# Preamble

Figures and tables in this document are often accompanied by a link to the Jupyter Notebook that generated them. These notebooks contain more information than can be conveyed in just the caption. These links look like, e.g: `2021-12-06__local_HH_dV_shape`. To visit these links, in the digital version of this document, just click them. In the paper version, convert them to an URL as follows: `https://tfiers.github.io/phd/nb/2021-12-06__local_HH_dV_shape.html`.

*Chapter 1*

---

# Introduction

---

- Problem statement

- Opportunity of voltage imaging, core idea of inverting the causal 'presynaptic spike $\rightarrow$ postsynaptic PSP' relation
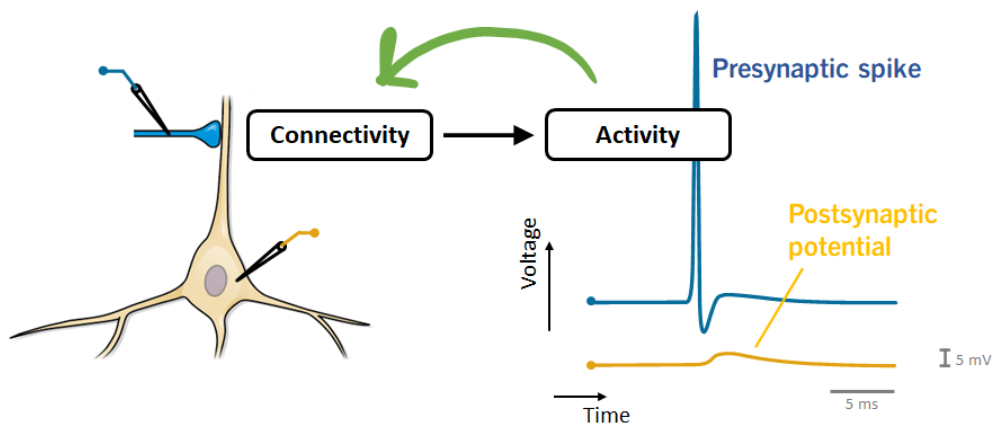
- Overview of thesis



Figure 1: **The causal link between neural connectivity and activity.** On the left, a cartoon of a synaptic connection. The axon of presynaptic neuron $M$ (in blue) impinges on postsynaptic neuron $N$ (in brown). The electrode icons indicate that their membrane voltages are recorded (shown on the right). A succesful spike in neuron $M$ will elicit a small but precisely-timed voltage bump in neuron $N$ (the postsynaptic potential, PSP). There is thus a causal relationship between 1) the existence of a connection $M \rightarrow N$ and 2) both neurons' membrane voltages. This causal relationship (black arrow) is exploited to perform network inference from voltage recordings (green arrow).
*Drawings adapted from Purves et al.'s "Neuroscience" textbook, 6th edition, 2018.*

## 1.1   Connectomics

- Motivation

- Ground-truth connectomics: tracing of electron microscopy and fluoresent injection imaging

- Necessity of connection *inference*

- Mention 'invasive' connection testing (stimulate one cell, record possible neighbours)

- Limitations of 'connectomics', and of inferred vs 'actual' connectomics.
  Terminology, e.g. 'functional connectomics'

## 1.2   Voltage imaging

- Technologies (from dyes to GEVIs)

- Specs: cell yield, tissue depth, recording duration, SNR, species

- ..and growth of these over time, and comparison with calcium imaging.
  To extrapolate how these might advance in the future

- Comparison with other recording techniques: ephys, calcium imaging, (and briefly mention coarser methods)

## 1.3   Network inference

- Working with events/spikes only, versus working with continuous signals; or a hybrid as here.

- Overview of the spikes-only methods

- The connectomics competition, and the findings about the best methods

- Mention other application domains, like gene regulatory networks

- Conclusion: connection inference by spikes is poor (Ila Fiete etc), and Vm imaging offers unique advantage (causality)

# Simulation details

In this chapter, we decribe our experimental setup: the neuron model we simulate, its inputs, and how we simulate voltage imaging.

## 2.1 The AdEx neuron

We choose to simulate the 'AdEx' point neuron model, or the 'adaptive exponential integrate-and-fire' neuron[BG05], with conductance-based synaptic currents. The AdEx neuron is a leaky-integrate-and-fire (LIF) neuron model, with two additions. First, the full upstroke of each spike is simulated, as an exponential runoff. Second, an extra dynamic variable is added: the adaptation current. This current allows the simulation of many non-linear effects of real neurons, like spike-rate adaptation and post-inhibitory rebound.

The AdEx model consists of two differential equations ((1) and (2)), and a discontinuous update after a spike is generated (3). One equation simulates the membrane voltage $V$, and one the adaptation current $w$:

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = -g_L(V - E_L) + g_L\Delta_T \exp\left(\frac{V - V_T}{\Delta_T}\right) - I_{\text{syn}} - w \quad (1)$$

$$\tau_w\frac{\mathrm{d}w}{\mathrm{d}t} = a(V - E_L) - w \quad (2)$$

Parameters are described in table 1.

The first term of equation (1) is the restorative force pulling the voltage back to the resting (or leak) potential $E_L$. The second, exponential term is what generates the spike upstrokes. It is practically zero over most of the sub-threshold regime, and only becomes large (and then very large) near the firing threshold. (This firing threshold is characterised further on).

$I_{\text{syn}}$ is the synaptic current, explained in section 2.4. We use the sign convention of inter alia Dayan & Abbott[1] where membrane currents are defined as positive when positive charges flow *out* of the cell. I.e.



Figure 2: A linear-plus-exponential model (red) fit to data from a cortical pyramidal neuron (black), from [Bad+08]

---

[1] [DA01], ch. 5.3, p. 162

a positive $I_{\text{syn}}$ decreases the membrane voltage (itself defined as the electric potential inside minus outside the cell).

The adaptation current $w$ decays exponentially to zero on its own ($-w$ in equation (2)), and is influenced by voltage deviations from equilibrium: for $a > 0$, $w$ acts on $V$ in the opposite direction of the deviation, and for $a < 0$, $w$ acts in the same direction. Izhikevich calls the former a resonant current, and the latter an amplifying one.[2]

[2] [Izh07], section 5.2.4. Note the different notation used; see the translation in table 2.

| Name | Description | Value |
|------|-------------|-------|
| $V$ | Membrane voltage | (in mV) |
| $w$ | Adaptation current | (in pA) |
| $C$ | Membrane capacitance | 104 pF |
| $g_L$ | Input / leak conductance | 4.3 nS |
| $E_L$ | Resting / leak potential | $-65$ mV |
| $\Delta_T$ | Threshold slope factor | 0.8 mV |
| $V_T$ | Location of minimum of $\frac{\mathrm{d}V}{\mathrm{d}t}$ | $-52$ mV |
| $\tau_w$ | Time constant of adaptation current | 88 ms |
| $a$ | Sensitivity of adaptation current to $V$ | $-0.8$ nS |
| $\theta$ | Spike definition threshold | 40 mV |
| $V_r$ | Reset voltage after spike | $-53$ mV |
| $b$ | Adaptation current bump after spike | 65 pA |

Table 1: Quantities and parameters of the AdEx neuron, equations (1) to (3). Values are from a model fit to a cortical regular spiking (RS) neuron, from [Nau+08]. By defining the location of $\frac{\mathrm{d}V}{\mathrm{d}t}$'s minimum, $V_T$ also co-determines the location of the firing threshold.

In this chapter, we will analyse $\frac{\mathrm{d}V}{\mathrm{d}t}$ as a function of $V$, i.e. analyse it as a dynamical system: will the voltage increase or decrease at the current voltage? For conciseness, we call this function $F(V)$. I.e. $F(V) = \frac{\mathrm{d}V}{\mathrm{d}t} =$ the right-hand-side of equation (1) here, scaled by $1/C$. We'll mostly analyse $F$ in the absence of synaptic and adaptation currents, i.e. for $I_{\text{syn}}$ and $w$ both zero. Figure 2 shows the $F(V)$ curve for an AdEx neuron fit to a real neuron. Figure 5 compares the $F(V)$ curve of an AdEx neuron with that of another two-dimensional neuron model, the Izhikevich neuron.

In addition to the two differential equations, the AdEx model also consists of an instantaneous reset condition. When the membrane voltage $V$ reaches a certain threshold $\theta$, a spike is recorded, $V$ is reset, and $w$ is increased:

$$\text{if } V > \theta \text{ then: } V \leftarrow V_r \tag{3}$$
$$w \leftarrow w + b$$

This bump of the adaptation current is what provides the spike rate adaptation: the more spikes the neuron has recently fired, the higher the adaptation current $w$ and the more it drives down the voltage, away from the firing threshold.

For a full description of the different firing patterns that the AdEx neuron can exhibit, for different parameter values (i.e, a characterisa-

tion of its bifurcations), see [Nau+08], "Firing Patterns in the Adaptive Exponential Integrate-and-Fire Model". For an even fuller description, including a derivation of two-dimensional neuron models from higher-dimensional, Hodgkin-Huxley-like models, see Eugene Izhikevich's book, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting* [Izh07].

In terms of parameter choice, we choose to simulate a cortical regular spiking (RS) neuron, i.e. a 'standard' excitatory neuron that does not display e.g. bursting or fast-spiking behaviour. We take our parameter values from [Nau+08][3], where they fitted an AdEx model to recordings from real cortical RS neurons injected with different step currents. Parameter values are listed in table 1.

[3] Section 6, and Table 1, row 3

Examples of the signals $V$ and $w$ that this model generates are given later, in figure 10 and figure 13.

## Analysis

Where are the fixed points of the dynamical system $\frac{dV}{dt} = F(V)$? I.e, where is $F = 0$? Like other neuron models, there are two fixed points: a stable one at the leak potential $E_L$, and an unstable one at the instantaneous firing threshold (which we'll call $E_T$).

We see in equation (1) (for $I_{syn} = 0$ and $w = 0$) that the leak potential $E_L$ is indeed a fixed point – or rather lies very close to one: the exponential term is negligibly small at $V = E_L$. The second fixed point has no straightforward expression. The exact solutions for $F$'s roots need the so called Lambert $W$ or 'product logarithm' functions $W_0$ and $W_{-1}$ (see figure 3). The roots are found at:

$$V = E_L - \Delta_T W_k \left( -\exp\left( \frac{E_L - V_T}{\Delta_T} \right) \right) \tag{4}$$

Figure 3: The Lambert $W$ functions for real numbers.

..where $k = 0$ gives the resting potential, and $k = -1$ the instantaneous firing threshold.[4] For our cortical RS neuron, this gives us an instantaneous threshold of $E_T = -49.6$ mV.

Also of interest – especially when comparing with the Izhikevich neuron later – is the slope of AdEx's $F(V)$, i.e. its derivative with respect to $V$:

$$\frac{dF}{dV} = \frac{d}{dV} \left( -g_L(V - E_L) + g_L \Delta_T \exp\left( \frac{V - V_T}{\Delta_T} \right) \right)$$

$$= -g_L + g_L \exp\left( \frac{V - V_T}{\Delta_T} \right) \tag{5}$$

This derivative is zero at $V = V_T$. I.e, unlike what is suggested by the '$T$' subscript, and the name 'effective threshold potential' given to it in the AdEx literature[BG05; Nau+08], $V_T$ is not the instantaneous threshold potential (it is not a zero of $F$), but rather the minimum of $F$ (it is the zero of $\frac{dF}{dV}$).

From equation (5), we also calculate the slope of $F$ at its two roots. At $V = E_L$, the slope (also known as the leak or input conductance

[4] Here too we see that the true resting potential is almost identical to $E_L$: $W_0$ passes through zero, and its argument is $\approx 0$, because the exponential's argument is negative. For $W_{-1}$ however, figure 3 shows that the second term is not negligible around 0.

here) is $-g_L$, plus a negligibly small exponential term. The negative sign shows that this is a stable fixed point: in a linearization of $F$ around this point, at voltages below the resting potential, $F$ (i.e. $\frac{dV}{dt}$) is positive and thus $V$ will increase. At voltages above $E_L$, $F$ is negative and $V$ will decrease. Small deviations on either side of the resting potential will thus decay back to this resting potential.

At the firing threshold $V = E_T$ (i.e. the second solution to equation (4)), the slope is:

$$ g_L \left( \exp \left( \frac{E_T - V_T}{\Delta_T} \right) - 1 \right) \tag{6} $$

..which is positive (as $E_T > V_T$), indicating that this is an unstable fixed point: a small deviation of the voltage above $E_T$ will blow up to infinity (i.e, a spike is generated).

## 2.2 Alternative neuron models

Why did we choose the AdEx model to simulate neuron voltages? In short, because it strikes a good balance between realism and complexity. We briefly consider here two alternative neuron models: the simple leaky-integrate-and-fire (LIF) neuron, and the more complex Hodgkin-Huxley (HH) neuron. In the next section, we go into more depth on a third alternative, the very similar Izhikevich neuron.

A simpler model than AdEx would be the well-known LIF neuron:

$$ C\frac{dV}{dt} = -g_L(V - E_L) - I_{\text{syn}} $$

$$ \text{if } V > \theta, \text{ then: } V \leftarrow V_r $$

As is apparent from comparing this with equations (1) and (3), the AdEx model is an extension of the LIF model. The LIF neuron lacks a simulation of the upstroke of spikes (the exponential term in equation (1)), and the slower time-scale adaptation current (equation (2)), which allows the simulation of many qualitatively different real neuron types. It is especially this first addition, the full upstroke simulation, that seems relevant in generating realistic voltage traces.

Would this thesis have been very different had we used LIF neurons instead? Probably not, though it might depend on the mean voltage level of the simulated neuron: if it is well below the firing threshold, both LIF and AdEx are linear (the exponential term is negligible), and they behave quasi identically. When a spike is generated in the AdEx model, the exponential feedback makes the upstroke very fast, and thus not many timesteps in the simulation are spent on it, versus the linear regime.

On the other hand, when the neuron would continuously teeter just below its firing threshold, the LIF and AdEx models do not behave

similarly. LIF's $F(V)$ curve is still fully linear, while AdEx's is not, and AdEx will behave more like a real neuron – see figure 2.

Another well-known alternative neuron model is the class of Hodgin-Huxley (HH)-like neurons. These models simulate the full trajectory of a spike: both its upstroke and its downstroke. Unfortunately they also have many free parameters. They also take a bit longer to simulate, being higher dimensional (having more differential equations), and containing many more exponential terms, which take the brunt of the time when numerically evaluating a differential expression.

Figure 4: Neuron models as dynamical systems: a comparison of the $F(V)$ curves of the Izhikevich and AdEx neurons. 'Experimental' is an AdEx neuron fit to a cortical pyramidal neuron (using data from [Bad+08]). The Izhikevich neuron's parameters were chosen to match the fixed points and the leak conductance. Arrows indicate whether the voltage will increase or decrease (1D flow field). $\rightarrow \bullet \leftarrow$ is a stable fixed point (resting potential), $\leftarrow$ o $\rightarrow$ is an unstable fixed point (spike threshold). $\dot{V} \equiv \frac{\mathrm{d}V}{\mathrm{d}t}$.

## 2.3 The Izhikevich neuron

Another alternative neuron model is the Izhikevich neuron, which is exceedingly similar to the AdEx neuron. These are the Izhikevich equations, using the same symbols as used before (in equations (1) to (3)):

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = k(V - E_L)(V - E_T) - I_{\mathrm{syn}} - w \tag{7}$$

$$\tau_w \frac{\mathrm{d}w}{\mathrm{d}t} = a(V - E_L) - w \tag{8}$$

$$\text{if } V > \theta, \text{ then:} \tag{9}$$
$$V \leftarrow V_r$$
$$w \leftarrow w + \Delta w$$

We have introduced two new parameters not present in the AdEx equations: the steepness of the parabola, $k$; and $E_T$, the instantaneous firing threshold (the firing threshold in the absence of any synaptic or adaptation currents).

The only difference with AdEx is in equation (7), where the $F(V)$ curve is not made by a linear plus exponential term, as in AdEx; but rather by a quadratic (a parabola). Its two zeros (the fixed points) are readily apparent, as $E_L$ and $E_T$.

### Correspondences with AdEx

In Izhikevich's book[5], different names are used for the same quanti-

5 [Izh07], section 5.2.4, equations 5.7 & 5.8

ties:

$$C\frac{dv}{dt} = k(v - v_r)(v - v_t) - u + I \tag{10}$$

$$\frac{du}{dt} = a(b(v - v_r) - u) \tag{11}$$

$$\text{if } v > v_{\text{peak}}, \text{ then:} \tag{12}$$
$$v \leftarrow c$$
$$u \leftarrow u + d$$

Table 2 compares both notation conventions.

| AdEx | Izh | Description | Units |
|------|-----|-------------|-------|
| $V$ | $v$ | Membrane voltage | V |
| $w$ | $u$ | Adaptation current | A |
| $\tau_w$ | $1/a$ | Time constant of adaptation current | s |
| $E_L$ | $v_r$ | Resting / leak potential | V |
| $V_r$ | $c$ | Reset voltage after spike | V |
| $a$ | $b$ | Sensitivity of adapt. current to $V$ | S |
| $b$ | $d$ | Adaptation current bump after spike | A |

Table 2: Different symbols used for the same quantities, in [BG05] and in most of this thesis ('AdEx') , and in [Izh07] ('Izh'). Membrane capacitance $C$ (in farad) is the same in both notations.

Beside these straightforward correspondences, there are some parameters in either model that have no direct equivalent in the other: $k$ and $v_t$ in Izhikevich, and $g_L$, $\Delta_T$, and $V_T$ in AdEx. For those, we'll look at the shape of Izhikevich's $F(V)$, as we've done for the AdEx neuron before.

First, the AdEx parameter $g_L$. This is the input conductance, a.k.a. the leak conductance, and the slope of $F(V)$ around the leak potential. We can find this same conductance for the Izhikevich neuron by taking the derivative with respect to $v$ of the right hand side of equation (10), at $w = 0$, $I_{\text{syn}} = 0$, and $v = v_r$. We find:

$$\frac{d}{dv}\left(k(v - v_r)(v - v_t)\right)\Big|_{v=v_r} = k(v_r - v_t) \tag{13}$$

(this value is negative: the leak potential is a stable fixed point. This corresponds to equation (1), where we find '$-g_L$'). Thus, our first nontrivial correspondence:

$$g_L = k(v_t - v_r) \tag{14}$$

We've seen that $V_T$ is the minimum of AdEx's $F$. The minimum of Izhikevich's $F$ is easily found as the average of the parabola's two zeros. I.e, $V_T$ corresponds to $(v_r + v_t)/2$.

Finally, $\Delta_T$ co-determines the slope of AdEx's $F$ at the firing threshold (equation (6)). Given that Izhikevich's $F$ is a parabola, with slopes equal in magnitude at both roots, we already know the firing

threshold slope: it is the same as the leak conductance, equation (14). Here, the AdEx model is more expressive than Izhikevich's: the slope of $\frac{dV}{dt}$ at the firing threshold can be independently tweaked from the leak conductance; in Izhikevich these two are clung together by the form of the quadratic equation.

## Comparison with AdEx

The Izhikevich and AdEx models are very similar. Their phase spaces are topologically identical: the adaptive current equation is identical (up to a renaming of the variables); and the $F(V)$-graph has the same shape, with two fixed points: a stable fixed point at the resting potential, and an unstable one at the firing threshold (figure 4).

They differ in the exact shape: Izhikevich's $F(V)$ is a parabola, while AdEx is the more realistic 'linear subthreshold, and then transitioning to an exponential' (see figures 2 and 4). As a result, Izhikevich neurons have an unrealistically slow spike upstroke, examples of which can be seen in figure 5.

A second issue is Izhikevich's subthreshold nonlinearity. The effects of this can be seen in figure 6. Positive input currents produce stronger responses than equally large negative input currents. This is explained by the quadratic $\frac{dV}{dt}$ shape: positive deviations are attenuated less, and negative deviations more, than a linear neuron would. Real and AdEx neurons do not suffer this assymetry.
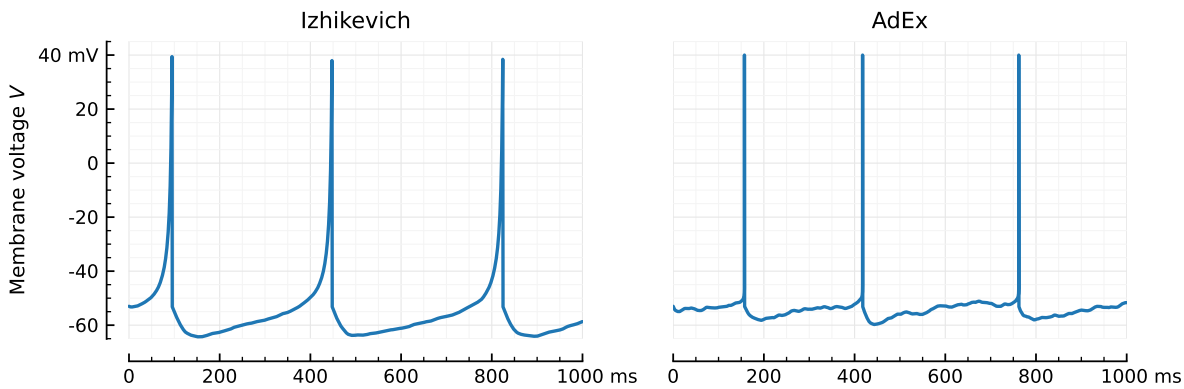


Figure 5: **Two neuron models behave differently for (near) identical parameters and input.**
The AdEx neuron's parameters are from [Nau+08], for a cortical regular spiking neuron. The Izhikevich neuron's parameters are copied from the AdEx neuron wherever they correspond directly. The other parameters are chosen so both models have the same resting and threshold potentials, and the same leak conductance, using the correspondences found earlier in this section. Both models receive EI-balanced synaptic input from 6500 Poisson spike trains with lognormal firing rates. The AdEx neuron was given stronger inputs ($\Delta g_{exc} = 12.2$ pS) than the Izhikevich neuron ($\Delta g_{exc} = 4$ pS), so as to obtain the same number of output spikes. (In both cases, $\Delta g_{inh} = 4 \, \Delta g_{exc}$). For more details, see 2023-06-23__Vm_traces_AdEx_Izh__Brian.

Figure 6: **The nonlinear response of Izhikevich neurons to sub-threshold input currents.**
'EIF' stands for exponential integrate-and-fire; it has the same $\frac{\mathrm{d}V}{\mathrm{d}t}$ as an AdEx neuron. Adaptation currents are negligibly small for both models in this test scenario.

This nonlinearity is not visible for small voltage deviations, which is what the postsynaptic potentials we are interested in in this thesis tend to be. There is however an effect of the neuron's average voltage: if this voltage is constantly on the higher side, then inputs – both negative and positive – will cause larger responses than if the median voltage was lower.

## 2.4 Synapse model

One as of yet unexplained term in our neuron model, equation (1), is the synaptic current $I_{\text{syn}}$. This is the following sum over all input synapses $i$ of the neuron:

$$I_{\text{syn}} = \sum_i g_i(V - E_i) \tag{15}$$

where $V$ is the global membrane voltage of the neuron, $E_i$ is the reversal potential of that synapse, and $g_i$ is the local synaptic conductance, which is modulated by presynaptic spikes.

For an excitatory synapse, $V < E_i$, making $g_i(V - E_i)$ negative, increasing the membrane voltage according to the sign convention for $I_{\text{syn}}$ in equation (1).

We simulate the synaptic conductances $g_i$ as exponentially decaying signals (with time constant $\tau_g$), and bump them up instantaneously on arrival of a presynaptic spike:

$$\frac{\mathrm{d}g_i}{\mathrm{d}t} = -g_i/\tau_g \tag{16}$$

On incoming presynaptic spike:

$$g_i \leftarrow g_i + \Delta g_i \tag{17}$$

Note that these are not the so called alpha-synapses. Those are two dimensional and also have an exponential rise, instead of just an exponential decay. (For an infinitely fast rise though, these models are of course the same). Simulating a full alpha synapse might increase the realism of our voltage traces, for a small simulation cost. We did not try this however. Foremost because alpha synapses fit to real data often have very fast rise times that are almost indistinguishable from instantaneous jumps.

For efficiency, we give all our excitatory synapses the same reversal potential, $E_{\text{exc}}$. Idem for the inhibitory synapses, with $E_{\text{inh}}$. This allows us to factor the synaptic current sum (equation (15)) as follows:

$$I_{\text{syn}} = (V - E_{\text{exc}}) \sum_{\text{exc } i} g_i + (V - E_{\text{inh}}) \sum_{\text{inh } i} g_i \tag{18}$$

The sums of conductance signals $g_i(t)$ can also be simplified. Say that the values of $g_i$ at $t = 0$ are $G_i$. The solution to equation (16) (at least in the time until a new presynaptic spike arrives) is then

$$g_i(t) = G_i\, e^{-t/\tau_g} \tag{19}$$

With this, and when all synapses have the same time constant $\tau_g$, the two sums in equation (18) can be factored as follows:[6]

$$\sum_i g_i(t) = \sum_i \left( G_i\, e^{-t/\tau_g} \right) = \left( \sum_i G_i \right) e^{-t/\tau_g} \tag{20}$$



1 nS

0

0    60 ms

Figure 7: Example synaptic conductance trace $g_1(t)$, with a single incoming spike at $t = 20$ ms.



1 nS

0

0    60 ms

Figure 8: Another example trace $g_2(t)$, with spikes at $t = 10$ ms and $30$ ms, and a smaller $\Delta g$.

[6] This is only valid in the time before any new spikes arrive. But the 'summability' still holds after a new spike. To see this, the given reasoning can be repeated, but simply with different values for the $G_i$ (all decayed by an amount $e^{-t_{\text{spike}}/\tau_g}$, and one increased by a bump $\Delta g_i$), and then redefining $t_{\text{spike}}$ to be $t = 0$.

This means that we need to only keep track of two conductance signals: $g_{\text{exc}}$ and $g_{\text{inh}}$, each the sum of all excitatory or all inhibitory synaptic conductances.

Our synaptic current sum then becomes simply:

$$I_{\text{syn}} = (V - E_{\text{exc}})\, g_{\text{exc}} + (V - E_{\text{inh}})\, g_{\text{inh}}, \qquad (21)$$

and we only need to simulate two differential equations, instead of one for every synapse:

$$\frac{\mathrm{d}g_{\text{exc}}}{\mathrm{d}t} = -g_{\text{exc}}/\tau_g \qquad (22)$$

$$\frac{\mathrm{d}g_{\text{inh}}}{\mathrm{d}t} = -g_{\text{inh}}/\tau_g,$$

where on arrival of a spike at synapse $i$ either $g_{\text{exc}}$ or $g_{\text{inh}}$ is instantaneously increased by a value $\Delta g_i$, depending on whether that synapse is excitatory or inhibitory.

We choose an excitatory reversal potential of $E_{\text{exc}} = 0$ mV, an inhibitory one of $E_{\text{inh}} = -80$ mV, and a time constant for the synaptic conductance decay of $\tau_g = 7$ ms. These values are rather arbitrary, but in line with other simulation studies (e.g. [Bre+07]).

## 2.5  Model summary

Combining equations (1) to (3) with equations (21) and (22), our complete AdEx point neuron with conductance-based synaptic current is:

$$C\frac{\mathrm{d}V}{\mathrm{d}t} = -g_L(V - E_L) + g_L\Delta_T \exp\left(\frac{V - V_T}{\Delta_T}\right)$$

$$\underbrace{-g_{\text{exc}}(V - E_{\text{exc}}) - g_{\text{inh}}(V - E_{\text{inh}})}_{-I_{\text{syn}}} - w$$

$$\tau_w\frac{\mathrm{d}w}{\mathrm{d}t} = a(V - E_L) - w$$

$$\tau_g\frac{\mathrm{d}g_{\text{exc}}}{\mathrm{d}t} = -g_{\text{exc}}$$

$$\tau_g\frac{\mathrm{d}g_{\text{inh}}}{\mathrm{d}t} = -g_{\text{inh}}$$

When $V > \theta$:     $V \leftarrow V_r$

$w \leftarrow w + b$

On input spike at..

exc. synapse $i$:     $g_{\text{exc}} \leftarrow g_{\text{exc}} + \Delta g_i$

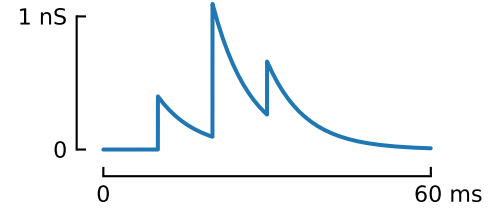inh. synapse $j$:     $g_{\text{inh}} \leftarrow g_{\text{inh}} + \Delta g_j$



Figure 9: A third synaptic conductance trace $g_3(t)$, with three input spikes at the same times and strengths as in figures 7 and 8. This signal is simulated independently, but turns out to be equal to the sum of the two others: $g_3(t) \equiv g_1(t) + g_2(t)$.
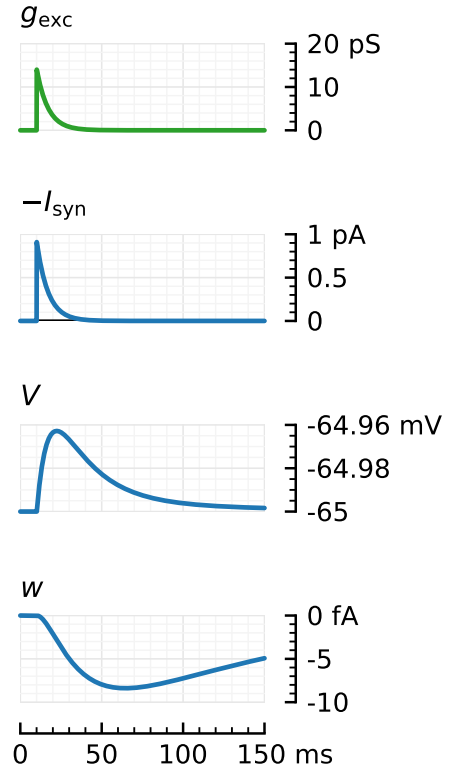


Figure 10: **Impulse response of the conductance-based AdEx neuron**, using a single excitatory input spike arriving at $t = 10$ ms, with a synaptic weight $\Delta g_{\text{exc}}$ of 14 pS. Parameters as in table 1, with $E_{\text{exc}} = 0$ mV and $\tau_g = 7$ ms. More details at 2023-07-26__AdEx_Nto1_we_I_syn.

We solve these equations numerically using first-order (Euler) integration, with a timestep $\Delta t$ of $0.1$ ms. This timestep is sufficiently small with respect to the different time constants in the model[7]. See section 2.8 below for more details on the numeric implementation.

Figure 10 shows the impulse response of this model, using a single incoming excitatory spike. The PSP bump is visible in the third panel. Note its tiny size, of about $0.04$ mV. Our task will be to detect this tiny signal, in a sea of voltage imaging noise and PSP bumps of other input neurons.

## 2.6   Input in the N-to-1 setup

In our simplest experimental setup, we simulate just one AdEx neuron. Its input is provided by an array of $N$ Poisson neurons, i.e. they each generate spike trains according to a Poisson process. We call this the 'N-to-1' setup.

### Log-normal Poisson spiketrains

The firing rates of real neurons often follow a long-tailed distribution: most neurons do not fire much at all, while a few fire a lot. We recapitulate this in the firing rates of our Poisson input neurons, by drawing their firing rates from a log-normal distribution. We look to the literature for realistic parameters for this distribution, namely to the modelling paper of Roxin et al. [Rox+11], and the experimental sources it cites [HDZ08; OCo+10].

Log-normal distributions are usually parametrized with $\mu$ and $\sigma$: the location and scale of the underlying normal distribution, i.e. after log-transforming the input domain. In the above sources however,
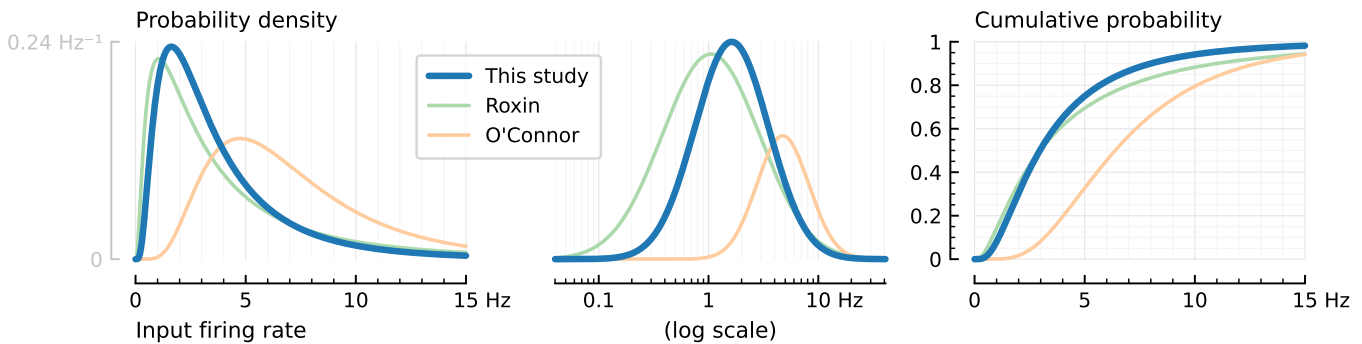


Figure 11: **Input spike trains are given log-normally distributed firing rates.**
The distribution used in our simulations is shown in bold. It has mean $\mu_x = 4$ Hz and variance of the underlying Gaussian $\sigma^2 = 0.6$. The two light distributions are from the literature. Note that the Roxin et al. distribution is slightly more heavy tailed than this study's: it has both more low firing and more high firing neurons.

the mean $\mu_x$ of the data distribution itself is given. We can find $\mu$, if $\sigma$ is known, as $\mu = \ln(\mu_x) - \sigma^2/2$.

Roxin et al. use a mean rate $\mu_x$ of 5 Hz and a variance of the logarithm of the rate $\sigma^2 = 1.04$ in their figure 2. Hromádka et al. [HDZ08] recorded neurons in the auditory cortex of awake rats during acoustic stimulation. They find a mean firing rate $\mu_x$ of 6.2 Hz and a median of 2.4 Hz; no numeric variances are given. O'Connor et al. [OCo+10] recorded neurons in the barrel (whisker) cortex of behaving mice. Ensembled over all layers of the cortex they recorded from, they report $\mu_x = 7.4$ Hz, $\sigma_x = 12.6$ Hz, a median of 1.5 Hz, and an interquartile range of 9.5 Hz. This corresponds to a Gaussian variance of $\sigma^2 = \ln(1 + \sigma_x^2/\mu_x^2) = 0.30$. The Roxin and O'Connor distributions are also shown in figure 11.

Given these data, we choose our parameters $\mu$ and $\sigma$ so that the distribution lies roughly halfway O'Connor's and Roxin's, while making sure our firing rates are rather low than high: as our idea for connection inference rests on the number of spikes that can be used to calculate spike-triggered averages (see chapter 3), we don't want to overestimate the number of available input spikes and obtain overly optimistic results.

Our log-normal distribution has a median of 2.96 Hz, very close to Roxin's 2.97 Hz. O'Connor's has 6.4 Hz, which is markedly different from the median of 1.5 Hz they reported (hinting that their data distribution might not be log-normal).

## Synaptic weights

We choose to simulate EI-balanced input, where we simulate four times as many excitatory as inhibitory input spiketrains; but with the inhibitory inputs four times as strong as the excitatory inputs.

I.e. if our total number of input spiketrains is $N = 6500$, we have $N_{exc} = 5200$ excitatory inputs, and $N_{inh} = 1300$ inhibitory inputs. In the N-to-1-setup, we give every input of the same type the same weight. I.e. in the model summary (section 2.5), $\forall i : \Delta g_i = \Delta g_{exc}$ and $\forall j : \Delta g_j = \Delta g_{inh}$. If we would thus choose $\Delta g_{exc} = 10$ pS, then, for EI-balanced input, $\Delta g_{inh}$ would be 40 pS.

To choose a value $\Delta g_{exc}$ for the excitatory input strength, we look at the desired output firing rate of our one simulated neuron. We want it to spike at a realistic, average rate. Thus, we try for it to have the same output firing rate as the mean input firing rate: $\mu_x = 4$ Hz. This rate is achieved by testing a range of different input drives $\Delta g_{exc}$: see figure 12. For $N = 6500$ inputs, we find an average output firing rate of 4.0 Hz at $\Delta g_{exc} = 15$ pS (and thus $\Delta g_{inh} = 60$ pS).

Two asides on figure 12. First, from the bottom panel, we see that the AdEx model, with parameters for a cortical RS (regular spiking) neuron, is a so called 'Type I' neuron:[8] the firing rate can be made arbitrarily low, and there is no discontinuous jump from 0 Hz to
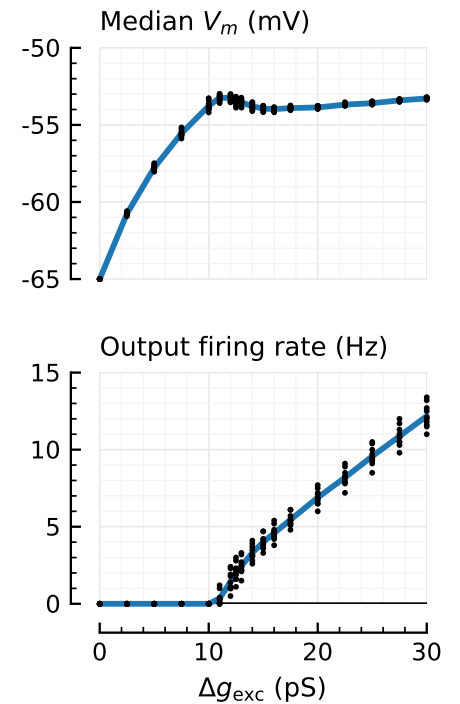


Figure 12: **Activity level of the output neuron in the N-to-1 setup, for increasing input strength $\Delta g_{exc}$.**
With $N = 6500$ EI-balanced inputs, and $\Delta g_{inh} = 4 \cdot \Delta g_{exc}$.
The blue line is the average over 10 different simulations (black dots), each with a different random seed for the input spike train generation.
Increasing input first heightens the voltage level without increasing the firing rate, and later heightens the firing rate without increasing the voltage level further.
$V_m$ = membrane voltage $V$.
Source: 2023-08-05__AdEx_Nto1_we_sweep.

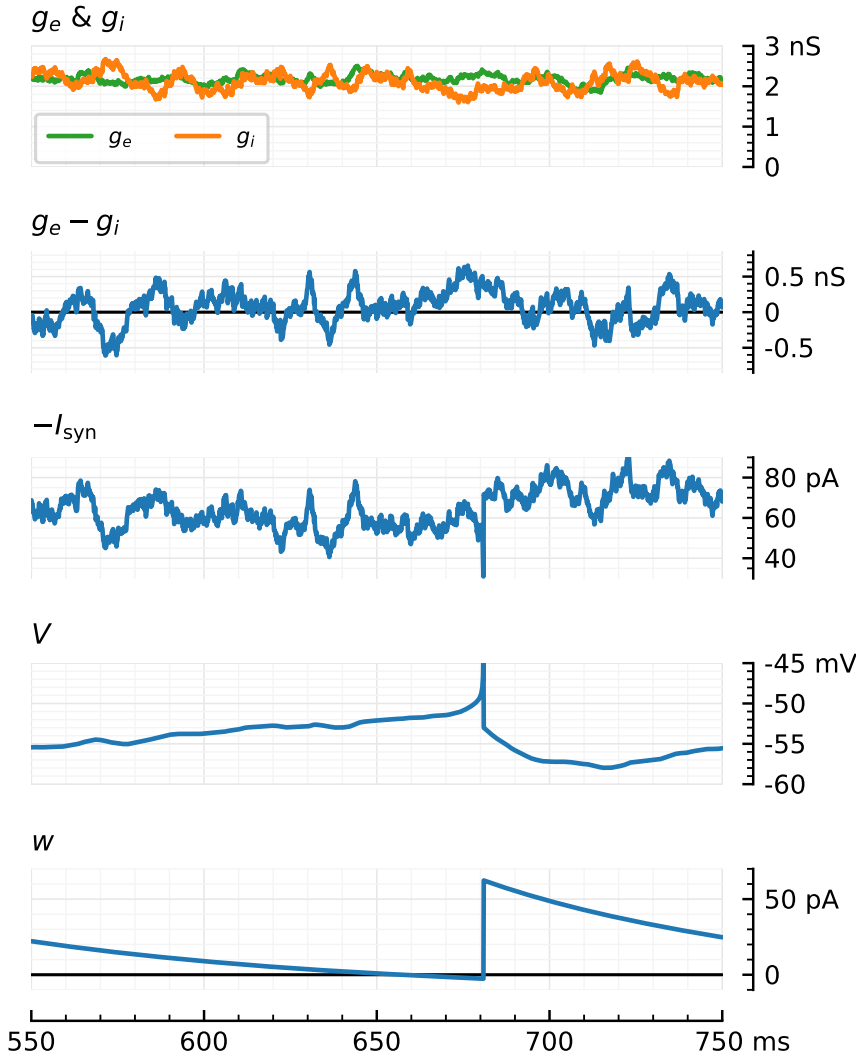[8] https://neuronaldynamics.epfl.ch/online/Ch4.S4.html

Figure 13: **The different signals in the conductance-based AdEx model**.
Example signal traces. N-to-1 setup, with $N = 6500$ EI-balanced input spiketrains, and $\Delta g_{\mathrm{exc}} = 15$ pS. Neuron parameters for a cortical regular spiking neuron.
Note that the synaptic current signal $-I_{\mathrm{syn}}$ is very similar to the difference of the excitatory and inhibitory synaptic conductances, as long as the membrane voltage $V$ does not vary too much; on a spike (here at $t \approx 680$ ms), $I_{\mathrm{syn}}$ spikes as well.
Also note that the neuron spikes at a time when the inhibitory input randomly falls below the excitatory input, for a certain time. Spikes are 'fluctuation driven' in this input regime.
$g_e \equiv g_{\mathrm{exc}}$ and $g_i \equiv g_{\mathrm{inh}}$.
The $V$ and $-I_{\mathrm{syn}}$ signals are cut-off by the plot boundaries at the spike time.
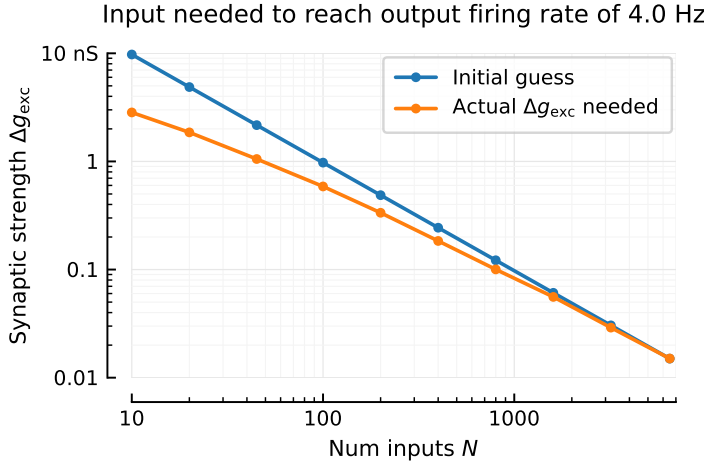Source: `2023-07-26__AdEx_Nto1_we_I_syn`.

some minimum firing rate.

Second, why does the output activity level increase for increasing excitatory input, if the inhibitory input becomes stronger by the same amount? The answer lies in the synaptic reversal potentials, which are $E_{\mathrm{exc}} = -80$ mV and $E_{\mathrm{inh}} = 0$ mV. The median membrane voltage (top panel in figure 12), at about $-60$ mV, lies closer to the inhibitory reversal potential than the excitatory potential, making the inhibitory term of the synaptic current $I_{\mathrm{syn}}$ weaker than the excitatory term (equation (21)).

The fact that this EI-balanced input leads to a net-excitatory effect can also be seen in figure 13: the signal $g_{\mathrm{exc}} - g_{\mathrm{inh}}$ hovers around zero, but the signal $I_{\mathrm{syn}}$ — even though it has largely the same shape — is nowhere near zero.

Later, we will compare the performance of different connection-detection algorithms across different numbers of inputs $N$. We choose the approximately evenly log-spaced sequence

$$N \in [10, 20, 45, 100, 200, 400, 800, 1600, 3200, 6500].$$

Figure 14: **Finding the right input strength in the N-to-1 setup, for different numbers of inputs** $N$.
For each $N$, the right input strength $\Delta g_{\text{exc}}$ was found using an iterative search procedure. For every $\Delta g_{\text{exc}}$ value evaluated, 10 different 10-second simulations were run (each with a different random seed for the input spiketrain generation). The average output firing rate of these 10 simulations was taken, and compared with the goal firing rate of $4$ Hz. $\Delta g_{\text{exc}}$ was then adjusted until the simulated firing rate was sufficiently close.
Source: `2023-08-05__AdEx_Nto1_we_sweep`.

For each of these $N$, we want the output firing rate to be the same, namely 4 Hz. To find the synaptic strengths $\Delta g_{\text{exc}}$ that accomplish this, we perform an iterative search.[9] As an initial guess, it would make sense for the required input strength to scale inversely proportional to the number of inputs. We find however that there is a slight deviation from this linear expectation (figure 14): the less inputs, the less strong each input should be to reach the same output firing rate, compared to a linear extrapolation from our finding of $\Delta g_{\text{exc}} = 15$ pS for $N = 6500$. E.g. the linear expectation for $N = 10$ would be $\Delta g_{\text{exc}} = 9.75$ nS ($= 15$ pS $\cdot\, 6500/10$), but we find we need only $\Delta g_{\text{exc}} = 2.83$ nS to reach an output firing rate of $4$ Hz.

[9] We use SciPy's `root_scalar` function, which uses 'Brent's method', which is like bisection but has faster convergence. We seed the algorithm with the bracket $[w_0/4, w_0 \cdot 4]$, where $w_0$ is the initial linear guess for $\Delta g_{\text{exc}}$. In about 8 iterations, we come within $0.01$ Hz of the desired $4$ Hz.

## 2.7 Voltage imaging

The signals detected by a light microscope in a voltage imaging setup are not the same as the real membrane voltage traces of which they are a reflection.

We model this lossy transformation by simply adding Gaussian noise to our simulated membrane voltage. As in the voltage imaging literature, we quantify the amount of this noise by a 'spike-SNR' measure (spike signal-to-noise ratio). This is defined as the height of an average spike relative to the standard deviation of the noise:

$$\text{spike-SNR} = \frac{\text{spike height}}{\sigma_{\text{noise}}} \tag{23}$$

In other words, if we call our output signal $y$, and with $V_i$ the voltage samples resulting from the numeric integration of equation (1), we have:

$$y_i = V_i + \varepsilon_i \tag{24}$$

$$\varepsilon_i \sim \mathcal{N}(0, \sigma_{\text{noise}})$$

A typical but conservative level of noise in voltage imaging recordings has a spike-SNR of 10. If we take as spike height the spike
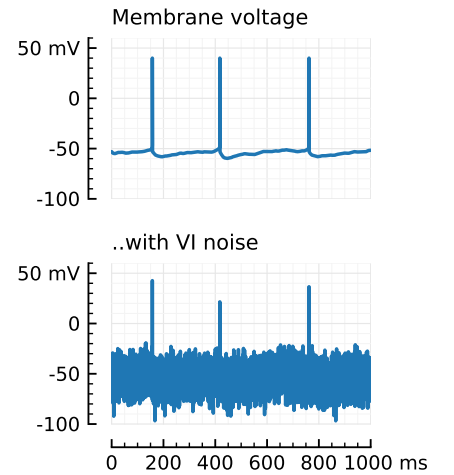


Figure 15: Simulated voltage trace (same as in figure 5, right), without and with 'voltage imaging' (VI) noise added ($\sigma_{\text{noise}} = 10.5$ mV).

detection threshold minus the leak potential ($\theta - E_L$, see table 1), we obtain a $\sigma_\text{noise}$ of 10.5 mV.

A more realistic model of the voltage imaging-transformation would also incorporate the exponential decay over time of the SNR (with a time constant of about 10 minutes), and the short-term 'smearing in time' of voltage indicators. The latter could be done by passing the voltage signal through a linear filter with some non-instantaneous impulse response.

## 2.8  Software implementation

To numerically simulate the model described in this chapter, different software implementations have been used throughout my PhD.

| Software | Language | Model | Simulation time | |
|----------|----------|-------|-----------------|---|
| Brian2 | Cython + Python | full | 55 | s |
| | | merged | 28 | s |
| | C++ | full | 29 | s |
| | | merged | 1.9 | s |
| own | Julia | full | 0.068 | s |

Note that the timings in table 3 do not include compilation times, which can be considerable.

*{to expand a bit}*

Table 3: **Time taken to simulate 10 seconds of the N-to-1 experiment** with N = 6500 Poisson inputs, and one conductance-based AdEx neuron. In the 'merged' models, only 200 Poisson spiketrains are independently simulated (instead of all 6500, as in the 'full' model). All other inputs are replaced by two single Poisson inputs (one excitatory and one inhibitory), with an equivalent effect. (That is, we use Brian's '`PoissonInput`'). Times do not include compilation time.
For details, see `2023-07-10__AdEx_Nto1_Brian_speedtest` (Cython), `2023-08-02__speedtest_brian_standalone_AdEx_Nto1` (C++), and `2023_08_09__Poisson_cquantile_upperbound` (Julia).

*Chapter 3*

# Spike-triggered averaging

As shown in figure 1, our idea for connection inference rests on the causal 'presynaptic spike' → 'postsynaptic voltage bump' link. I.e. we want to know for which neuron pairs a spike in one is reliably followed by such a bump in the other. The problem is that these bumps (the postsynaptic potentials or PSPs) are minute, and are easily drowned out by other PSPs, spikes, and voltage imaging noise.

So, as is often done in neuroscience, we take the *average* over many instantiations, so as to hopefully find a signal in the noise. Specifically, we take spike-triggered averages, or **STA**s, of neurons' voltage traces. If there is a connection from neuron '1' to neuron '2', then an STA of neuron 2's voltage imaging signal, based on neuron 1's spikes, would hopefully show the PSP.

To use this idea for an actual connection test, we look specifically at the height of an STA, and compare it to a distribution of STA heights that we'd expect were the two neurons not connected. This is illustrated in figure 17.
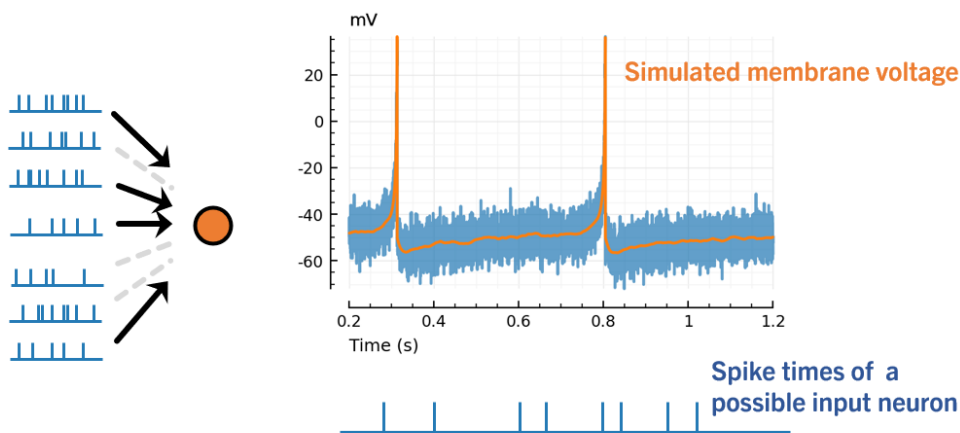
Figure 16: **The 'N-to-1' problem.**
*Left*: A neuron $N$ (orange circle), and the spike trains of other neurons in the network (blue).  Some of these other neurons impinge directly on $N$ (black arrows), while others are not (directly) connected (dashed gray lines). Given only neuron $N$'s voltage signal and the other neurons' spiketrains, we want to detect the direct inputs, while rejecting the not-directly-connected spiketrains.
*Right*: The simulated membrane voltage of the impinged-upon neuron (orange), and the same signal with Gaussian noise added, to simulate a voltage imaging signal (blue). Underneath the plot, one of the possible input spiketrains, time-aligned to the voltage signal. This alignment is used later to extract spike-triggered windows from the voltage signal.
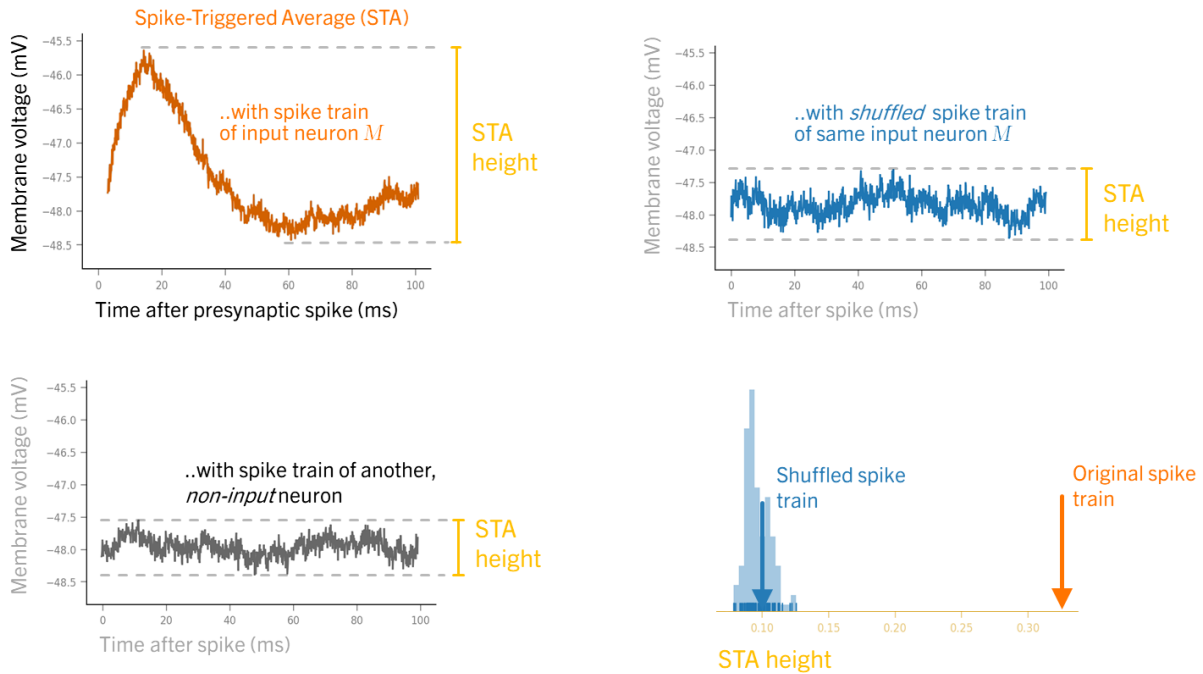
Figure 17: **A simple connection test: STA height with shuffle control.**
The spikes of a possible input neuron are aligned to the voltage trace of the neuron of interest $N$, as in figure 16. For every such spike, a 100-ms long window is cut out of the voltage of $N$. The average of all these windows is called the spike-triggered average (STA).
*Left*: Two example STAs of neuron $N$'s membrane voltage: one for an actually connected input neuron, $M$ (top, orange); and one for a non-input neuron (below, gray). Given an STA signal $x$, we will use its height $h = \max(x) - \min(x)$ (also known as 'peak-to-peak' or 'ptp') to test whether two neurons are connected.
*Right*: An STA of $N$'s membrane voltage using a shuffled version of $M$'s spike times (which is made by randomly permuting the inter-spike-intervals of $M$). This 'shuffled STA height' provides a control for the STA height connection test statistic: "what do we expect the STA height to be if there is *no* connection $M{\rightarrow}N$". By calculating different such shuffles, we obtain a null-distribution for the STA height test statistic. And by comparing the real STA height to this distribution, we can calculate a $p$-value. Here, the real STA is larger than all shuffle controls, of which there are 100. So $p < 0.01$, and at $\alpha = 0.05$, we conclude there is indeed a connection $M{\rightarrow}N$.

# New connection inference methods

## 4.1 STA Template correlation

- Idea for the two-pass test

- Template examples. Both ideal and template found with first-pass (STA height-only)

- Peformance for different N, & comparison with previous method

## 4.2 Linear regression of the upstroke

- Non-STA method: concatenated individual windows as (X, y)

- Examples of pooled windows, and fits

- Mention the problem of unknown transmission delays

- Perfomance for different N, & comparison with previous methods

## 4.3 Fitting a full STA model

- Model design

- Iterative model fitting

- Problem of overfitting, and parameter-constraints to solve it

- An advantage: fit parameters (like transmission delay and time constants) are biologically  meaningful

- Perfomance for different N, & comparison with previous methods

## 4.4 Clustering, & Hierarchical model fitting

- (Time-permitting)

## 4.5 Zhou/Cai's 'Spike-triggered regression'

- (Time-permitting)

## 4.6 Computational cost

- Timings of each method, extrapolation for larger number of tested connections

## 4.7 Summary

- Conclusions of the N-to-1 experiment

- Leadup to the network experiments: what we could not yet test (the problem of indirect connections, as e.g. identified in the connectomics challenge)

*Chapter* 5

# Network model

- Network connectivity, E/I balance, raster plots

- Too many possible connections to test them all → Subsampling

- Performance of last chapter's methods

- If time: experiment with a network that is less densely connected than our current fully-random one. Why? To better examine the effect of indirect connections / colliders (For the current connectivity, there are too many of those. But in a more realistic, 'localized' network, there are less, and so it seems easier to isolate and examine their effect).

*Chapter* $6$

# Discussion

## 6.1  Summary & conclusions

..

## 6.2  Future work

- Test on real data

- Direct comparison with spikes-only methods

- More complexity in the testing setup: different transmission delays and time constants per synapse / neuron, plus:

- Short term synaptic plasticity. Bursting. Oscillations.

- Simulate different brain areas (different cell types and connectivity patterns). Simulate the same area, but in different states (up vs down, e.g.)

- New connection test method to try: something deep learning-based (we have infinite training data, given our simulation)

# References

[Bad+08]   Laurent Badel, Sandrine Lefort, Thomas K. Berger, Carl C. H. Petersen, Wulfram Gerstner, and Magnus J. E. Richardson. "Extracting Non-Linear Integrate-and-Fire Models from Experimental Data Using Dynamic I−V Curves". In: *Biological Cybernetics* (Nov. 15, 2008). DOI: 10.1007/s00422-008-0259-4.

[BG05]   Romain Brette and Wulfram Gerstner. "Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity". In: *Journal of Neurophysiology* (Nov. 2005). DOI: 10.1152/jn.00686.2005.

[Bre+07]   Romain Brette et al. "Simulation of Networks of Spiking Neurons: A Review of Tools and Strategies". In: *Journal of Computational Neuroscience* (Dec. 1, 2007). DOI: 10.1007/s10827-007-0038-6.

[DA01]   Peter Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Computational Neuroscience. The MIT Press, 2001. 460 pp. ISBN: 978-0-262-04199-7.

[HDZ08]   Tomáš Hromádka, Michael R. DeWeese, and Anthony M. Zador. "Sparse Representation of Sounds in the Unanesthetized Auditory Cortex". In: *PLOS Biology* (Jan. 29, 2008). DOI: 10.1371/journal.pbio.0060016.

[Izh07]   Eugene M. Izhikevich. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. Cambridge, Mass.: The MIT Press, 2007. 464 pp. ISBN: 978-0-262-51420-0.

[Nau+08]   Richard Naud, Nicolas Marcille, Claudia Clopath, and Wulfram Gerstner. "Firing Patterns in the Adaptive Exponential Integrate-and-Fire Model". In: *Biological Cybernetics* (Nov. 15, 2008). DOI: 10.1007/s00422-008-0264-7.

[OCo+10]   Daniel H. O'Connor, Simon P. Peron, Daniel Huber, and Karel Svoboda. "Neural Activity in Barrel Cortex Underlying Vibrissa-Based Object Localization in Mice". In: *Neuron* (Sept. 23, 2010). DOI: 10.1016/j.neuron.2010.08.026. pmid: 20869600.

[Rox+11]    Alex Roxin, Nicolas Brunel, David Hansel, Gianluigi Mongillo, and Carl van Vreeswijk. "On the Distribution of Firing Rates in Networks of Cortical Neurons". In: *Journal of Neuroscience* (Nov. 9, 2011). DOI: 10.1523/JNEUROSCI.1677-11.2011. pmid: 22072673.