsymbol{C}^{(j)} \in \mathbb{R}^{n_j \times p}$ , such that  $\log \widetilde{\boldsymbol{\lambda}}_t^{(j)} = \mu_t^{(j)} \mathbf{1}_{n_j} + \boldsymbol{C}^{(j)} \boldsymbol{x}_t^{(j)} = \left(\mathbf{1}_{n_j}, \boldsymbol{C}^{(j)}\right) \cdot \left(\mu_t^{(j)}, \boldsymbol{x}_t^{'(j)}\right)'$ . Denote  $\widetilde{\boldsymbol{C}}^{(j)} = \left(\mathbf{1}_{n_j}, \boldsymbol{C}^{(j)}\right), \widetilde{\boldsymbol{x}}_t^{(j)} = \left(\mu_t^{(j)}, \boldsymbol{x}_t^{'(j)}\right)', \widetilde{\boldsymbol{x}}^{(j)} = \left(\widetilde{\boldsymbol{x}}_1^{'(j)}, \dots, \widetilde{\boldsymbol{x}}_T^{'(j)}\right)', \widetilde{\boldsymbol{A}}^{(j)} = diag(f^{(j)}, \boldsymbol{A}^{(j)}), \widetilde{\boldsymbol{b}}^{(j)} = \left(g^{(j)}, \boldsymbol{b}^{'(j)}\right)'$  and  $\widetilde{\boldsymbol{Q}}^{(j)} = diag(\boldsymbol{\Sigma}^{(j)}, \boldsymbol{Q}^{(j)})$ . The full conditional distribution  $P(\widetilde{\boldsymbol{x}}^{(j)}|\dots) = P(\widetilde{\boldsymbol{x}}^{(j)}|\{\widetilde{\boldsymbol{y}}_t\}_{t=1}^T, \widetilde{\boldsymbol{C}}^{(j)}, \widetilde{\boldsymbol{A}}^{(j)}, \widetilde{\boldsymbol{b}}^{(j)}, \widetilde{\boldsymbol{Q}}^{(j)})$  is approximated by a global Laplace approximation, i.e.  $P(\widetilde{\boldsymbol{x}}^{(j)}|\dots) \approx N_{(p+1)T}(\widetilde{\boldsymbol{x}}^{(j)}|\boldsymbol{\mu}_{\widetilde{\boldsymbol{x}}^{(j)}}, \boldsymbol{\Sigma}_{\widetilde{\boldsymbol{x}}^{(j)}})$ , with  $\boldsymbol{\mu}_{\widetilde{\boldsymbol{x}}^{(j)}} = \arg\max_{\widetilde{\boldsymbol{x}}^{(j)}} P(\widetilde{\boldsymbol{x}}^{(j)}|\dots)$  and  $\boldsymbol{\Sigma}_{\widetilde{\boldsymbol{x}}^{(j)}}) = -\left(\nabla\nabla\log P(\widetilde{\boldsymbol{x}}^{(j)}|\dots)\right|_{\widetilde{\boldsymbol{x}}^{(j)}=\boldsymbol{\mu}_{\widetilde{\boldsymbol{x}}^{(j)}}})^{-1}$ . The log full conditional distribution  $h(\widetilde{\boldsymbol{x}}^{(j)}) = \log P(\widetilde{\boldsymbol{x}}^{(j)}|\dots)$  is given by:

$$h = \text{const} + \sum_{t=1}^T \widetilde{\boldsymbol{C}}^{\prime(j)} \left( \widetilde{\boldsymbol{y}}_t - \widetilde{\boldsymbol{\lambda}}_t \right) - \frac{1}{2} (\widetilde{\boldsymbol{x}}_1^{(j)} - \widetilde{\boldsymbol{x}}_0^{(j)})' \widetilde{\boldsymbol{Q}}_0^{(j)} (\widetilde{\boldsymbol{x}}_1^{(j)} - \widetilde{\boldsymbol{x}}_0^{(j)}) - \sum_{t=2}^T \frac{1}{2} (\widetilde{\boldsymbol{x}}_t^{(j)} - \widetilde{\boldsymbol{A}}^{(j)} \widetilde{\boldsymbol{x}}_{t-1}^{(j)})' \widetilde{\boldsymbol{Q}}^{(j)} (\widetilde{\boldsymbol{x}}_t^{(j)} - \widetilde{\boldsymbol{A}}^{(j)} \widetilde{\boldsymbol{x}}_{t-1}^{(j)})$$

It should be  $b_t^{(j)}$ , but not  $b^{(j)}$ ! Fix that and then edit the writing!