Supplementary material: Poisson-randomised DirBN: large mutation is needed in Dirichlet belief networks

Xuhui Fan
UNSW Data Science Hub, and School of Mathematics and Statistics, University of New South Wales
xuhui.fan@unsw.edu.au

Bin Li
Shanghai Key Laboratory of IIP, School of Computer Science, Fudan University
libin@fudan.edu.cn

Yaqiong Li
Australian Artificial Intelligence Institute, University of Technology, Sydney
yaqiong.li@student.uts.edu.au

Scott A. Sisson
UNSW Data Science Hub, and School of Mathematics and Statistics, University of New South Wales
scott.sisson@unsw.edu.au

1 Integer variable sampling for deep stochastic integer networks

In this section, we present the detail calculation of integer variable sampling for deep stochastic integer networks. The other random variables are the same as Pois-DirBN.

Sampling $C_{iki'}^{(l+\frac{1}{2})}$ The posterior probability involving $C_{iki'}^{(l+\frac{1}{2})}, \boldsymbol{\pi}_i^{(l)}, \boldsymbol{\pi}_{i'}^{(l+1)}$ is:

$$
\begin{aligned}
& P(C_{iki'}^{(l+\frac{1}{2})}, \boldsymbol{\pi}_i^{(l)}, \boldsymbol{\pi}_{i'}^{(l+1)}|-) \\
\propto & P(\boldsymbol{\pi}_i^{(l)}|\boldsymbol{C}^{(l-\frac{1}{2})}, \boldsymbol{D}^{(l-\frac{1}{2})}) P(C_{ik\cdot}^{(l+\frac{1}{2})}|\boldsymbol{\pi}_i^{(l)}) P(D_{ik\cdot}^{(l+\frac{1}{2})}|\boldsymbol{\pi}_i^{(l)}) P(\boldsymbol{\pi}_{i'}^{(l+1)}|\boldsymbol{C}^{(l+\frac{1}{2})}, \boldsymbol{D}^{(l+\frac{1}{2})}) P(C_{i'k\cdot}^{(l+\frac{3}{2})}|\boldsymbol{\pi}_{i'}^{(l+1)}) P(D_{i'k\cdot}^{(l+\frac{3}{2})}|\boldsymbol{\pi}_{i'}^{(l+1)}) \\
\propto & \left[\prod_k [\pi_{ik}^{(l)}]^{\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki}^{(l-\frac{1}{2})} + \sum_{k''} D_{ik''k}^{(l-\frac{1}{2})} - 1}\right] \left[\frac{(\sum_k C_{ik\cdot}^{(l+\frac{1}{2})})!}{C_{ik\cdot}^{(l+\frac{1}{2})}!} \prod_k [\pi_{ik}^{(l)}]^{C_{ik\cdot}^{(l+\frac{1}{2})}}\right] \left[\prod_k [\pi_{ik}^{(l)}]^{D_{ik\cdot}^{(l+\frac{1}{2})}}\right] \\
& \cdot \left[\frac{\Gamma(\sum_k(\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})}))}{\Gamma(\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})})} \prod_k [\pi_{i'k}^{(l+1)}]^{\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})} - 1}\right] \\
& \cdot \left[\prod_k [\pi_{i'k}^{(l+1)}]^{C_{i'k\cdot}^{(l+\frac{3}{2})} + D_{i'k\cdot}^{(l+\frac{3}{2})}}\right] e^{-M} \frac{M^{\sum_k C_{ik\cdot}^{(l+\frac{1}{2})}}}{(\sum_k C_{ik\cdot}^{(l+\frac{1}{2})})!} \cdot \frac{C_{ik\cdot}^{(l+\frac{1}{2})}!}{C_{iki'}^{(l+\frac{1}{2})}!} [\omega_{ii'}]^{C_{iki'}^{(l+\frac{1}{2})}} \\
\propto & \prod_k [\pi_{ik}^{(l)}]^{\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki}^{(l-\frac{1}{2})} + \sum_{k''} D_{ik''k}^{(l-\frac{1}{2})} + C_{ik\cdot}^{(l+\frac{1}{2})} + D_{ik\cdot}^{(l+\frac{1}{2})} - 1} \prod_k [\pi_{i'k}^{(l+1)}]^{\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})} + C_{i'k\cdot}^{(l+\frac{3}{2})} + D_{i'k\cdot}^{(l+\frac{3}{2})} - 1} \\
& \cdot \left[\frac{\Gamma(\sum_k(\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})}))}{\Gamma(\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})})}\right] \cdot \frac{[M\omega_{ii'}]^{C_{iki'}^{(l+\frac{1}{2})}}}{C_{iki'}^{(l+\frac{1}{2})}!}
\end{aligned} \tag{1}
$$

Preprint. Under review.

By integrating out $\boldsymbol{\pi}_i^{(l)}, \boldsymbol{\pi}_{i'}^{(l+1)}$, we get the posterior ratio for $C_{iki'}^{(l+\frac{1}{2})}$:

$$
\begin{aligned}
&P(C_{iki'}^{(l+\frac{1}{2})}|-) = \int_{\boldsymbol{\pi}_i^{(l)}, \boldsymbol{\pi}_{i'}^{(l+1)}} P(C_{iki'}^{(l+\frac{1}{2})}, \boldsymbol{\pi}_i^{(l)}, \boldsymbol{\pi}_{i'}^{(l+1)}|-) d\boldsymbol{\pi}_i^{(l)} d\boldsymbol{\pi}_{i'}^{(l+1)} \\
\propto & \frac{[M\omega_{ii'}]^{C_{iki'}^{(l+\frac{1}{2})}}}{C_{iki'}^{(l+\frac{1}{2})}!} \cdot \frac{\Gamma(\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki}^{(l-\frac{1}{2})} + \sum_{k''} D_{ik''k}^{(l-\frac{1}{2})} + C_{ik\cdot}^{(l+\frac{1}{2})} + D_{ik\cdot}^{(l+\frac{1}{2})})}{\Gamma(\sum_k [\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki}^{(l-\frac{1}{2})} + \sum_{k''} D_{ik''k}^{(l-\frac{1}{2})} + C_{ik\cdot}^{(l+\frac{1}{2})} + D_{ik\cdot}^{(l+\frac{1}{2})}])} \\
&\cdot \left[\frac{\Gamma(\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})} + C_{i'k\cdot}^{(l+\frac{3}{2})} + D_{i'k\cdot}^{(l+\frac{3}{2})})}{\Gamma(\sum_k [\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})} + C_{i'k\cdot}^{(l+\frac{3}{2})} + D_{i'k\cdot}^{(l+\frac{3}{2})}])} \right] \\
&\cdot \left[\frac{\Gamma(\sum_k (\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})}))}{\Gamma(\alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})})} \right]
\end{aligned} \tag{2}
$$

Thus, when $0 < l < L-1$, the posterior ratio of $C_{iki'}^{(l+\frac{1}{2})}$ is:

$$
P(C_{iki'}^{(l+\frac{1}{2})}|-) \propto \frac{[\omega_{ii'} M_c]^{C_{iki'}^{(l+\frac{1}{2})}}}{C_{iki'}^{(l+\frac{1}{2})}!} \frac{\Gamma(\sum_{k''} v_{k''}^{(0)})}{\Gamma(v_k^{(0)})} \cdot \frac{\Gamma(v_k^{(1)})}{\Gamma(\sum_{k''} v_{k''}^{(1)})} \cdot \frac{\Gamma(v_k^{(2)})}{\Gamma(\sum_{k''} v_{k''}^{(2)})}
$$

where $v_k^{(0)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})}$, $v_k^{(1)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki}^{(l-\frac{1}{2})} + \sum_{k''} D_{ik''k}^{(l-\frac{1}{2})} + \sum_{i''} C_{iki''}^{(l+\frac{1}{2})} + \sum_{k''} D_{ikk''}^{(l+\frac{1}{2})}$,$v_k^{(2)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(l+\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(l+\frac{1}{2})} + \sum_{i''} C_{i'ki''}^{(l+\frac{3}{2})} + \sum_{k''} D_{i'kk''}^{(l+\frac{3}{2})}$.

We can use similar techniques as in the main paper, to augment the second term $\frac{\Gamma(\sum_{k''} v_{k''}^{(0)})}{\Gamma(v_k^{(0)})}$ in the r.h.s. of the above equation. For the third and fourth term, we can use Beta random variables to augment these Gamma ratios.

When $l = 0$, the posterior ratio of $C_{iki'}^{(\frac{1}{2})}$ is:

$$
P(C_{iki'}^{(\frac{1}{2})}|-) \propto \frac{[\omega_{ii'} M_c]^{C_{iki'}^{(\frac{1}{2})}}}{C_{ik'i'}^{(\frac{1}{2})}!} \frac{\Gamma(\sum_{k''} v_{k''}^{(0)})}{\Gamma(v_k^{(0)})} \cdot \frac{\Gamma(v_k^{(1)})}{\Gamma(\sum_{k''} v_{k''}^{(1)})} \cdot \frac{\Gamma(v_k^{(2)})}{\Gamma(\sum_{k''} v_{k''}^{(2)})}
$$

where $v_k^{(0)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(\frac{1}{2})}$, $v_k^{(1)} = \alpha_k^{(\pi)} + \sum_{i''} C_{iki''}^{(\frac{1}{2})} + \sum_{k''} D_{ikk''}^{(\frac{1}{2})}$,$v_k^{(2)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(\frac{1}{2})} + \sum_{i''} C_{i'ki''}^{(\frac{3}{2})} + \sum_{k''} D_{i'kk''}^{(\frac{3}{2})}$.

When $l = L-1$, the posterior ratio of $C_{iki'}^{(L-\frac{1}{2})}$ is:

$$
P(C_{iki'}^{(L-\frac{1}{2})}|-) \propto \frac{[\omega_{ii'} M_c]^{C_{iki'}^{(L-\frac{1}{2})}}}{C_{iki'}^{(L-\frac{1}{2})}!} \frac{\Gamma(\sum_{k''} v_{k''}^{(0)})}{\Gamma(v_k^{(0)})} \cdot \frac{\Gamma(v_k^{(1)})}{\Gamma(\sum_{k''} v_{k''}^{(1)})} \cdot \frac{\Gamma(v_k^{(2)})}{\Gamma(\sum_{k''} v_{k''}^{(2)})}
$$

where $v_k^{(0)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(L-\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(L-\frac{1}{2})}$, $v_k^{(1)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki}^{(l-\frac{3}{2})} + \sum_{k''} D_{ik''k}^{(l-\frac{3}{2})} + \sum_{i''} C_{iki''}^{(L-\frac{1}{2})} + \sum_{k''} D_{ikk''}^{(L-\frac{1}{2})}$,$v_k^{(2)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki'}^{(L-\frac{1}{2})} + \sum_{k''} D_{i'k''k}^{(L-\frac{1}{2})} + X_{i'k}$.

Sampling $D_{ikk'}^{(l+\frac{1}{2})}$ Similarly, when $0 < l < L-1$, the posterior ratio of $D_{ikk'}^{(l+\frac{1}{2})}$ is expressed as:

$$
P(D_{ikk'}^{(l+\frac{1}{2})}|-) \propto \frac{[\phi_{kk'} M_d]^{D_{ikk'}^{(l+\frac{1}{2})}}}{D_{ikk'}^{(l+\frac{1}{2})}!} \cdot \frac{\Gamma(\sum_{k''} v_{k''}^{(0)})}{\Gamma(v_{k'}^{(0)})} \cdot \frac{\Gamma(v_k^{(1)})}{\Gamma(\sum_{k''} v_{k''}^{(1)})} \cdot \frac{\Gamma(v_{k'}^{(2)})}{\Gamma(\sum_{k''} v_{k''}^{(2)})}
$$

where $v_{k'}^{(0)} = \alpha_{k'}^{(\pi)} + \sum_{i''} C_{i''k'i}^{(l+\frac{1}{2})} + \sum_{k''} D_{ik''k'}^{(l+\frac{1}{2})}$, $v_k^{(1)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki}^{(l-\frac{1}{2})} + \sum_{k''} D_{ik''k}^{(l-\frac{1}{2})} + \sum_{i''} C_{iki''}^{(l+\frac{1}{2})} + \sum_{k''} D_{ikk''}^{(l+\frac{1}{2})}$,$v_{k'}^{(2)} = \alpha_{k'}^{(\pi)} + \sum_{i''} C_{i''k'i}^{(l+\frac{1}{2})} + \sum_{k''} D_{ik''k'}^{(l+\frac{1}{2})} + \sum_{i''} C_{ik'i''}^{(l+\frac{3}{2})} + \sum_{k''} D_{ik'k''}^{(l+\frac{3}{2})}$.

When $l = 0$, the posterior ratio of $D_{ikk'}^{(\frac{1}{2})}$ is:

$$
P(D_{ikk'}^{(\frac{1}{2})}|-) \propto \frac{[\phi_{kk'} M_d]^{D_{ikk'}^{(\frac{1}{2})}}}{D_{ikk'}^{(\frac{1}{2})}!} \cdot \frac{\Gamma(\sum_{k''} v_{k''}^{(0)})}{\Gamma(v_{k'}^{(0)})} \cdot \frac{\Gamma(v_k^{(1)})}{\Gamma(\sum_{k''} v_{k''}^{(1)})} \cdot \frac{\Gamma(v_{k'}^{(2)})}{\Gamma(\sum_{k''} v_{k''}^{(2)})}
$$

where $v_{k'}^{(0)} = \alpha_{k'}^{(\pi)} + \sum_{i''} C_{i''k'i}^{(\frac{1}{2})} + \sum_{k''} D_{ik''k'}^{(\frac{1}{2})}$, $v_k^{(1)} = \alpha_k^{(\pi)} + \sum_{i''} C_{iki''}^{(\frac{1}{2})} + \sum_{k''} D_{ikk''}^{(\frac{1}{2})}$,$v_{k'}^{(2)} = \alpha_{k'}^{(\pi)} + \sum_{i''} C_{i''k'i}^{(\frac{1}{2})} + \sum_{k''} D_{ik''k'}^{(\frac{1}{2})} + \sum_{i''} C_{ik'i''}^{(\frac{3}{2})} + \sum_{k''} D_{ik'k''}^{(\frac{3}{2})}$.

When $l = L - 1$, the posterior ratio of $D_{ikk'}^{(L-\frac{1}{2})}$ is:

$$P(D_{ikk'}^{(L-\frac{1}{2})}|-) \propto \frac{[\phi_{kk'} M_d]^{D_{ikk'}^{(L-\frac{1}{2})}}}{D_{ikk'}^{(L-\frac{1}{2})}!} \cdot \frac{\Gamma(\sum_{k''} v_{k''}^{(0)})}{\Gamma(v_{k'}^{(0)})} \cdot \frac{\Gamma(v_k^{(1)})}{\Gamma(\sum_{k''} v_{k''}^{(1)})} \cdot \frac{\Gamma(v_{k'}^{(2)})}{\Gamma(\sum_{k''} v_{k''}^{(2)})}$$

where $v_{k'}^{(0)} = \alpha_{k'}^{(\pi)} + \sum_{i''} C_{i''k'i}^{(L-\frac{1}{2})} + \sum_{k''} D_{ik''k'}^{(L-\frac{1}{2})}$, $v_k^{(1)} = \alpha_k^{(\pi)} + \sum_{i''} C_{i''ki}^{(L-\frac{3}{2})} + \sum_{k''} D_{ik''k}^{(L-\frac{3}{2})} + \sum_{i''} C_{iki''}^{(L-\frac{1}{2})} + \sum_{k''} D_{ikk''}^{(L-\frac{1}{2})}$,$v_{k'}^{(2)} = \alpha_{k'}^{(\pi)} + \sum_{i''} C_{i''k'i}^{(L-\frac{1}{2})} + \sum_{k''} D_{ik''k'}^{(L-\frac{1}{2})} + X_{ik'}$.

2 Running time results

Per-iteration run time: Figure 1 shows the running time per iteration for the four datasets. As we can see, our inference algorithm scales linearly with the number of layers. We might need around 15 hours for 1 000 MCMC iterations on the *PPI* data.

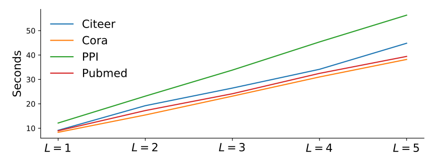

Figure 1: Per-iteration run time of Pois-DirBN-CD on the four real datasets.