Предположим такой порождающий процесс:

1) Сэмплируем из факторного анализатора (**FA**): $h^3 \sim N(m + Fs, \Psi)$, где $s \sim N(0, I)$
2) Сэмплируем из *условной* машины Больцмана (**CDBM**): $p(v, h^1, h^2|h^3)$

,где

$$P(v, h^1, h^2, h^3) = \frac{1}{\mathcal{Z}} \exp\left(v^T W^1 h^1 + h^{1^T} W^2 h^2 + h^{2^T} W^3 h^3 + a^T v + b^T h^1 + c^T h^2 - \frac{1}{2}(h^3 - d)^T (h^3 - d)\right)$$

$$P(v, h^1, h^2|h^3) = \frac{P(v, h^1, h^2, h^3)}{\sum_{v,h^1,h^2} P(v, h^1, h^2, h^3)} = \frac{P(v, h^1, h^2, h^3)}{P(h^3)}$$
$$= \frac{1}{\mathcal{Z}(\mathbf{h^3})} \exp\left(v^T W^1 h^1 + h^{1^T} W^2 h^2 + h^{2^T} W^3 \mathbf{h^3} + a^T v + b^T h^1 + c^T h^2\right)$$

Какое распределение соответствует этому порождающему процессу?

$$P(v, h^1, h^2, h^3, s) = \mathrm{CDBM}(v, h^1, h^2|h^3)\mathrm{FA}(h^3, s)$$

Но $\mathrm{CDBM}(v, h^1, h^2|h^3)$ имеет знаменатель $\mathcal{Z}(\mathbf{h^3})$, зависящий от h^3 и не поддающийся вычислению.

Таким образом, мы не можем использовать вариационный вывод для получения аппроксимации апостериорного распределения

$$q(h^1, h^2, h^3, s|v) = q(h^1)q(h^2)q(h^3)q(s)$$
$$\log q(h^3) = \mathbb{E}_{q(h^1)q(h^2)q(s)}\{\log P(v, h^1, h^2, h^3, s)\} + \mathrm{const}$$

,так как не получится посчитать ожидание $\mathbb{E}_{q(h^1)q(h^2)q(s)}\{-\log \mathcal{Z}(\mathbf{h^3})\}$

Сэмплирование по Гиббсу $P(h^3| \dots)$ тоже неясно как делать из-за знаменателя. В статье формулу (11) я немного не понял: почему там $P(h_n^2|h_n^3)$ первый множитель, а не $P(h_n^3|h_n^2)$?

В статье они комбинируют направленную и ненаправленную модель. Для двух ненаправленных моделей можно бы было просто перемножить «факторы» и получить *совместное* распределение, я так понимаю. А в этом случае неясно как делать *вывод* и обучение в таких моделях.