

1. **Algorithm 1:** Use the Gibbs sampling algorithm to approximate (using samples) the posterior distribution  $p(\boldsymbol{\pi}, \mathbf{Z}, \boldsymbol{\Theta} \mid \mathbf{X})$ . Derive the posterior.

$$\underline{p(\boldsymbol{\pi} \mid \mathbf{Z})} :$$

From Lecture 3:

$$p(\boldsymbol{\pi} \mid \mathbf{Z}) = p(\boldsymbol{\pi} \mid \{z_n\}_{n=1}^N) = \text{Dir}(\boldsymbol{\pi} \mid \alpha_1 + m_1, \dots, \alpha_K + m_K), \text{ with } m_k = \sum_{n=1}^N [z_n = k]$$

$$\underline{p(\boldsymbol{\theta}_k \mid \mathbf{X}, \mathbf{Z})} :$$

$$p(\boldsymbol{\theta}_k \mid \mathbf{X}, \mathbf{Z}) = \frac{p(\boldsymbol{\theta}_k, \mathbf{X} \mid \mathbf{Z})}{p(\mathbf{X} \mid \mathbf{Z})} \propto p(\boldsymbol{\theta}_k, \mathbf{X} \mid \mathbf{Z}) = p(\mathbf{X} \mid \boldsymbol{\theta}_k, \mathbf{Z})p(\boldsymbol{\theta}_k \mid \mathbf{Z})$$

$$p(\mathbf{X} \mid \boldsymbol{\theta}_k, \mathbf{Z}) = p(\{\mathbf{x}_n\}_{n=1}^N \mid \boldsymbol{\theta}_k, \{z_n\}_{n=1}^N) = \prod_{n=1}^N p(x_n \mid \boldsymbol{\theta}_k, z_n)$$

$$p(\boldsymbol{\theta}_k \mid \mathbf{Z}) = p(\boldsymbol{\theta}_k \mid \gamma) = \text{Dir}(\boldsymbol{\theta}_k \mid \gamma)$$

$$\begin{aligned} \Rightarrow p(\boldsymbol{\theta}_k \mid \mathbf{X}, \mathbf{Z}) &\propto \text{Dir}(\boldsymbol{\theta}_k \mid \gamma) \prod_{n=1}^N p(x_n \mid \boldsymbol{\theta}_k, z_n) \\ &= \frac{1}{B(\gamma)} \prod_{i=1}^{|I|} \theta_{ki}^{\gamma_{ki}-1} \prod_{n=1}^N \prod_{j=1}^{W_n} \text{Cat}(x_{nj} \mid \boldsymbol{\theta}_k)^{[z_n=k]} \\ &= \frac{1}{B(\gamma)} \prod_{i=1}^{|I|} \theta_{ki}^{\gamma_{ki}-1} \prod_{n=1}^N \prod_{j=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{[x_{nj}=m][z_n=k]} \\ &= \frac{1}{B(\gamma)} \prod_{j=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{\gamma_{ki}-1 + \sum_{n=1}^N \sum_{j=1}^{W_n} [x_{nj}=m][z_n=k]} \end{aligned}$$

$$c_{km} := \sum_{n=1}^N \sum_{j=1}^{W_n} [x_{nj} = m][z_n = k]$$

$$\Rightarrow p(\boldsymbol{\theta}_k \mid \mathbf{X}, \mathbf{Z}) \propto \frac{1}{B(\gamma)} \prod_{m=1}^{|I|} \theta_{km}^{\gamma_{km} + c_{km} - 1} \propto \text{Dir}(\boldsymbol{\theta}_k \mid \gamma_{k1} + c_{k1}, \dots, \gamma_{k|I|} + c_{k|I|})$$

$$\Rightarrow p(\boldsymbol{\theta}_k \mid \mathbf{X}, \mathbf{Z}) = \text{Dir}(\boldsymbol{\theta}_k \mid \gamma_k + \mathbf{c}_k)$$

$$\underline{p(z_n = k \mid \mathbf{x}_n, \boldsymbol{\pi}, \boldsymbol{\Theta})} :$$

$$p(z_n = k \mid \mathbf{x}_n, \boldsymbol{\pi}, \boldsymbol{\Theta}) = \frac{p(\mathbf{x}_n, z_n = k \mid \boldsymbol{\pi}, \boldsymbol{\Theta})}{p(\mathbf{x}_n \mid \boldsymbol{\pi}, \boldsymbol{\Theta})} = \frac{p(\mathbf{x}_n \mid z_n = k, \boldsymbol{\Theta})p(z_n = k \mid \boldsymbol{\pi})}{p(\mathbf{x}_n \mid \boldsymbol{\pi}, \boldsymbol{\Theta})}$$

$$p(z_n = k \mid \boldsymbol{\pi}) = \pi_k$$

$$p(\mathbf{x}_n \mid z_n = k, \boldsymbol{\theta}_k) = \prod_{i=1}^{W_n} \text{Cat}(x_{ni} \mid \boldsymbol{\theta}_k) = \prod_{i=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{[x_{ni}=m]} = \prod_{m=1}^{|I|} \theta_{km}^{\sum_{i=1}^{W_n} [x_{ni}=m]}$$

$$\begin{aligned} \Rightarrow p(z_n = k \mid \mathbf{x}_n, \boldsymbol{\pi}, \boldsymbol{\Theta}) &= \frac{\pi_k \prod_{i=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{[x_{ni}=m]}}{p(\mathbf{x}_n \mid \boldsymbol{\pi}, \boldsymbol{\Theta})} \\ &= \frac{\pi_k \prod_{i=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{[x_{ni}=m]}}{\sum_{k'=1}^K \pi_{k'} \prod_{j=1}^{W_n} \prod_{m'=1}^{|I|} \theta_{k'm'}^{[x_{nj}=m']}} \\ &= \frac{\pi_k \prod_{m=1}^{|I|} \theta_{km}^{\sum_{i=1}^{W_n} [x_{ni}=m]}}{\sum_{k'=1}^K \pi_{k'} \prod_{m'=1}^{|I|} \theta_{k'm'}^{\sum_{j=1}^{W_n} [x_{nj}=m']}} \end{aligned}$$

$$\underline{\log p(z_n = k \mid \mathbf{x}_n, \boldsymbol{\pi}, \boldsymbol{\Theta})} :$$

$$\log p(z_n = k \mid \mathbf{x}_n, \boldsymbol{\pi}, \boldsymbol{\Theta}) = \log p(\mathbf{x}_n \mid z_n = k, \boldsymbol{\Theta}) + \log p(z_n = k \mid \boldsymbol{\pi}) - \log p(\mathbf{x}_n \mid \boldsymbol{\pi}, \boldsymbol{\Theta})$$

$$\log p(\mathbf{x}_n \mid z_n = k, \boldsymbol{\Theta}) = \sum_{m=1}^{|I|} \log \theta_{km} \sum_{i=1}^{W_n} [x_{ni} = m]$$

$$\log p(z_n = k \mid \boldsymbol{\pi}) = \log \pi_k$$

$$\begin{aligned} \log p(\mathbf{x}_n \mid \boldsymbol{\pi}, \boldsymbol{\Theta}) &= \log \left( \sum_{k'=1}^K \pi_{k'} \prod_{m'=1}^{|I|} \theta_{k'm'}^{\sum_{j=1}^{W_n} [x_{nj}=m']} \right) \\ &= \log \left( \sum_{k'=1}^K p(\mathbf{x}_n \mid z_n = k', \boldsymbol{\Theta}) p(z_n = k' \mid \boldsymbol{\pi}) \right) \\ &= \log \left( \sum_{k'=1}^K \exp(\log p(\mathbf{x}_n \mid z_n = k', \boldsymbol{\Theta}) + \log p(z_n = k' \mid \boldsymbol{\pi})) \right) \end{aligned}$$

2. **Algorithm 2:** Use the (collapsed) Gibbs sampling algorithm to approximate (using samples) the posterior distribution  $p((\mathbf{Z}, \boldsymbol{\Theta}) \mid \mathbf{X})$ . Derive the posterior.

$p(\boldsymbol{\theta}_k \mid \mathbf{X}, \mathbf{Z})$  : From Exercise 1.1:

$$p(\boldsymbol{\theta}_k \mid \mathbf{X}, \mathbf{Z}) = \text{Dir}(\boldsymbol{\theta}_k \mid \gamma_k + \mathbf{c}_k)$$

$p(z_n = k \mid \mathbf{x}_n, \mathbf{Z}_{-n}, \boldsymbol{\Theta})$  :

$$p(z_n = k \mid \mathbf{x}_n, \mathbf{Z}_{-n}, \boldsymbol{\Theta}) = \frac{p(z_n = k, \mathbf{x}_n \mid \mathbf{Z}_{-n}, \boldsymbol{\Theta})}{p(\mathbf{x}_n \mid \mathbf{Z}_{-n}, \boldsymbol{\Theta})} \propto p(\mathbf{x}_n \mid z_n = k, \mathbf{Z}_{-n}, \boldsymbol{\Theta}) p(z_n = k \mid \mathbf{Z}_{-n}, \boldsymbol{\Theta})$$

$$\begin{aligned} p(\mathbf{x}_n \mid z_n = k, \mathbf{Z}_{-n}, \boldsymbol{\Theta}) &= p(\mathbf{x}_n \mid z_n = k, \boldsymbol{\Theta}) = \prod_{i=1}^{W_n} \text{Cat}(x_{ni} \mid \boldsymbol{\theta}_k) \\ &= \prod_{i=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{[x_{ni}=m]} = \prod_{m=1}^{|I|} \theta_{km}^{\sum_{i=1}^{W_n} [x_{ni}=m]} \end{aligned}$$

$$\begin{aligned} p(z_n = k \mid \mathbf{Z}_{-n}) &= \int p(z_n = k, \boldsymbol{\pi} \mid \mathbf{Z}_{-n}) d\boldsymbol{\pi} \\ &= \int p(z_n = k \mid \boldsymbol{\pi}) p(\boldsymbol{\pi} \mid \{z_i\}_{i \neq n}) d\boldsymbol{\pi} \\ &= \int \prod_{l=1}^K \pi_l^{[z_n=l]} p(\boldsymbol{\pi} \mid \{z_i\}_{i \neq n}) d\boldsymbol{\pi} \\ &= \int \prod_{l=1}^K \pi_l^{[z_n=l]} \text{Dir}(\boldsymbol{\pi} \mid \alpha_1 + \tilde{m}_1, \dots, \alpha_K + \tilde{m}_K) d\boldsymbol{\pi}, \text{ with } \tilde{m}_k = \sum_{i \neq n} [z_i = k] \\ &= \int \prod_{l=1}^K \pi_l^{[z_n=l]} \frac{1}{B(\boldsymbol{\alpha} + \tilde{\mathbf{m}})} \prod_{k=1}^K \pi_k^{\alpha_k + \tilde{m}_k - 1} d\boldsymbol{\pi} \\ &= \int \frac{1}{B(\boldsymbol{\alpha} + \tilde{\mathbf{m}})} \prod_{l=1}^K \pi_l^{\alpha_l + \tilde{m}_l + [z_n=l] - 1} d\boldsymbol{\pi} \\ &= \int \frac{1}{B(\boldsymbol{\alpha} + \tilde{\mathbf{m}})} \frac{B(\tilde{\boldsymbol{\alpha}})}{B(\tilde{\boldsymbol{\alpha}})} \prod_{l=1}^K \pi_l^{\alpha_l + \tilde{m}_l + [z_n=l] - 1} d\boldsymbol{\pi}, \text{ with } \tilde{\alpha}_k = \alpha_k + \tilde{m}_k + [z_n = k] \\ &= \int \frac{B(\tilde{\boldsymbol{\alpha}})}{B(\boldsymbol{\alpha} + \tilde{\mathbf{m}})} \text{Dir}(\boldsymbol{\pi} \mid \tilde{\boldsymbol{\alpha}}) d\boldsymbol{\pi} \\ &= \frac{B(\tilde{\boldsymbol{\alpha}})}{B(\boldsymbol{\alpha} + \tilde{\mathbf{m}})} \end{aligned}$$

$$p(z_n = k \mid \mathbf{Z}_{-n}) = \frac{B(\tilde{\alpha})}{B(\alpha + \tilde{m})} = \frac{\prod_{l=1}^K \Gamma(\tilde{\alpha})}{\Gamma\left(\sum_{k=1}^K \tilde{\alpha}_{\tilde{k}}\right)} \cdot \frac{\Gamma\left(\sum_{k'=1}^K \alpha_{k'} + \tilde{m}_{k'}\right)}{\prod_{l'=1}^K \Gamma(\alpha_{l'} + \tilde{m}_{l'})}$$

$$\sum_{\tilde{k}=1}^K \tilde{\alpha}_{\tilde{k}} = \sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}} + [z_n = \tilde{k}] = \sum_{\tilde{k}=1}^K [z_n = \tilde{k}] + \sum_{\tilde{k}=1}^K (\alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}) = 1 + \sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}$$

$$\Rightarrow p(z_n = k \mid \mathbf{Z}_{-n}) = \frac{\prod_{l=1}^K \Gamma(\tilde{\alpha})}{\Gamma\left(1 + \sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}\right)} \cdot \frac{\Gamma\left(\sum_{k'=1}^K \alpha_{k'} + \tilde{m}_{k'}\right)}{\prod_{l'=1}^K \Gamma(\alpha_{l'} + \tilde{m}_{l'})}$$

$$\Gamma(x+1) = x\Gamma(x) \quad \Rightarrow \quad \Gamma\left(1 + \sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}\right) = \Gamma\left(\sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}\right) \sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}$$

$$\begin{aligned} \Rightarrow p(z_n = k \mid \mathbf{Z}_{-n}) &= \frac{\prod_{l=1}^K \Gamma(\tilde{\alpha})}{\Gamma\left(\sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}\right) \sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}} \cdot \frac{\Gamma\left(\sum_{k'=1}^K \alpha_{k'} + \tilde{m}_{k'}\right)}{\prod_{l'=1}^K \Gamma(\alpha_{l'} + \tilde{m}_{l'})} \\ &= \frac{1}{\sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}} \frac{\prod_{l=1}^K \Gamma(\alpha_l + \tilde{m}_l + [z_n = l])}{\prod_{l'=1}^K \Gamma(\alpha_{l'} + \tilde{m}_{l'})} \\ &= \frac{1}{\sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}} \frac{\Gamma(\alpha_k + \tilde{m}_k + 1) \prod_{l=1, l \neq k}^K \Gamma(\alpha_l + \tilde{m}_l + 0)}{\Gamma(\alpha_k + \tilde{m}_k) \prod_{l'=1, l' \neq k}^K \Gamma(\alpha_{l'} + \tilde{m}_{l'})} \\ &= \frac{1}{\sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}} \frac{\Gamma(\alpha_k + \tilde{m}_k + 1) \prod_{l=1, l \neq k}^K \Gamma(\alpha_l + \tilde{m}_l)}{\Gamma(\alpha_k + \tilde{m}_k) \prod_{l'=1, l' \neq k}^K \Gamma(\alpha_{l'} + \tilde{m}_{l'})} \end{aligned}$$

$$\Gamma(x+1) = x\Gamma(x) \quad \Rightarrow \quad \Gamma(1 + \alpha_k + \tilde{m}_k) = \Gamma(\alpha_k + \tilde{m}_k)(\alpha_k + \tilde{m}_k)$$

$$\begin{aligned} \Rightarrow p(z_n = k \mid \mathbf{Z}_{-n}) &= \frac{1}{\sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}} \frac{\Gamma(\alpha_k + \tilde{m}_k)(\alpha_k + \tilde{m}_k)}{\Gamma(\alpha_k + \tilde{m}_k)} \\ &= \frac{\alpha_k + \tilde{m}_k}{\sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \tilde{m}_{\tilde{k}}} = \frac{\alpha_k + \sum_{i=1, i \neq n}^N [z_i = k]}{\sum_{\tilde{k}=1}^K \alpha_{\tilde{k}} + \sum_{\tilde{k}=1}^K \sum_{j=1, j \neq n} [z_j = \tilde{k}]} \end{aligned}$$

$$\sum_{\tilde{k}=1}^K \sum_{j=1, j \neq n} [z_j = \tilde{k}] = \sum_{j=1, j \neq n}^N \sum_{\tilde{k}=1}^K [z_j = \tilde{k}] = \sum_{j=1, j \neq n}^N 1 = N - 1$$

$$\Rightarrow p(z_n = k \mid \mathbf{Z}_{-n}) = \frac{\alpha_k + \sum_{i=1, i \neq n}^N [z_i = k]}{N - 1 + \sum_{l=1}^K \alpha_l}$$

$$\underline{p(z_n = k \mid \mathbf{x}_n, \mathbf{Z}_{-n}, \Theta) :}$$

$$\begin{aligned} p(z_n = k \mid \mathbf{x}_n, \mathbf{Z}_{-n}, \Theta) &\propto p(\mathbf{x}_n \mid z_n = k, \mathbf{Z}_{-n}, \Theta) p(z_n = k \mid \mathbf{Z}_{-n}) \\ &= p(\mathbf{x}_n \mid z_n = k, \mathbf{Z}_{-n}, \Theta) \frac{\alpha_k + \sum_{i=1, i \neq n}^N [z_i = k]}{N - 1 + \sum_{l=1}^K \alpha_l} \\ &= \left( \prod_{m=1}^{|I|} \theta_{km}^{\sum_{i=1}^{W_n} [x_{ni} = m]} \right) \frac{\alpha_k + \sum_{i=1, i \neq n}^N [z_i = k]}{N - 1 + \sum_{l=1}^K \alpha_l} \end{aligned}$$

$$\underline{\log p(z_n = k \mid \mathbf{x}_n, \mathbf{Z}_{-n}, \Theta) :}$$

$$\log p(z_n = k \mid \mathbf{x}_n, \mathbf{Z}_{-n}, \Theta) \propto \log p(\mathbf{x}_n \mid z_n = k, \mathbf{Z}_{-n}, \Theta) + \log p(z_n = k \mid \mathbf{Z}_{-n})$$

$$\log p(\mathbf{x}_n \mid z_n = k, \mathbf{Z}_{-n}, \Theta) = \sum_{m=1}^{|I|} \log(\theta_{km}) \sum_{i=1}^{W_n} [x_{ni} = m]$$

$$\log p(z_n = k \mid \mathbf{Z}_{-n}) = \log \left( \alpha_k + \sum_{i=1, i \neq n}^N [z_i = k] \right) - \log \left( N - 1 + \sum_{l=1}^K \alpha_l \right)$$

3. **Algorithm 3:** Use the (collapsed) Gibbs sampling algorithm to approximate (using samples) the posterior distribution  $p(\mathbf{Z} \mid \mathbf{X})$ . Derive the posterior.

$$\underline{p(z_n = k \mid X, \mathbf{Z}_{-n})} :$$

$$\begin{aligned} p(z_n = k \mid X, \mathbf{Z}_{-n}) &= p(z_n = k \mid \mathbf{x}_n, X_{-n}, \mathbf{Z}_{-n}) \propto p(z_n = k, \mathbf{x}_n \mid X_{-n}, \mathbf{Z}_{-n}) \\ &= p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) p(z_n = k \mid \cancel{X_{-n}}, \mathbf{Z}_{-n}) \\ &= p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) p(z_n = k \mid \mathbf{Z}_{-n}) \end{aligned}$$

$$\underline{p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n})} :$$

$$\begin{aligned} p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) &= \int p(\mathbf{x}_n, \boldsymbol{\theta}_k \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) d\boldsymbol{\theta}_k \\ &= \int p(\boldsymbol{\theta}_k \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) p(\mathbf{x}_n \mid z_n = k, \boldsymbol{\theta}_k, \cancel{X_{-n}}, \cancel{\mathbf{Z}_{-n}}) d\boldsymbol{\theta}_k \end{aligned}$$

From Exercise 1.1:

$$p(\boldsymbol{\theta}_k \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) = \text{Dir}(\boldsymbol{\theta}_k \mid \boldsymbol{\gamma}') \quad , \text{ where } \gamma'_m = \gamma_m + \sum_{i \neq n} \sum_{j=1}^{W_n} [x_{ij} = m] [z_i = k]$$

$$p(\mathbf{x}_n \mid z_n = k, \boldsymbol{\theta}_k) = \prod_{i=1}^{W_n} \text{Cat}(x_{ni} \mid \boldsymbol{\theta}_k) = \prod_{i=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{[x_{ni}=m]}$$

$$\begin{aligned} \Rightarrow p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) &= \int \text{Dir}(\boldsymbol{\theta}_k \mid \boldsymbol{\gamma}') \prod_{i=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{[x_{ni}=m]} d\boldsymbol{\theta}_k \\ &= \int \frac{1}{B(\boldsymbol{\gamma}')} \prod_{m'=1}^{|I|} \theta_{km'}^{\gamma'_{m'}-1} \prod_{i=1}^{W_n} \prod_{m=1}^{|I|} \theta_{km}^{[x_{ni}=m]} d\boldsymbol{\theta}_k \\ &= \int \frac{1}{B(\boldsymbol{\gamma}')} \prod_{m=1}^{|I|} \theta_{km}^{\gamma'_m-1+\sum_{i=1}^{W_n} [x_{ni}=m]} d\boldsymbol{\theta}_k \\ &= \int \frac{1}{B(\boldsymbol{\gamma}')} \prod_{m=1}^{|I|} \theta_{km}^{\gamma'_m+s_m-1} d\boldsymbol{\theta}_k \quad , \text{ where } s_m = \sum_{i=1}^{W_n} [x_{ni} = m] \\ &= \int \frac{B(\boldsymbol{\gamma}' + \mathbf{s})}{B(\boldsymbol{\gamma}')} \text{Dir}(\boldsymbol{\theta}_k \mid \boldsymbol{\gamma}' + \mathbf{s}) d\boldsymbol{\theta}_k \\ &= \frac{B(\boldsymbol{\gamma}' + \mathbf{s})}{B(\boldsymbol{\gamma}')} \end{aligned}$$

$$p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) = \frac{B(\boldsymbol{\gamma}' + \mathbf{s})}{B(\boldsymbol{\gamma}')} = \frac{\prod_{m=1}^{|I|} \Gamma(\gamma'_m + s_m)}{\Gamma\left(\sum_{m'=1}^{|I|} \gamma'_{m'} + s_{m'}\right)} \cdot \frac{\Gamma\left(\sum_{l=1}^{|I|} \gamma'_l\right)}{\prod_{l'=1}^{|I|} \Gamma(\gamma'_{l'})}$$

$$\begin{aligned} \Gamma(x+1) &= x\Gamma(x) \quad \Rightarrow \text{For } c \in \mathbb{N}: \quad \Gamma(x+c) = (x+c-1)\Gamma(x+c-1) \\ &= (x+c-1)(x+c-2)\Gamma(x+c-2) \\ &= \dots = \left(\prod_{i=1}^c (x+c-i)\right) \Gamma(x) \\ &= \left(\prod_{i=0}^{c-1} (x+i)\right) \Gamma(x) \end{aligned}$$

$$\begin{aligned} \Rightarrow p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) &= \frac{\prod_{m=1}^{|I|} (\Gamma(\gamma'_m) \prod_{i=0}^{s_m-1} (\gamma'_m + i))}{\Gamma\left(\sum_{m'=1}^{|I|} \gamma'_{m'} + s_{m'}\right)} \cdot \frac{\Gamma\left(\sum_{l=1}^{|I|} \gamma'_l\right)}{\prod_{l'=1}^{|I|} \Gamma(\gamma'_{l'})} \\ &= \frac{\left(\prod_{m=1}^{|I|} \Gamma(\gamma'_m)\right) \left(\prod_{m=1}^{|I|} \prod_{i=0}^{s_m-1} (\gamma'_m + i)\right)}{\Gamma\left(\sum_{m'=1}^{|I|} \gamma'_{m'} + s_{m'}\right)} \cdot \frac{\Gamma\left(\sum_{l=1}^{|I|} \gamma'_l\right)}{\prod_{l'=1}^{|I|} \Gamma(\gamma'_{l'})} \\ &= \prod_{m=1}^{|I|} \prod_{i=0}^{s_m-1} (\gamma'_m + i) \cdot \frac{\Gamma\left(\sum_{l=1}^{|I|} \gamma'_l\right)}{\Gamma\left(\sum_{m'=1}^{|I|} \gamma'_{m'} + s_{m'}\right)} \end{aligned}$$

$$\begin{aligned} \sum_{m'=1}^{|I|} \gamma'_{m'} + s_{m'} &= \sum_{m'=1}^{|I|} \gamma'_{m'} + \sum_{i=1}^{W_n} [x_{ni} = m'] = \left(\sum_{m'=1}^{|I|} \gamma'_{m'}\right) + \sum_{i=1}^{W_n} \sum_{m'=1}^{|I|} [x_{ni} = m'] \\ &= \left(\sum_{m'=1}^{|I|} \gamma'_{m'}\right) + \sum_{i=1}^{W_n} 1 = W_n + \sum_{m'=1}^{|I|} \gamma'_{m'} \end{aligned}$$

$$\begin{aligned} \Rightarrow p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) &= \prod_{m=1}^{|I|} \prod_{i=0}^{s_m-1} (\gamma'_m + i) \cdot \frac{\Gamma\left(\sum_{l=1}^{|I|} \gamma'_l\right)}{\Gamma\left(W_n + \sum_{m'=1}^{|I|} \gamma'_{m'}\right)} \\ &= \prod_{m=1}^{|I|} \prod_{i=0}^{s_m-1} (\gamma'_m + i) \cdot \frac{\Gamma\left(\sum_{l=1}^{|I|} \gamma'_l\right)}{\Gamma\left(\sum_{m'=1}^{|I|} \gamma'_{m'}\right) \prod_{i=0}^{W_n-1} (i + \sum_{m'=1}^{|I|} \gamma'_{m'})} \\ &= \frac{\prod_{m=1}^{|I|} \prod_{i=0}^{s_m-1} (\gamma'_m + i)}{\prod_{i=0}^{W_n-1} (i + \sum_{m'=1}^{|I|} \gamma'_{m'})} \end{aligned}$$

$$\underline{p(z_n = k \mid \mathbf{Z}_{-n})} :$$

From Exercise 1.2:

$$p(z_n = k \mid \mathbf{Z}_{-n}) = \frac{\alpha_k + \sum_{i=1, i \neq n}^N [z_i = k]}{N - 1 + \sum_{l=1}^K \alpha_l}$$

$$\begin{aligned} \Rightarrow p(z_n = k \mid X, \mathbf{Z}_{-n}) &= p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) p(z_n = k \mid \mathbf{Z}_{-n}) \\ &= \frac{\prod_{m=1}^{|I|} \prod_{i=0}^{s_m-1} (\gamma'_m + i)}{\prod_{i=0}^{W_n-1} (i + \sum_{m'=1}^{|I|} \gamma'_{m'})} \cdot \frac{\alpha_k + \sum_{i=1, i \neq n}^N [z_i = k]}{N - 1 + \sum_{l=1}^K \alpha_l} \end{aligned}$$

$$\underline{\log p(z_n = k \mid X, \mathbf{Z}_{-n})} :$$

$$\Rightarrow \log p(z_n = k \mid X, \mathbf{Z}_{-n}) = \log p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) + \log p(z_n = k \mid \mathbf{Z}_{-n})$$

$$\begin{aligned} \log p(\mathbf{x}_n \mid z_n = k, X_{-n}, \mathbf{Z}_{-n}) &= \sum_{m=1}^{|I|} \sum_{i=0}^{s_m-1} \log(\gamma'_m + i) - \sum_{i=0}^{W_n-1} \log(i + \sum_{m'=1}^{|I|} \gamma'_{m'}) \\ \log p(z_n = k \mid \mathbf{Z}_{-n}) &= \log \left( \alpha_k + \sum_{i=1, i \neq n}^N [z_i = k] \right) - \log \left( N - 1 + \sum_{l=1}^K \alpha_l \right) \end{aligned}$$

$$\begin{aligned} \Rightarrow \log p(z_n = k \mid X, \mathbf{Z}_{-n}) &= \sum_{m=1}^{|I|} \sum_{i=0}^{s_m-1} \log(\gamma'_m + i) - \sum_{i=0}^{W_n-1} \log(i + \sum_{m'=1}^{|I|} \gamma'_{m'}) \\ &\quad + \log \left( \alpha_k + \sum_{i=1, i \neq n}^N [z_i = k] \right) - \log \left( N - 1 + \sum_{l=1}^K \alpha_l \right) \end{aligned}$$