## Q2 Equivalent networks

$$\vec{a}^{(1)} = W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}$$

$$\vec{a}^{(2)} = W^{(2)}\vec{a}^{(1)} + \vec{b}^{(2)}$$

$$\vec{a}^{(3)} = W^{(3)}\vec{a}^{(2)} + \vec{b}^{(3)}$$

for network in Figure 2.a

$$\begin{aligned}\vec{a}^{(3)} &= W^{(3)}\vec{a}^{(2)} + \vec{b}^{(3)} \\ &= W^{(3)}\left(W^{(2)}\vec{a}^{(1)} + \vec{b}^{(2)}\right) + \vec{b}^{(3)} \\ &= W^{(3)}\left[W^{(2)}\left(W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}\right) + \vec{b}^{(2)}\right] + \vec{b}^{(3)} \\ &= \underbrace{W^{(1)}W^{(2)}W^{(3)}}_{\hookrightarrow \tilde{W}}\vec{a}^{(0)} + \underbrace{W^{(2)}W^{(3)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}}_{\hookrightarrow \tilde{b}}\end{aligned}$$

so for network in Figure 2.b

$$\tilde{W} = W^{(1)}W^{(2)}W^{(3)}\vec{a}^{(0)} \qquad \tilde{b} = W^{(2)}W^{(3)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}$$