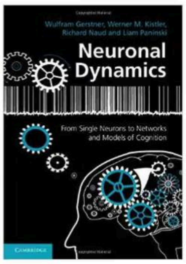
Computational Neuroscience: Neuronal Dynamics of Cognition

Synaptic Plasticity and Learning

Wulfram Gerstner

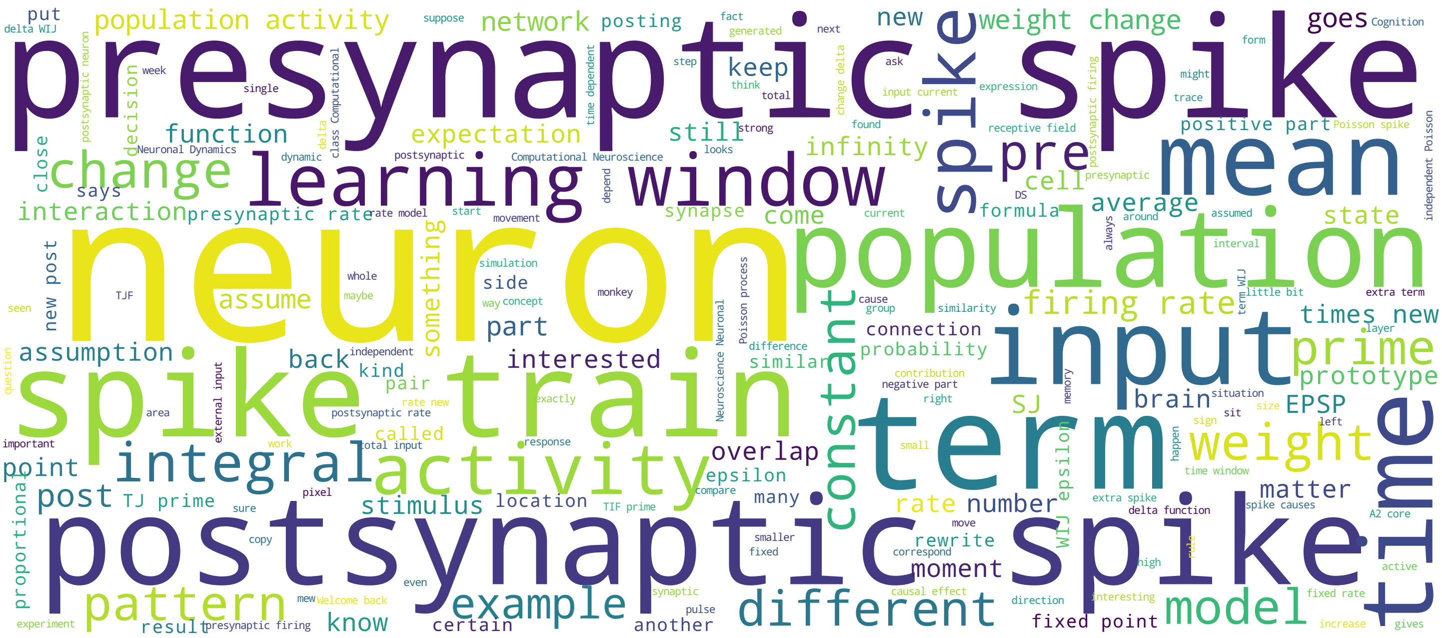
EPFL, Lausanne, Switzerland

Reading for plasticity:
NEURONAL DYNAMICS
- Ch. 19.1-19.3



Cambridge Univ. Press

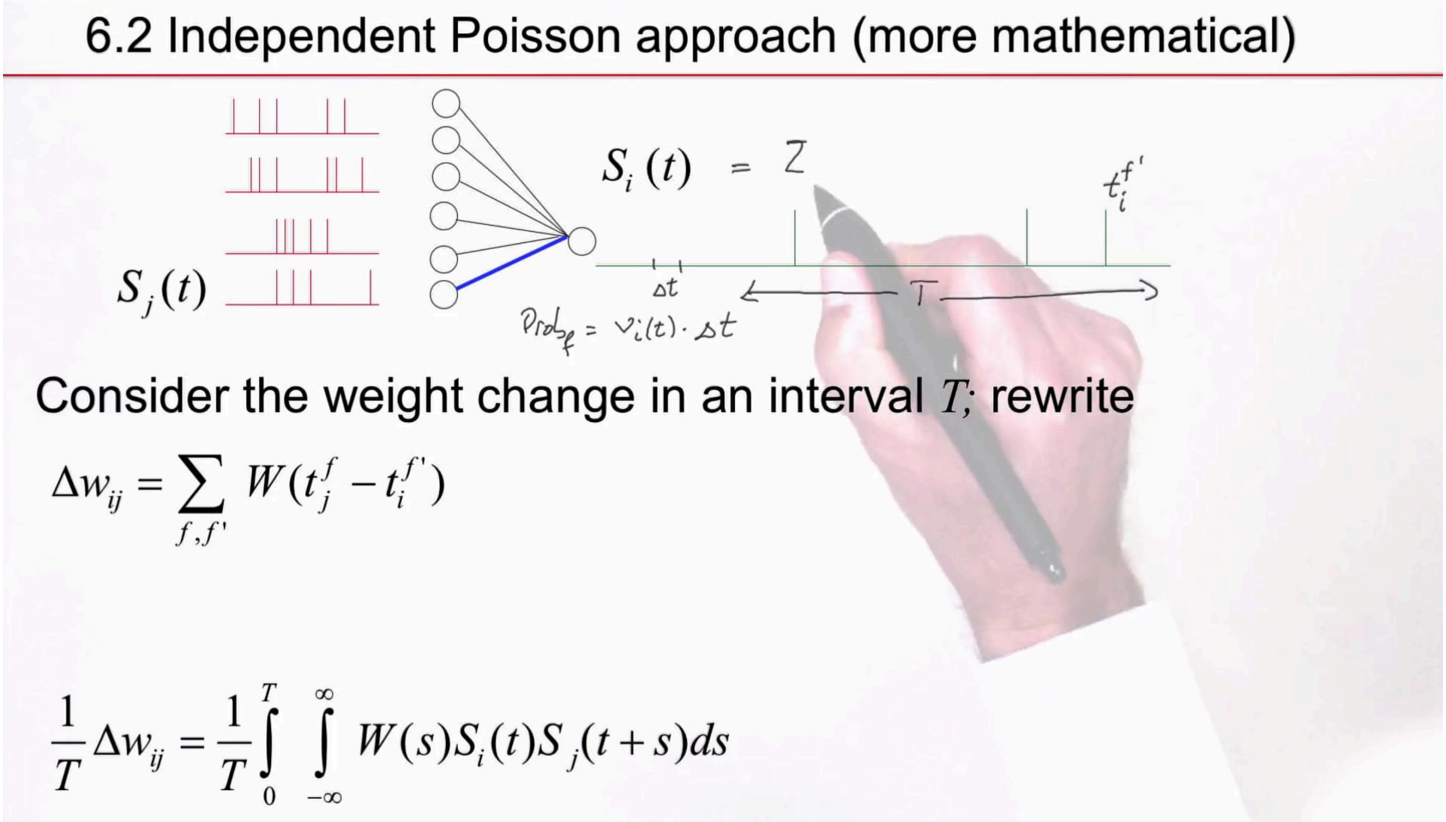
1. **Synaptic plasticity: aims**
2. **Classification of plasticity**
3. **Model of short-term plasticity**
4. **Models of long-term plasticity**
5. **Spike-Timing Models of plasticity**
6. **From spiking models to rate models**
   - intuitive
   - independent Poisson
   - model with EPSPs
7. **Triplet STDP model**
   - Relation to experiments
   - Relation to BCM rule
8. **Online learning of memories**



Search MOOC

Video

EPFL

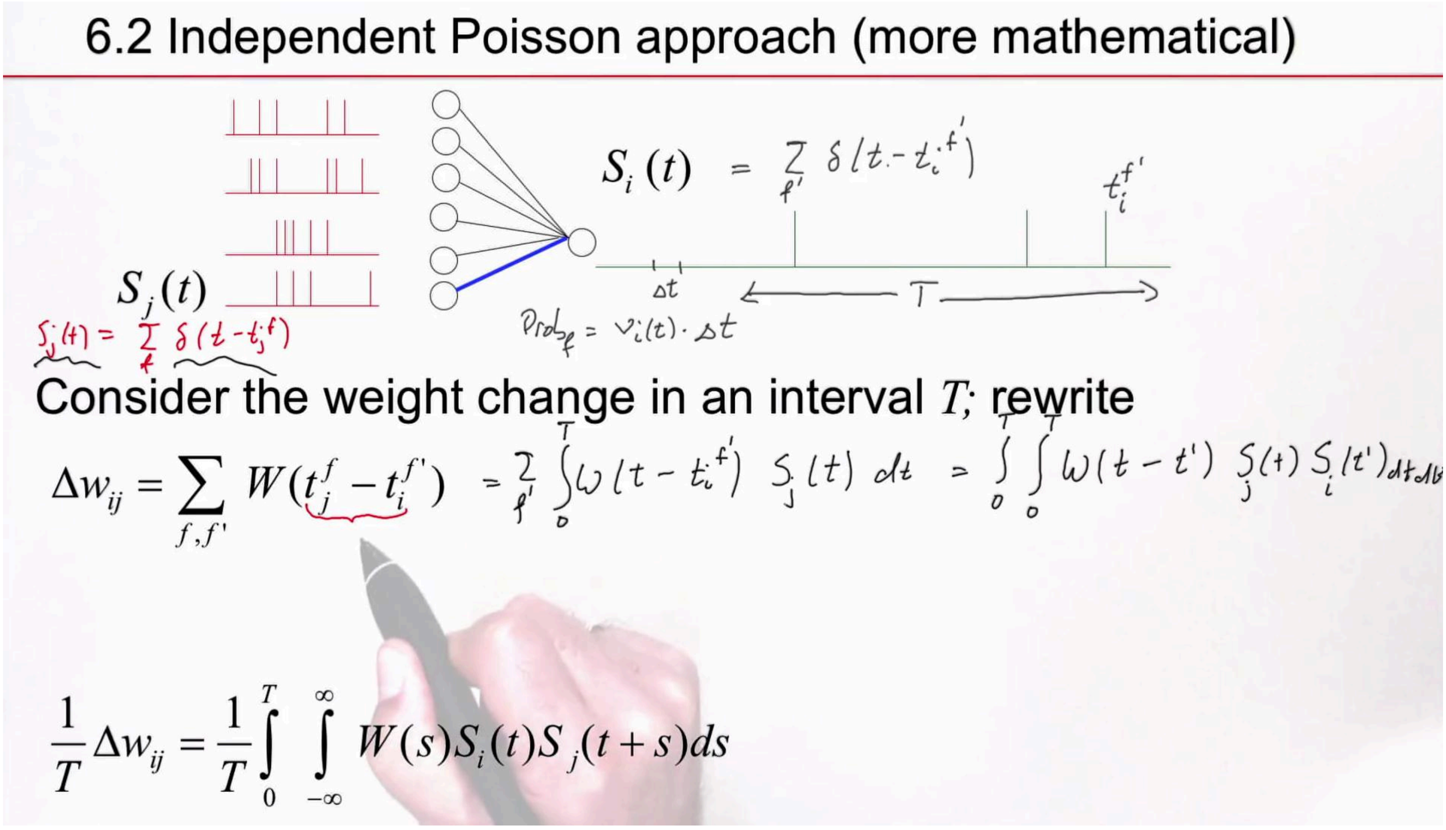


Welcome back to the class, Computational Neuroscience, Neuronal Dynamics of Cognition. This video presents a mathematical detour and should be interesting for those who want to learn more about the relation between STDP and rate-based, Hebbian models. Those not interested in Math can go straight to part 7. So, now let's redo the same kind of argument but from a more mathematical perspective. So, assume that the input is given by Poisson spike trains. I assume that the output is also given by a Poisson spike train, so the probability that it will find a spike in a very small time window of length, Delta T is the rate times the Delta T. This is the probability of firing in this interval Delta T. Now, we are interested in the weight change over a much longer time window, which I call capital T. So, we consider the weight change in this interval, and this is the sum of all spikes of the presynaptic neuron of the posting of the neuron that occur in this window, capital T. Let's first rewrite this expression. And to do that, I use this notion of a spike train. I have here my spikes; the spikes sit at some time, which I call TIF prime. So this would be F prime equal one first spike, second spike, third spike.

Notes

Summary

0m 01s



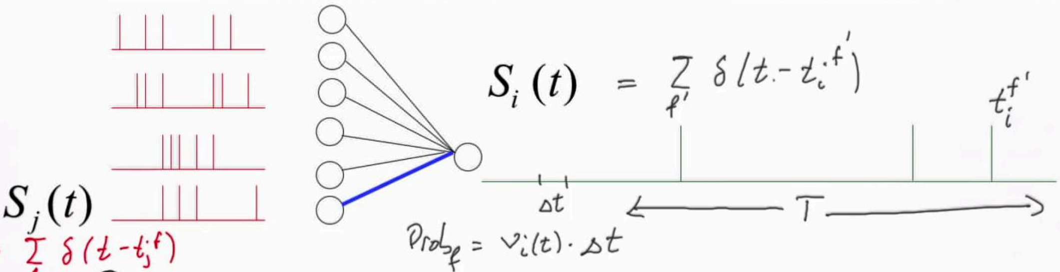
So despite train is the sum of all these pulses, T minus, TIF prime, sum over F prime. Similarly, for the presynaptic spike train, I can write it as SJ of T, sum of all the pulses, sum over F Delta of T minus, TJF. Now. I use this expression, now to rewrite this expression, sum of all pairs of spikes. So, let's first keep the postsynaptic spike fixed and consider the presynaptic spikes. I can rewrite this by saying this is W of T minus, TIF prime, integral S of TDT integral from zero to T. And I kept the postsynaptic spikes. So I still have to sum over F prime. Now, these are delta functions, so if you block this back in, you see that you get back to this formula, you take the red term here with these delta functions, you insert it here, you find this back, I can do something very similar for the postsynaptic spike train. I insert this as well. So let me write this as W off, T minus,T prime. S of T, S of T prime. Now I'm missing the indices. So here I had replaced the presynaptic spike train. So this is SJ of T, here I had replaced the postsynaptic spike train and here I integrated from zero to T and I here integrate also from zero to T and this goes over DT, DT prime. Now let's look at the timing difference.

Notes

Summary

1m 45s

6.2 Independent Poisson approach (more mathematical)



$$S_j(t) = \sum_f \delta(t - t_j^f)$$

Consider the weight change in an interval T; rewrite

$$\Delta w_{ij} = \sum_{f,f'} W(\underbrace{t_j^f - t_i^{f'}}_{s}) = \sum_{f'} \int_0^T W(t - t_i^{f'})\, S_j(t)\, dt = \int_0^T \int_0^T W(\underbrace{t - t'}_{s})\, S_j(t)\, S_i(t')\, dt\, dt'$$

$$t - t' = s$$
$$t = s + t'$$

$$\frac{1}{T}\Delta w_{ij} = \frac{1}{T} \int_0^T \int_{-\infty}^{+\infty} W(s)\, S_i(t')\, S_j(t'+s)\, ds\, dt'$$

$$\frac{1}{T}\Delta w_{ij} = \frac{1}{T} \int_0^T \int_{-\infty}^{\infty} W(s) S_i(t) S_j(t+s)\, ds\, dt$$

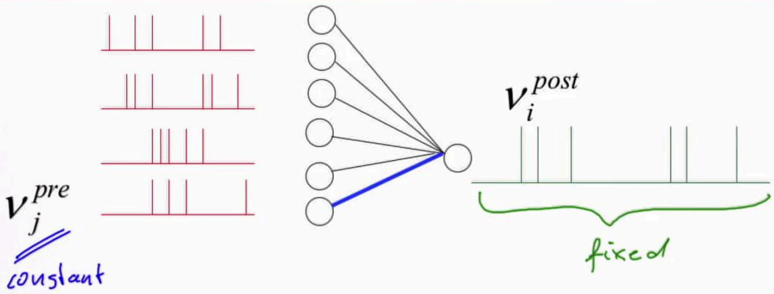
So this is a small S time difference between pre and postsynaptic spike. Let's introduce this here. So I have a weight change which I can now write taking this formula here, FW of S and I will keep this term here. I have SI of T prime and if I have T Minus, T prime equal S, then I can write TS, S plus, T prime, which means this SJ is S and the integral now goes over S. So, I've here the integral from zero to T but I assume that T is very large so it doesn't really matter whether for one of the integrals I put the bounce to infinity. So I will take the integral over S from minus infinity to plus infinity , and then I keep the integral over T,T prime and this is my change W from zero to T. Final steps I renamed T prime as T and I write a one over T in front here. Why is this interesting to have a line over T? Because I have here still the integral over my interval. If I am interested in the mean change per unit of time, I should divide by one over T. And this gives my final formula the weight change, delta WIJ over capital T is the integral over the learning window. And here I have two different spike trains.

Notes

Summary

4m 04s

6.2 Independent Poisson approach (more mathematical)



Assume presynaptic spikes are generated by Poisson process with rate ν_j^{pre} (fixed) ✓

Assume postsynaptic spikes are generated by Poisson process with rate ν_i^{post} (fixed, independent of presynaptic spikes!!!!) ✓

What is the expected change of weights in a time T? $\frac{\langle \Delta w_{ij} \rangle}{T}$

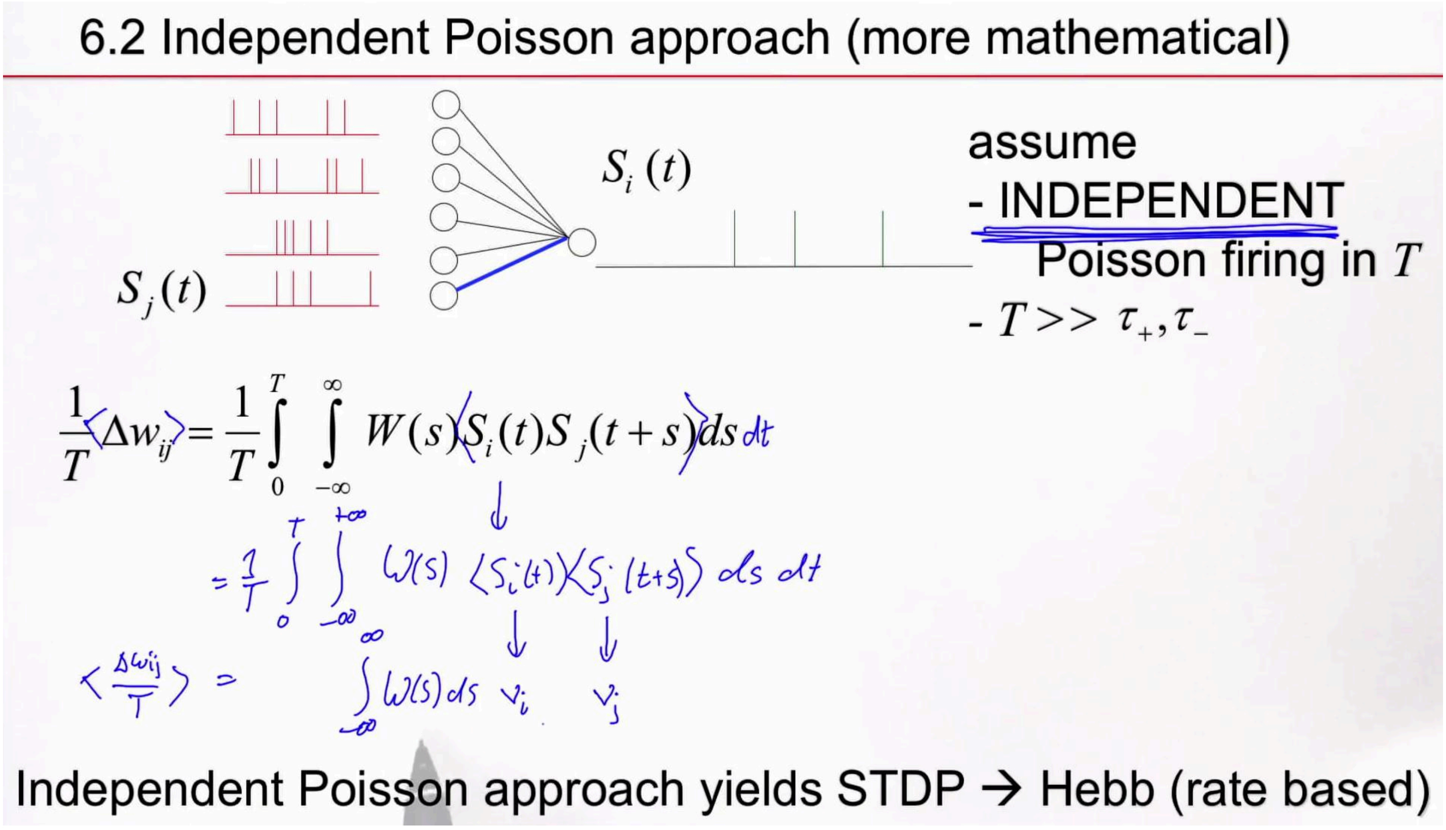
$(T >> \tau_+, \tau_-)$

So far this was a exercise in rewriting. Now, we come to the real Poisson model. What are our assumptions? We assume that the presynaptic spikes are generated by a Poisson process. So spikes are independently generated with a certain rate, the presynaptic rate new J. Similarly, I assume that the postsynaptic spikes are generated independently and with rate new post. And this weight similarly, the presynaptic rate is constant in time. So the presynaptic rate are generated by a Poisson process, that's my assumption. The posting of the spikes are generated by a Poisson impulses with also fixed rate. And this fixed rate is independent of presynaptic spikes. Now this is a strange assumption and we'll be using it later, but for the moment just keep this. And then I ask, what's the expected change of weights? Delta WIJ over a time window T in expected means I take the expectation signs.

Notes

Summary

6m 08s

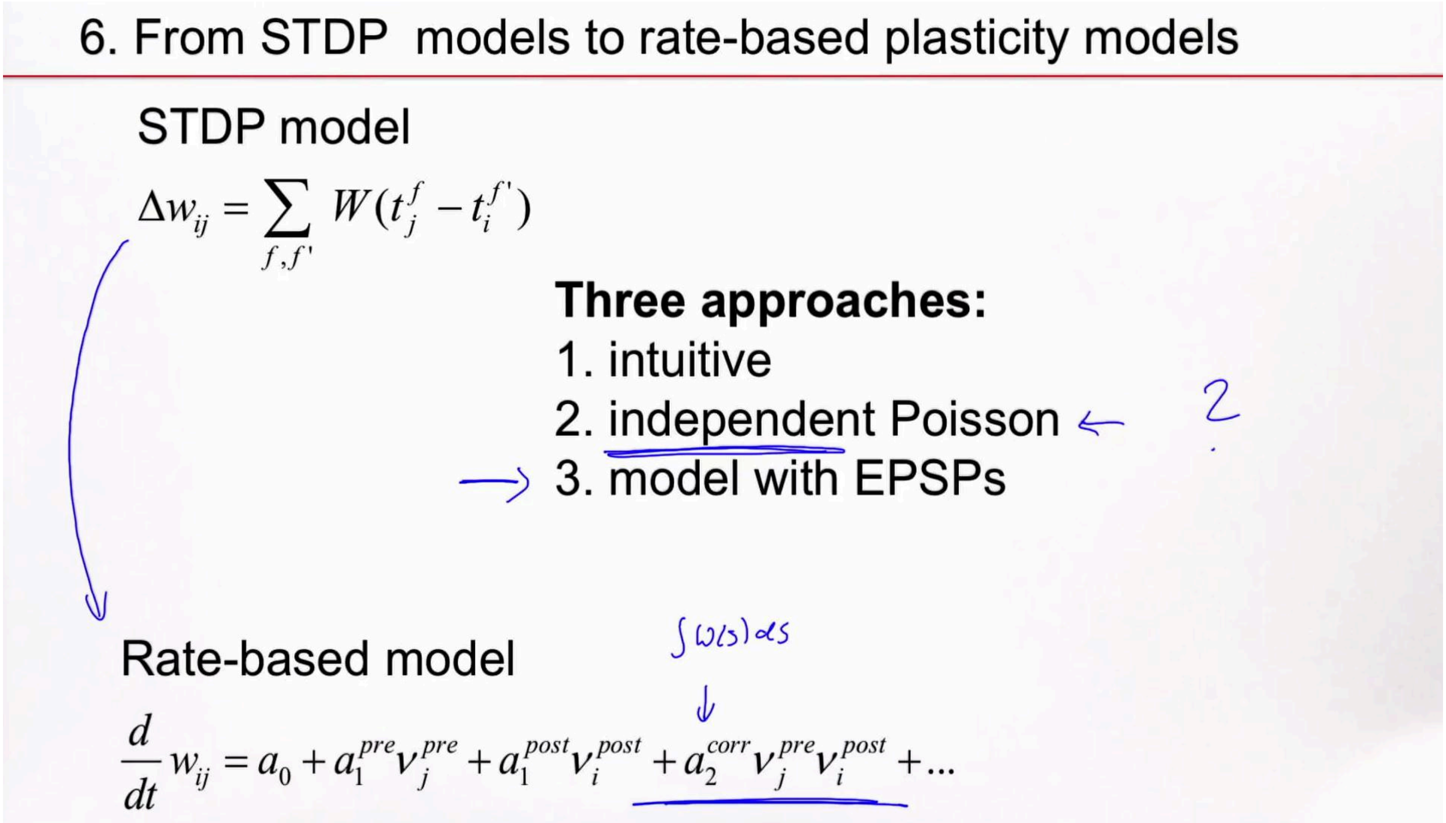


So that's the program. Now we had on the previous slide, the result that the weight change is this. So we now assume that all the findings are independent Poisson process, and we take the averages. So let's look at the averaging. This is the averages. The averages go over the whole interval, but this is a fixed determinative window. So the real averaging is just over the spike trains. And now you see that what matters are the correlations between presynaptic spikes and postsynaptic spikes. However, I have assumed that presynaptic spikes and postsynaptic spikes are independent. So this is S I of T. This is SJ of T plus S. T, S, T, T, W of S minus infinity to plus infinity zeroed T. One over T. Moreover I assumed that these are fixed rates, constant rate. So this is the rate, new I, the posting that we rate. This is the rate new J, all the time dependence has gone. So I can take the integral directly over W of S, T, S from minus infinity to plus infinity. And then I have another integral from zero to T over something that's constant in time.

Notes

Summary

7m 24s



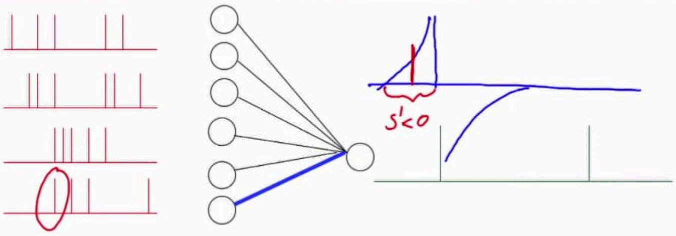
So this just cancels out, which means my weight change, delta WIJ over T expected is presynaptic rate, postsynaptic rate times the integral over the learning window, thus we made the transition from a STDP model to a rate based model and what we found is that under the assumption of independent Poisson processes, we get a term that corresponds to the correlation term with a pre-factor A2 core, which is equal to the integral over the learning window. Now, there's something very strange here because we started off with a Hebbian idea saying that well the presynaptic spike that contribute to firing the postsynaptic ones are the important ones to change the weight. So, there's these causal interaction, if a presynaptic spike arrives, it actually causes a change in the neuron. So, it's very strange to assume that I have independent Poisson processes and therefore in the next part I will look at a model where presynaptic spikes cause EPSPs, postsynaptic potentials. So that there is no longer an assumption of independence between pre and postsynaptic firing.

Notes

Summary

9m 19s

6.3 Short Detour: Writing styles for STDP models



$$\Delta w_{ij} = \sum_{f,f'} W(t_j^f - t_i^{f'})$$

$$\frac{1}{T}\Delta w_{ij} = \frac{1}{T}\int_0^T \int_{-\infty}^{\infty} W(+s') S_i(t) S_j(t+s') ds$$

s' =

$$\frac{d}{dt} w_{ij} = S_i(t)\int_0^{\infty} W_+(-s) S_j(t-s) ds + S_j(t)\int_0^{\infty} W_-(s) S_i(t-s) ds$$

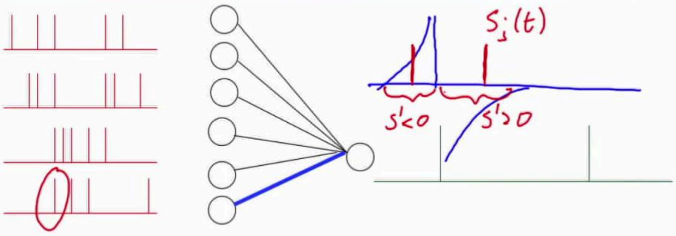
Before we start with the model that includes EPSPs, let's change our writing styles. I've already explained that we can make the transition from a notation sum of pairs to a notation where I have spike trains, postsynaptic spike train, and presynaptic spike train. I just mentioned that in Hebbian learning, the causal relations between pre and postsynaptic spikes are important. So let's, let's look at this, again, I have my learning window centered at the postsynaptic spike. It has a positive part and it has a negative part. Now, if you assume that one of these presynaptic spikes, this one here arrives at the postsynaptic arrives just before the postsynaptic spike. Then I can think of this postsynaptic spike as the one that triggers the change if pre before post. Then the change Delta W is in this kind of formula is triggered at the moment of the postsynaptic spike. And so this, I can pull this out and say the changes happen at the moment of the postsynaptic spike under the condition that the presynaptic spike is before the postsynaptic one. But it's before the postsynaptic one, if S prime is smaller than zero, so this part here is S prime smaller than zero or I have a notation that say T minus S there as prime is minus S.

Notes

Summary

10m 36s

6.3 Short Detour: Writing styles for STDP models



$$\Delta w_{ij} = \sum_{f,f'} W(t_j^f - t_i^{f'})$$

$$\frac{1}{T}\Delta w_{ij} = \frac{1}{T}\int_0^T \int_{-\infty}^{\infty} W(+s')S_i(t)S_j(t+s')ds$$

$s' = -s$

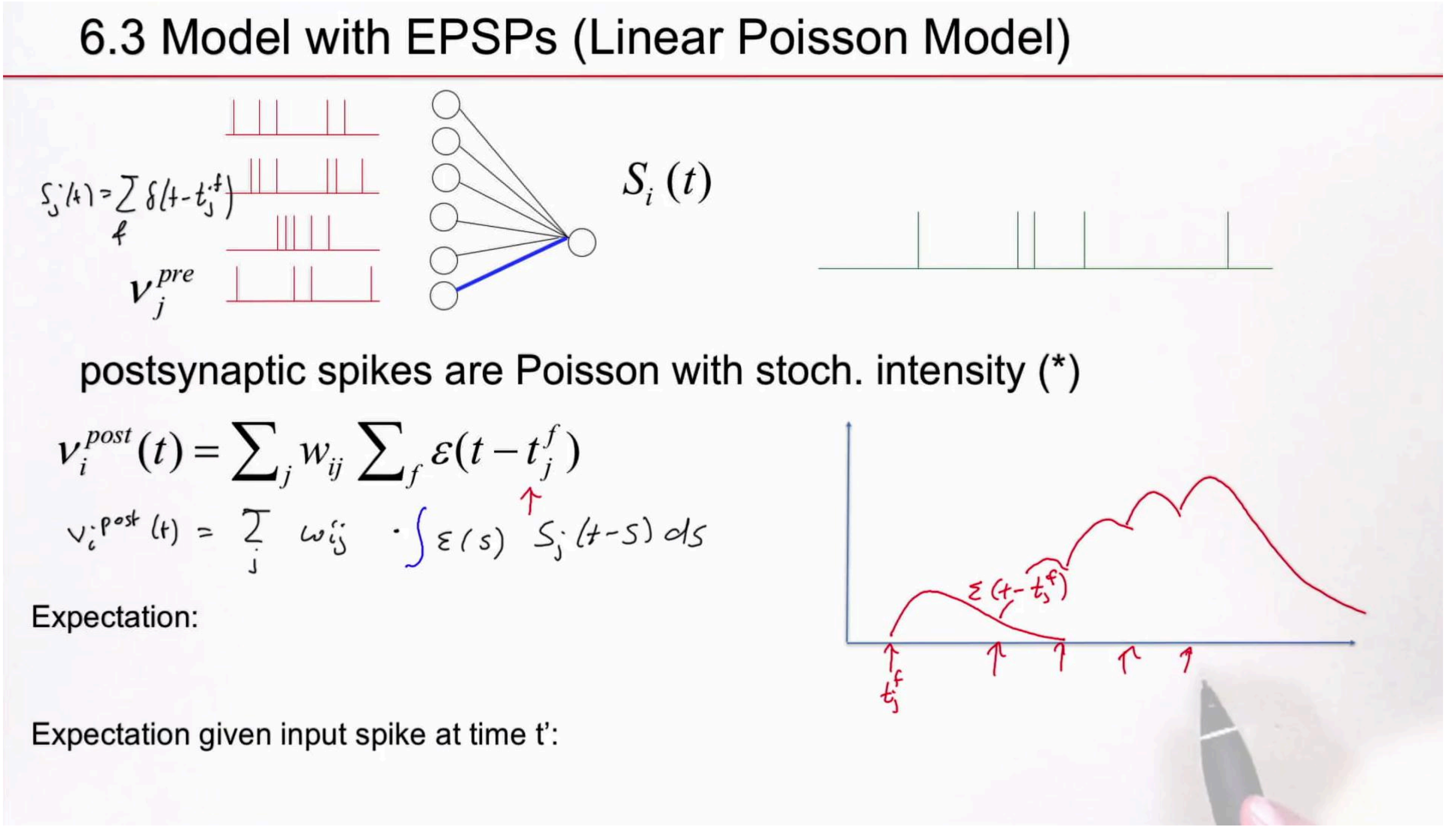
$$\frac{d}{dt}w_{ij} = S_i(t)\int_0^{\infty} \underbrace{W_+(-s)S_j(t-s)ds}_{\text{pre - before - post}} + S_j(t)\int_0^{\infty} W_-(s)S_i(t-s)ds$$

So, this part here has all the conditions pre before post. Similarly I might consider a situation where I have a presynaptic spike here and I have a postsynaptic spike just before. Then I can consider that this spike here, this presynaptic spike causes the weight change, so I pull out this term for the condition that S prime is larger than zero. Now the posting of the spike occurred first and then I have a term that says the update happens at the moment of presynaptic spike. So why did I discuss this? Because you see that they are terms that really consider the situations pre before post. And this is the situation, where presynaptic spike could have participated in triggering the postsynaptic spike.

Notes

Summary

12m 23s

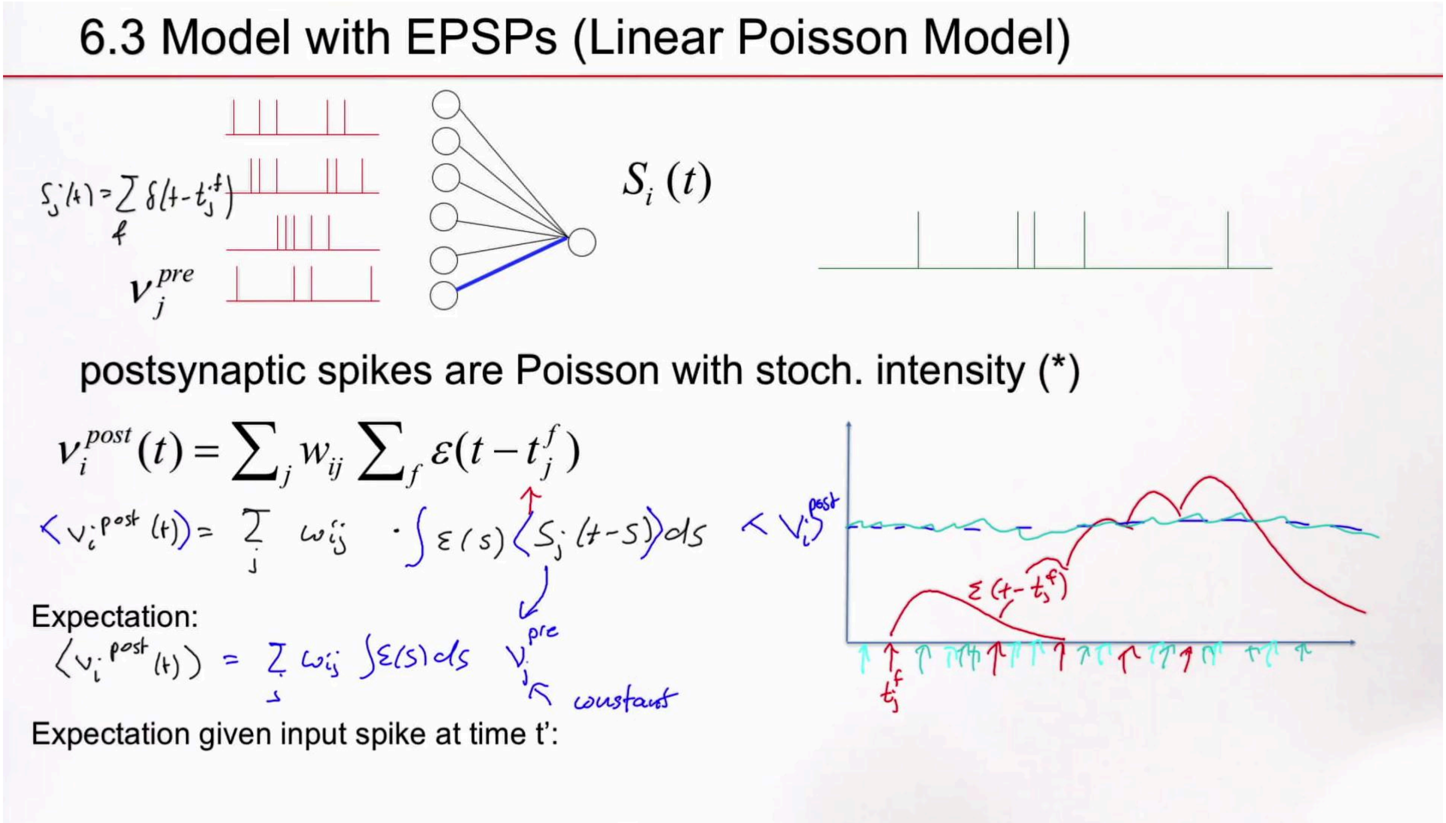


And this is what we are going to formalize next. So, here I have my presynaptic spikes and I have postsynaptic spikes. The difference now is that I say, each presynaptic spike causes a post, causes an ESP. So let me explain this. I have my presynaptic spikes at time TJF for example, there is a presynaptic spike here it causes an EPSP, which I call epsilon of T minus TJF. This was the first spike. Then I will have other spikes and basically these formulas says, I just add up the contributions of the different presynaptic spikes. Now the firing rate of the postsynaptic neuron is now time-dependent. The firing rate goes up, goes down and goes up again, goes up, goes up and so forth. To rewrite the postsynaptic firing rate, I used the presynaptic spike train. I write epsilon of S, S, J of T minus SDS sum over J, WIJ. Where the presynaptic spike train is the sum, over all presynaptic spikes of delta T minus TJF. So if you block the definition of the spike train back in, you get to this formula that I had before, you will be interested in the expectation. So this new post, this is this firing rate, this time-dependent firing rate for this one, very specific realization of the input spikes.

Notes

Summary

13m 32s

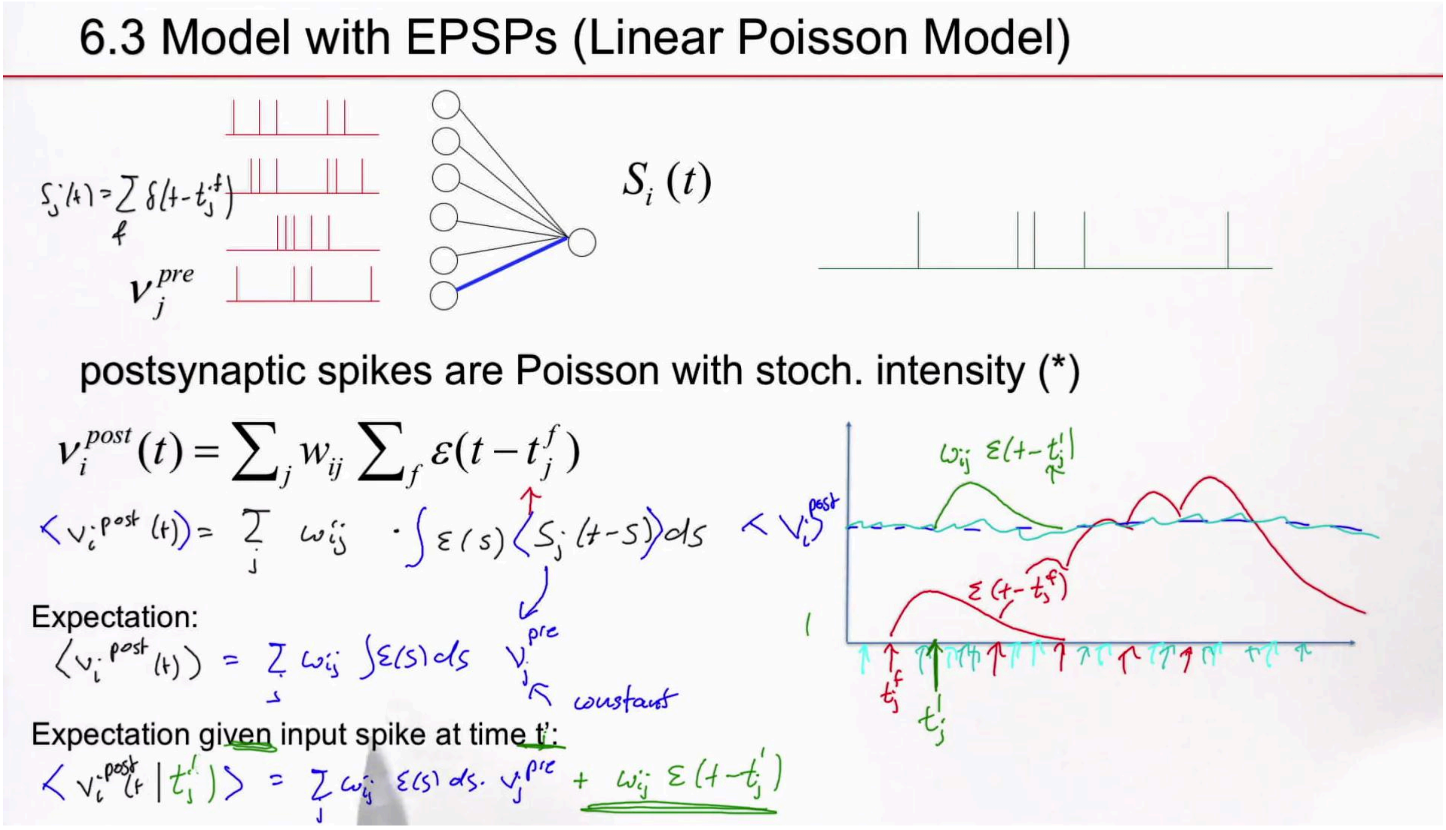


But now we can ask, well, what should we expect? What happens on average? So now I look at the expected postsynaptic rate Well where does the expectations sign go? Let's do this, here I have my expectation. This is a fixed weight. This is a fixed time course of the EPSP; the only stochastic part is the realization of this spike train. That's one of the realization, there are many others, now the expectation of this is justifying rate, this will be the presynaptic firing rate. And I still assume that the presynaptic firing rate is constant. Now since its constant, I can take out the integral and I have sum over J, WIJ integral epsilon of S. DS. So the expected rate is constant and proportional to the presynaptic firing rate. So that means in expectation, this is my postsynaptic rate, so in expectation, I don't see anything of these causal effects that each presynaptic spike causes a postsynaptic spike. So over many different realizations, sometimes the spike train is like this. Sometimes the spike train is like this, and have all these different realizations. This is what I expect to see. Now, where is the causality? Let's change the perspective.

Notes

Summary

15m 24s

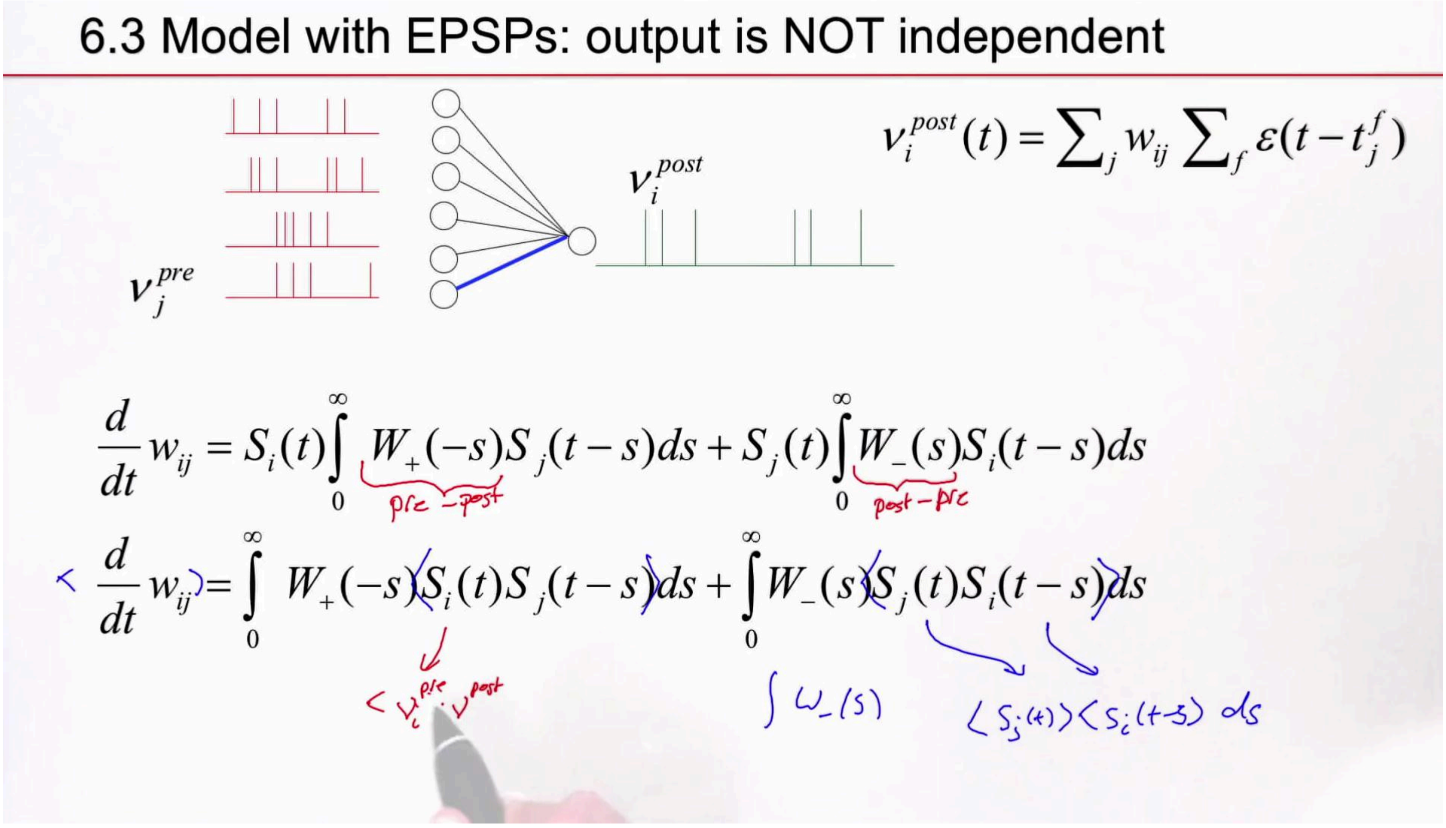


Say, I now, so on average the spikes come wherever they come, but I know for sure that there was one spike here, so there was one presynaptic spike here. Then on average, I will have this trace here, but since I know that have this extra spike, that actually means I know for sure that this extra spike causes an EPSP. And this EPSP changes the firing rate by an amount WIJ epsilon of T minus T prime. Where this is the presynaptic spike of my neuron J. So, if I look at the expectation, new I post of T given the specific spike that I know for sure that has arrived at time TJF, look at the expansion of this. Then I see my normal trace, which I can copy from above WIJ epsilon of S, TS times new J pre and then I have these extra bump. So, I write plus WIJ, epsilon of T minus TJ prime, TJ prime, so I should change this here. What I meant is this specific spike at time, TJ prime, so we have to distinguish between the normal expectation where I just might say I don't know what happened, so spikes arrive whenever they want and the expectation given an input spike at time TJ prime. And this extra spike makes a causal contribution to the postsynaptic firing rate.

Notes

Summary

17m 17s



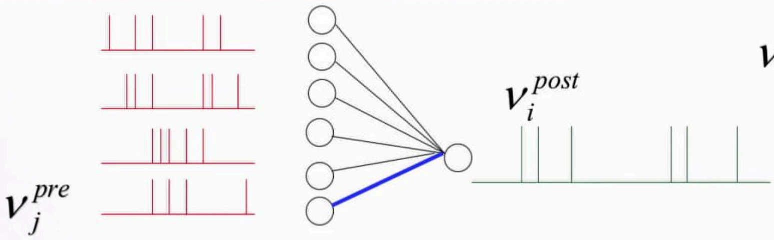
Since this is the same term, I can also write this in the form. This is the normal expectation plus the extra term WIJ epsilon of T minus TJ prime. And this is now what you're going to use on the next slide. So two slides ago I've already introduced the different notation styles. This notation style says for minus S this is the condition pre before post. This is where the causal effects can take place. This is the condition of post before pre. If the postsynaptic spike comes before the presynaptic ones, then there cannot be any causal relation anyway. Now we are interested in the expected weight change, so let's take the expectation signs. This is the fixed positive part of the learning window. The expectation goes to this. This is the fixed negative part of the learning window. The expectation goes to the spike trains. Now since this is post before pre there cannot be any additional causal information so this part is similar to what we had seen before is just SJ of T averaged times SI of T minus S, averaged. The interesting part happens over here, I have a part that just says well there's a contribution of new I pre times new post each time expectation.

Notes

Summary

19m 11s

6.3 Model with EPSPs: output is NOT independent



$$\nu_i^{post}(t) = \sum_j w_{ij} \sum_f \varepsilon(t - t_j^f)$$

$$\frac{d}{dt} w_{ij} = S_i(t)\int_0^\infty \underbrace{W_+(-s)S_j(t-s)}_{pre-post}ds + S_j(t)\int_0^\infty \underbrace{W_-(s)S_i(t-s)}_{post-pre}ds$$

$$\left\langle \frac{d}{dt} w_{ij}\right\rangle = \int_0^\infty W_+(-s)\langle S_i(t)S_j(t-s)\rangle ds + \int_0^\infty W_-(s)\langle S_j(t)S_i(t-s)\rangle ds$$

$$= \int_0^\infty W_+(-s)\,\langle \nu_j^{pre}\rangle\langle \nu_i^{post}\rangle + \int w_{ij} W(-s)\varepsilon(s)ds \cdot \nu_j^{pre} + \int_0^\infty W_-(s) \quad \langle S_j(t)\rangle\langle S_i(t-s)\rangle ds$$

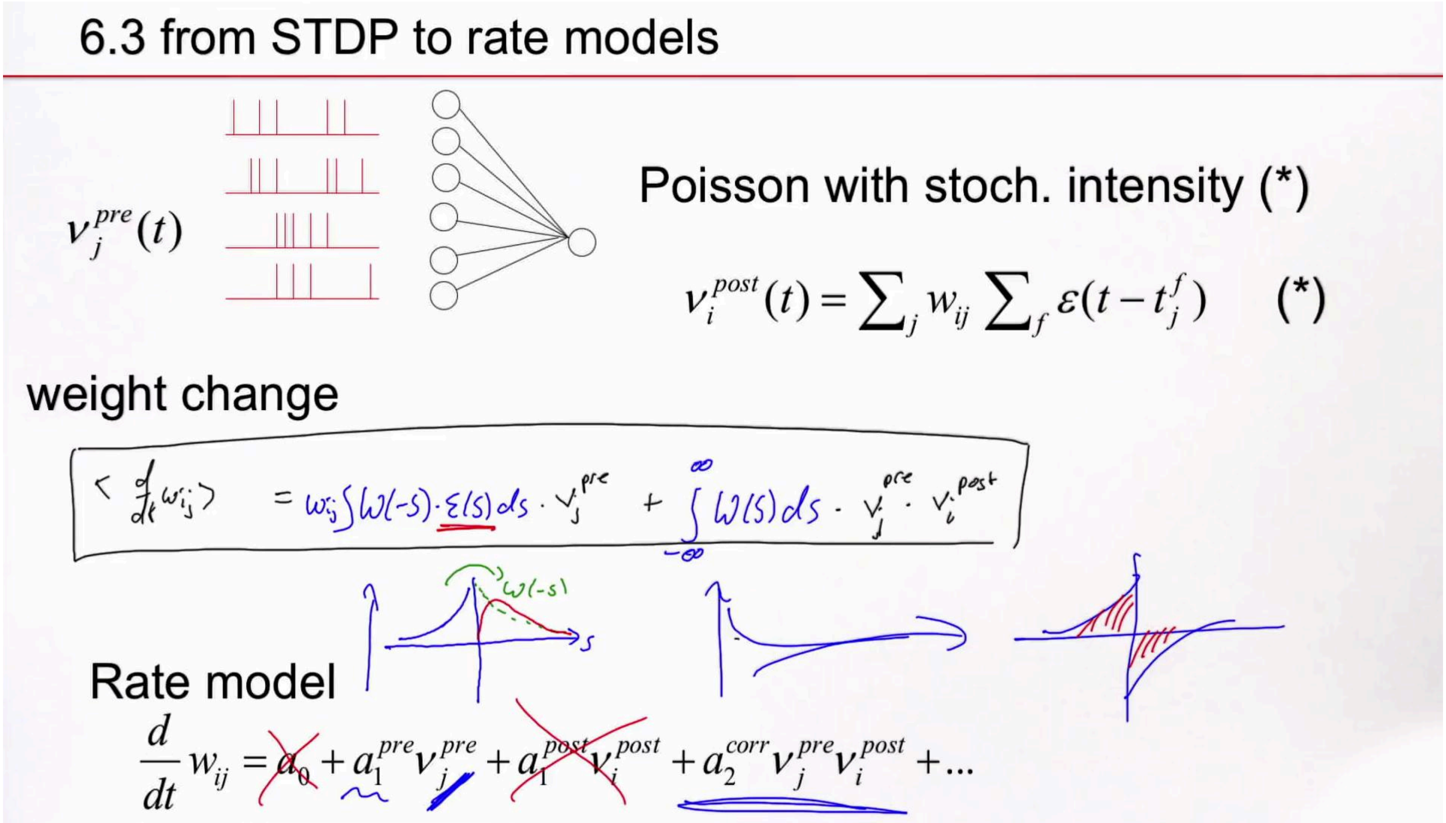
$$\left\langle \frac{d}{dt} w_{ij}\right\rangle = w_{ij}\int W(-s)\varepsilon(s)\cdot \nu_j^{pre} + \int_{-\infty}^{\infty} W(s)\,ds \cdot \nu_j^{pre} \cdot \nu_i^{post}$$

This is where I split it up the same way I would do it here. I would say this is new J pre, new I post and so I see that the indices are not correct. Post is IJ is pre. But I had here the extra green part that says, but look for any extra spike, I have this term WIJ epsilon after firing times. And I have to treat this extra term together with the learning window. So I will have here the normal term, if this is constant, this is constant, I just close the learning window. You have minus S from zero to infinity times post and then I have the extra part that says, okay W of minus S, I have epsilon of S, DS but these term was for one single spike, one single given presynaptic spike. I will have more spikes on the presynaptic side if the firing rate is higher. So I have to multiply here in addition with a new J pre. So let me sum up from the green part, I have a term WIJ integral, W of minus S, epsilon of S times new J pre. And then from the other parts I have the integral of the positive part of the learning window with minus S. I have the integral over negative part of the learning window with plus S. So this goes back to a full integral from minus infinity to plus infinity, double of S, DS times new J, times new post. So this is my total expected weight change. DDT, WIJ expected is this term here and this term here. Now this looks like a rate model, so let's compare this.

Notes

Summary

21m 02s

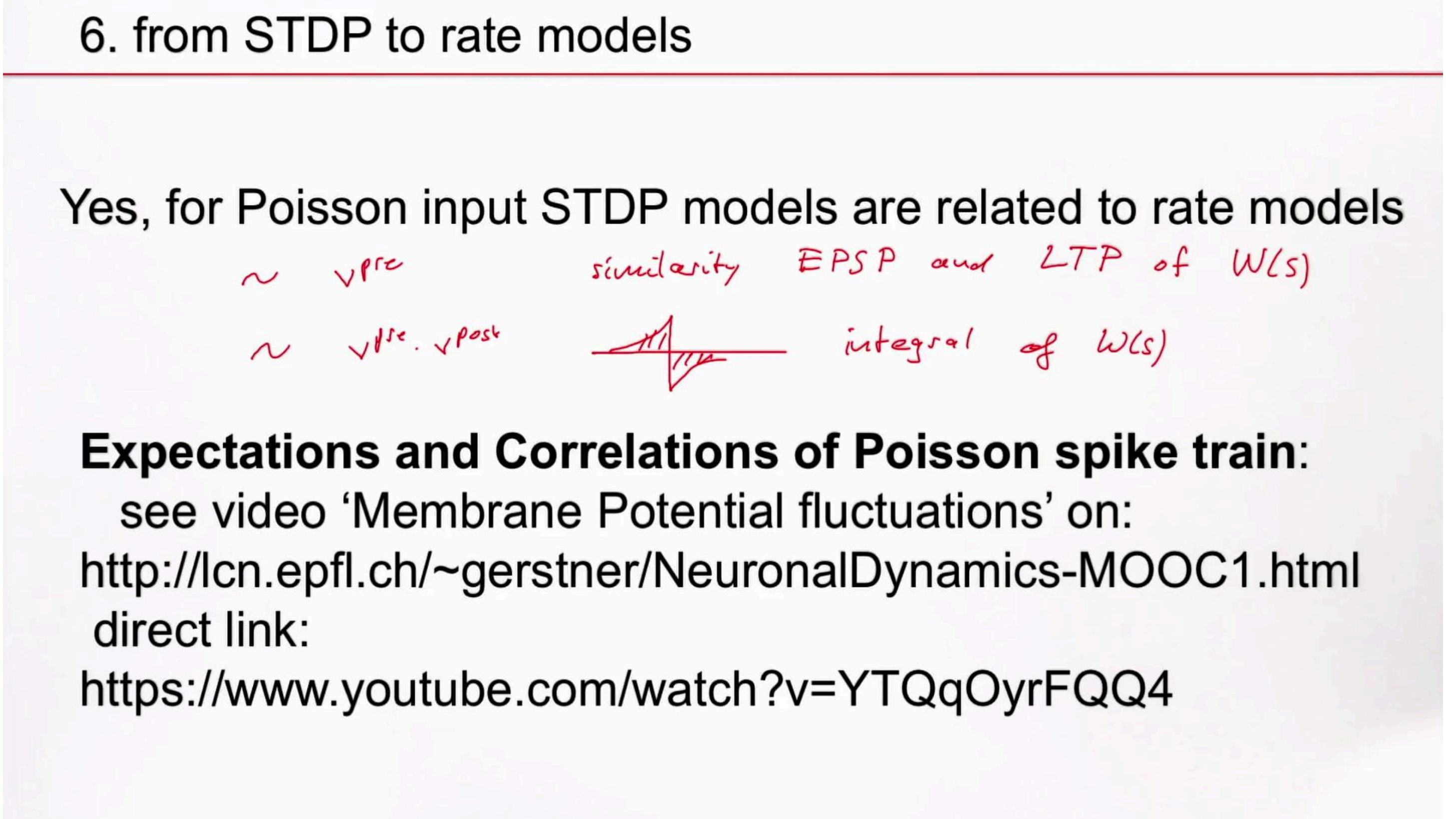


My general rate model I have terms of this form. What we found is that this factor is the integral over the learning window WF, W of SES, so this term is A2 core. And then we found an extra term that was proportional to this spike arrival rate, which means that this A 1 pre is my extra term, which is W of minus S, epsilon of S, DS. Then we did not find this term and we did not find this term. So I have two terms. DDT, WIJ in expectation is this terms times new J pre less this term UJ pre-new I post. So this causal effect is proportional to this weight here. Let me interpret those terms. So these term here A2 core is the integral over the learning window. Positive parts minus negative parts. This term here looks like this, you plot the learning window and what matters is the minus side. So this is S, this is negative side. And then you compare this with epsilon. So if I take this and I brought this back as minus S, then this is like flipping the learning window around. So flip this around and you have this contribution. So this is W of minus S and then we have also the EPSP which looks like this. So what matters here for this term is the similarity between the EPSP and the flipped over learning window, the positive part of the learning window that flipped over to the other side.

Notes

Summary

23m 23s



So let me summarize this. Yes, for Poisson input, STDP models are indeed related to rate models. And I find two terms. I find a term that is proportional to the presynaptic rate. And for this term what matters is the similarity between the EPSP and the LTP part of the learning window. And then I have another term which is new pre times new post. And what matters for this term is the integral of the learning window. Now, for the mathematical parts, I used expectations, Delta notation for the spike trains, correlations of Poisson spike trains. And there's additional information on this on a separate video.

Notes

Summary

25m 54s