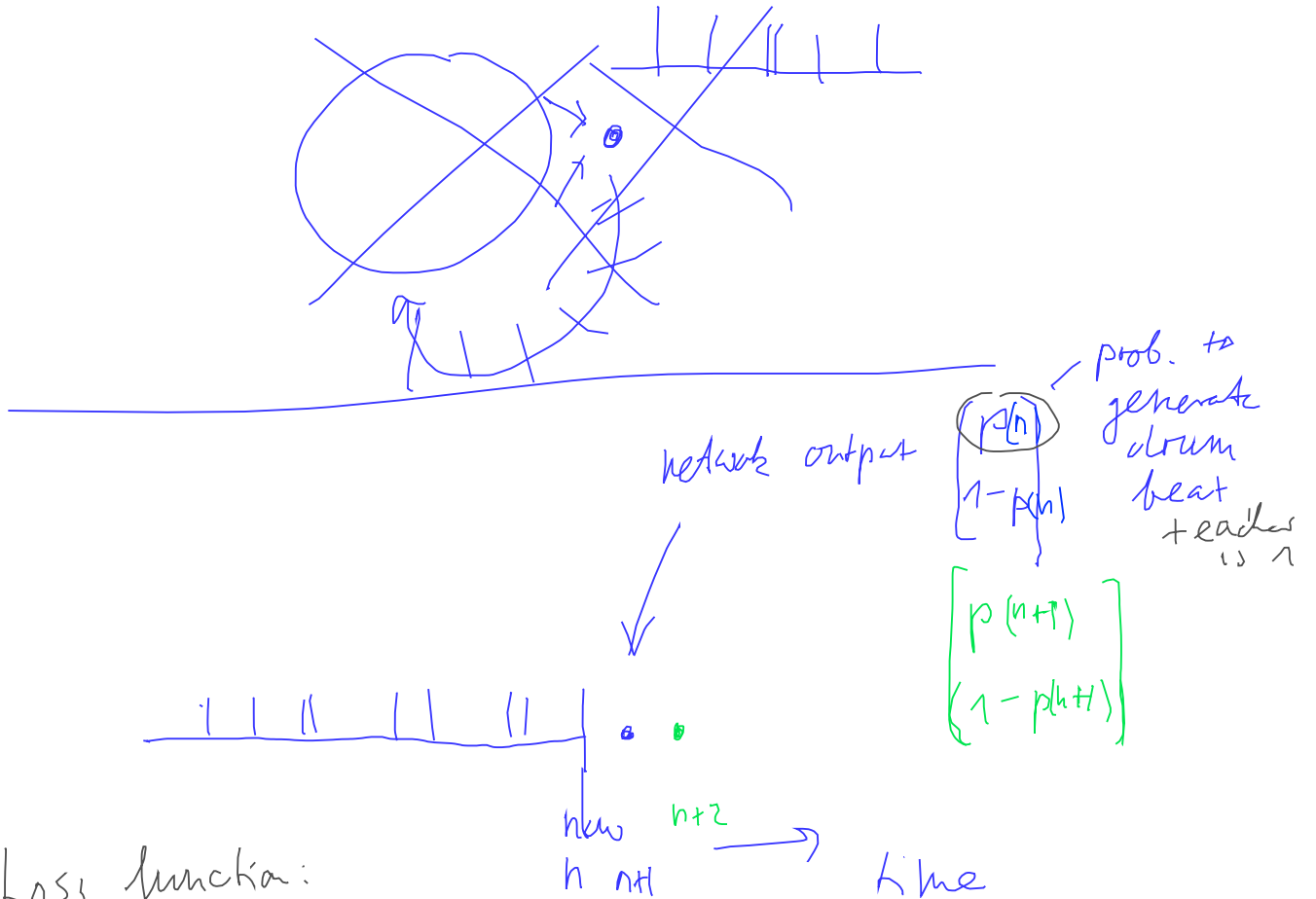


training



Loss function:

$$L_n = \begin{cases} \log p(n) & \text{if teacher is 1} \\ -\log(1-p(n)) & \text{if teacher is 0} \end{cases} \quad \text{network cycles}$$

Use in the end
average over L_n
 ~~$\log(n)$~~

Mixed decision making

$$\begin{bmatrix} p(k) \\ 1-p(k) \end{bmatrix} \rightarrow \begin{bmatrix} p(k)^a \\ (1-p(k))^a \end{bmatrix} \rightarrow \begin{bmatrix} \frac{p(k)^a}{p(k)^a + (1-p(k))^a} \\ \frac{(1-p(k))^a}{p(k)^a + (1-p(k))^a} \end{bmatrix}$$

