

Computational Neuroscience Notes

Matthew Young

May 2023

1 Integrate-and-Fire Models

- These class of models assume that spikes of a given neuron have roughly the same form and are more concerned about the frequency and timing of spikes.
- In general, neuron models where action potentials are described as events are called **Integrate-and-Fire (IF)** models.
- The simplest of the IF models is the **Leaky Integrate-and-Fire** model, characterised by:
 1. a linear differential equation to describe the evolution of the membrane potential;
 2. a threshold for spike firing.

1.1 Integration of Inputs

- A neuron is like an RC circuit. It is surrounded by a cell membrane acting as an insulator.
 1. When a short current pulse $I(t)$ is injected, the cell membrane gets charged.
 2. The cell membrane acts like a capacitor of capacity C .
 3. As the insulator is not perfect, the charge will slowly leak over time through the cell membrane. The membrane is characterised by a finite leak resistance R .
- Analysing the circuit, the law of current conservation gives:

$$I(t) = I_R + I_C$$

- Using Ohm's law, the resistive current is $I_R = u_R/R$ where $u_R = u - u_{rest}$ is the voltage across the resistor. Whereas, the capacitive current is $I_C = C du/dt$ from the definition of capacity. Thus

$$I(t) = \frac{u(t) - u_{rest}}{R} + C \frac{du}{dt}$$

- Introducing the time constant $\tau_m = RC$ of the "leaky integrator" yields

$$\tau_M \frac{du}{dt} = -[u(t) - u_{rest}] + RI(t)$$

- We refer to u as the **membrane potential** and τ_M as the **membrane time constant** of the neuron.
- The solution to the differential equation above with initial condition $u(t_0) = u_{rest} + \Delta u$ is

$$u(t) - u_{rest} = \Delta u \exp\left(-\frac{t - t_0}{\tau_m}\right)$$

1.2 Pulse Input

- The amplitude of the voltage response of a leaky integrator depends only on the total charge $q = \int I(t)dt$, i.e. area under the curve, and not the height of the current pulse.

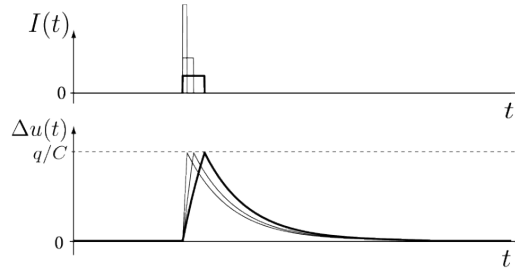


Figure 1: Short pulses and total charge delivered on the membrane.

1.3 The Threshold for Spike Firing

- The firing time $t^{(f)}$ in the leaky integrate-and-fire model is defined by a threshold criterion

$$t^{(f)} : u(t^f) = \vartheta$$

- Immediately after $t^{(f)}$, the potential is reset to a new value $u_r < \vartheta$.
- We write $t_u^{(f)}$ as the firing times of neuron i where $f = 1, 2, \dots$ are the spike labels.
- The **spike train** of a neuron i may be denoted as the sequence of firing times

$$S_i(t) = \sum_f \delta(t - t_i^{(f)})$$

where $\delta = 0$ for $x \neq 0$ and $\int_{-\infty}^{\infty} \delta(x)dx = 1$.

- Under a constant current, regular spikes can be observed. Increasing the current also increases the firing rate.

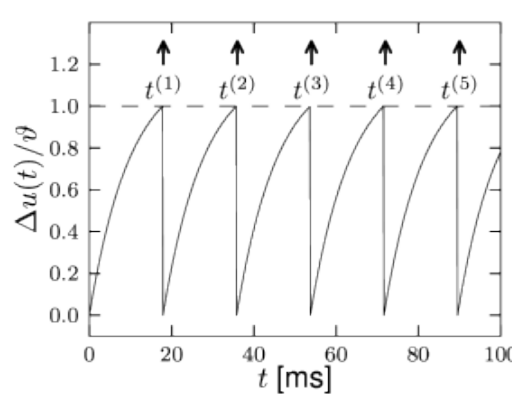


Figure 2: Time course of membrane potential driven by constant input current $I_0 = 1.5$. After a spike, the potential is reset to $u_r = u_{rest}$.

- Plotting du/dt against u , we find for $I = 0$ that u_{rest} is a stable fixed point. Whereas, for $I = I^{const} > 0$, there are no fixed points.

2 The Hodgkin-Huxley Model

2.1 Nernst potential

- It is known in thermodynamics that the probability of a molecule to take a state of energy E is proportional to the Boltzmann factor $p(E) \propto \exp(-E/kT)$.

- The energy of positive ions at location x is given by $E(x) = qu(x)$ where q is the charge in a static electric field and $u(x)$ is the potential at x .
- Since there is a large number of ions, we may interpret the probability as an ion density. Denote by $n(x)$ the ion density at point x . The relation between the density at point x_1 and point x_2 is

$$\frac{n(x_1)}{n(x_2)} = \exp \left[-\frac{qu(x_1) - qu(x_2)}{kT} \right]$$

- A difference in electric potential $\Delta u = u(x_1) - u(x_2)$ generates a difference in ion density.
- Solving the above equation for Δu , we find that, at equilibrium, the concentration difference generates a voltage

$$\Delta u = \frac{kT}{q} \ln \frac{n_2}{n_1}$$

called the **Nernst potential**.

2.2 Reversal Potential

- Embedded in the cell membrane are specific proteins which act as ion gates.
 1. **ion pumps** actively transport ions from one side to the other;
 2. **ion channels** allow the passage of ions.
- In neurons, there is more sodium outside the cell than inside, while the reverse is true for potassium.
- At equilibrium, the difference in concentration in sodium ions causes a Nernst potential E_{Na} . When Δu is smaller than E_{Na} , more Na^+ ions flow into the cell through sodium ion channels. If it is larger, than ions flow out.
- The direction of current is thus reversed when the voltage Δu passes E_{Na} . Here, E_{Na} is called the **reversal potential**.
- Measured experimentally, we find $E_K < u_{rest} < E_{Na}$. Therefore, in the stationary state, active ion pumps balance the flow of ions in and out the cell.

2.3 Hodgkin-Huxley Model

- The neuron cell can be expressed as a circuit as follows
 1. The semipermeable cell membrane acts as a capacitor.
 2. When an input current $I(t)$ is injected into the cell, it either charges the capacitor or leaks through the channels.
 3. The sodium and potassium channels have leak resistances R_{Na} and R_K respectively. The unspecific channel accounting for leaks has resistance R .
 4. The Nernst potential for each ion type are represented by batteries with voltages E_{Na} , E_K , and E_L .

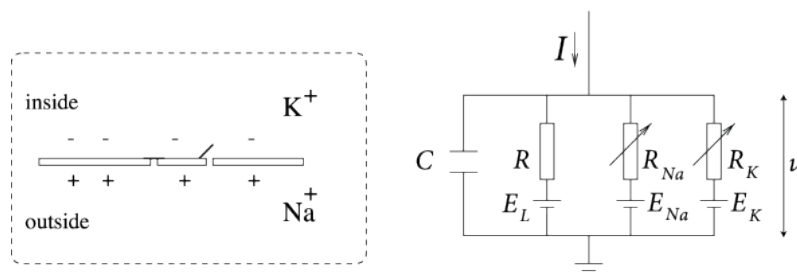


Figure 3: Circuit corresponding to the Hodgkin-Huxley model.

- By the conservation of electric charge, the applied current can be split in a capacitive current I_C and further components I_k passing through the ion channels

$$I(t) = I_C(t) + \sum_k I_k(t)$$

- From the definition of capacitance, we find $I_C = C du/dt$, and so

$$C \frac{du}{dt} = - \sum_k I_k(t) + I(t)$$

- Since the resistance of the leakage channel is voltage-independent, and the voltage at the leak resistor is $u - E_L$, Ohm's law yields

$$I_L = g_L(u - E_L)$$

where $g_L = 1/R$ is the conductance of the leakage channel.

- Similarly, for the other ion channels, we find

$$C \frac{du}{dt} = g_{Na}(E_{Na} - u) + g_K(E_K - u) + g_L(E_L - u) + I(t)$$

However, the g_{Na} and g_K are voltage and time dependent depending on the number of open channels. When all are open, then the maximum conductance \bar{g}_{Na} and \bar{g}_K are observed.

- Using variables m, n, h to describe the probability that a channel is open at any moment,

$$C \frac{du}{dt} = \bar{g}_{Na} m^3 h (E_{Na} - u) + \bar{g}_K n^4 (E_K - u) + g_L (E_L - u)$$

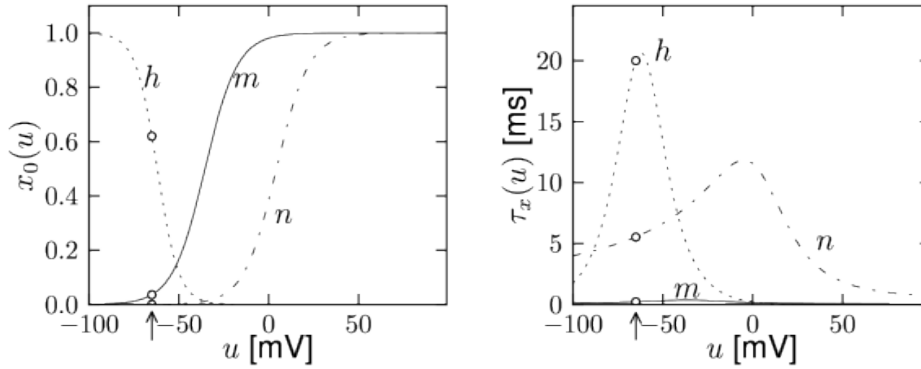


Figure 4: Voltage dependence of the gating variables m, n, h .

- To reduce the model to a 2D system, we assume m is fast enough and reaches its steady state early, i.e. $m = m_0$. Whereas $n \sim 1 - h =: w$. The system is thus

$$C \frac{du}{dt} = \bar{g}_{Na} m_0^3 (1 - w) (E_{Na} - u) + \bar{g}_K w^4 (E_K - u) + g_L (E_L - u)$$

$$\tau_w \frac{dw}{dt} = -w + w_0$$

3 2D Neuron Models

3.1 Fitzhugh-Nagumo Model

- For small $\epsilon > 0$ and $b_0, b_1 > 0$, the model is given by

$$\dot{u} = u - \frac{1}{3}u^3 - w + I$$

$$\dot{w} = \epsilon(b_0 + b_1 u - w)$$

- When $I = 0$, there exists a stable fixed point (u_{rest}).
- When $I > 0$, there exists an unstable fixed point and a limit cycle.
- As the behaviour of the system changes from a stable to a limit cycle, this is called a **Hopf bifurcation**.

3.2 Morris-Lecar Model

- Similar to the 2D Hodgkin-Huxley model, but with sigmoid functions and no exponents.
- It is given by

$$C \frac{du}{dt} = -g_1 \hat{m}_0(u)(u - u_1) - g_2 \hat{w}(u - u_2) - g_L(u - u_L) + I$$

$$\frac{d\hat{w}}{dt} = -\frac{1}{\tau(u)} [\hat{w} - w_0(u)]$$

where we have the activations

$$\hat{m}_0(u) = \frac{1}{2} \left[1 + \tanh \left(\frac{u - u_1}{u_2} \right) \right], \quad w_0(u) = \frac{1}{2} \left[1 + \tanh \left(\frac{u - u_3}{u_4} \right) \right]$$

$$\tau(u) = \frac{\tau_w}{\cosh \left(\frac{u - u_3}{2u_4} \right)}$$

- When $I = 0$, there exists a stable, a saddle, and an unstable fixed point. The saddle can be interpreted as the threshold.
- When $I > 0$ and the two fixed points merge, there exists one saddle and one unstable fixed point.
- When $I > 0$ and only one fixed point remains, there is only an unstable fixed point.

3.3 Adaptive Exponential Integrate-and-fire Model

- Integrate and fire but with additional exponential term.
- It is given by

$$\tau_m \frac{du}{dt} = -[u - u_{rest}] + \Delta t \exp \left(\frac{u - \theta_m}{\Delta t} \right) - R w + R I(t)$$

$$\tau_w \frac{dw}{dt} = a(u - u_{rest}) - w$$

- Together with this, whenever $u = \theta_m$, we update $u \mapsto u_{rest}$ and $w \mapsto w + b$.
- It is able to capture the behaviour of neurons pausing after some number of spikes before continuing. A larger b means the pauses are more frequent, whereas a smaller one allows for multiple burst before a pause.

4 Spatial Structure of Neurons

4.1 The Cable Equation

- It describes the flow of current across the dendrite. It is similar to that of inside the cell membrane but allows for a longitudinal flow and thereby the spatial dynamics.

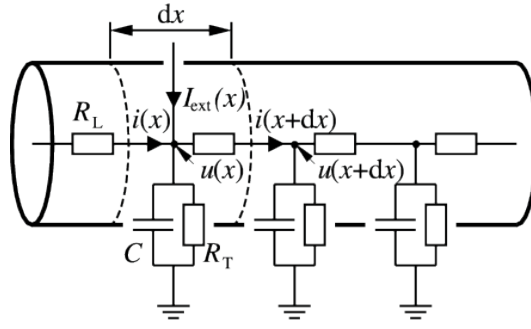


Figure 5: Part of a dendrite and the corresponding circuit diagram.

- By Ohm's law, we have

$$u(t, x + dx) - u(t, x) = R_L i(t, x)$$

- By the conservation of current, i.e. what comes in must come out

$$I_{ext}(t, x) + i(t, x) = i(t, x + dx) + C \frac{du}{dt}$$

or equivalently,

$$i(t, x + dx) - i(t, x) = I_{ext}(t, x) - C \frac{du}{dt}$$

- We assign specific quantities in terms of unit lengths:

$$R_L = r_L dx, \quad C = c dx, \quad I_{ext}(t, x) = i_{ext}(t, x) dx, \quad I_{ion}(t, x) = i_{ion}(t, x) dx$$

- Substituting, dividing by dx , and taking the limit $dx \rightarrow 0$ leads to

$$\begin{aligned} \frac{\partial}{\partial x} u(t, x) &= r_L i(t, x) \\ \frac{\partial}{\partial x} i(t, x) &= c \frac{\partial}{\partial t} u(t, x) + \sum_{ion} i_{ion}(t, x) - i_{ext}(t, x) \end{aligned}$$

- Taking the derivative of above and substituting below,

$$\frac{\partial^2}{\partial x^2} u(t, x) = cr_L \frac{\partial}{\partial t} u(t, x) + r_L \sum_{ion} i_{ion}(t, x) - r_L i_{ext}(t, x)$$

4.2 Model of Synapse

- The synaptic current is given by

$$I_{syn}(t) = g_{syn}(t)(u(t) - E_{syn})$$

where we model the synaptic conductance following an exponential decay

$$g_{syn}(t) = \sum_f \bar{g}_{syn} \exp\left(-\frac{t-t^f}{\tau}\right) O(t-t^f)$$

$$O(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

4.3 Modelling Noise

4.3.1 Noise in the Input

- Consider adapting the integrate-and-fire model and adding a deterministic noise or some other stochastic noise, e.g. white noise, colored noise

$$\tau_m \frac{du}{dt} = f(u) + RI^{det} + RI^{noise}$$

4.3.2 Stochastic spike arrival

- We can induce a weighted current every time a presynaptic neuron spikes

$$\tau_m \frac{du}{dt} = f(u) + RI^{det} + \sum_k \sum_{t_k^f} w_k \delta(t - t_k^f)$$

4.3.3 Noise in the output

- We may allow spiking before reaching a threshold, yielding some probability of spiking

$$s = f(u(t) - \theta)$$

where the probability is higher the closer the potential is to the threshold.

5 Characterisation of neural activity

5.1 Rate as spike count

- We count the number of spikes n_{sp} and denote the average as the rate

$$\nu = \frac{n_{sp}}{T}$$

- To consider deviations from trial to trial, we devise the Fano factor

$$F = \frac{\langle \Delta n_{sp}^2 \rangle}{\langle \Delta n_{sp} \rangle}$$

which is the variance divided by the mean.

5.2 Rate as spike density

- The average spike count may be unreliable when the duration of spiking is not very long.
- Consider instead a time delta common to all k trials. Measuring the spike density,

$$\rho = \frac{1}{\Delta t} \frac{1}{K} \sum_k n_k(t, t + \Delta t)$$

- Alternatively, in terms of the spike train, we have

$$n_k^{sp} = \int_t^{t+\Delta t} s_k(t') dt'$$

5.3 Rate as population activity

- Assume neurons are equivalent, then we average across the neurons

$$A = \frac{1}{\Delta t} \frac{n_{act}(t, t + \Delta t)}{N}$$

or, alternatively in terms of spike trains,

$$A(t) = \frac{1}{\Delta t} \frac{1}{N} \int_t^{t+\Delta t} \sum_k \sum_{t_k^f} \delta(t' - t_k^f) dt'$$

5.4 Interspike interval

- Measure the intervals between spikes and derive their distribution

$$P_0(s) = P(t^f + s | t^f)$$

where

$$\int_t^{t+\Delta t} P(t' | t^f) dt'$$

is the probability that the next spike occurs in the interval $[t, t + \Delta t]$.

- The mean and variance are given by

$$\langle s \rangle = \int_0^\infty s P_0(s) ds$$

$$\langle \Delta s^2 \rangle = \int_0^\infty s^2 P_0(s) ds - \langle s \rangle^2$$

- To determine the regularity of the intervals, we consider the coefficient of variation c_v

$$c_v = \frac{\langle \Delta s^2 \rangle}{\langle s \rangle^2}$$

- Note that a Poisson process has $c_v = 1$. And so, if $c_v > 1$ then it is less regular than a Poisson process, otherwise it is more regular.
- If spiking is periodic, then $c_v = 0$, whereas $c_v > 1$ if it is bursting.

5.5 Neural Networks

5.6 Neural Networks as a Population Description

- We assume that the network is homogeneous, that is:
 1. identical neurons
 2. identical inputs
 3. statistically homogeneous connectivity
- The stationary population activity is the mean firing rate of a single neuron and is thus

$$A_0 = \nu_i$$

where we note A_0 is a spatial average and ν_i is a temporal average.

5.6.1 Rate neuron description

- We call a population of homogeneous neurons interconnected as a **rate neuron**.
- We may adapt the firing rate of a single neuron under the IF model to include both noise and time dependence.
- Originally, we have the gain function

$$\nu_i = g(I_0) = \tau \ln \left(\frac{RI_0}{RI_0 - (\theta - u_{rest})} \right)^{-1}$$

for $RI_0 > \theta - u_{rest}$.

- Adding noise, the original equation becomes

$$\tau \frac{du}{dt} = -[u - u_{rest}] + RI_0 + \sigma \eta$$

where at each time step, we draw random numbers from a Gaussian with variance σ^2 . We can abstract this curve as a smoothed sigmoid gain function

$$\nu_i = g_\sigma(I_0) = \frac{\alpha}{1 + e^{-\beta(I_0 - \gamma)}}$$

- Indeed at the steady state, we have

$$\nu = g_{\sigma}(I_0)$$

- The time-dependent firing rate can be modelled as a low-pass filter of the steady state

$$\tau \frac{d\nu}{dt} = -\nu + g_{\sigma}(I(t))$$

5.6.2 Rate neuron with self-connection

- Suppose a neuron is connected to itself with weight w , then the dynamics of its firing rate is governed by

$$\dot{\nu} = -\nu + g(w\nu + I)$$

- Depending on the weight, we either get a single stable point or a bistability together with an unstable point. This bistability allows for a model of working memory.

5.6.3 Two inhibitory rate-neurons coupled

- A presynaptic neuron is inhibitory if it induces a negative current.
- The firing rates are governed by

$$\begin{aligned}\dot{\nu}_1 &= -\nu_1 + g(-w\nu_2 + I) \\ \dot{\nu}_2 &= -\nu_2 + g(-w\nu_1 + I)\end{aligned}$$

- This yields three fixed points: two stable on opposite ends and one unstable. In either of the stable fixed points, the firing rate of one neuron is high whereas the other is low. This exhibits an inhibitory behaviour between two neurons.

5.6.4 One inhibitory and one excitatory neurons coupled

- The firing rates are governed by

$$\begin{aligned}\dot{\nu}_E &= -\nu_E + g(w_{EE}\nu_E - w_{EI}\nu_I + I^E) \\ \dot{\nu}_I &= -\nu_I + g(w_{IE}\nu_E - w_{II}\nu_I + I^I)\end{aligned}$$

- This yields one stable fixed point. As shifting the nullclines left or right either increases or decreases both neuron activity at the same time, this explains the correlation of inhibitory and excitatory firing rates observed in practice.

5.6.5 Continuous Rate Network

- Also known as a **ring network**, involves neurons encoding for visual stimuli with preferences on the orientation of some stimulus.
- That is, a neuron may prefer an orientation somewhere between 0 and 2π radians. To account for the cyclic nature of angles, we connect neurons together in a ring.
- Denoting the weight of the connection from neuron a to b as $w_{ab} = w(\theta_a, \theta_b)$, we have the continuous rate equation

$$\tau \frac{d\nu(\theta)}{dt} = -\nu(\theta) + g \left[\int_{-\pi}^{\pi} w(\theta, \theta') \nu(\theta') d\theta' + I(\theta) \right]$$

- Often we use $w(\theta, \theta') = \cos(\theta - \theta')$ as for $|\theta - \theta'|$ small (big), w is positive (negative). This leads to inhibitory behaviour when a neuron is completely different in orientation and an excitatory behaviour with similar orientation neurons.

5.6.6 Original Ring Network

- Under the continuous rate network, we may present a stimulus: an oriented bar with degree θ_0 and model the current coming into a neuron i with preferred orientation θ_i as

$$I_i(\theta_0) = c[(1 - \epsilon) + \epsilon \cos(2(\theta_i - \theta_0))]$$

where c denotes the image contrast and ϵ is how strong the input is modulated.

- Two observations from experimental data can then be explained:
 1. **contrast invariance** - increasing the contrast of the stimulus, the tuning curve (firing rate function in terms of orientation) does not get broader. Even though the input current to the neurons become broader, the firing rate does not. A form of inhibition is responsible for this invariance and is known as the **iceberg effect**.
 2. **sustained activity** - after removing an induced stimulus, neurons persist to fire with the same pattern. Some form of bistability is responsible where there is a stable fixed point at a high firing rate.

5.7 Neural Networks of spiking models

- We aim to adapt IF models for a large number of neurons and hopefully address experimental results showing low spiking rate but high variance - akin to a Poisson process.

5.7.1 Balanced Network

- Neurons are assumed to
 1. be identical
 2. receive the same current
 3. have statistically uniform connectivity
- The number of excitatory and inhibitory neurons N_E and N_I are very large.
- Each neuron gets $K_E = K$, $K_I = \gamma K$ number of incoming synapses.
- Given r_X, r_E, r_I as the firing rate of the external input, excitatory neurons, and inhibitory neurons respectively, the neurons receive a mean input

$$\begin{aligned}\mu_E &= K[w_{EE}r_X + w_{EE}r_E - \gamma w_{EI}r_I] \\ \mu_I &= K[w_{IE}r_X + w_{IE}r_E - \gamma w_{II}r_I]\end{aligned}$$

and input variance

$$\begin{aligned}\sigma_E^2 &= K[w_{EE}^2r_X + w_{EE}^2r_E + \gamma w_{EI}^2r_I] \\ \sigma_I^2 &= K[w_{IE}^2r_X + w_{IE}^2r_E + \gamma w_{II}^2r_I]\end{aligned}$$

- Note that if we keep the weights of the network the same but scale it ten times bigger, we should expect that the output is ten times more frequent. However, this is not the case experimentally. We observe low firing rate with high variance. So, we attempt to scale the synaptic weights based on the number of incoming synapses K .
- If we consider a **weak coupling**, that is, scale the synaptic weights relative to the number of connections, i.e. $w \mapsto w/K$, then we have

$$\begin{aligned}\mu_E &= w_{EE}(r_X + r_E) - \gamma w_{EI}r_I \\ \sigma_E^2 &= \frac{1}{\sqrt{K}} \sqrt{w_{EE}^2(r_X + r_E) + w_{EI}^2r_I}\end{aligned}$$

- While the mean does not grow with K , the variance tends to zero suggesting a regular fire rate which is inconsistent with data.

- If we consider a **strong coupling**, that is scale $w \mapsto w/\sqrt{K}$, we get

$$\mu_E = \sqrt{K}(w_{EE}(r_X + r_E) - \gamma w_{EI}r_I)$$

$$\sigma_E^2 = \sqrt{w_{EE}^2(r_x + r_E) + w_{EI}^2 r_I}$$

- While the variance does not shrink with K , the mean firing rate grows with K which is yet again inconsistent with data.
- However, if the above term is of a certain order, then this is not a problem:

$$w_{EE}(r_X + r_E) - \gamma w_{EI}r_I \sim O\left(\frac{1}{\sqrt{K}}\right)$$

- The above LHS is called the **balanced state**.
- The idea is that the excitatory and inhibitory currents are both large but cancel each other out, thereby leading to low mean firing rate but the variance of which is the sum of the variance of two incredibly large currents.

6 Learning and Memory

- Three types of learning
 - **unsupervised learning**: learn statistics of inputs
 - **reinforcement learning**: learn with a reward
 - **supervised learning**: have reward signals at any point

6.1 Unsupervised learning

- **Hebb principle**: "Who fire together, wire together".
- In experiments, it is observed that by firing a presynaptic and postsynaptic neuron together multiple times, the weight of the postsynaptic response becomes bigger. This phenomenon is called **potentiation**. If it lasts longer than 30 minutes than it is called **long-term potentiation**.

6.1.1 Hebbian learning in rate-based formalism

- Consider a presynaptic neuron with firing rate x_i and a postsynaptic neuron with firing rate y , with a synaptic weight of w_i connecting the two. The dynamics of the weight can change by

$$\frac{dw_i}{dt} = F(w_i, x_i, y)$$

for all i , where F is only concerned with other local neurons x_j . That is, the weight isn't influenced by neurons that are far from it.

- Expanding F in a Taylor series,

$$\frac{dw_i}{dt} = c_0(w_i) + c_1^{pre}(w_i)x_i + c_1^{post}(w_i)y + c_2^{post}(w_i)y^2 + c_2^{post}(w_i)x_i^2 + \underbrace{c_{11}^{corr}(w_i)x_i y_i}_{\text{correlation term}} + O(x^3)$$

- By only considering a pure Hebbian term, i.e. the correlation, we have

$$\frac{dw_i}{dt} = c_{11}^{corr} x_i y$$

- Since w_i is changing much slower than x_i and y , we can take the temporal average

$$\frac{dw_i}{dt} = c_{11}^{corr} \langle x_i y \rangle$$

- In a linear rate-framework, the postsynaptic firing rate y is the weighted average of the incoming firing rates and thus

$$y = \sum_j w_j x_j$$

and so

$$\frac{dw_i}{dt} = c_{11}^{corr} \sum_j \langle x_i x_j \rangle w_j$$

- Define $Q_{ij} = \langle x_i x_j \rangle$ as an entry to the input correlation matrix \mathbf{Q} . Then we have, for all the weights,

$$\frac{d\vec{w}}{dt} = c_{11}^{corr} \mathbf{Q} \vec{w}$$

- However, this only continues to grow over time which is physically not correct. to address this we can consider
 - introduce a **hard bound** $0 < w_i < w^{max}$, but all weights will become fixed at w^{max} inhibiting any possible form of learning.
 - induce some **depression**

6.1.2 Covariance rule

- We assume that only when the presynaptic and postsynaptic neurons spike together more than their average do their weights strengthen, i.e.

$$\tau_w \frac{d\vec{w}}{dt} = (x - \theta_x)(y - \theta_y)$$

where $\theta_x = \langle \vec{x} \rangle$, $\theta_y = \langle y \rangle$ are the temporal averages.

- Writing as entries on a matrix,

$$C_{ij} = (x_j - \langle x \rangle)(x_i - \langle x \rangle)$$

we get

$$\tau_w \frac{dw_i}{dt} = \sum_j (x_j - \langle x \rangle) w_j (x_i - \langle x \rangle)$$

$$\tau_w \frac{d\vec{w}}{dt} = \mathbf{C} \vec{w}$$

6.1.3 BCM learning rule

- Consider adding depression on the Hebbian learning rule

$$\tau \frac{d\vec{w}}{dt} = \vec{x} y (y - \theta)$$

- When we have two incoming x_1, x_2 with $x_1 \implies y < \theta$ and $x_2 \implies y > \theta$, then we find that $w_1 \rightarrow 0$ and $w_2 \rightarrow w^{max}$. This phenomenon is called **selectivity**.
- However, if both x_1, x_2 are induce $y < \theta$, e.g. when θ is absurdly high, then the postsynaptic neuron becomes a silent neuron. This is not good as it costs energy to maintain these weights and connections despite the neuron having the inability to function.
- To account for this, we make θ dynamic by adding the updating rule

$$\tau_\theta \frac{d\theta}{dt} = \frac{y^2}{y_{target}} - \theta$$

- Now, even if x_1, x_2 both induce $y < \theta$, the threshold adjusts itself by sliding until one of the presynaptic rates are greater than it, thus allowing for selectivity once again.
- The average post-synaptic rate has a steady state of y_{target} as later $\langle y \rangle = \theta$.
- Selectivity in neurons is beneficial in the development of infants, e.g. certain neurons learn to look for certain important features.

6.1.4 Spike-timing dependent plasticity (STDP)

- Experimentally, we find that if a postsynaptic neuron fires as soon as a presynaptic neuron fires, then their synaptic weight increases. Otherwise, it decreases.
- To model this, we may add a trace that decays after each pre and post spike. We compare at each pre spike if a non-zero post spike exists and vice-versa. Denote by x and y the amount of trace for the pre and post spike.

$$\begin{aligned}\tau_+ \frac{dx}{dt} &= -x + \delta(t - t^{pre}) \\ \tau_- \frac{dy}{dt} &= -y + \delta(t - t^{post})\end{aligned}$$

- Then we can update the weight as

$$\frac{dw}{dt} = A_+ x \delta(t - t^{post}) - A_- y \delta(t - t^{pre})$$

- This allows us to learn two things:
 1. **Responding to earlier stimulus** - when multiple presynaptic neurons fire sequentially to the same postsynaptic neuron, we find that the increase in weights shift the postsynaptic response earlier to the point where the latter presynaptic neuron weights are substantially lower than that of the first one.
 2. **Learning sequences** - when we have a series of neurons that are hooked up linearly, we find that the weights strengthen in one direction.

6.2 Supervised learning

- The goal of this learning paradigm is to learn associations between inputs \vec{x}^μ and target output y_t^μ . That is to have the produced output $y^\mu \rightarrow y_t^\mu$.
- Here, the actual output is binary, either a spike or no spike.
- Given this network of inputs (presynaptic neurons) and a single output (postsynaptic neuron), we have a structure called a **perceptron**.
- The corresponding output is given by

$$y^\mu = O \left[\sum_i w_i x_i^\mu - b \right]$$

where O is the Heaviside step function and b is the firing threshold.

- The error of the prediction can be given by

$$E = (y_t - y)^2$$

and the derivative with respect to a weight w_i is

$$\frac{dE}{dw_i} = 2(y_t - y)x_i$$

- The discrete and continuous learning rules are thus

$$\begin{aligned}\vec{w} &= \vec{w} + \alpha \vec{x}(y_t - y) \\ \tau_w \frac{d\vec{w}}{dt} &= \vec{x}(y_t - y)\end{aligned}$$

where α is called the **learning rate**.

- Geometrically, the perceptron defines a hyperplane. If we plot the data points and label them accordingly, then the task is essentially to linearly separate the two clusters.
- For example, we can learn AND and OR gates as they are linearly separable. To tackle nonlinearly separable data such as XOR gates, we stack layers of perceptrons. In this case, the second layer could learn the concept of an OR gate and a NON-AND gate and therefore learn an AND gate as an output neuron to therefore produce an XOR.

6.3 Reinforcement Learning

6.3.1 Rescola-Wagner Rule

- Given a stimulus u (either zero or one) and a reward r , we have the expected reward

$$v = wu$$

where w is the weight to be learned.

- The goal is to minimize the error

$$E = (r - v)^2$$

- Measuring the change in the error in terms of the weights

$$\frac{\partial E}{\partial w} \propto \underbrace{(r - v)}_{\delta} u$$

where δ is called the **prediction error**.

- The update rule is thus given by

$$w \mapsto w + \epsilon \delta u$$

where ϵ is the learning rate.

6.3.2 Blocking experiment

- When an animal is trained to expect:

$$\begin{aligned} s_1 &\rightarrow r \\ s_1 + s_2 &\rightarrow r \end{aligned}$$

for stimuli s_1 and s_2 , the animal does not associate s_2 as leading to a reward.

- This is consistent with the rule as when s_1 is fully learned, then $w_1 = 1$ leading to an error of zero. As such, $\delta = 0$ and there is no learning, i.e. no updating of weight w_2 .

6.3.3 Secondary conditioning

- When an animal is trained to expect:

$$\begin{aligned} s_1 &\rightarrow r \\ s_1 &\rightarrow s_2 \quad (\text{no reward}) \end{aligned}$$

the animal associates both stimuli as leading to a reward.

- However, this is inconsistent with the learning rule as if $s_1 \rightarrow r$, then it learns $w_1 = 1$. However, since we can have $s_1 + s_2$ but no reward, then it learns that $w_2 := -w_1$.

6.3.4 Temporal difference (TD) learning

- This improves upon the previous learning rule by taking time into account.
- Define $v(t)$ as the expected value of the total future reward from time t . We can estimate this as

$$v(t) \sim \left\langle \sum_{\tau=0}^{T-t} r(t+\tau) \right\rangle$$

where T is the total time of the experiment.

- We can model this as a weighted sum of the temporal inputs

$$v(t) = \sum_{\tau=0}^t w(\tau) u(t-\tau)$$

- The corresponding error is thus

$$E = \left(\underbrace{\sum_{\tau} r(t + \tau) - v(t)}_{\delta(t)} \right)^2$$

- However, this is in terms of the future which is inaccessible. Instead, we write the recurrence relation

$$\underbrace{\sum_{\tau=0}^{T-t} r(t + \tau)}_{v(t)} = r(t) + \underbrace{\sum_{\tau=0}^{T-t-1} r(t + 1 + \tau)}_{v(t+1)}$$

and so we have the learning rules

$$\begin{aligned} \delta(t) &= r(t) + v(t+1) - v(t) \\ w(\tau) &\rightarrow w(\tau) + \epsilon \delta(t) u(t - \tau) \end{aligned}$$

7 Hopfield network

- We take inspiration from spins of electromagnets in physics and consider weights equal to the product of spins.
- We use $h_i(t)$ to denote the current at time t .

7.1 Pattern storing

- Given N binary neurons, with $s_i = \pm 1$, the task of the network is to learn M different kinds of patterns p_i^μ .
- That is, we want the neurons to be at p_i^μ and stay there, i.e. it is the fixed point

$$s_i(t) = s_i(t + \Delta t) = p_i^\mu$$

- We have the probability of neuron i firing

$$P(s_i(t + \Delta t) = 1 \mid h_i(t)) = g \left(\sum_j w_{ij} s_j(t) \right)$$

and assign the weights as

$$w_{ij} = c \sum_{\mu=1}^M p_i^\mu p_j^\mu$$

for some $c > 0$, under a Hebbian-like learning scheme.

7.2 Pattern retrieval

- We want to converge to the fixed point corresponding to the pattern μ which is most similar to the initial condition.
- Define the overlap of the target pattern and the prediction as

$$m^\mu(t) = \frac{1}{N} \sum_i p_i^\mu s_i(t)$$

- Then the current in terms of the overlap is

$$h_i(t) = \sum_j w_{ij} s_j(t) = c \sum_{j=1}^N \sum_{\mu=1}^M p_i^\mu p_j^\mu s_j(t) = cN \sum_{\mu=1}^M p_i^\mu m^\mu(t)$$

- Taking $c = 1/N$, we thus have the probability of a neuron i firing

$$P(s_i(t + \Delta t) | h_i(t)) = g \left(\sum_{\mu=1}^M p_i^\mu m^\mu(t) \right)$$

that is, the dynamics of retrieval are determined by the overlap of patterns.

kkkjhhghjkkjhgllllllllkfkjgktyeb kklklkopopooklkgggffvb cdszasd