3.3 Complicated likelihood for leaky units on a tree

Let $u_1 \downarrow$ and $u_2 \downarrow$ be the children of node u, then we get the following

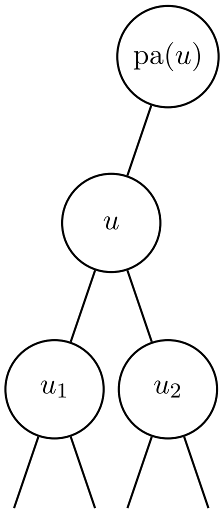

Figure 1: Binary tree starting at parent of u, branches below u_1, u_2 denotes subtrees

Let $p_{Z_{u\downarrow}} = \sum_{Z_{u\downarrow}} p(X_u, X_{u\downarrow}, Z_{u\downarrow}|Z_u, Z_{\text{pa}(u)})$ where $\sum_{Z_{u\downarrow}}$ denotes the summation over all latent variables below u, not including u, then

$$
\begin{aligned}
p_{Z_{u\downarrow}} &= \sum_{Z_{u\downarrow}} p(X_u, X_{u_1}, X_{u_2}, X_{u_1\downarrow}, X_{u_2\downarrow}, Z_{u_1\downarrow}, Z_{u_2\downarrow} | Z_u, Z_{\mathrm{pa}(u)}, Z_{u_1}, Z_{u_2}) p(Z_{u_1}) p(Z_{u_1\downarrow}) \\
&= \sum_{Z_{u\downarrow}} p(X_u | Z_u, Z_{\mathrm{pa}(u)}, Z_{u_1}, Z_{u_2}) p(Z_{u_1}) p(Z_{u_2}) p(X_{u_1}, X_{u_1\downarrow}, Z_{u_1\downarrow} | Z_{u_1}, Z_u) p(X_{u_2}, X_{u_2\downarrow}, Z_{u_2\downarrow} | Z_{u_2}, Z_u) \\
&= \sum_{Z_{u\downarrow}} \Big[p(X_u | Z_u, Z_{\mathrm{pa}(u)}, Z_{u_1}, Z_{u_2}) p(Z_{u_1}) p(Z_{u_2}) \\
&\qquad p(X_{u_1}, X_{u_1\downarrow}, Z_{u_1\downarrow} | Z_{u_1}, Z_u) p(X_{u_2}, X_{u_2\downarrow}, Z_{u_2\downarrow} | Z_{u_2}, Z_u) \Big] \\
&= \sum_{Z_{u_1}, Z_{u_2}} \Big[p(X_u | Z_u, Z_{\mathrm{pa}(u)}, Z_{u_1}, Z_{u_2}) p(Z_{u_1}) p(Z_{u_2}) \\
&\qquad \Big(\sum_{Z_{u_1\downarrow}} p(X_{u_1}, X_{u_1\downarrow}, Z_{u_1\downarrow} | Z_{u_1}, Z_u) \Big) \Big(\sum_{Z_{u_2\downarrow}} p(X_{u_2}, X_{u_2\downarrow}, Z_{u_2\downarrow} | Z_{u_2}, Z_u) \Big) \Big] \\
&= \sum_{Z_{u_1}, Z_{u_2}} \Big[p(X_u | Z_u, Z_{\mathrm{pa}(u)}, Z_{u_1}, Z_{u_2}) p(Z_{u_1}) p(Z_{u_2}) \Big(p_{Z_{u_1\downarrow}} \Big) \Big(p_{Z_{u_2\downarrow}} \Big) \Big] \\
&= \sum_{Z_{u_1}, Z_{u_2}} \Big[\mathcal{N}\Big(X_u | (1-\alpha)\mu_{Z_u} + \frac{\alpha}{3}(\mu_{Z_{u_1}} + \mu_{Z_{u_2}} + \mu_{Z_{\mathrm{pa}(u)}}) \Big) \cdot \\
&\qquad \cdot \pi(Z_{u_1}) \pi(Z_{u_2}) \Big(p_{Z_{u_1\downarrow}} \Big) \Big(p_{Z_{u_2\downarrow}} \Big) \Big]
\end{aligned}
$$

On a more compact form we have thus shown that

$$
\begin{aligned}
p_{Z_{u\downarrow}} &= \sum_{Z_{u\downarrow}} p(X_u, X_{u\downarrow}, Z_{u\downarrow} | Z_u, Z_{\mathrm{pa}(u)}) && (1) \\
&= \sum_{Z_{u_1}, Z_{u_2}} \Big[\mathcal{N}\Big(X_u | (1-\alpha)\mu_{Z_u} + \frac{\alpha}{3}(\mu_{Z_{u_1}} + \mu_{Z_{u_2}} + \mu_{Z_{\mathrm{pa}(u)}}) \Big) \cdot && (2) \\
&\qquad \cdot \pi(Z_{u_1}) \pi(Z_{u_2}) \Big(p_{Z_{u_1\downarrow}} \Big) \Big(p_{Z_{u_2\downarrow}} \Big) \Big] && (3)
\end{aligned}
$$

Which shows the recursion.

We have thus ended up with two new subproblems $p_{Z_{u_1\downarrow}}$ and $p_{Z_{u_2\downarrow}}$ from $p_{Z_{u\downarrow}}$ which can be divided into subproblems continuously until the leaves are reached.

It is important to note that $p_{Z_{u_1\downarrow}}$ is independent of Z_{u_2} and similarly $p_{Z_{u_2\downarrow}}$ is independent of Z_{u_1}. It is thus only necessary to compute $p_{Z_{u_1\downarrow}}$ when summing over Z_{u_1}. When summing over Z_{u_2} the value for $p_{Z_{u_1\downarrow}}$ can be computed once, stored and then be reused. Applying this for all subproblems results in a linear algorithm for computing $p(X)$.

Special case: root

Let u denote the root and u_1, u_2 denote its children.

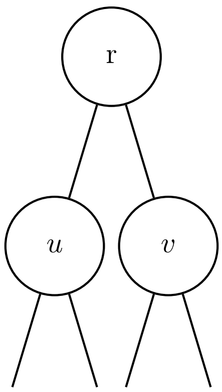

Figure 2: Binary tree starting at root r, branches below u, v denotes subtrees

$$
\begin{aligned}
p(X) = \sum_{Z_u, Z_{u_1}, Z_{u_2}} & \Big[p(X_u | Z_{u_1}, Z_{u_2}, Z_u) p(Z_{u_1}, Z_{u_2}, Z_u) \\
& \Big(\sum_{Z_{u_1\downarrow}} p(X_{u_1}, X_{u_1\downarrow}, Z_{u_1\downarrow} | Z_{u_1}, Z_u) \Big) \Big(\sum_{Z_{u_2\downarrow}} p(X_{u_2}, X_{u_2\downarrow}, Z_{u_2\downarrow} | Z_{u_2}, Z_u) \Big) \Big] \\
= \sum_{Z_u, Z_{u_1}, Z_{u_2}} & \Big[\mathcal{N}\Big(X_u | (1-\alpha)\mu_{Z_u} + \frac{\alpha}{2}(\mu_{Z_{u_1}} + \mu_{Z_{u_2}}) \Big) \pi(Z_u)\pi(Z_{u_1})\pi(Z_{u_2}) \quad (4) \\
& \Big(\sum_{Z_{u_1\downarrow}} p(X_{u_1}, X_{u_1\downarrow}, Z_{u_1\downarrow} | Z_{u_1}, Z_u) \Big) \Big(\sum_{Z_{u_2\downarrow}} p(X_{u_2}, X_{u_2\downarrow}, Z_{u_2\downarrow} | Z_{u_2}, Z_u) \Big) \Big] \quad (5)
\end{aligned}
$$

Special case: leaf

Let u_1 denote a *leaf node*, it thus has no children which implies that $Z_{u_1\downarrow} = \emptyset$.
Using previous definition of $p_{Z_{u\downarrow}}$ yields

$$
\begin{aligned}
p_{Z_{u_1\downarrow}} &= \sum_{Z_{u_1\downarrow}} p(X_{u_1}, X_{u_1\downarrow}, Z_{u_1\downarrow} | Z_{u_1}, Z_{\mathrm{pa}(u_1)}) \\
&= p(X_{u_1} | Z_{u_1}, Z_{\mathrm{pa}(u_1)}) \\
&= \mathcal{N}\Big(X_{u_1} | (1-\alpha)\mu_{Z_{u_1}} + \alpha\mu_{Z_{\mathrm{pa}(u_1)}} \Big) \quad (6)
\end{aligned}
$$