2) $h(u) = \frac{1-e^{-u}}{1+e^{-u}}$ $\tanh(u) = \frac{e^{u}-e^{-u}}{e^{u}+e^{-u}}$

a) $\frac{d}{du}\left[\frac{1-e^{-u}}{1+e^{-u}}\right] = \frac{\partial}{\partial u}\left[\frac{2}{1+e^{-u}} - 1\right]$

$= \frac{2}{(1+e^{-u})^2}$ $\boxed{\lim_{u \to \pm 0}\left[\frac{2}{(1+e^{-u})^2}\right] = \frac{1}{2}}$

$\frac{\partial}{\partial u}\left[(e^{u}-e^{-u})(e^{u}+e^{-u})^{-1}\right] = (e^{u}+e^{-u})(e^{u}+e^{-u})^{-1} +$

$(e^{u}-e^{-u})(-1)(e^{u}-e^{-u})(e^{u}+e^{-u})^{-2})$

$= 1 - \frac{(e^{u}-e^{-u})^2}{(e^{u}+e^{-u})^2}$ $\boxed{\lim_{u \to \pm 0}\left[1 - \left(\frac{e^{u}-e^{-u}}{e^{u}+e^{-u}}\right)^2\right] = 1}$

therefore, $\tanh(u)$ is deeper than $h(u)$ around 0

b) How to make slope deeper.

make 2^{nd} term in derivative go to 0 faster $\Rightarrow$

$\tanh(u) \Rightarrow \frac{e^{u/k} - e^{-u/k}}{e^{u}+e^{-u}}$, where $0<k<1$

c) May want slope to be steeper in cases where one wants something more like a delta function i.e. in cases where the on/off switching is more like a switch

5) $\left[E = \dfrac{\frac{1}{2} \sum^{N} \sum^{K} (z_{kn} - t_{kn})^2}{NK} \right]$

$$z_k = h(v_k) = \frac{2}{1+e^{-v_k}} - 1$$

$$\frac{\partial E}{\partial b_{jk}} = \left(\frac{\partial E}{\partial z_k}\right)\left(\frac{\partial z_k}{\partial v_k}\right)\left(\frac{\partial v_k}{\partial b_{jk}}\right)$$

$$v_k = b_{0k} + \sum^{J} b_{jk} y_j$$

$$y_j = h(u_j) = \frac{2}{1+e^{u_j}} - 1$$

$$\left[u_j = a_{0j} + \sum^{I} a_{ij} x_j \right]$$

$$\frac{\partial v_k}{\partial b_{jk}} = \frac{\partial}{\partial b_{jk}}\left(b_{0k} + \sum^{J} b_{jk} y_j\right)$$

$$= \begin{cases} 1 & j=0 \\ y_j & j>0 \end{cases}$$

$$\frac{\partial z_k}{\partial v_k} = \frac{\partial}{\partial v_k}\left(\frac{2}{1+e^{-v_k}} - 1\right)$$

$$= \frac{2}{(1+e^{-v_k})^2}$$

$$\frac{\partial E}{\partial z_k} = \frac{\partial}{\partial z_k}\left[\frac{1}{2NK} \sum^{N} \sum^{K} (z_{kn} - t_{kn})^2\right] = \cancel{\frac{1}{2NK}}$$

$$= \frac{1}{NK} \sum^{N} (z_{kn} - t_{kn})$$

$$\frac{\partial E}{\partial b_{jk}} = \frac{2}{NK} \frac{\sum^{N} z_{kn} - t_{kn}}{(1+e^{-v_k})^2} \begin{cases} 1 & j=0 \\ y_j & j>0 \end{cases}$$

3 (cont)

$$\frac{\partial E}{\partial a_{ij}} = \left(\sum_{k=1}^{K} \frac{\partial E}{\partial z_k} \frac{\partial z_k}{\partial v_k} \frac{\partial v_k}{\partial y_j}\right) \frac{\partial y_j}{\partial u_j} \frac{\partial u_j}{\partial a_{ij}}$$

$$\sum_{k=1}^{K} \left[\frac{2}{NK} \frac{b_{jk}}{(1+e^{-v_k})^2} (z_k - t_k) \right]$$

$$\frac{\partial y_j}{\partial u_j} = \frac{2}{1+e^{-u_j}}$$

$$\frac{\partial u_j}{\partial a_{ij}} = \begin{cases} 1 & i = 0 \\ x_j & i > 0 \end{cases}$$

$$\frac{\partial E}{\partial a_{ij}} = \left[\sum^{K} \frac{2}{NK} \frac{(b_{jk})(z_k - t_k)}{(1+e^{-v_k})^2} \right] \left(\frac{2}{(1+e^{-u_j})^2} \right) \begin{cases} 1 & i = 0 \\ x_j & i > 0 \end{cases}$$

$$\frac{\partial E}{\partial c_{ik}} = \frac{\partial E}{\partial z_k} \frac{\partial z_k}{\partial w_k} \frac{\partial w_k}{\partial c_{ik}}$$

$$w_k = c_{0k} + \sum^{I} b_{ik} t$$

$$\frac{\partial E}{\partial c_{ik}} = \frac{2}{NK} \cdot \frac{\sum^{N} z_{kn} - t_{kn}}{(1+e^{-w_k})^2} \cdot \begin{cases} 1 & i = 0 \\ y_i & i > 0 \end{cases}$$

4)

$$\tanh(u) = \frac{e^{u}-e^{-u}}{e^{u}+e^{-u}} \qquad \frac{\partial}{\partial u}\tanh(u) = 1 - \left(\frac{e^{u}-e^{-u}}{e^{u}+e^{-u}}\right)^2$$

$$\frac{\partial E}{\partial b_{jk}} = \left(\frac{\partial E}{\partial z_k}\right)\left(\frac{\partial z_k}{\partial v_k}\right)\left(\frac{\partial v_k}{\partial b_{jk}}\right)$$

$$= \frac{2}{NK}\left(\sum^{N} z_{kn}-t_{kn}\right)\left(1-\tanh(v_k)\right)\begin{cases} 1 & j=0 \\ y_j & j>0 \end{cases}$$

$$\frac{\partial E}{\partial a_{ij}} = \left(\sum^{K} \frac{2}{NK} b_{jk}\left[1-\tanh(v_k)\right]\right)\left[1-\tanh(u_k)\right]\begin{cases} 1 & j=0 \\ x_j & j>0 \end{cases}$$

$$\sum^{N}(z_k-t_k)$$

$$\frac{\partial E}{\partial c_{ik}} = \frac{2}{NK}\sum^{N}(z_k-t_k)\left[1-\tanh(\omega_k)\right]\begin{cases} 1 & i=0 \\ y_i & i>0 \end{cases}, \quad \omega_k = c_{0k}+\sum^{I} c_{ik}$$