C.4. Discrete Setting.

Consider the following generative model:

$$b_k \sim \text{Bern}(p), \tag{68}$$

$$\lambda_k \mid b_k \sim \begin{cases} \delta_0 & \text{if } b_k = 0 \\ g(\cdot) & \text{otherwise} \end{cases} \tag{69}$$

$$y_k \mid b_k, \lambda_k, c_k \sim \begin{cases} \delta_0 & \text{if } b_k = 0 \\ F_{c_k \lambda_k}(\cdot) & \text{otherwise.} \end{cases} \tag{70}$$

We augment the data, sampling $\tilde{y}_k \sim F_{(\bar{c}-c_k)\lambda_k}(\tilde{y}_k)$, s.t. marginally, $\bar{y}_k = y_k + \tilde{y}_k \sim F_{\bar{c}\lambda_k}(\bar{y}_k)$.

If $\bar{y}_k > 0$, $b_k = 1$. For $\bar{y}_k = 0$,

$$P(b_k = 1 \mid \bar{y}_k = 0, \bar{c}) = \frac{P(b_k = 1)P(\bar{y}_k = 0 \mid \bar{c}, b_k = 1)}{P(b_k = 1)P(\bar{y}_k = 0 \mid \bar{c}, b_k = 1) + P(b_k = 0)P(\bar{y}_k = 0 \mid \bar{c}, b_k = 0)} \tag{71}$$

$$= \frac{P(b_k = 1) \int P(0, \lambda_k \mid b_k = 1, \bar{c}) d\lambda_k}{P(b_k = 1) \int P(0, \lambda_k \mid b_k = 1, \bar{c}) d\lambda_k + P(b_k = 0)P(\bar{y}_k = 0 \mid \bar{c}, b_k = 0)} \tag{72}$$

$$= \frac{pf(\bar{c}, 0)}{pf(\bar{c}, 0) + (1 - p)} \tag{73}$$

$$\equiv \tilde{p}. \tag{74}$$

That is, when $y_k = 0$, the complete conditional is only a function of p and $\bar{c}$ and does not depend on k. Therefore, for all of the classes $\{k : \bar{y}_k = 0\}$ the b_k, in their complete conditional, are i.i.d. We can then sample them cheaply.