Elementary Channels for Quantization, Neural Networks and Molecular Information Exchange

Rudolf Mathar
joint work with Gholamreza Alirezaei

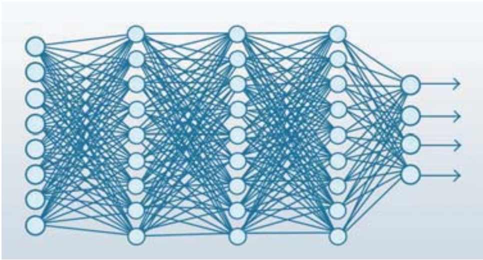
Institute for Theoretical Information Technology
Faculty of Electrical Engineering and Information Technology

Chair for Theoretical Information Technology | RWTH AACHEN UNIVERSITY

Some Questions

- The brain seems to be a biological neural network. Is its functionality really understood?

Some Questions



- The brain seems to be a biological neural network. Is its functionality really understood?
- Artificial neural networks, used in deep learning, show great successes. Why exactly?
- Are the nodes in an optimized artificial neural network trained in a way to maximize information flow?
- Maximize biological neurons (local) information flow?

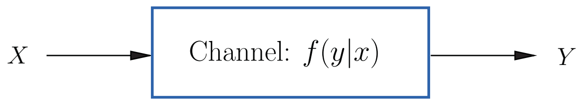
Overview

- Basic concepts from information theory
- Cortical neurons and neural networks
- Artificial neurons and neural networks (ANN)
- Channels describing neurons in ANN
- Binary threshold neurons
- Molecular information exchange

Overview

- Basic concepts from information theory
- Cortical neurons and neural networks
- Artificial neurons and neural networks (ANN)
- Channels describing neurons in ANN
- Binary threshold neurons
- Molecular information exchange

Channels



How much information can you get across a channel?

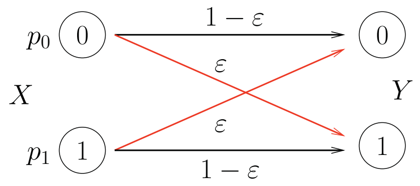
$$H(X) = -\int f(x) \log f(x) dx$$ **entropy** of X

$$H(Y \mid X)$$ **conditional entropy** of Y given X

$$I(X;Y) = H(Y) - H(Y \mid X)$$ **mutual information** between X and Y

$$C = \max_{P^X} I(X;Y)$$ **capacity** of the channel

Example: Binary Symmetric Channel (BSC)



Mutual Information:

$$I(X;Y) = H\big(p_0(1-\varepsilon) + p_1\varepsilon, \varepsilon p_0 + (1-\varepsilon)p_1\big) - H(\varepsilon, 1-\varepsilon)$$

The maximum of $I(X;Y)$ over all input distributions (p_0, p_1) is attained at the uniform distribution $(p_0^*, p_1^*) = (0.5, 0.5)$.

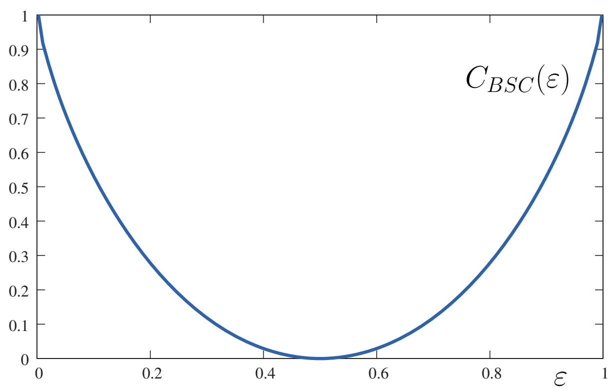
Capacity:

$$C = 1 + (1-\varepsilon)\log_2(1-\varepsilon) + \varepsilon \log_2 \varepsilon$$

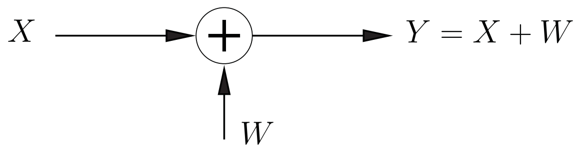
Example: Binary Symmetric Channel (BSC)

Capacity as a function of the error probability ε:

$$C_{BSC}(\varepsilon) = 1 + (1-\varepsilon)\log_2(1-\varepsilon) + \varepsilon \log_2 \varepsilon)$$



Example: Gaussian Channel (AWGN)



Average power constraint: $E[X^2] \leq M$

Noise distribution is zero-mean Gaussian:

$$W \sim f_W(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-x^2/2\sigma^2\right)$$

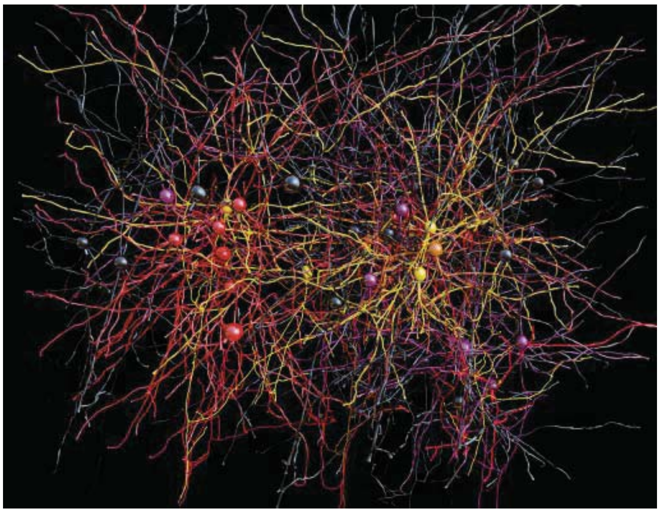
Capacity:

$$C = \frac{1}{2} \ln\left(1 + M/\sigma^2\right) = \frac{1}{2} \ln(1 + SNR)$$

Overview

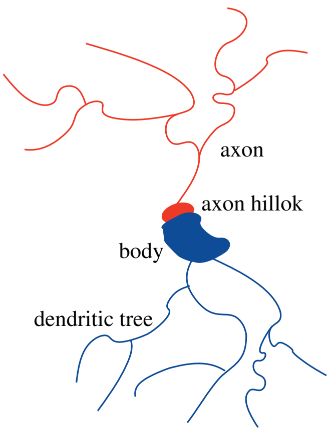
- Basic concepts from information theory
- Cortical neurons and neural networks
- Artificial neurons and neural networks (ANN)
- Channels describing neurons in ANN
- Binary threshold neurons
- Molecular information exchange

Cortical Neural Networks



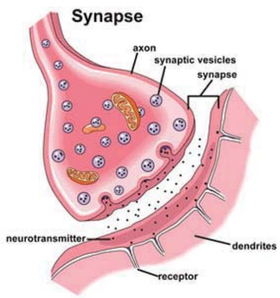
Neuroscientists have constructed a network map of connections between cortical neurons, traced from a 100 terabytes 3D data set. The data were created by an electron microscope in nanoscopic detail, allowing every one of the "wires" to be seen, along with their connections. Some of the neurons are color-coded according to their activity patterns in the living brain. (credit: Clay Reid, Allen Institute; Wei-Chung Lee, Harvard Medical School; Sam Ingersoll, graphic artist)

A Typical Cortical Neuron



- The axon branches to contact other neurons.
- A dendritic tree collects input from other neurons.
- Axons contact dendritic trees at synapses and inject spikes of activity.
- An axon hillock generates outgoing spikes whenever enough charge has flowed in at synapses to depolarize the cell membrane.

Synapses



- When a spike travels along an axon and arrives at a synapse, vesicles of a transmitter chemical are released.
- The transmitter molecules diffuse through the synaptic cleft and bind to receptor molecules in the membrane of the post-synaptic neuron.
- This opens up holes that allow specific ions to cross.

Synapses

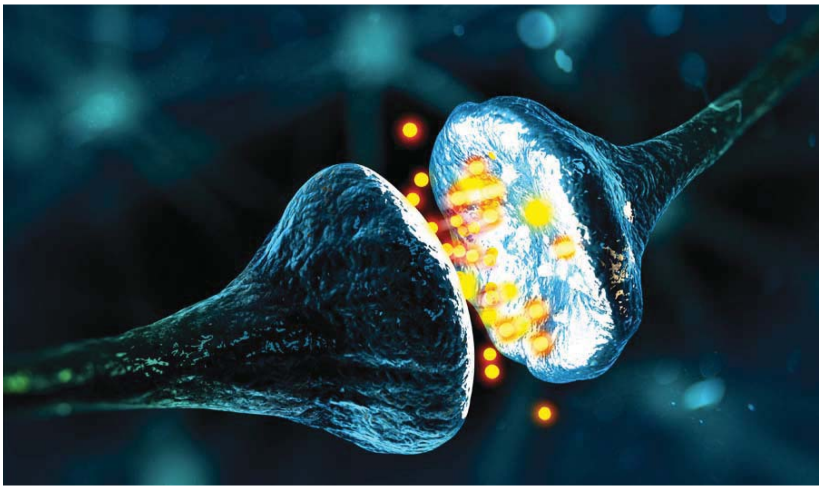
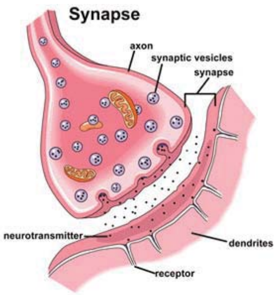


Image Credits: Inside SCIENCE, Andrii Vodolazhskyi

Synapses



- The effectiveness of the synapse can be changed by
  - varying the number of vesicles of the transmitter
  - varying the number of receptor molecules
- Synapses are slow, but
  - they are very small and very low-power
  - they adapt using locally available signals

How (most people think) the brain works

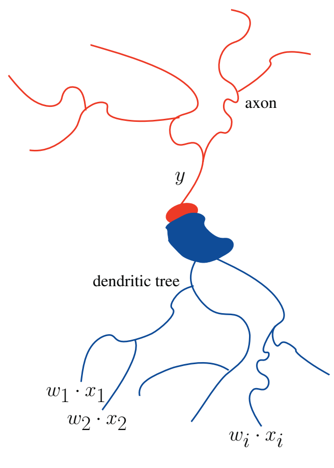
- Each neuron receives input from other neurons
  - A few neurons are connected to receptors.
  - Neurons use spikes to communicate.
- The effect of each input line on the neuron is controlled by synaptic weights
  - Weights can be positive or negative.
- The synaptic weights **adapt** so that the whole network learns to perform useful computations
  - Recognizing objects, understanding language, making plans, controlling the body
- Humans have about 10^{11} neurons each with about 10^4 weights
  - Computations in parallel in a short time, huge bandwidth

Overview

- Basic concepts from information theory
- Cortical neurons and neural networks
- Artificial neurons and neural networks (ANN)
- Channels describing neurons in ANN
- Binary threshold neurons
- Molecular information exchange

Idealizing Neurons

Substitute spikes by real values x_i, model intensities by weights w_i.

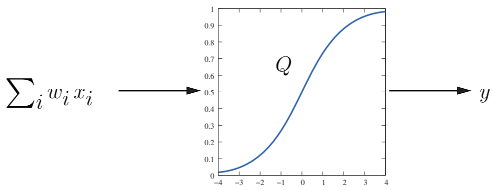


$$y = Q\Big(\sum_i w_i \cdot x_i\Big)$$

What types of **activation functions** Q are appropriate?

Sigmoid Neurons

- First compute a weighted sum of the inputs.
- Send out a sigmoid function of the weighted sum.

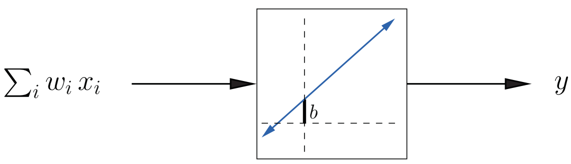


$$y = Q\Big(\sum_i w_i x_i\Big), \qquad Q(z) = \frac{1}{1+e^{-z}},\ z \in \mathbb{R}$$

- The logistic function with convenient derivatives

Linear Neurons

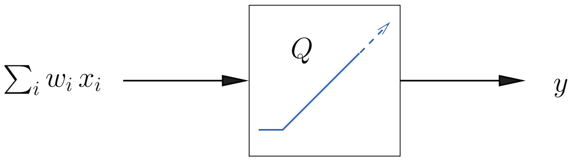
- First compute a weighted sum of the inputs.
- Send out a linear transformation of the input.



$$y = Q\Big(\sum_i w_i x_i\Big), \qquad Q(z) = az + b$$

Rectified Linear Neurons

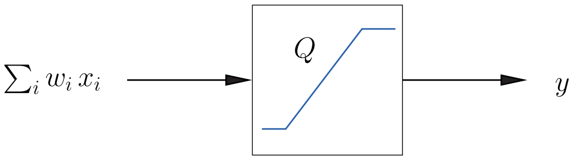
- First compute a weighted sum of the inputs.
- Send out a rectified linear function of the weighted.



$$y = Q\Big(\sum_i w_i x_i\Big), \qquad Q(z) = \begin{cases} 0, & \text{if } z < 0 \\ z, & \text{if } z \geq 0 \end{cases}$$

Censoring Neurons

- First compute a weighted sum of the inputs.
- Send out a censored linear function of the weighted sum.

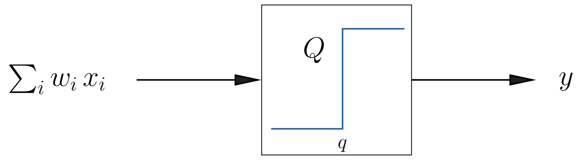


$$y = Q\Big(\sum_i w_i x_i\Big), \qquad Q(z) = \begin{cases} 0, & \text{if } z < 0 \\ z, & \text{if } 0 \leq z < 1 \\ 1, & \text{if } z \geq 1 \end{cases}$$

Binary Threshold Neurons

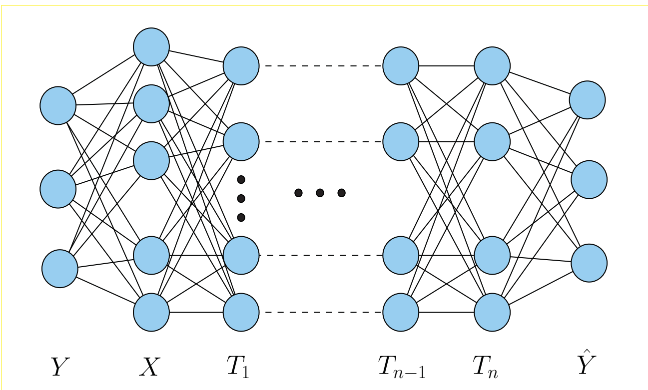
McCulloch-Pitts (1943) (influenced Von Neumann)

- First compute a weighted sum of the inputs.
- Send out a fixed size spike of activity if the weighted sum exceeds a threshold q.



$$y = Q\Big(\sum_i w_i x_i\Big), \qquad Q(z) = \begin{cases} 0, & \text{if } z < q \\ 1, & \text{if } z \geq q \end{cases}$$

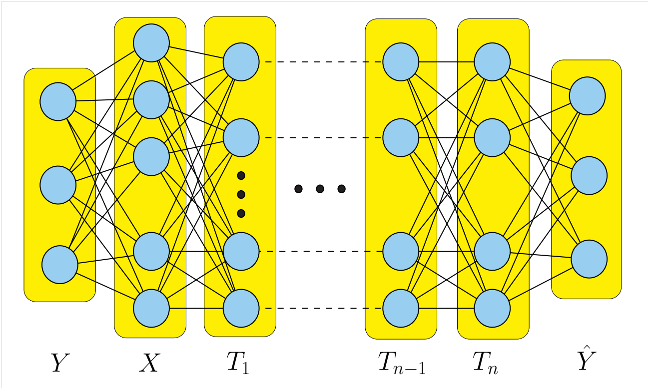
Deep Neural Networks - an Information Theoretic View



Information is passed from input X to output $\hat{Y}$ layer by layer.
The network forms a sequence of consecutive channels:

$Y, X, T_1, \ldots, T_{n-1}, T_n$ (a Markov chain)

Deep Neural Networks - an Information Theoretic View



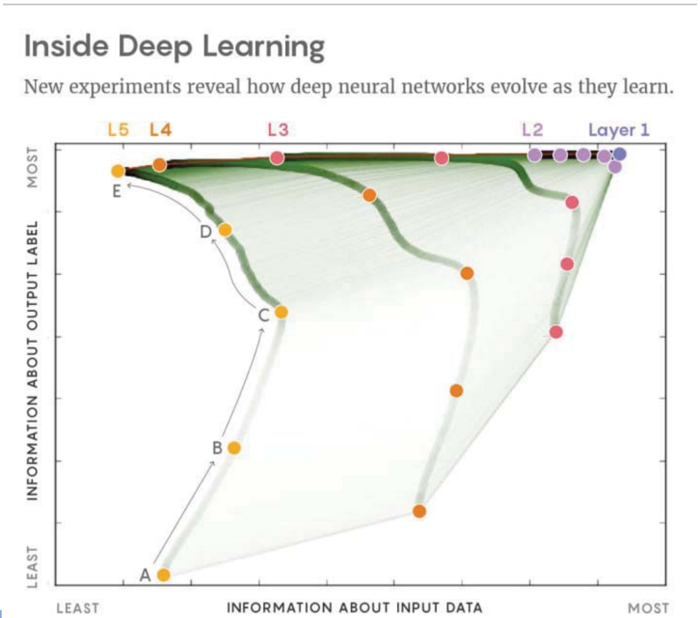
Consider $I(X;T_1) \geq \cdots \geq I(X;T_n) \geq I(X;\hat{Y})$

Naftali Tishby [2017]: When optimizing parameters of the DNN

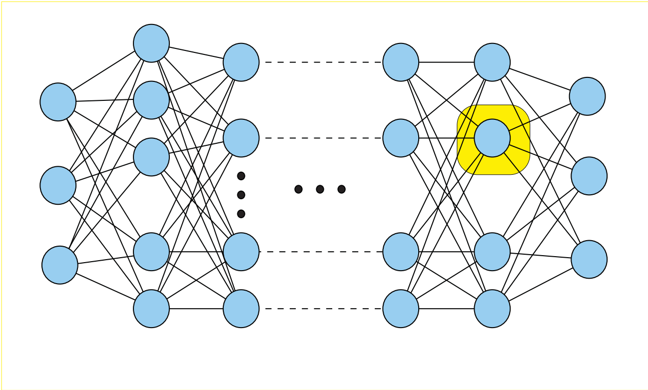
- $I(X;T_j)$ first increases, then decreases,
- $I(T_j;Y)$ tends to its max with the number of iterations.

Deep Neural Networks - Naftali Tishby's Result

$I(T_j;Y)$

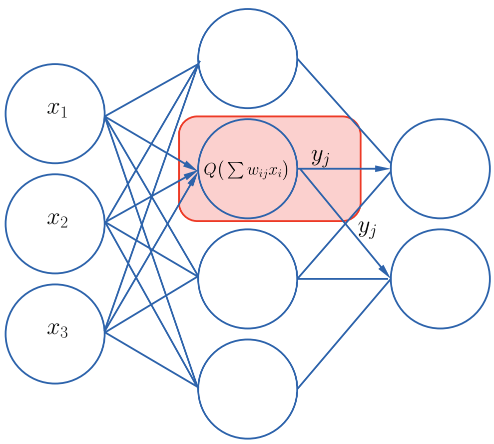


Deep Neural Networks - an Information Theoretic View



In the following consider nodes in isolation.
Is there a typical information theoretic behavior?

Artificial Neural Network

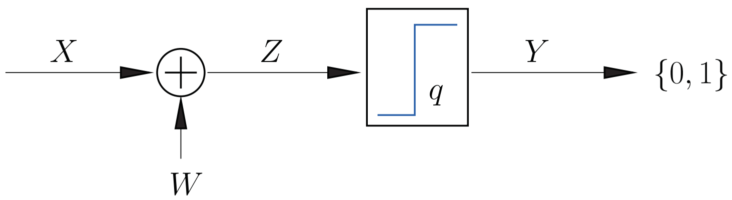


Neurons can be considered as channels.

Overview

- Basic concepts from information theory
- Cortical neurons and neural networks
- Artificial neurons and neural networks (ANN)
- Channels describing neurons in ANN
- Binary threshold neurons
- Molecular information exchange

Binary Threshold Neurons = Quantization Channel



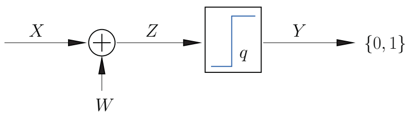
$$Y = Q(X + W), \qquad Q(s) = \begin{cases} 1, & \text{if } s > q \\ 0, & \text{otherwise} \end{cases}$$

X, W stochastically independent r.v.

$X \sim F$, input with cdf F

$W \sim \Phi$, noise with cdf Φ

Quantization Channel



Weighted self information

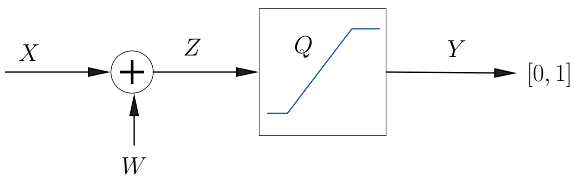
$$\rho(p) = -p \log p, \quad , 0 \leq p \leq 1$$

Mutual Information

$$\begin{aligned} I(X;Y) = &+ \rho\Big(\int \Phi(q-x)dF(x)\Big) \\ &- \int \rho(\Phi(q-x)dF(x)) \\ &+ \rho\Big(1 - \int \Phi(q-x)dF(x)\Big) \\ &- \int \rho\big(1-\Phi(q-x)\big)dF(x) \end{aligned}$$

RWTH AACHEN UNIVERSITY

Censored Channel



$$Y = Q(X+W), \qquad Q(z) = \begin{cases} 0, & \text{if } z < 0 \\ z, & \text{if } 0 \leq z < 1 \\ 1, & \text{if } z \geq 1 \end{cases}$$

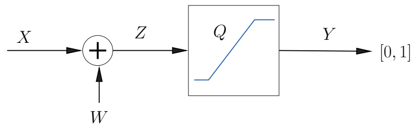
X, W stochastically independent r.v.

$X \sim F$, input with cdf F

$W \sim \Phi$, noise with cdf φ

RWTH AACHEN UNIVERSITY

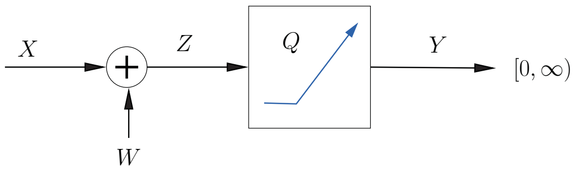
Censored Channel



Mutual Information

$$
\begin{aligned}
I(X;Y) = &+ \rho\Big(\int \Phi(0 - x)dF(x)\Big) \\
&- \int \rho(\Phi(0 - x)dF(x)) \\
&+ \rho\Big(1 - \int \Phi(1 - x)dF(x)\Big) \\
&- \int \rho(1 - \Phi(1 - x))\,dF(x) \\
&+ \int_0^1 \rho\Big(\int \varphi(y - x)dF(x)\Big)dy \\
&- \int_0^1 \int \rho(\varphi(y - x))\,dF(x)dy
\end{aligned}
$$

Rectified Linear Channel



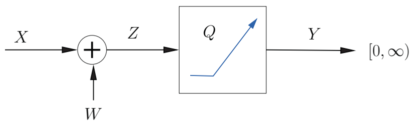
$$
Y = Q(X + W), \qquad Q(z) = \begin{cases} 0, & \text{if } z < 0 \\ z, & \text{if } z \geq 0 \end{cases}
$$

X, W stochastically independent r.v.

$X \sim F$, input with cdf F

$W \sim \Phi$, noise with cdf Φ and density φ

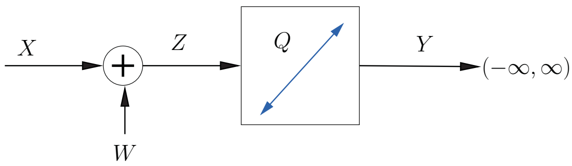
Rectified Linear Channel



Mutual Information

$$
\begin{aligned}
I(X;Y) = &+ \rho\Big(\int \Phi(0 - x)dF(x)\Big) \\
&- \int \rho(\Phi(0 - x)dF(x)) \\
&+ \int_0^\infty \rho\Big(\int \varphi(y - x)dF(x)\Big)dy \\
&- \int_0^\infty \int \rho(\varphi(y - x))\,dF(x)dy
\end{aligned}
$$

Linear Channel



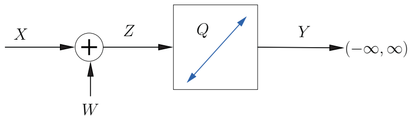
$$Y = Q(X + W), \qquad Q(z) = z, \quad z \in \mathbb{R}$$

X, W stochastically independent r.v.

$X \sim F$, input with cdf F

$W \sim \Phi$, noise with cdf Φ and density φ

Linear Channel

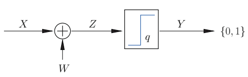


Mutual Information

$$I(X;Y) = +\int_{-\infty}^{\infty} \rho\Big(\int \varphi(y-x)dF(x)\Big)dy$$
$$-\int_{-\infty}^{\infty}\int \rho\big(\varphi(y-x)\big)dF(x)dy$$

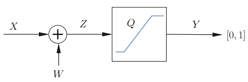
Mutual Information of Neuron Channels

Binary threshold



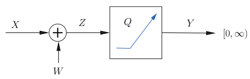
$$+\rho\Big(\int \Phi(q-x)dF(x)\Big)$$
$$-\int \rho(\Phi(q-x)dF(x))$$
$$+\rho\Big(1-\int \Phi(q-x)dF(x)\Big)$$
$$-\int \rho(1-\Phi(q-x))\,dF(x)$$

Censoring



$$+\rho\Big(\int \Phi(0-x)dF(x)\Big)$$
$$-\int \rho(\Phi(0-x)dF(x))$$
$$+\rho\Big(1-\int \Phi(1-x)dF(x)\Big)$$
$$-\int \rho(1-\Phi(1-x))\,dF(x)$$
$$+\int_0^1 \rho\Big(\int \varphi(y-x)dF(x)\Big)dy$$
$$-\int_0^1\int \rho(\varphi(y-x))\,dF(x)dy$$

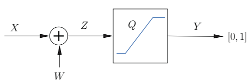
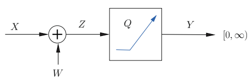
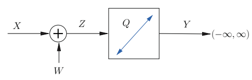
Rectified linear



$$+\rho\Big(\int \Phi(0-x)dF(x)\Big)$$
$$-\int \rho(\Phi(0-x)dF(x))$$
$$+\int_0^{\infty} \rho\Big(\int \varphi(y-x)dF(x)\Big)dy$$
$$-\int_0^{\infty}\int \rho(\varphi(y-x))\,dF(x)dy$$

- Determining capacity means to optimize each over the input distribution F.
- We focus on the binary threshold case.

Mutual Information of Neuron Channels

Censoring	Rectified linear	Linear
		

Censoring

$$+\rho\Big(\int \Phi(0-x)dF(x)\Big)$$
$$-\int \rho(\Phi(0-x)dF(x))$$
$$+\rho\Big(1-\int \Phi(1-x)dF(x)\Big)$$
$$-\int \rho(1-\Phi(1-x))dF(x)$$
$$+\int_0^1 \rho\Big(\int \varphi(y-x)dF(x)\Big)dy$$
$$-\int_0^1 \int \rho(\varphi(y-x))dF(x)dy$$

Rectified linear

$$+\rho\Big(\int \Phi(0-x)dF(x)\Big)$$
$$-\int \rho(\Phi(0-x)dF(x))$$
$$+\int_0^\infty \rho\Big(\int \varphi(y-x)dF(x)\Big)dy$$
$$-\int_0^\infty \int \rho(\varphi(y-x))dF(x)dy$$

Linear

$$+\int_{-\infty}^\infty \rho\Big(\int \varphi(y-x)dF(x)\Big)dy$$
$$-\int_{-\infty}^\infty \int \rho(\varphi(y-x))dF(x)dy$$

- Determining capacity means to optimize each over the input distribution F.
- We focus on binary threshold case.

Overview

- Basic concepts from information theory
- Cortical neurons and neural networks
- Artificial neurons and neural networks (ANN)
- Channels describing neurons in ANN
- Binary threshold neurons
- Molecular information exchange

Capacity of the Binary Threshold Neuron

Results:

- The peak-power constrained capacity-achieving input has discrete support.
- This input is concentrated on two mass points.
- Input probabilities are known in closed form.
- Channel capacity is known in closed form.
- The optimum threshold is unknown in general, however, we will discuss it for a relevant case in neuro science.

Capacity of the Binary Threshold Neuron

Fix quantization threshold q, set $\gamma_i = \Phi(q - x_i)$.

The capacity-achieving distribution in the class $\mathcal{D}_{pp}$ of all distributions with finite support $[x_1, x_m] \subset \mathbb{R}$ is given by the two-point distribution concentrated on $\{x_1, x_m\}$ with probabilities

$$p_1^\star = \frac{1 - (1 + a^s)\gamma_m}{(1 + a^s)(\gamma_1 - \gamma_m)} \quad \text{and} \quad p_m^\star = \frac{(1 + a^s)\gamma_1 - 1}{(1 + a^s)(\gamma_1 - \gamma_m)}$$

Further,

$$C(\gamma_1, \gamma_m) = \max_{\boldsymbol{p}} I(\boldsymbol{p}, \gamma) = \log_a\big(1 + a^s\big) - (s + t)$$

with $\quad s = \frac{h(\gamma_1) - h(\gamma_m)}{\gamma_1 - \gamma_m}$ and $t = \frac{\gamma_1 h(\gamma_m) - \gamma_m h(\gamma_1)}{\gamma_1 - \gamma_m}$,

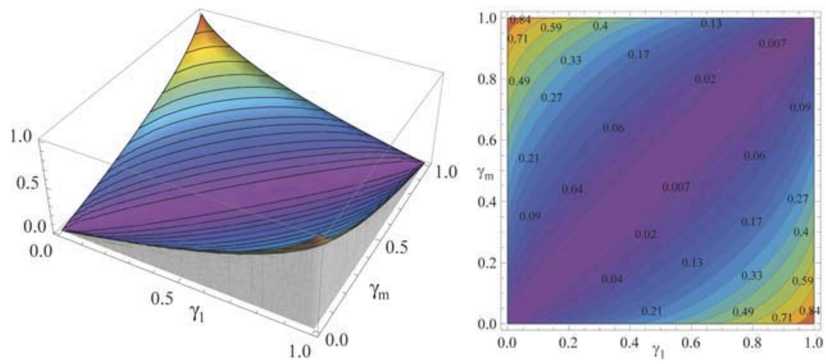
and $\quad h(p) = -p \log_a(p) - (1 - p) \log_a(1 - p)$.

Capacity as a function: $C(\gamma_1, \gamma_m)$

Theorem. $C(\gamma_1, \gamma_m)$ as a function of $(\gamma_1, \gamma_2) \in [0,1]^2$ has the following properties.

- $C(\gamma_1, \gamma_m)$ is symmetric in the sense that
  $C(\gamma_1, \gamma_m) = C(1 - \gamma_m, 1 - \gamma_1)$.
- $C(\gamma_1, \gamma_m)$ is a strictly increasing function of γ_1 and a strictly decreasing function of γ_m, $0 \leq \gamma_m < \gamma_1 \leq 1$.
- $C(\gamma_1, \gamma_m)$ is a convex function of $(\gamma_1, \gamma_m) \in [0,1]^2$. It is even strictly convex whenever $\gamma_1 \neq \gamma_m$.

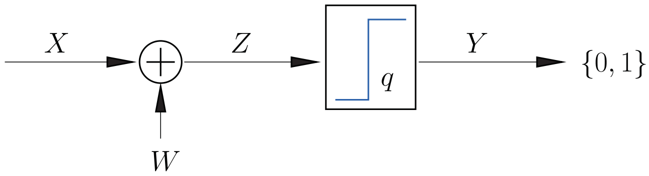
Visualizing $C(\gamma_1, \gamma_m)$



Recall $\gamma_1 = \Phi(q - x_1)$, $\gamma_m = \Phi(q - x_m)$

Fundamental question: What is the optimum threshold?

Optimum Threshold



Solve

$$\max_{q \in \mathbb{R}} C(\gamma_1, \gamma_m)$$

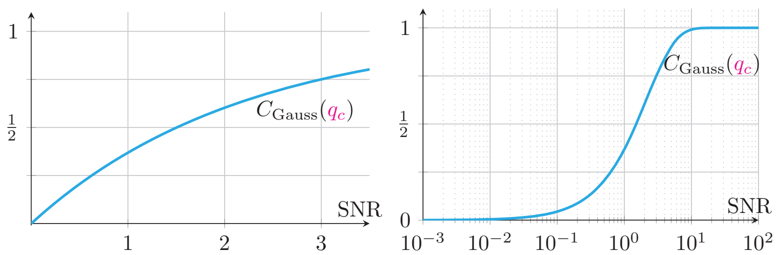
$$\text{such that } \gamma_i = \Phi(q - x_i), i = 1, m$$

If the noise distribution is symmetric and has zero mean, is the capacity optimizing threshold $q^* = (x_1 + x_m)/2$?

Gaussian Noise

Yes, for Gaussian noise!
Maximum capacity $C(q^\star)$ is attained at the centroid $q_c = \frac{x_1 + x_m}{2}$
and is a strictly increasing function of SNR.

$$C(\text{SNR}) = \log(2) - h\Big(\frac{1}{2}\text{erfc}\big(\sqrt{\tfrac{\text{SNR}}{2}}\big)\Big)$$



Rectangular Noise

Distribution function with support $[u_\ell, u_r]$:

$$\Phi_{\text{Rect}}(w) = \begin{cases} 0\,, & w \le u_\ell\,, \\ \frac{w-u_\ell}{u_r-u_\ell}\,, & u_\ell < w \le u_r\,, \\ 1\,, & u_r < w\,, \end{cases}$$

Density function:

$$\varphi_{\text{Rect}}(w) = \begin{cases} \frac{1}{u_r-u_\ell}\,, & u_\ell < w \le u_r\,, \\ 0\,, & \text{otherwise,} \end{cases}$$

Rectangular Noise

Using symmetry, monotonicity and convexity (see above):

- If $u_r - u_\ell < x_m - x_1$, then
  $C^* = \log_a 2$ is attained at any

  $$q^* \in [x_1 + u_r, x_m + u_\ell].$$

  with $p_1^* = p_m^* = 1/2$.

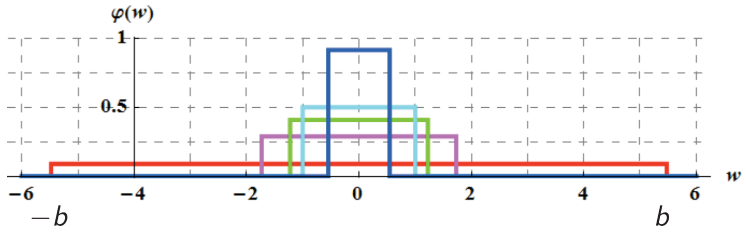
- If $u_r - u_\ell \ge x_m - x_1$, then
  $C^* = h_a(\bar{\gamma}) - p_1^*\, h_a(\gamma_1) + p_2^*\, h_a(\gamma_m)$ is attained at

  $$q^* \in \{x_1 + u_r, x_m + u_\ell\}$$

  with $p_1^* = \frac{\bar{\gamma}-\gamma_m}{\gamma_1-\gamma_m}$, $p_m^* = \frac{\gamma_1-\bar{\gamma}}{\gamma_1-\gamma_m}$.

Rectangular Noise

Consider input $x_1 = -1$, $x_m = 1$ (power 1) and noise with pdf

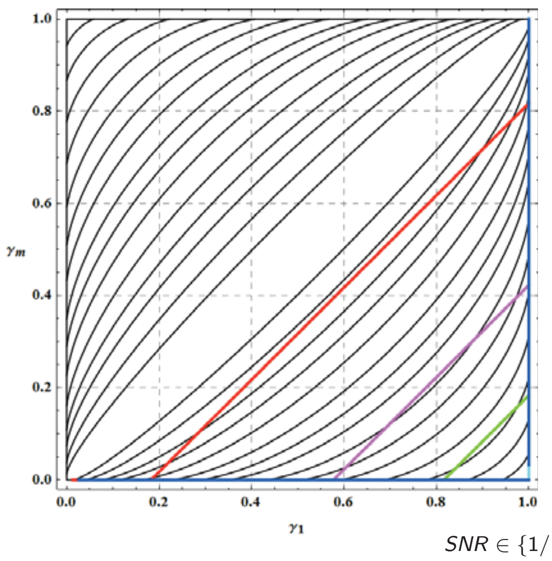


such that

$$SNR = \frac{3}{b^2} \in \{1/10, 1, 2, 3, 10\}$$

Rectangular Noise

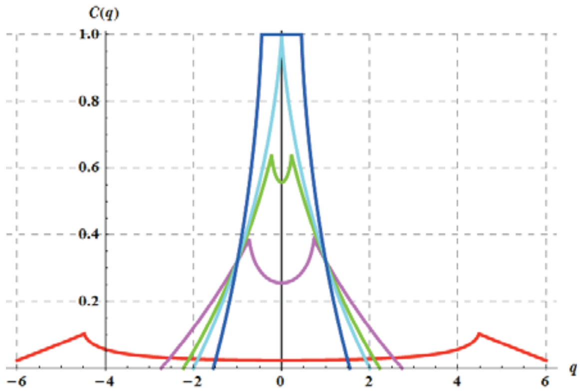
$\max_{q \in \mathbb{R}} \; C(\gamma_1, \gamma_m)$ such that $\gamma_i = \Phi(q - x_i), \; i = 1, 2$



$SNR \in \{1/10, 1, 2, 3, 10\}$

Rectangular Noise

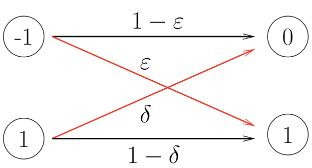
Capacity as a function of threshold q



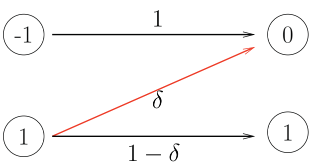
$SNR \in \{1/10, 1, 2, 3, 10\}$

Rectangular Noise

Any q between the optimum "peaks" yields a BAC.

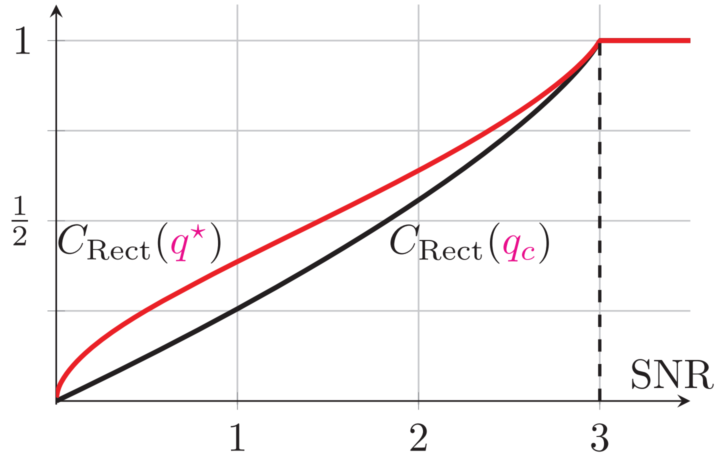


The optimum q^* yields a Z-channel of higher capacity.



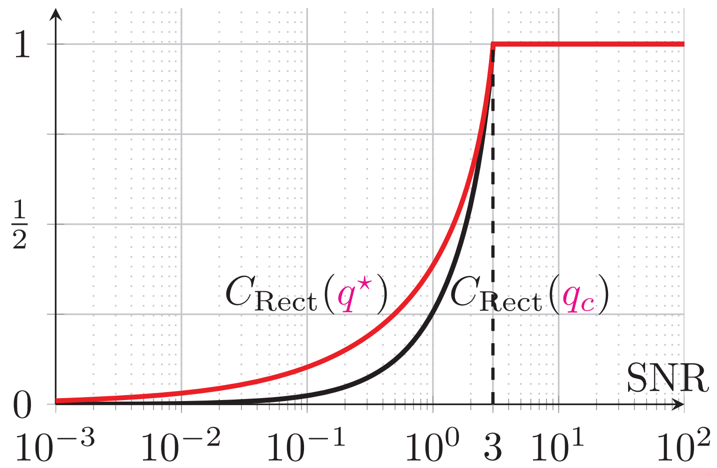
Hence, Z-channels evolve in a rather natural way as optimum one-bit quantization channels.

Rectangular Noise



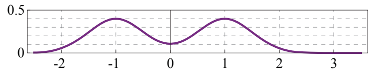
Capacity as a function of SNR.

Rectangular Noise

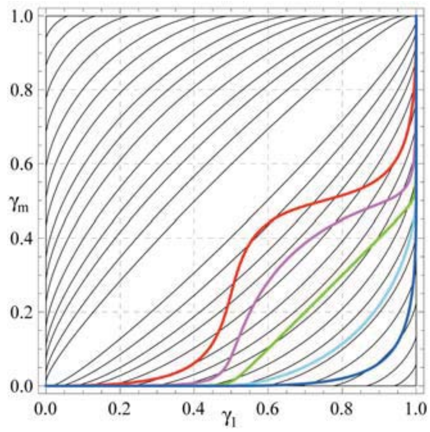


Capacity as a function of SNR (logarithmic).

Mixture of Two Gaussians Noise



Gaussian mixture, variance 0.25 each



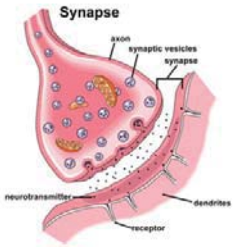
Contours of $C(\gamma_1, \gamma_m)$

$x_m - x_1 = 1, 1.5, 2, 2.5, 3$

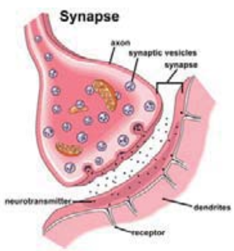
Overview

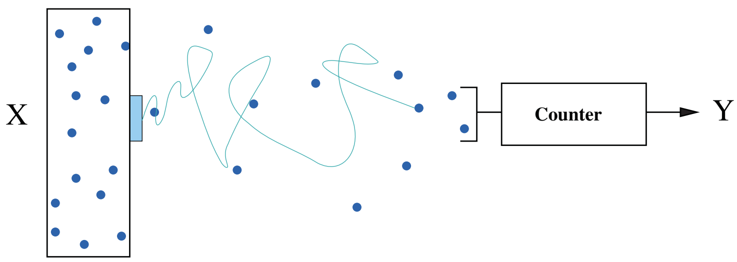
- Basic concepts from information theory
- Cortical neurons and neural networks
- Artificial neurons and neural networks (ANN)
- Channels describing neurons in ANN
- Binary threshold neurons
- Molecular information exchange

Molecular Information Exchange

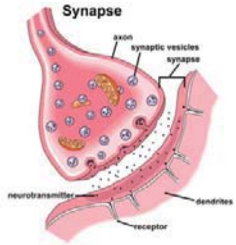


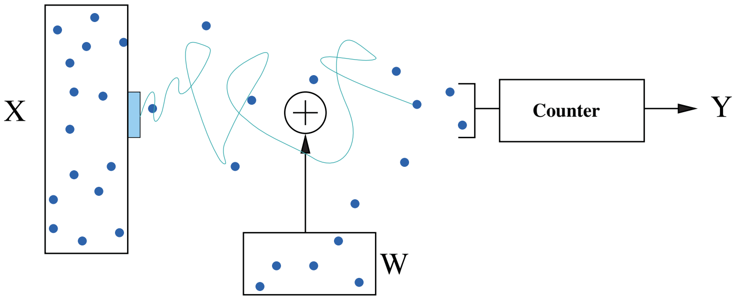
Molecular Information Exchange



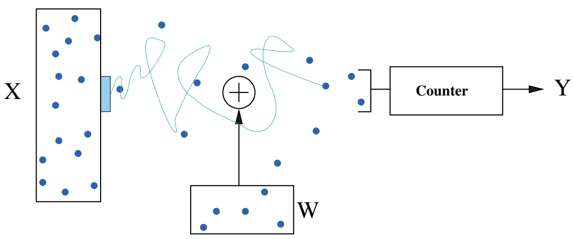


Molecular Information Exchange





Molecular Information Exchange



System model for a Poisson diffusion channel with parameter α:

$$X = N \cdot S,\ S \sim \begin{pmatrix} x_1 & \cdots & x_m \\ p_1 & \cdots & p_m \end{pmatrix},\ 0 \le x_1, \dots, x_m \le 1$$

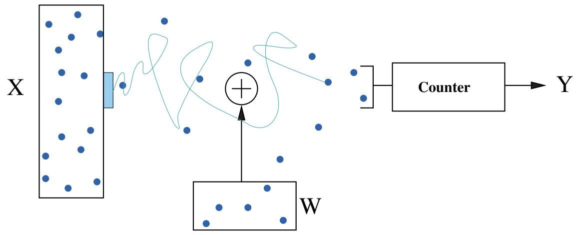
$$W \sim \text{Poi}(\lambda) \text{ independent of } X$$

It holds

$$P^{Y|S=x_i} = \text{Poi}(\alpha N x_i + \lambda) = \text{Poi}(\gamma_i)$$

$$P(Y = k) = \sum_{i=1}^{m} p_i\, e^{-\gamma_i} \frac{\gamma_i^k}{k!} = q_k$$

Molecular Information Exchange



Mutual information: $(\gamma_i = \alpha N x_i + \lambda)$

$$H(Y) = -\sum_{k=0}^{\infty} q_k \ln q_k = H\Big(\sum_{i=1}^{m} p_i \operatorname{Poi}(\gamma_i)\Big)$$

$$H(Y|X) = -\sum_{i=1}^{m} p_i \, H\big(\operatorname{Poi}(\gamma_i)\big)$$

$$I(X;Y) = H(Y) - H(Y|X) = H\Big(\sum_{i=1}^{m} p_i \operatorname{Poi}(\gamma_i)\Big) - \sum_{i=1}^{m} p_i \, H\big(\operatorname{Poi}(\gamma_i)\big)$$

Molecular Information Exchange

Various representations of mutual information:

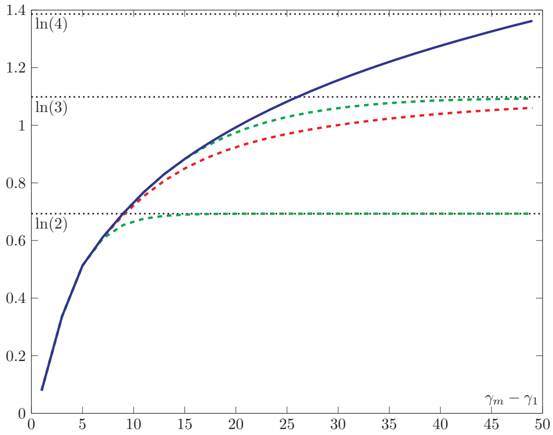
$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H\Big(\sum_{i=1}^{m} p_i \operatorname{Poi}(\gamma_i)\Big) - \sum_{i=1}^{m} p_i \, H\big(\operatorname{Poi}(\gamma_i)\big) \\ &= 1 - D\Big(\sum_{i=1}^{m} p_i \operatorname{Poi}(\gamma_i) \,\Big\|\, \operatorname{Poi}(1)\Big) - \sum_{i=1}^{m} p_i \, \gamma_i (1 - \ln \gamma_i) \\ &= \sum_{k=0}^{\infty} \Bigg[\rho\Big(\sum_{i=1}^{m} p_i \, e^{-\gamma_i} \frac{\gamma_i^k}{k!}\Big) - \sum_{i=1}^{m} p_i \, \rho\Big(e^{-\gamma_i} \frac{\gamma_i^k}{k!}\Big) \Bigg] \end{aligned}$$

with $\rho(z) = -z \ln z$

Determining capacity of Poisson channels is notoriously hard.

Molecular Information Exchange

Numerical calculation of the capacity as a fuction of the support span:



blue: general capacity

green: 3 support points (opt. positions)

red: 3 support points (2 extreme, 1 in the middle)

green: 2 support points (2 extreme)

Still to be done ...

- Do artifical neurons optimize capacity when learning?
- Do biological neurons do the same?
- Can information theory speed up learning of ANNs?
- ANNs with Feedback / cross channels / unlayered structure?
- Capacity-achieving distribution of the molecular information channel?
- Many technical questions ...

Thanks for your attention!

Questions or Comments?

Papers are available from

http://www.ti.rwth-aachen.de/publications/
mathar@ti.rwth-aachen.de