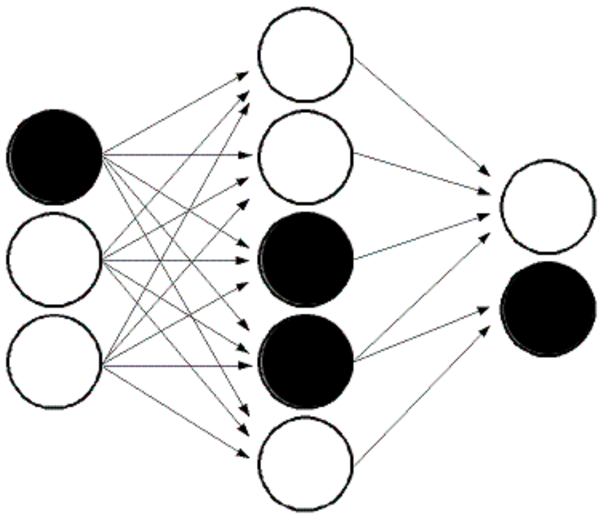
In the name of god

Probabilistic Spiking Neural Networks
- Intro -

Ali Fathi

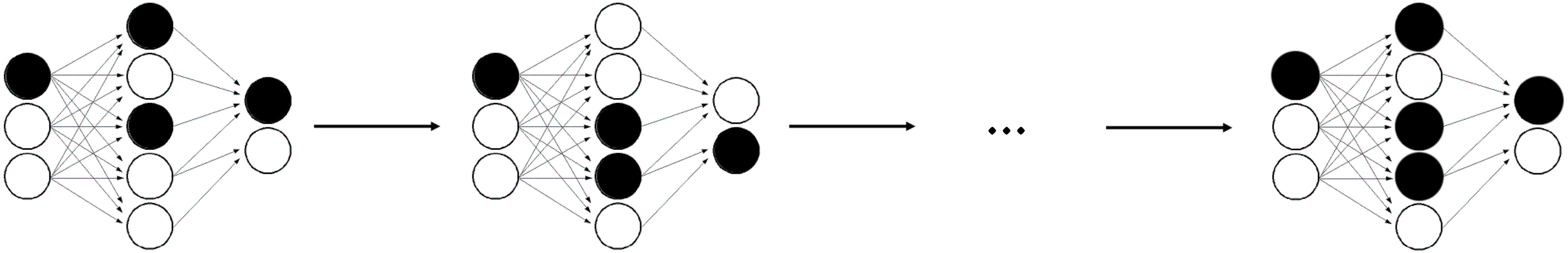
Supervoser: Dr. Salehkaleybar

Binary Networks



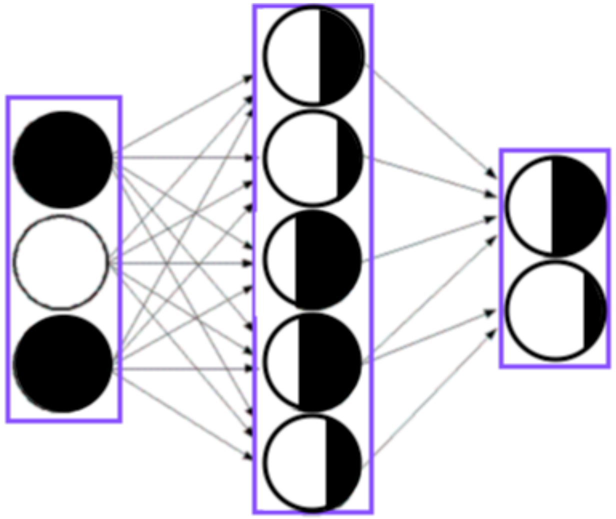
Probabilistic View

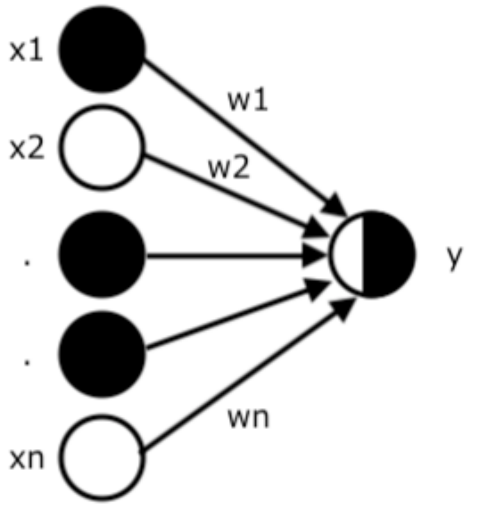
Boltzmann Machine



Probabilistic Layer Model

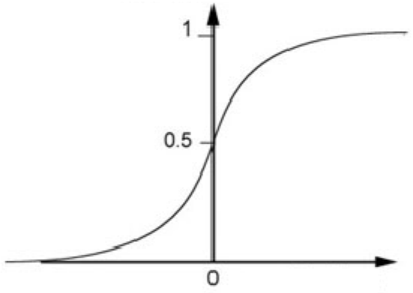
- A Basic Compositional Model for Spiking Neural Networks; Nancy Lynch, Cameron Musco; 2018

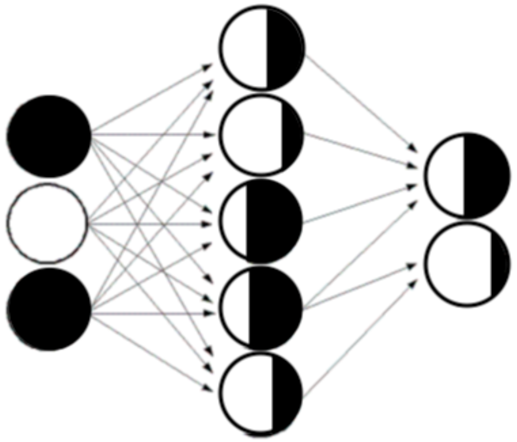


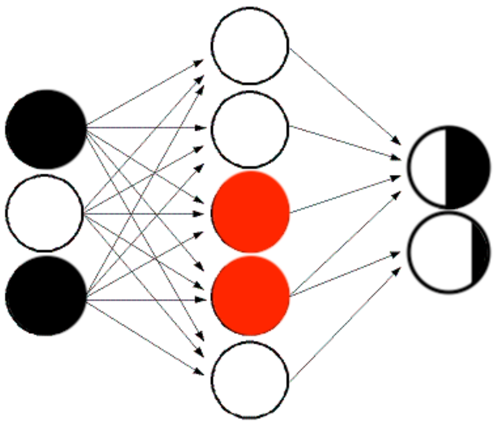


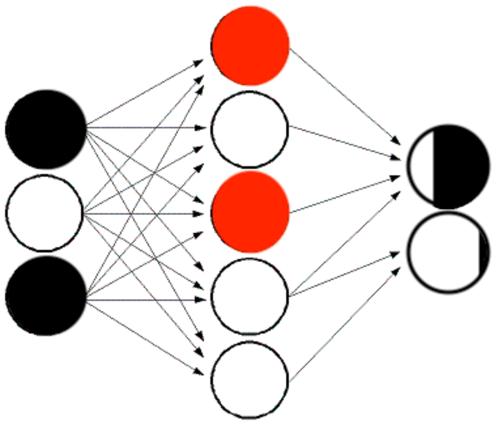
$$pot_y = \sum w_i x_i$$

$$p_y = \sigma(pot_y)$$

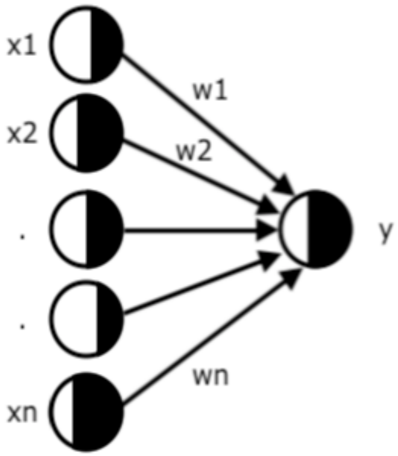








؟


x1
x2
.
.
xn
w1
w2
wn
y

$$X_s = [00 \dots 000,$$

$$00 \dots 001,$$

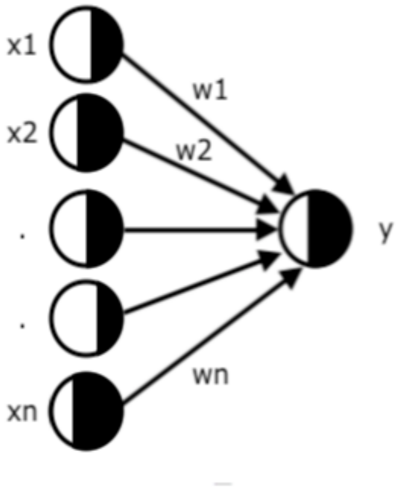
$$00 \dots 010,$$

$$00 \dots 011,$$

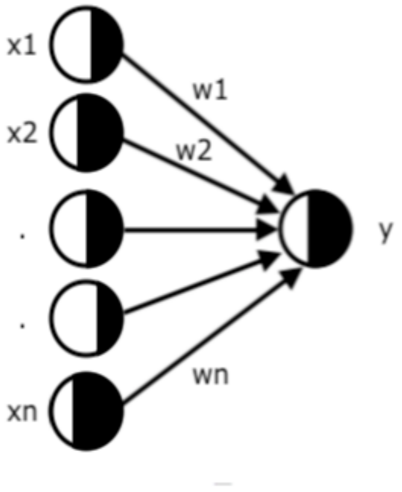
$$\vdots$$

$$11 \dots 110,$$

$$11 \dots 111]$$



$$X_s = [00\ldots000,$$
$$00\ldots001,$$
$$00\ldots010,$$
$$00\ldots011,$$
$$.$$
$$.$$
$$.$$
$$11\ldots110,$$
$$11\ldots111]$$

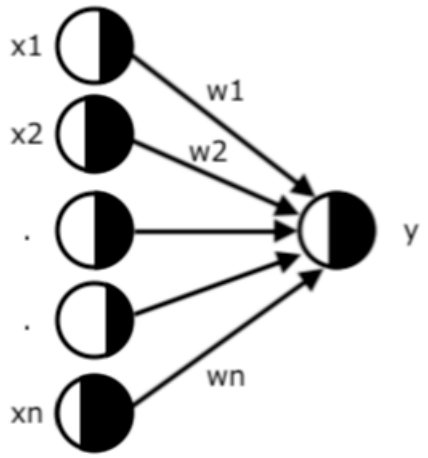


$$p_y = \sigma(\boldsymbol{w}^T.\boldsymbol{x}_0),$$
$$\sigma(\boldsymbol{w}^T.\boldsymbol{x}_1),$$
$$\sigma(\boldsymbol{w}^T.\boldsymbol{x}_2),$$
$$\sigma(\boldsymbol{w}^T.\boldsymbol{x}_3),$$
$$.$$
$$.$$
$$.$$
$$\sigma(\boldsymbol{w}^T.\boldsymbol{x}_{M-\mathbf{1}}),$$
$$\sigma(\boldsymbol{w}^T.\boldsymbol{x}_M)]$$

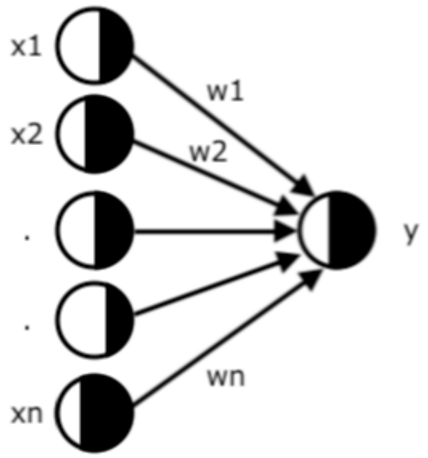
$$
\begin{aligned}
X_s = [&00\ldots000,\\
&00\ldots001,\\
&00\ldots010,\\
&00\ldots011,\\
&\vdots\\
&11\ldots110,\\
&11\ldots111]
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{P} = [&\mathbb{P}(00\ldots000),\\
&\mathbb{P}(00\ldots001),\\
&\mathbb{P}(00\ldots010),\\
&\mathbb{P}(00\ldots011),\\
&\vdots\\
&\mathbb{P}(11\ldots110),\\
&\mathbb{P}(11\ldots111)]
\end{aligned}
$$

$$
\begin{aligned}
p_y = \ &\sigma(\boldsymbol{w}^T.\boldsymbol{x}_0),\\
&\sigma(\boldsymbol{w}^T.\boldsymbol{x}_1),\\
&\sigma(\boldsymbol{w}^T.\boldsymbol{x}_2),\\
&\sigma(\boldsymbol{w}^T.\boldsymbol{x}_3),\\
&\vdots\\
&\sigma(\boldsymbol{w}^T.\boldsymbol{x}_{M-\mathbf{1}}),\\
&\sigma(\boldsymbol{w}^T.\boldsymbol{x}_M)]
\end{aligned}
$$

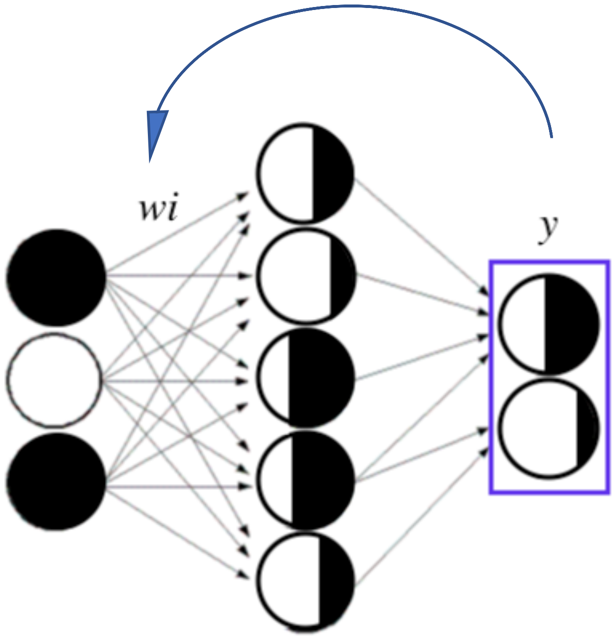


$$p_y = \sum_{\boldsymbol{x} \in X_S} \mathbb{P}(\boldsymbol{x}).\,\sigma(\boldsymbol{w}^T.\,\boldsymbol{x})$$



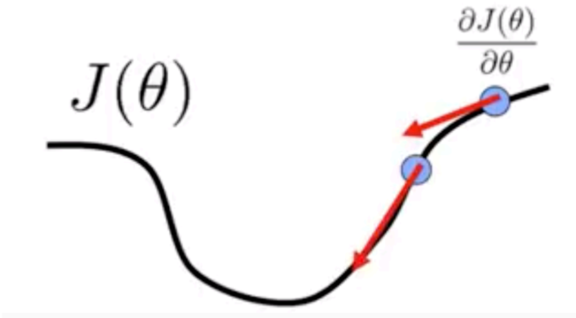
$$\mathbb{P}(\boldsymbol{x}) = q_1 q_2 \ldots q_n$$

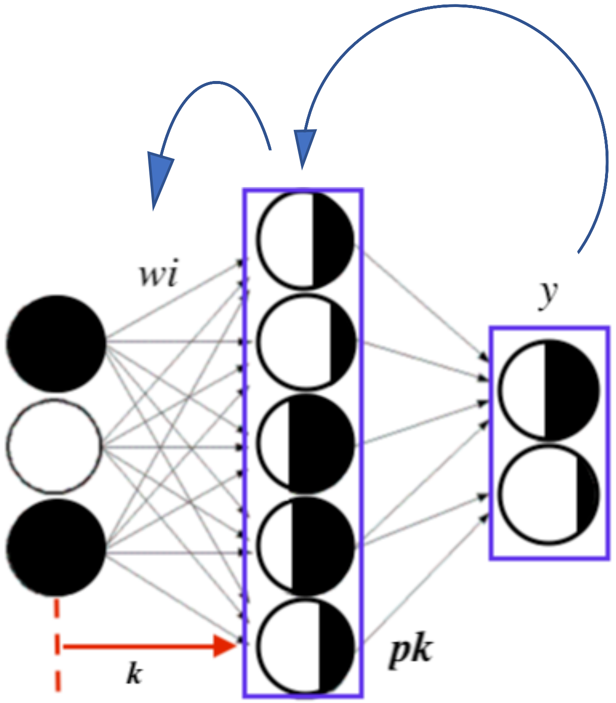
$$q_i = \begin{cases} p_i & x_i = 1 \\ 1 - p_i & x_i = 0 \end{cases}$$



$$Loss = Loss(\boldsymbol{p_y})$$

$$\frac{\partial Loss}{\partial w_i} = ?$$

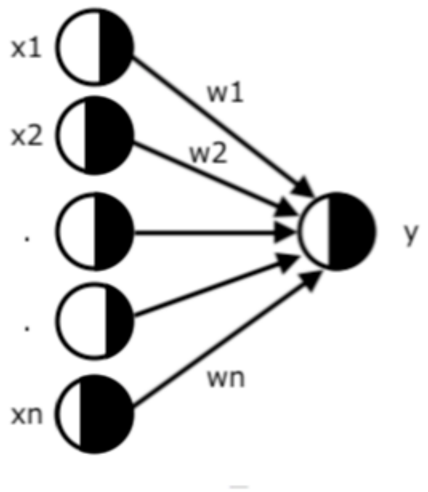




$$Loss = Loss(\boldsymbol{p_y}) \qquad \boldsymbol{p_y} = f(\boldsymbol{p_k})$$

$$\frac{\partial Loss}{\partial w_i} = ? \qquad \frac{\partial Loss}{\partial w_i} = \frac{\partial Loss}{\partial \boldsymbol{p_k}} \cdot \frac{\partial \boldsymbol{p_k}}{\partial w_i}$$

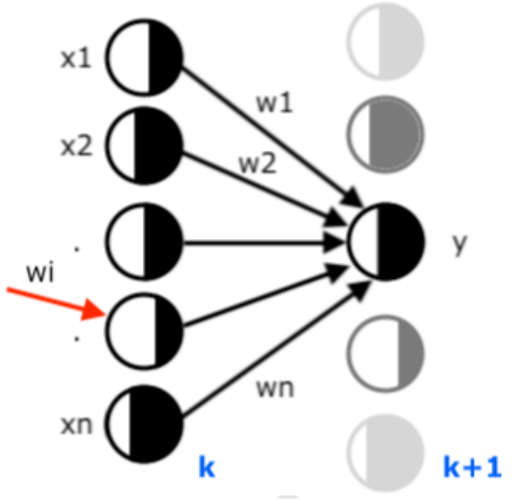
$$\frac{\partial Loss}{\partial w_i} = \frac{\partial Loss}{\partial \boldsymbol{p_k}} . \frac{\partial \boldsymbol{p_k}}{\partial w_i}$$



$$p_y = \sum_{\boldsymbol{x} \in X_s} \mathbb{P}(\boldsymbol{x}) . \sigma(\boldsymbol{w}^T . \boldsymbol{x})$$

$$\frac{\partial p_y}{\partial \boldsymbol{w}} = \sum_{\boldsymbol{x} \in X_s} \boldsymbol{x} . [\mathbb{P}(\boldsymbol{x}) . \sigma'(\boldsymbol{w}^T . \boldsymbol{x})]$$

$$\frac{\partial Loss}{\partial w_i} = \frac{\partial Loss}{\partial \boldsymbol{p_k}} . \frac{\partial \boldsymbol{p_k}}{\partial w_i}$$



$$f_{y_j}(\boldsymbol{w_j}^T.\boldsymbol{x}) = \begin{cases} \sigma(\boldsymbol{w_j}^T.\boldsymbol{x}) & y_j = 1 \\ 1 - \sigma(\boldsymbol{w_j}^T.\boldsymbol{x}) & y_j = 0 \end{cases}$$

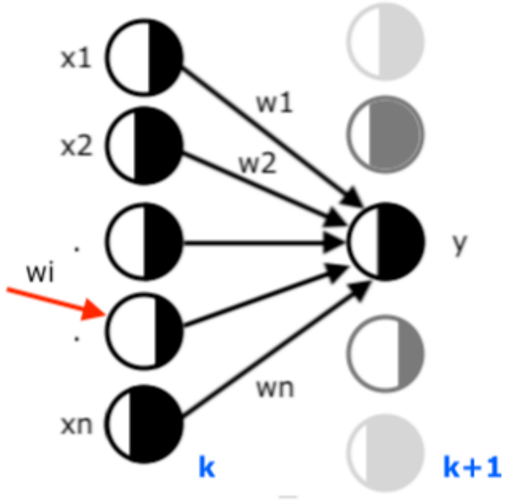
$$\frac{\partial f_{y_j}(\boldsymbol{w_j}^T.\boldsymbol{x})}{\partial \boldsymbol{w_j}} = (-1)^{(1-y_j)} . \frac{\partial \sigma(\boldsymbol{w_j}^T.\boldsymbol{x})}{\partial \boldsymbol{w_j}}$$

$$= (-1)^{(1-y_j)} . \boldsymbol{x} . \sigma'(\boldsymbol{w}_l^T.\boldsymbol{x})$$

$$p_{\boldsymbol{y}} = \sum_{\boldsymbol{x} \in X_s} \mathbb{P}(X = \boldsymbol{x}) . \prod_{j=1}^{m} f_{y_j}(\boldsymbol{w_j}^T.\boldsymbol{x})$$

$$\frac{\partial p_{\boldsymbol{y}}}{\partial \boldsymbol{w}_l} = \sum_{\boldsymbol{x} \in X_s} \boldsymbol{x} . [\mathbb{P}(X = \boldsymbol{x}) . (-1)^{(1-y_j)} . \sigma'(\boldsymbol{w}_l^T.\boldsymbol{x}) . \prod_{\substack{j=1 \\ j \neq l}}^{m} f_{y_j}(\boldsymbol{w_j}^T.\boldsymbol{x})]$$

$$\frac{\partial Loss}{\partial w_i} = \frac{\partial Loss}{\partial \boldsymbol{p_k}} \cdot \frac{\partial \boldsymbol{p_k}}{\partial w_i}$$

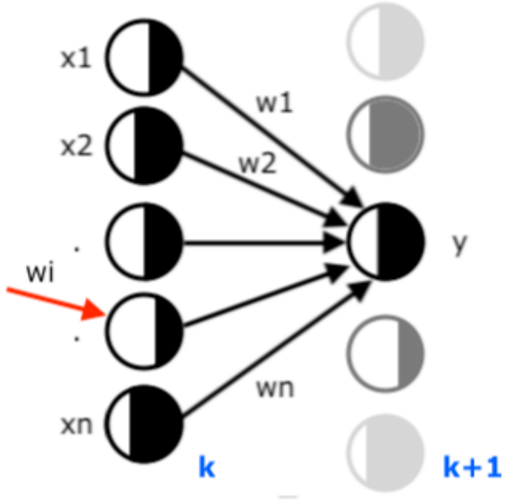


$$\frac{\partial Loss}{\partial \boldsymbol{p_k}} = \frac{\partial Loss}{\partial \boldsymbol{p_{k+1}}} \cdot \frac{\partial \boldsymbol{p_{k+1}}}{\partial \boldsymbol{p_k}}$$

$$p_y = \sum_{\boldsymbol{x} \in X_S} \mathbb{P}(\boldsymbol{x}) . \sigma(\boldsymbol{w}^T . \boldsymbol{x})$$

$$\frac{\partial p_y}{\partial p_{x_i}} = \sum_{\boldsymbol{x} \in X_S} \frac{\partial \mathbb{P}(\boldsymbol{x})}{\partial p_{x_i}} . \sigma(\boldsymbol{w}^T . \boldsymbol{x})$$

$$\frac{\partial Loss}{\partial w_i} = \frac{\partial Loss}{\partial \boldsymbol{p_k}} . \frac{\partial \boldsymbol{p_k}}{\partial w_i}$$



$$\mathbb{P}(\boldsymbol{x}) = q_1 q_2 \dots q_m$$

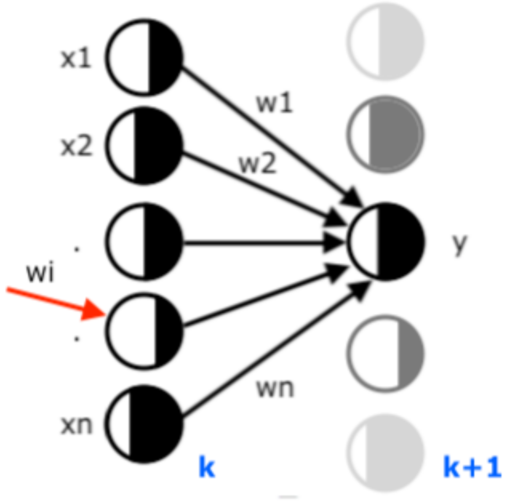
$$\frac{\partial Loss}{\partial \boldsymbol{p_k}} = \frac{\partial Loss}{\partial \boldsymbol{p_{k+1}}} . \frac{\partial \boldsymbol{p_{k+1}}}{\partial \boldsymbol{p_k}}$$

$$q_i = \begin{cases} p_i & x_i = 1 \\ 1 - p_i & x_i = 0 \end{cases}$$

$$\frac{\partial p_y}{\partial p_{x_i}} = \sum_{\boldsymbol{x} \in X_s} \frac{\partial \mathbb{P}(\boldsymbol{x})}{\partial p_{x_i}} . \sigma(\boldsymbol{w}^T . \boldsymbol{x})$$

$$= \sum_{\substack{\boldsymbol{x} \in X_s \\ x_i = 1}} \mathbb{P}_{-i}(\boldsymbol{x}) . \sigma(\boldsymbol{w}^T . \boldsymbol{x}) - \sum_{\substack{\boldsymbol{x} \in X_s \\ x_i = \boldsymbol{0}}} \mathbb{P}_{-i}(\boldsymbol{x}) . \sigma(\boldsymbol{w}^T . \boldsymbol{x})$$

$$\frac{\partial Loss}{\partial w_i} = \frac{\partial Loss}{\partial \boldsymbol{p_k}} \cdot \frac{\partial \boldsymbol{p_k}}{\partial w_i}$$

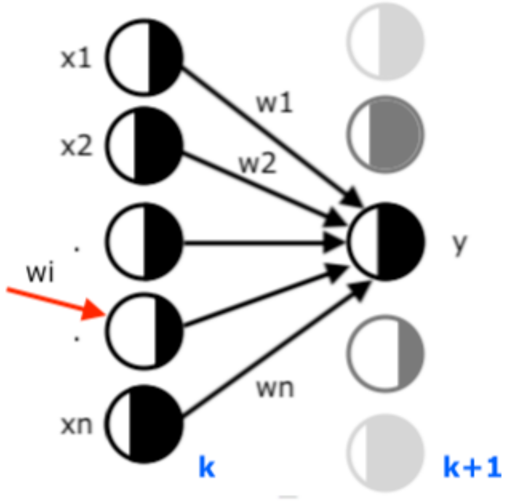


$$\frac{\partial Loss}{\partial \boldsymbol{p_k}} = \frac{\partial Loss}{\partial \boldsymbol{p_{k+1}}} \cdot \frac{\partial \boldsymbol{p_{k+1}}}{\partial \boldsymbol{p_k}}$$

$$p_{\boldsymbol{y}} = \sum_{\boldsymbol{x} \in X_s} \mathbb{P}(X = \boldsymbol{x}) \cdot \prod_{j=1}^{m} f_{y_j}(\boldsymbol{w_j}^T \cdot \boldsymbol{x})$$

$$\frac{\partial p_{\boldsymbol{y}}}{\partial p_{\boldsymbol{x}}} = \sum_{\boldsymbol{x} \in X_s} \frac{\partial \mathbb{P}(X = \boldsymbol{x})}{\partial p_{\boldsymbol{x}}} \cdot \prod_{j=1}^{m} f_{y_j}(\boldsymbol{w_j}^T \cdot \boldsymbol{x})$$

$$\frac{\partial Loss}{\partial w_i} = \frac{\partial Loss}{\partial \boldsymbol{p_k}} \cdot \frac{\partial \boldsymbol{p_k}}{\partial w_i}$$



$$\mathbb{P}(\boldsymbol{x}) = q_1 q_2 \dots q_m$$

$$\frac{\partial Loss}{\partial \boldsymbol{p_k}} = \frac{\partial Loss}{\partial \boldsymbol{p_{k+1}}} \cdot \frac{\partial \boldsymbol{p_{k+1}}}{\partial \boldsymbol{p_k}}$$

$$q_i = \begin{cases} p_i & x_i = 1 \\ 1 - p_i & x_i = 0 \end{cases}$$

$$\frac{\partial p_{\boldsymbol{y}}}{\partial p_{\boldsymbol{x}}} = \sum_{\boldsymbol{x} \in X_s} \frac{\partial \mathbb{P}(X = \boldsymbol{x})}{\partial p_{\boldsymbol{x}}} \cdot \prod_{j=1}^{m} f_{y_j}(\boldsymbol{w_j}^T . \boldsymbol{x})$$

$$= \sum_{\substack{\boldsymbol{x} \in X_s \\ x_i = 1}} \mathbb{P}_{-i}(\boldsymbol{x}) . \prod_{j=1}^{m} f_{y_j}(\boldsymbol{w_j}^T . \boldsymbol{x}) - \sum_{\substack{\boldsymbol{x} \in X_s \\ x_i = \boldsymbol{0}}} \mathbb{P}_{-i}(\boldsymbol{x}) . \prod_{j=1}^{m} f_{y_j}(\boldsymbol{w_j}^T . \boldsymbol{x})$$

In the name of god

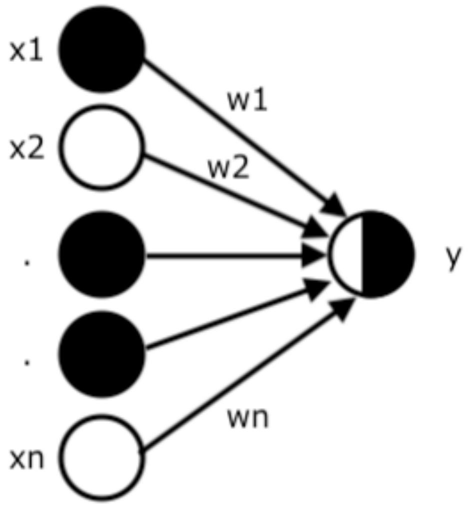
Designing and Analysis of Learning Algorithms for Probabilistic Spiking Neural Networks

- Final Part -

Ali Fathi

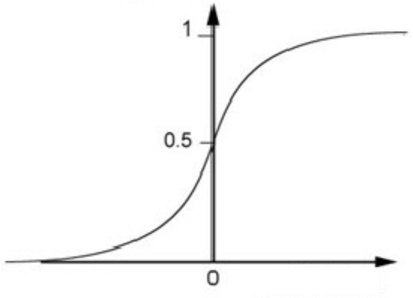
Supervisor: Dr. Salehkaleybar

Recap.



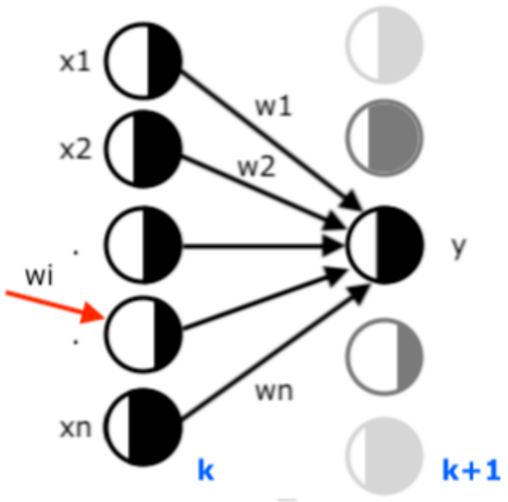
$$pot_y = \sum w_i x_i$$

$$p_y = \sigma(pot_y)$$



$$p_{\boldsymbol{y_j}}(\boldsymbol{x}) = \mathbb{P}\big(y_j = 1 |\, X = \boldsymbol{x}\big) = \sigma(\boldsymbol{w_j}^T.\boldsymbol{x})$$

$$f_{y_j}(\boldsymbol{w_j}^T.\boldsymbol{x}) = \begin{cases} p_{\boldsymbol{y_j}}(\boldsymbol{x}) & y_j = 1 \\ 1 - p_{\boldsymbol{y_j}}(\boldsymbol{x}) & y_j = 0 \end{cases}$$



$$? = \prod_{j=1}^{m} \sum_{\boldsymbol{x} \in X_S} \mathbb{P}(X = \boldsymbol{x}).\, f_{y_j}(\boldsymbol{w_j}^T.\boldsymbol{x})$$

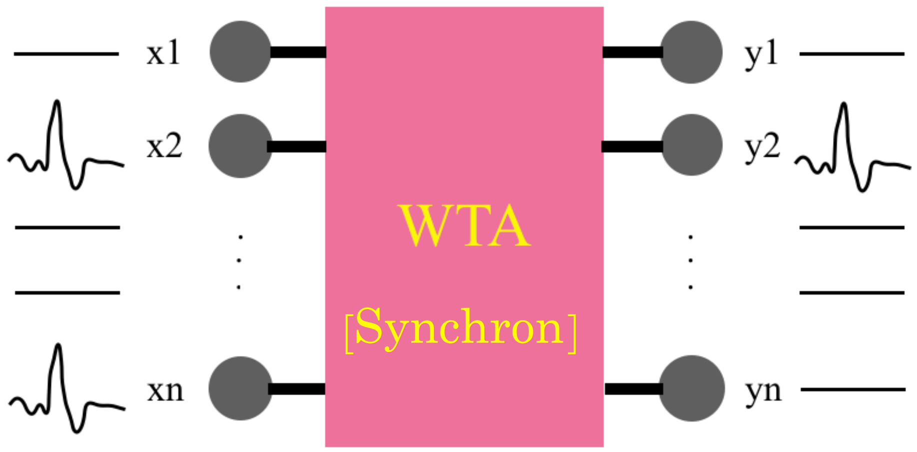
$$\mathbb{P}(Y = \boldsymbol{y}) = \sum_{\boldsymbol{x} \in X_S} \mathbb{P}(X = \boldsymbol{x}).\, \mathbb{P}(Y = \boldsymbol{y} \mid X = \boldsymbol{x})$$

$$= \sum_{\boldsymbol{x} \in X_S} \mathbb{P}(X = \boldsymbol{x}).\prod_{j=1}^{m} \mathbb{P}\big(Y_j = \boldsymbol{y_j} \mid X = \boldsymbol{x}\big)$$

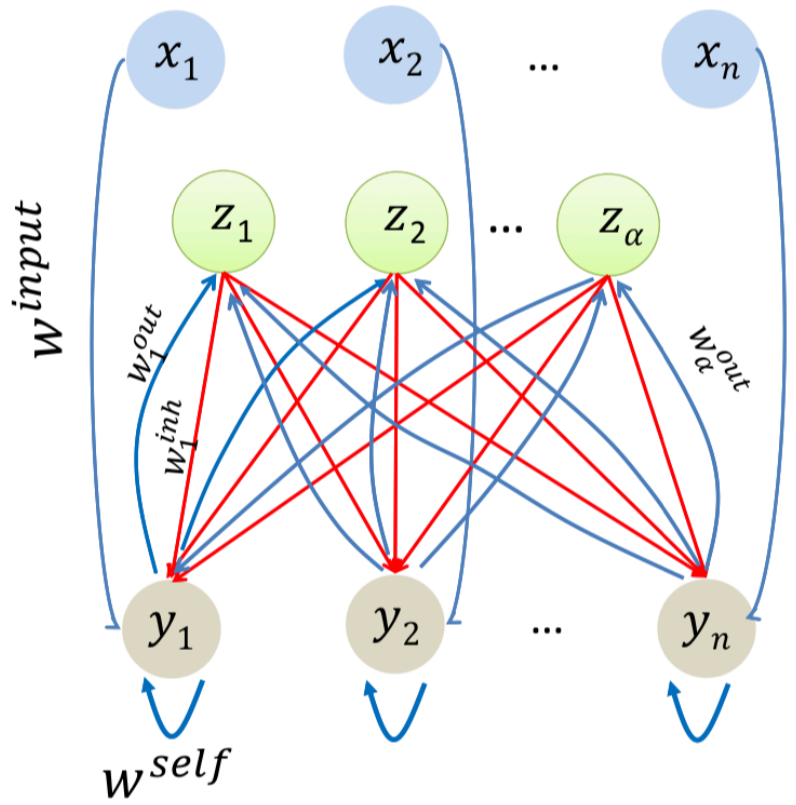
$$= \sum_{\boldsymbol{x} \in X_S} \mathbb{P}(X = \boldsymbol{x}).\prod_{j=1}^{m} f_{y_j}(\boldsymbol{w_j}^T.\boldsymbol{x})$$

WTA Network

(Winner Takes All)



Leader selection, Communication channels, Brain performance

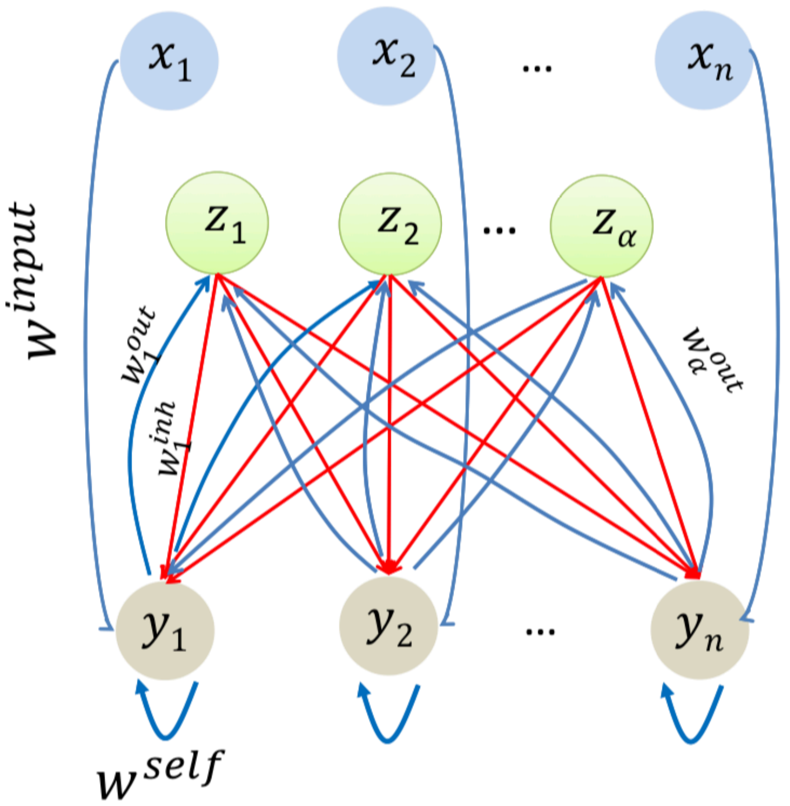


x_1
x_2
...
x_n
w^{input}
z_1
z_2
...
z_α
w_1^{out}
w_1^{inh}
w_α^{out}
y_1
y_2
...
y_n
w^{self}

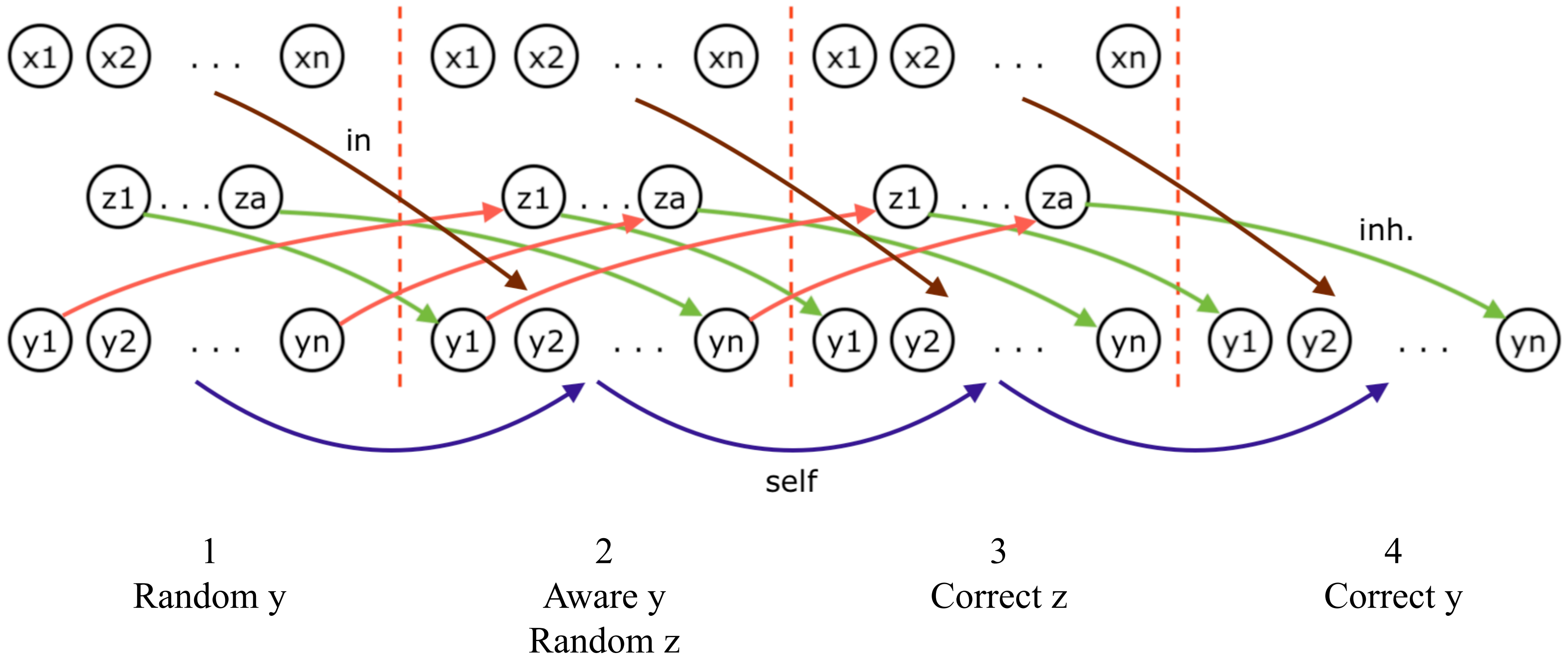
Setting Weights by Hand [2]

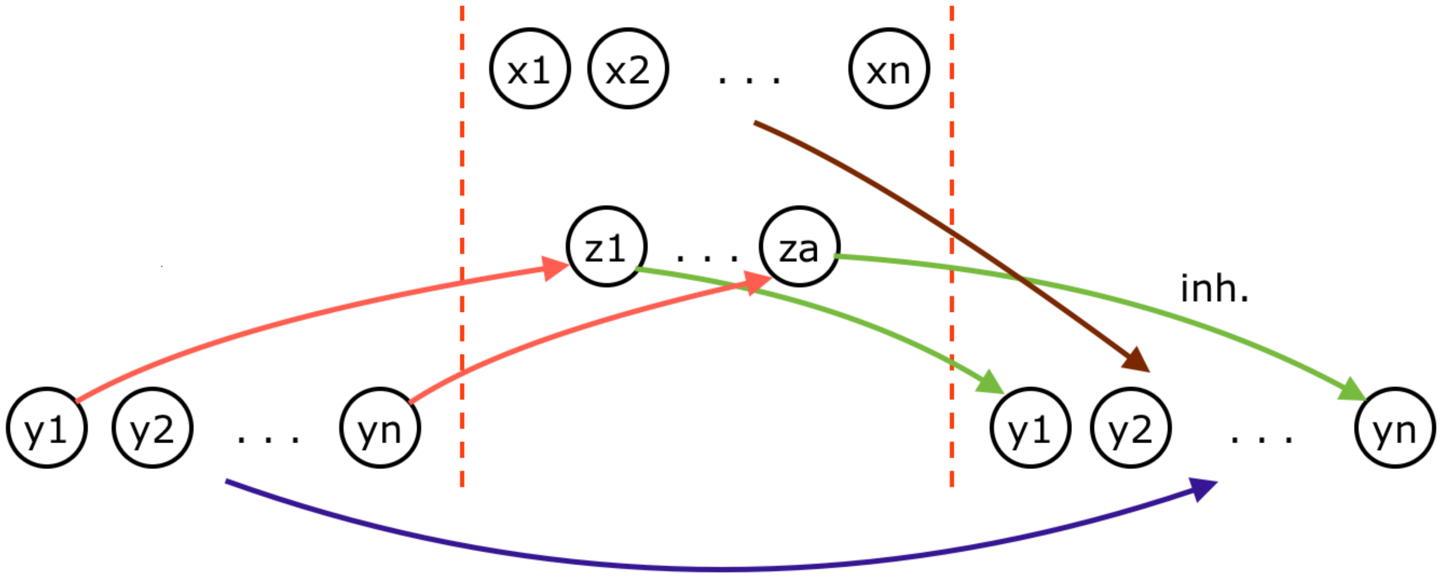
$$\alpha = 2 \in O(1) \;\rightarrow\; \mathcal{HT} = O(\log n)$$

$$\alpha \in O(\log n) \;\rightarrow\; \mathcal{HT} = O(1)$$

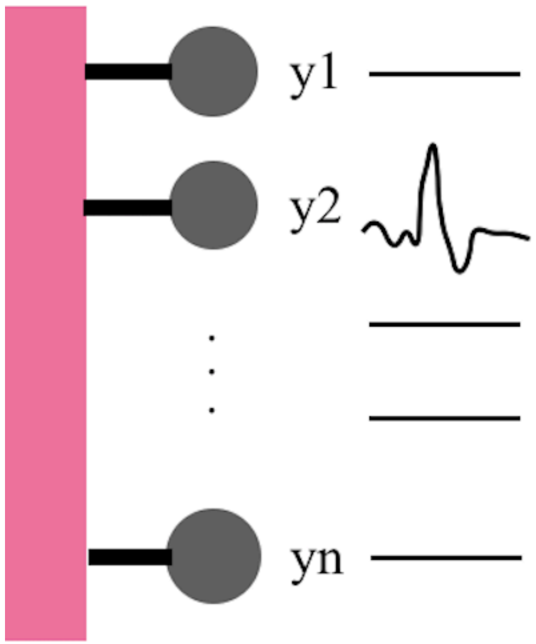


Expansion in Time


x1
x2
. . .
xn
in
z1
. . .
za
inh.
y1
y2
yn
self
1
Random y
2
Aware y
Random z
3
Correct z
4
Correct y



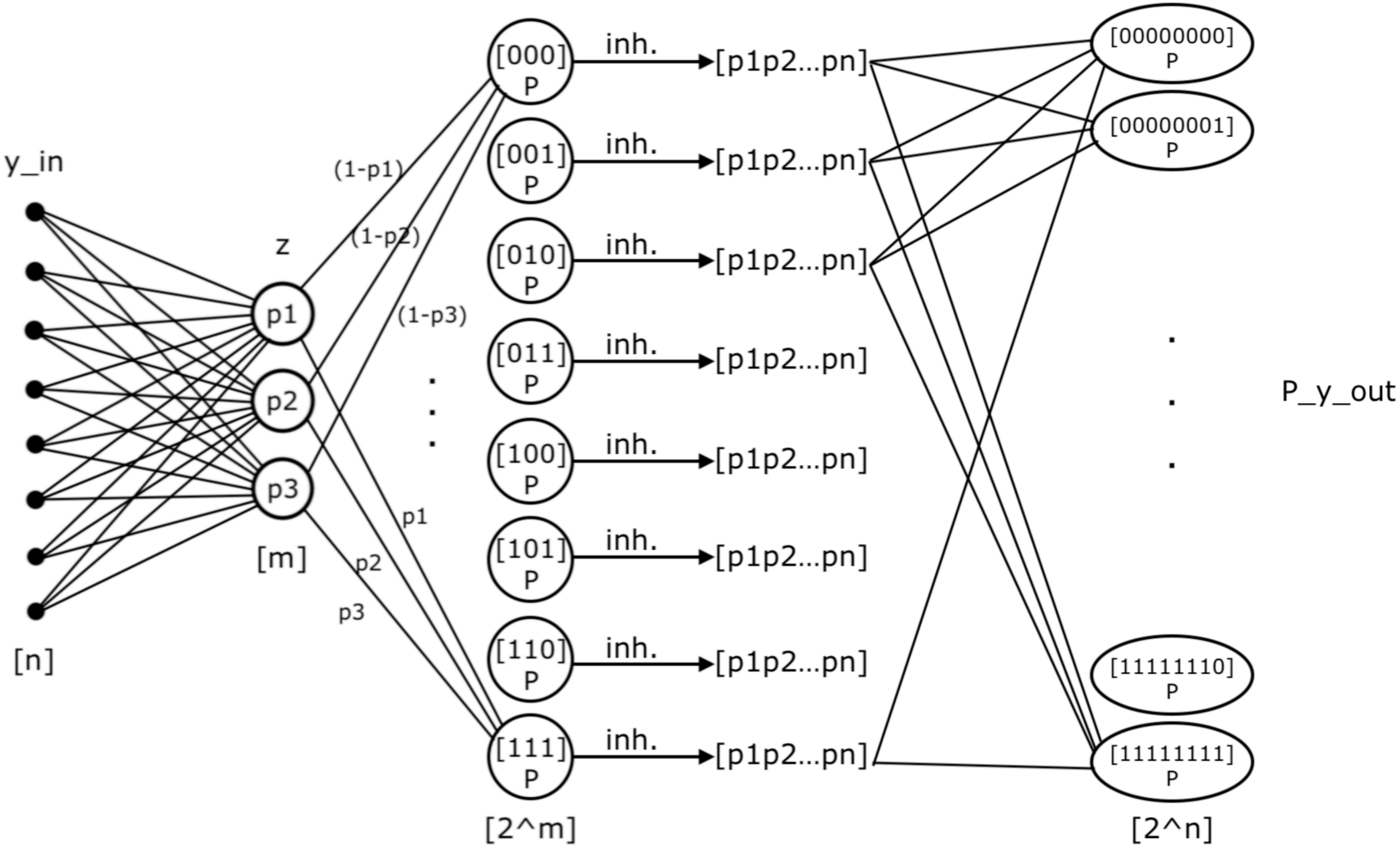
$$\mathbb{P}(\boldsymbol{z}) = p_1 p_2 \dots p_\alpha$$



$$Marginal\ Loss = \sum_{x_i y_{i=0}} p_i^2 + (\sum_{x_i y_{i=1}} p_i - 1)^2$$

Complete Network Model

(Poison hemlock mpdel!)



y_in
[n]
z
p1
p2
p3
[m]
(1-p1)
(1-p2)
(1-p3)
p1
p2
p3
[000] P
[001] P
[010] P
[011] P
[100] P
[101] P
[110] P
[111] P
[2^m]
inh.
[p1p2...pn]
[00000000] P
[00000001] P
[11111110] P
[11111111] P
[2^n]
P_y_out

Linear w.r.t. prev. layer neurons!

$$\mathbb{P}(Y = \boldsymbol{y}) = \sum_{\boldsymbol{x} \in X_S} \mathbb{P}(X = \boldsymbol{x}). \prod_{j=1}^{n} f_{y_j}(\boldsymbol{w_j}^T . \boldsymbol{x})$$

$$\boxed{\Rightarrow \overrightarrow{\mathbb{P}(Y)} = \mathrm{W}. \overrightarrow{\mathbb{P}(X)}}$$

$$\overrightarrow{\mathbb{P}_m(X)} = \begin{pmatrix} \mathbb{P}(X = 000 \ldots 0) \\ \mathbb{P}(X = 000 \ldots 1) \\ \ldots \\ \mathbb{P}(X = 111 \ldots 1) \end{pmatrix}, \qquad \overrightarrow{\mathbb{P}_n(Y)} = \begin{pmatrix} \mathbb{P}(Y = 000 \ldots 0) \\ \mathbb{P}(Y = 000 \ldots 1) \\ \ldots \\ \mathbb{P}(Y = 111 \ldots 1) \end{pmatrix}$$

$$W_{2^n\times 2^m} =$$

$$= \begin{pmatrix} \prod_{j=1}^{n}(1-\sigma(\boldsymbol{w_j}^T.\boldsymbol{x_0})) & \cdots & \prod_{j=1}^{n}(1-\sigma(\boldsymbol{w_j}^T.\boldsymbol{x_{2^m}})) \\ \vdots & \ddots & \vdots \\ \prod_{j=1}^{n}\sigma(\boldsymbol{w_j}^T.\boldsymbol{x_0}) & \cdots & \prod_{j=1}^{n}\sigma(\boldsymbol{w_j}^T.\boldsymbol{x_{2^m}}) \end{pmatrix}$$

$$[W_{2^n\times 2^m}]_{(a,b)} = \prod_{j=1}^{n} f_{a_j}(\boldsymbol{w_j}^T.\boldsymbol{x}_b)$$

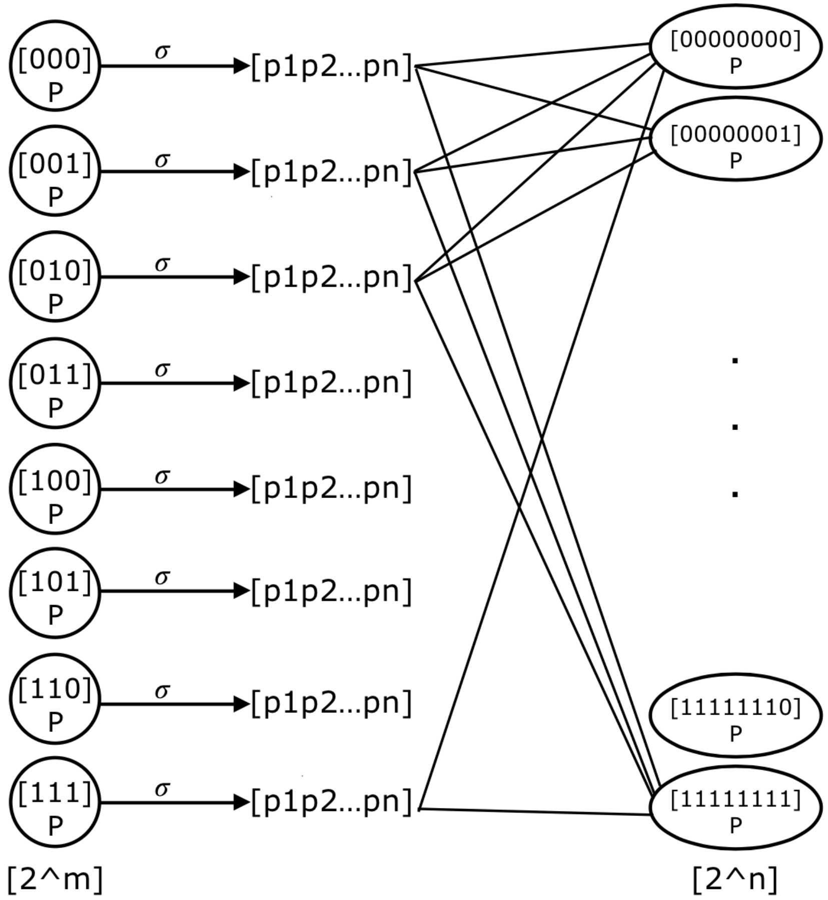
$$\boldsymbol{x}_b = bin(b)$$
$$a_j = bin(a)[j]$$

$$f_{a_j}(\boldsymbol{w_j}^T.\boldsymbol{x}) = \begin{cases} \sigma(\boldsymbol{w_j}^T.\boldsymbol{x}) & y_j = 1 \\ 1-\sigma(\boldsymbol{w_j}^T.\boldsymbol{x}) & y_j = 0 \end{cases}$$

$\underline{n\times m}$

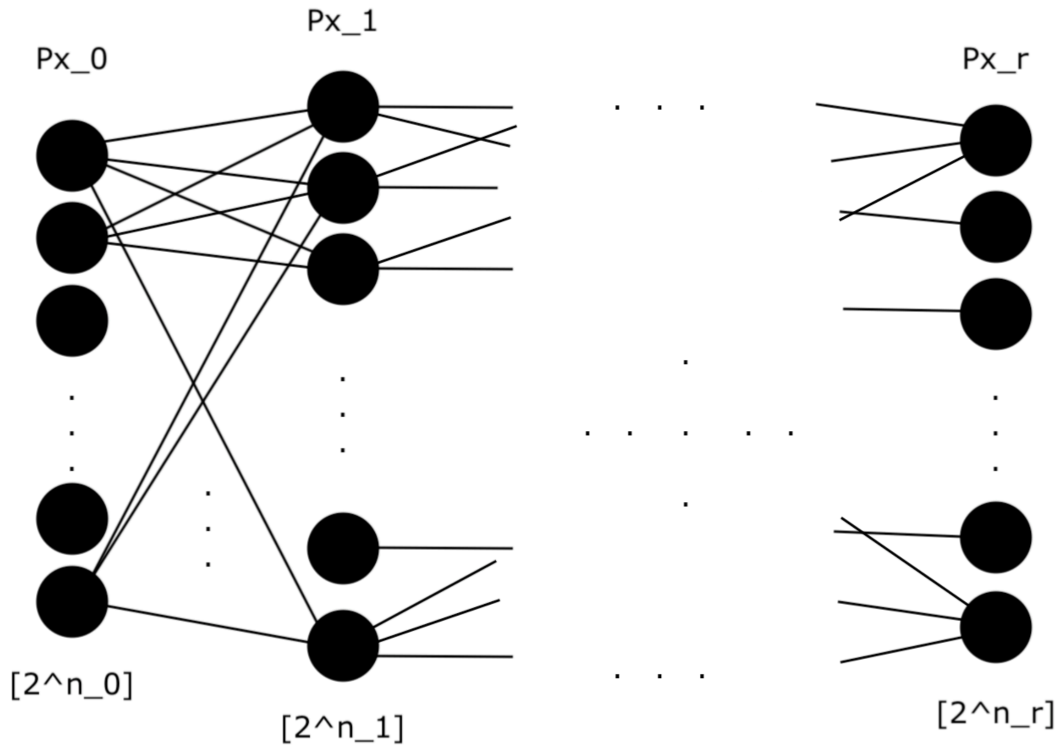
$\underline{free\ parameters}$

$\underline{(\boldsymbol{w}_j'\boldsymbol{s})}$



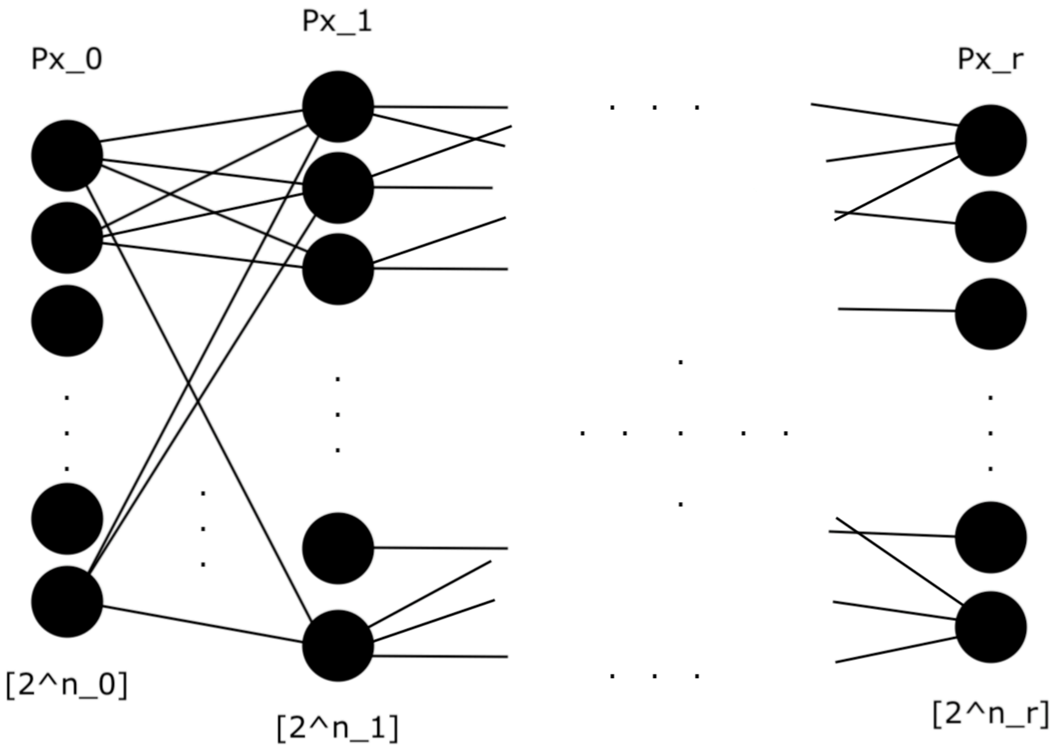
$$\overrightarrow{\mathbb{P}(X_r)} = W_{2^{n_r} \times 2^{n_{r-1}}} \ldots W_{2^{n_1} \times 2^{n_0}} . \overrightarrow{\mathbb{P}(X_0)}$$

$$\overrightarrow{\mathbb{P}(X_0 = x_{in})} = \mathbb{I}[X_0 = x_{in}]$$



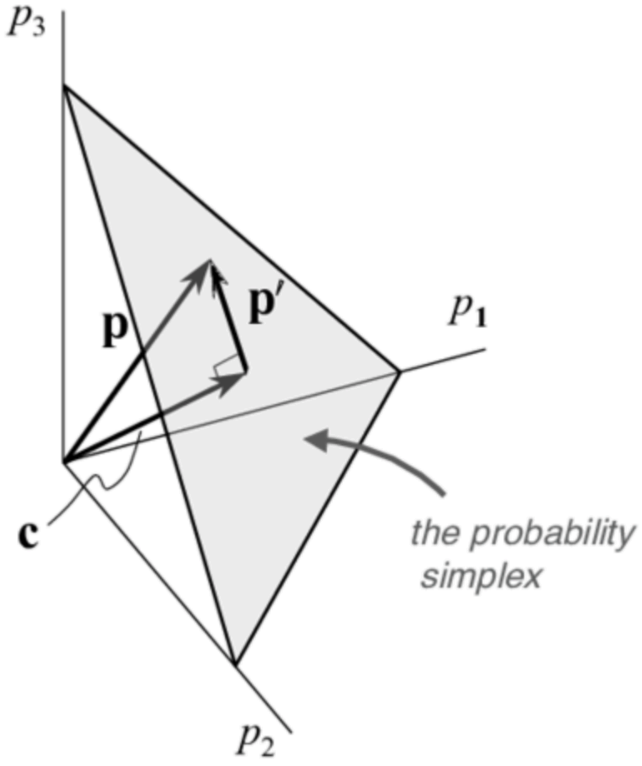
Conditional Probability Matrix!

$$W_{2^{n_i} \times 2^{n_j}} = \mathbb{P}\left(X_i = \boldsymbol{x_i} | X_j = \boldsymbol{x_j}\right)$$



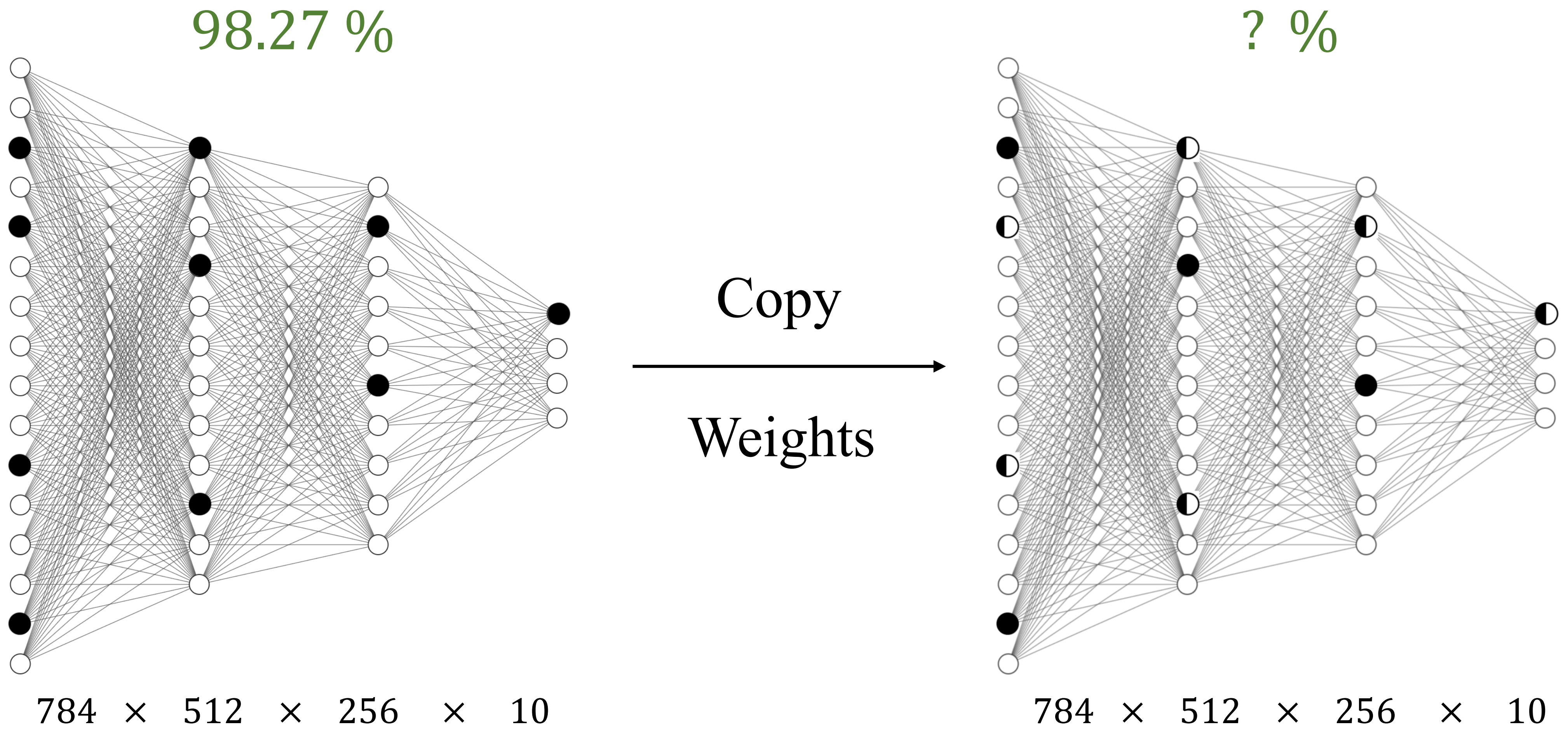
Transition Probability Matrix

$$\overrightarrow{\mathbb{P}(Y)} = \mathrm{W}.\overrightarrow{\mathbb{P}(X)}$$



Naïve Learning!

(On MNIST)



98.27 %
784 × 512 × 256 × 10
Copy
Weights
? %
784 × 512 × 256 × 10

98.15 % !

```
Test accuracy: 98.27 %
Probabilistic Accuracy: 0.9815
```

Future Works Suggestions

- Studying computational power of the network with linear transition matrix

- Studying the networks training time

- Methods to train

References

- A Basic Compositional Model for Spiking Neural Networks; Nancy Lynch, Cameron Musco; 2018

- Computational Tradeoffs in Biological Neural Networks: Self-Stabilizing Winner-Take-All Networks; Nancy Lynch, Cameron Musco, Merav Parter; 2016

Thanks!