

Modelling the role of local neurons in
weakening odor onset transients in the
antennal lobe of *Drosophila*
Schnagel Lab Theory Rotation - Always be faking data

Gerick Lee

2016-09-02

1 Introduction

In the antennal lobe of *Drosophila*, olfactory processing is performed by three neuron types: Olfactory Receptor Neurons (ORNs), which bind to odors and transduce this signal into a spiking output, forming synapses with inhibitory Local Neurons (LNs), as well as excitatory Projection Neurons (PNs), which transmit the signal to the rest of the fly brain. LNs receive projections from many ORNs in a non-specific manner, and project to the presynaptic side of the ORN-PN synapse in a similarly non-specific fashion [7]. PNs are more selective, connecting only to ORNs with a specific odor preference (further details of antennal lobe olfactory circuitry can be found in Wilson [9]).

While ORN responses to odor can vary in a number of ways, a common feature is a sharp onset transient which decays during sustained odor presentation [5]. In contrast, this transient is much weaker in PNs, with the reduction in transient strength being modulated by inhibition [7].

In the weakly electric fish, the work of Bell and colleagues has shown that ganglion cells can learn and project a inhibitory negative image[1–4]. This negative image is used to cancel the artifact generated when the fish sends its pulse into its environment. Inspired by this, we hypothesized that the LNs might be used to weaken onset transients; a learning mechanism would allow the fly to self-organize its olfactory behavior rather than depending on genetic or other pre-programming.

To test this, we used a model of ORNs similar to that proposed by Nagel and Wilson [5]. Both linear (convolutional) models and dynamical firing rate models were created, and were fed into a delay-line LN model, where

increases in the odor strength (not the ORNs) would generate a single firing rate increase after a fixed delay. These firing rates were then fed into a model of the Projection Neurons (PNs), which used a simple anti-Hebbian learning rule to adjust the weights of the inhibitory neurons.

2 Groundwork

As a first step, we needed to generate reliable gaussian filters (the linear ORN models were generated using either a difference-of-gaussians or gaussian-by-line multiplication). Figure 1 shows a test signal (**A**) and the frequency representation of that signal, with an overlapping gaussian filter (**B**). The filter has a Full-Width-at-Half-Max (FWHM) of 2 Hz, which can be seen in the figure.

2.1 Odor

As an input for our model system, we generated a simulated odor as a zero vector with the length of the desired simulation (with a simulated sample rate of 1000 Hz). we then generated simulated step odor presentations by setting consecutive windows to a fixed value (in this case 0.5), for a period corresponding to a desired presentation length (usually 1 to 1.5 seconds, or 1000 – 1500 samples). An example section of an odor stimulus is shown as Figure 2.

3 ORN Modelling

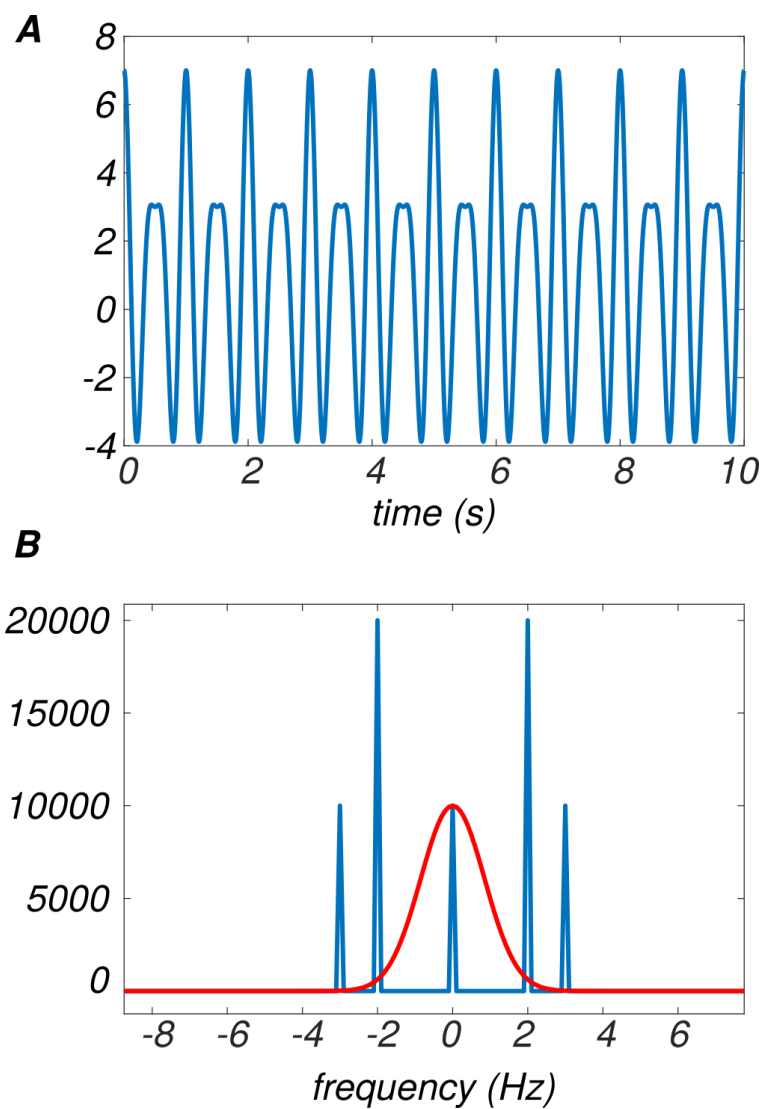
For our investigation of PN responses, we first needed a model of ORNs. We began with a linear filter model before moving to a dynamical model, inspired by the success of the model in Nagel and Wilson [5].

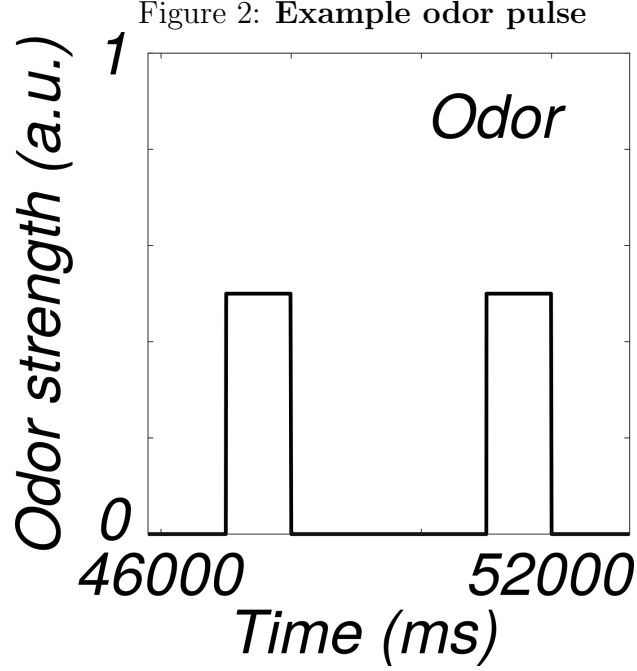
3.1 Linear Filter Models

3.1.1 Difference-of-Gaussian Model

To model ORNs, we first used a linear filter model using both Difference-of-Gaussians (DoG) and line-gaussian multiplication as the filters in the frequency domain. Our first attempt used a DoG filter with five parameters (governing the width of each gaussian, as well as their offset from the signal, and each other). Figure 3 shows an example curve, with the accompanying

Figure 1: **A:** Example signal with non-zero components at the DC, 2Hz, and 3Hz. **B:** Fourier Transform of signal (blue), as well as (amplitude-adjusted) filter (red) with $\text{FWHM} = 2$ Hz.





odor. Note that it is possible to rectify the response at zero, so as to better capture the asymmetries observed *in vivo*.

3.1.2 Gaussian-Line Model

To reduce the number of free parameters, we replaced the DoG filter with one multiplying a gaussian filter with a line. This model was also rectified at zero, to avoid negative firing rates (and return more realistic ORN responses). Three example responses (two onset, one offset - all of which were generated using the same function [on and offset transient-type responses are determined by the same parameters]) are shown in Figure 4.

3.2 Dynamic Model

For greater accuracy, we switched to a dynamic model of receptor function. This model was governed by the following equation:

$$\frac{dR^*}{dt} = k_{on}[O] - (k_{on}[O] + k_{off})R^*, \quad (1)$$

where R^* is the proportion of bound receptor (between 0 and 1), $[O]$ is the odor concentration, and there are two scalar rate constants k_{on} and k_{off} . The

Figure 3: **Difference-of-Gaussian based linear filter model** Odor pulse (in arbitrary units) is shown in blue.

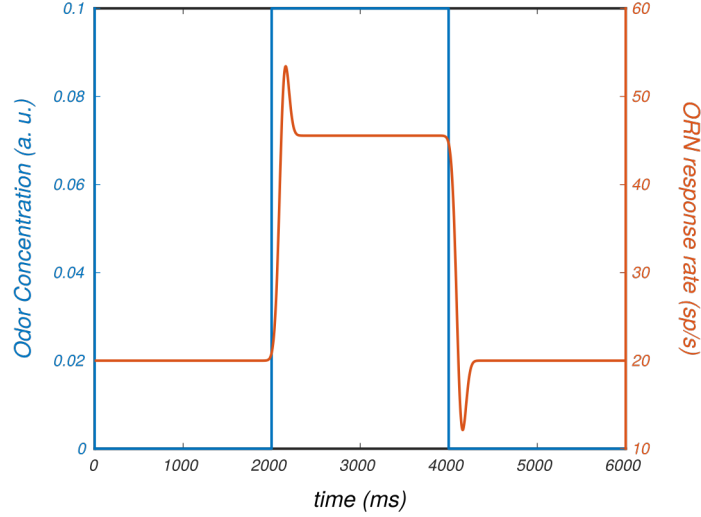


Figure 4: **Variability in ORN firing rates: gaussian-by-line model.** Odor strength (in black, at bottom) is in arbitrary units.

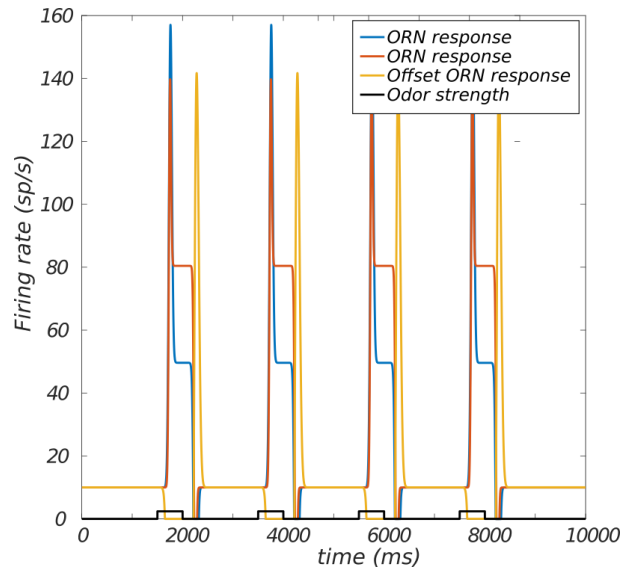
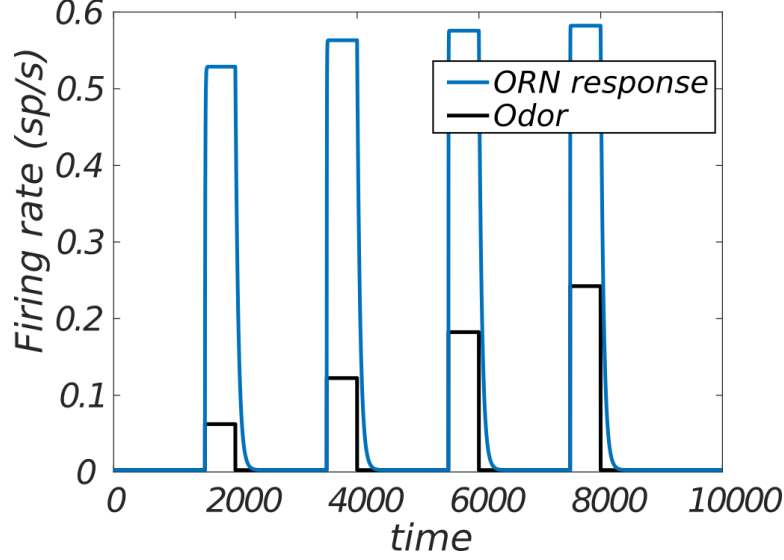


Figure 5: **ORN firing rates: Basic dynamic model.** Odor strength (in black, at bottom) is in arbitrary units, though relative magnitudes are preserved.



output could be multiplied by an arbitrary scalar to better reflect real-world firing rates.

In this model (Figure 5) activation alone varies with concentration (as in [5]), resulting from the fact that only k_{on} is affected by odor concentration. This model also saturates with increasing concentration (as a result of the limited range of R^*).

Unsatisfied with the shape of the ORN response in this model, we began to experiment with various inactivation terms. We first used a paradigm in which bound receptor (R^*) could vary between an active and inactive (R_i) state (data not shown). We then converged on a model in which inactive receptor recycled back to the unbound state (R). We modeled this with the following equations:

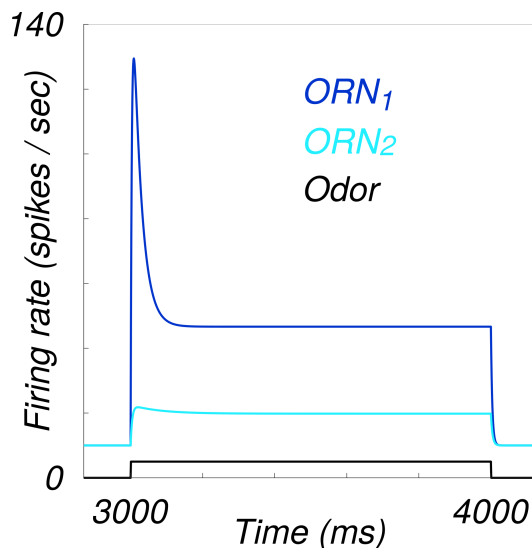
$$\frac{dR^*}{dt} = k_{on}[O] + (k_{re} - k_{on}[O]) R_i - (k_{on}[O] + k_{off} + k_{de}) R^*, \quad (2)$$

$$\frac{dR_i}{dt} = k_{de} R^* - k_{re} R_i, \quad (3)$$

where R_i was the inactive state, k_{de} was the constant of inactivation (from R^* to R_i), and k_{re} was the constant for the conversion from R_i to R .

This model (Figure 6) kept the onset-concentration dependence, and was able to mimic the various shapes of (onset-transient response) curves seen in

Figure 6: **ORN firing rates: Dynamic model with one-way inactivation.** Odor strength (in black, at bottom) is in arbitrary units.



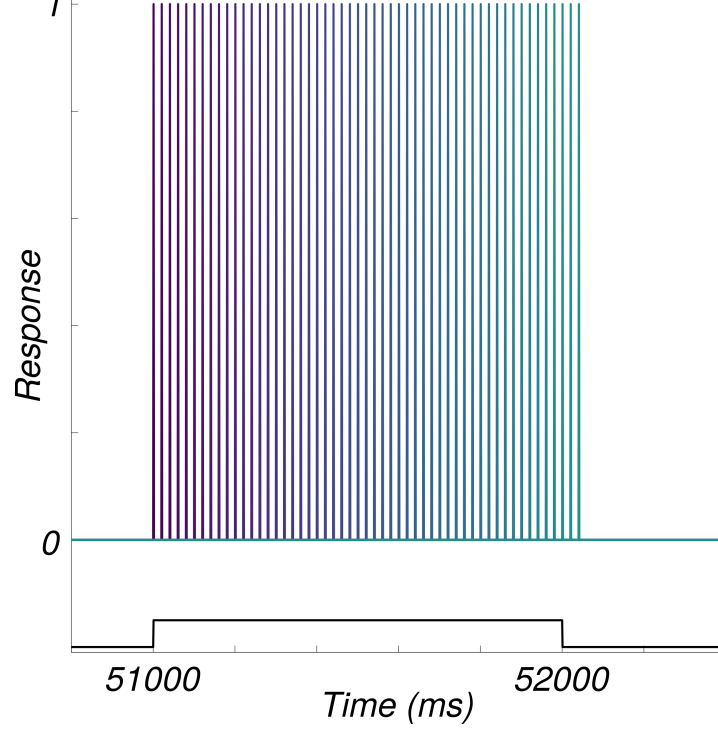
[5]. These neurons were the ones used for later modelling (if future models require it, offset neurons can be generated using the linear filter model, but none were used for the LN or PN simulations discussed here).

4 LN model

To model LNs, we used a two-stage model. The first step was to generate a series of "basis-functions", where each LN was temporally linked to the onset of odors (importantly, because odors were always simulated as idealized, consistent step functions, we used odors, not ORN responses, to trigger LN basis functions). Each LN basis function was evenly spaced from other LNs, and together, they tiled the range of step-durations used. Two types of basis functions were used: The first type was a set of simple delta functions (depicted in Figure 7), where each time bin had exactly one responsive LN delta function assigned to it (leading to an unrealistically large number of simulated LNs). The second type was a set of tightly overlapping gaussian functions (reducing the number of required LNs to a more physiologically plausible 200 (reviewed in [6]). Gaussian peaks were spaced 6 bins apart, and are depicted in Figures 8. For both types of basis set, a delay was added between odor onset and the first LN basis function.

After the basis set was generated, the second step was to project the ORN

Figure 7: **LN basis set: Delta function with uniform delay.** Odor strength (in black, at bottom) is in arbitrary units. Each spike corresponds to a different LN.



responses from all simulated ORNs on to the LN population, to mirror the all-to-all ORN to LN projections seen *in vivo* [9]. This projection is depicted in Equation 4.

$$\vec{r}_{LN}(t) = \sum_i^{N_{ORN}} \vec{b}_{LN}(t)^T * \vec{r}_{ORN}(t), \quad (4)$$

using $\vec{r}_{LN}(t)$ as the firing rate of all LNs at time t , $\vec{b}_{LN}(t)$ as the value of the basis set (either delta functions or gaussians) at time t , and $\vec{r}_{ORN}(t)$ as the firing rate of all ORNs in the system at time t . This results in a vector with a length corresponding to the number of ORNs. The firing rate was then the sum of this vector. Examples of LN firing rates are shown in Figures 9 and 10 for the delta and gaussian basis functions, respectively.

Figure 8: **LN basis set: Evenly spaced gaussians.** Odor strength (in black) is in arbitrary units.

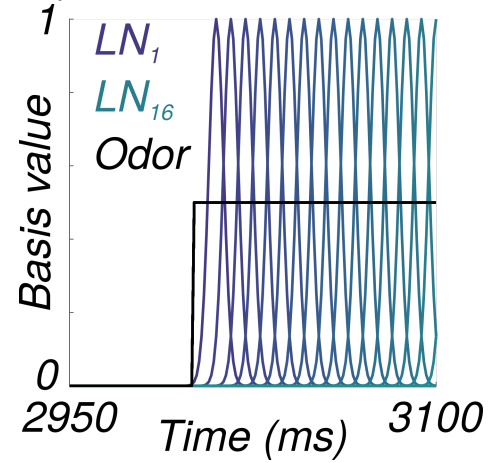


Figure 9: **LN responses: ORN projection onto a delta basis set.** For display purposes, an evenly spaced subsampling of LNs is shown. Odor strength (in black) is in arbitrary units.

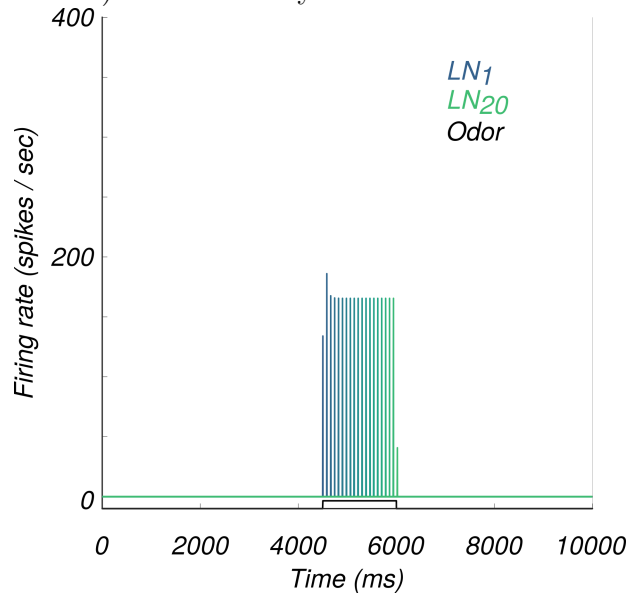
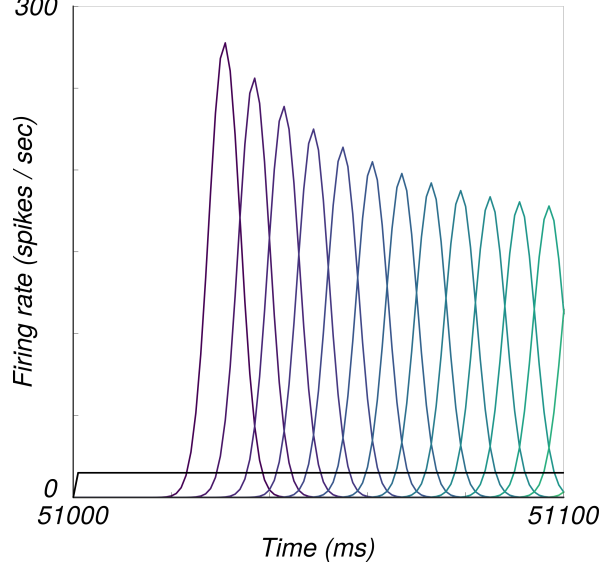


Figure 10: **LN responses: ORN projection onto a gaussian basis set.** Twelve LNs are shown. Odor strength (in black) is in arbitrary units.



5 PN model

To test our main hypothesis, we used a simple learning rule governing the inhibition from LNs to PNs. We modeled the inputs to PNs as a simple additive combination of ORN and LN inputs, with a PN receiving exactly one ORN input (one-to-one ORN to PN connectivity) and LN inputs from the entire population of LNs (all-to-all LN to PN connectivity). We modelled the firing rate of the i^{th} PN as:

$$r_{PN_i}(t) = r_{ORN_i}(t-1) + \left(\beta * \vec{r}_{LN}(t-1)^T * \vec{w} \right), \quad (5)$$

where $\vec{r}_{PN_i}(t)$ was the firing rate of the i^{th} PN at time t , $r_{ORN_i}(t-1)$ was the firing rate of the i^{th} ORN at time $t-1$, β was a scale factor to weaken the degree of inhibition (β was set to values below one: $1E-6$ was used for the delta function basis, and $1E-7$ was used for the gaussian basis). $\vec{r}_{LN}(t-1)$ was the vector of LN firing rates at time $t-1$, and \vec{w} was the vector of weights, which was initialized at -0.3 , and was restricted to the range $[-\infty, 0]$.

5.1 Learning rule (weight adjustment)

Weight adjustments for the inputs to a given ORN were modeled with the following rule:

$$\frac{d\vec{w}}{dt} = \vec{w}(t) - \alpha * \vec{r}_{LN}(t) * \sum_n^{N_{ORN}} r_{ORN_i}(t) - \bar{r}_{ORN_n}(t - \tau : t), \quad (6)$$

where α was a constant learning rate, set to 0.1. Changes for a given weight were driven by LN activity which coincided with deviations in the i^{th} ORN firing rate from the ORN population mean across the preceding τ bins (τ was generally set to 200 ms): Note that the mean ORN firing rate included all ORNs, and not just the one whose weights were being adjusted (also of note is that the $r_{ORN_i}(t)$ term does not change with the index n).

For a given LN weight to update, two conditions must be met: The LN must be active (both basis sets used a baseline firing rate of zero), and the respective ORN must have a firing rate deviating from the average across the previous τ time bins. If either value is zero, the weight will not change. The magnitude of the weight adjustment was predominantly driven by the difference between the ORN rate and the population mean rate.

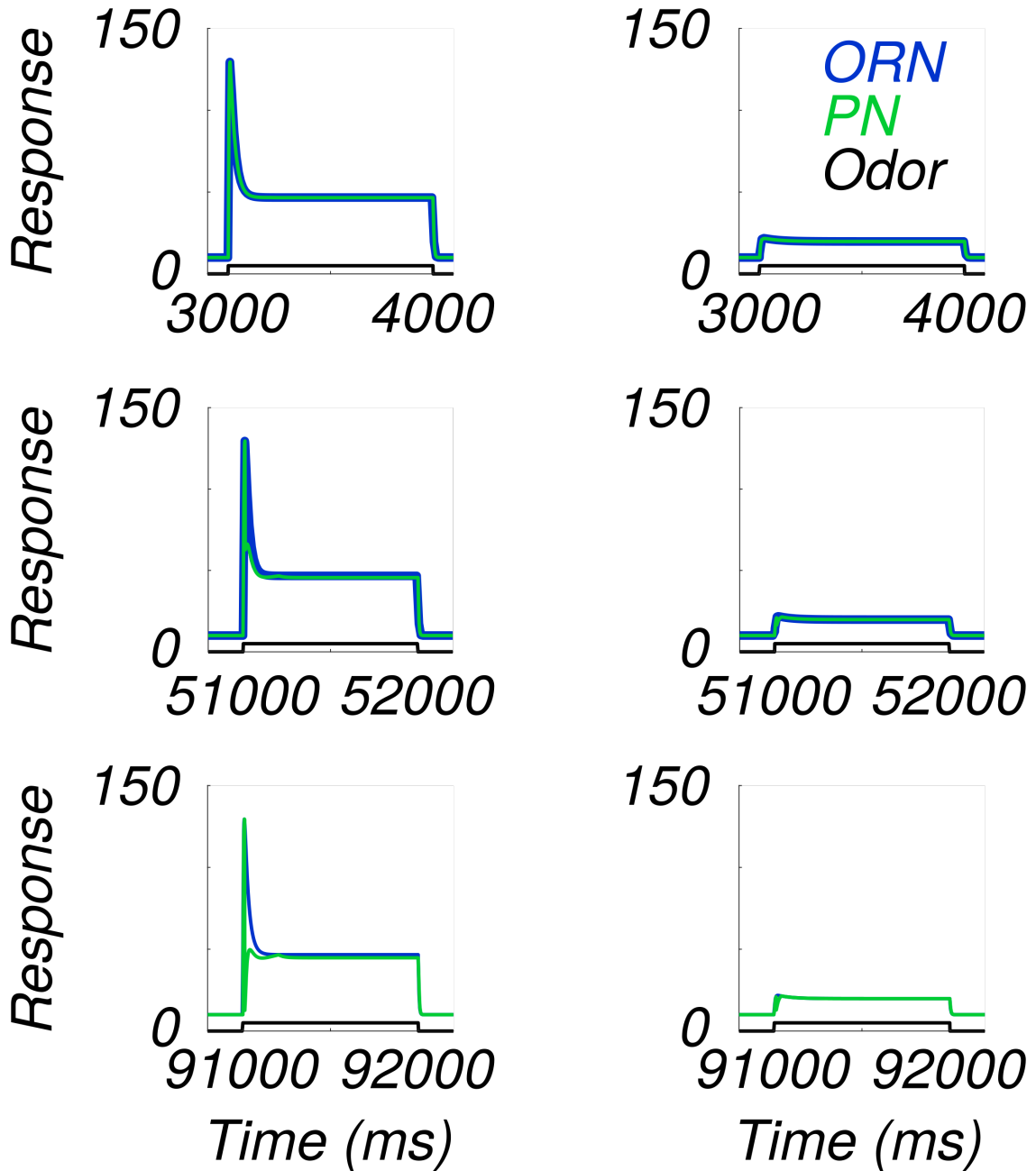
6 PN results

Two sample PNs, from a set of four, and using the delta basis set for the LN population are shown in Figure 11. For reference, ORN inputs are also shown. The top row depicts the PN response at the first odor presentation. At that time, weights have yet to change from their initial values. As a result, the initial PN response closely reflects the dynamics of the ORN input.

The middle row shows the response to the thirteenth odor presentation, and the bottom row depicts the twenty-third odor presentation (presentations for display were chosen arbitrarily). Because of the imposed delay in the model, the initial ORN transient cannot be inhibited - LNs are not yet active for the first ten ms of the response. Following this delay, for both the middle and bottom row, the effects of LN inhibition can be seen in the weakened onset response - the onset transient is no longer as strong as it is in the ORN inputs.

Figure 12 depicts PN responses in the same manner as in Figure 11, but using the gaussian basis set for the LNs. Results are qualitatively similar, with a noticeable difference being the weakening of the early onset transient - the non-zero width of the gaussians shortens the lag (the first LN gaussian is centered at the tenth ms after onset - the same as the first delta function).

Figure 11: **PN responses, early, middle, and late - delta LN basis set** The each column depicts three different responses of a different ORN-PN pair to an odor pulse. The left column has a strong onset transient with fast decay, the right column has a weaker onset transient and slower decay. The first row is the first odor presentation, the middle row is the thirteenth, and the bottom row is the twenty-third. The ORN responses are the same within a given column. Odor strength (in black) is in arbitrary units.



Regardless of the LN basis set used, the PN response rate fluctuates around the eventual steady state, before slowly converging to this steady state. Before this steady state time, inhibition strength fluctuates, driven by the relation between recent ORN firing rates $r_{ORN}(t)$ and the rolling average $\bar{r}_{ORN_n}(t - \tau : t)$.

These slow fluctuations in PN firing rate end once the ORN firing rate has been relatively stable for τ time bins (ORN rates continuously decay, but the change is negligibly small). Deviations from the mean ORN firing rate are small; hence the weights are too weak to substantially inhibit the ORN inputs.

6.1 Weights

Figure 13 shows the weights after training for both basis sets. Because weights were unrestricted, they continue to grow for each odor presentation. Weight adjustments are partially a function of the deviation of the input ORN firing rate (that is, from a single ORN) from the population average across the last τ ms. Hence the weights during the initial odor response grow at a much faster pace than those from the latter part of the odor stimulus.

7 Discussion

We set out to investigate whether a system implementing a simple learning rule could learn to weaken onset transients, and generate a circuit output which more closely matched the original input signal. Our model succeeded at generating a simulated PN signal that captures a number of the features seen in actual PN responses [7], and it did so using a minimal set of parameters. To claim that this model represents the solution used *in vivo* would be a stretch (and is almost certainly untrue - weight values are extraordinarily high, and continue to grow across repeat presentations); nevertheless, it is possible that the principles driving our model - namely a Hebbian weight adjustment scheme, whereby LN inhibition learns to account for onset transients by becoming more negative - may be functionally similar to those seen in biology. The above-mentioned example of the weakly electric fish is evidence of a similar solution to a parallel problem. More physiological and modeling work will need to be done before any strong claims can be made; for now, it suffices to say that we have failed to disprove this model as a potential solution to the problem.

Figure 12: **PN responses at three time points - gaussian LN basis set** The left column depicts a model ORN with a strong onset transient; the right column depicts an ORN model with a weak onset transient. Odor strength (in black) is in arbitrary units.

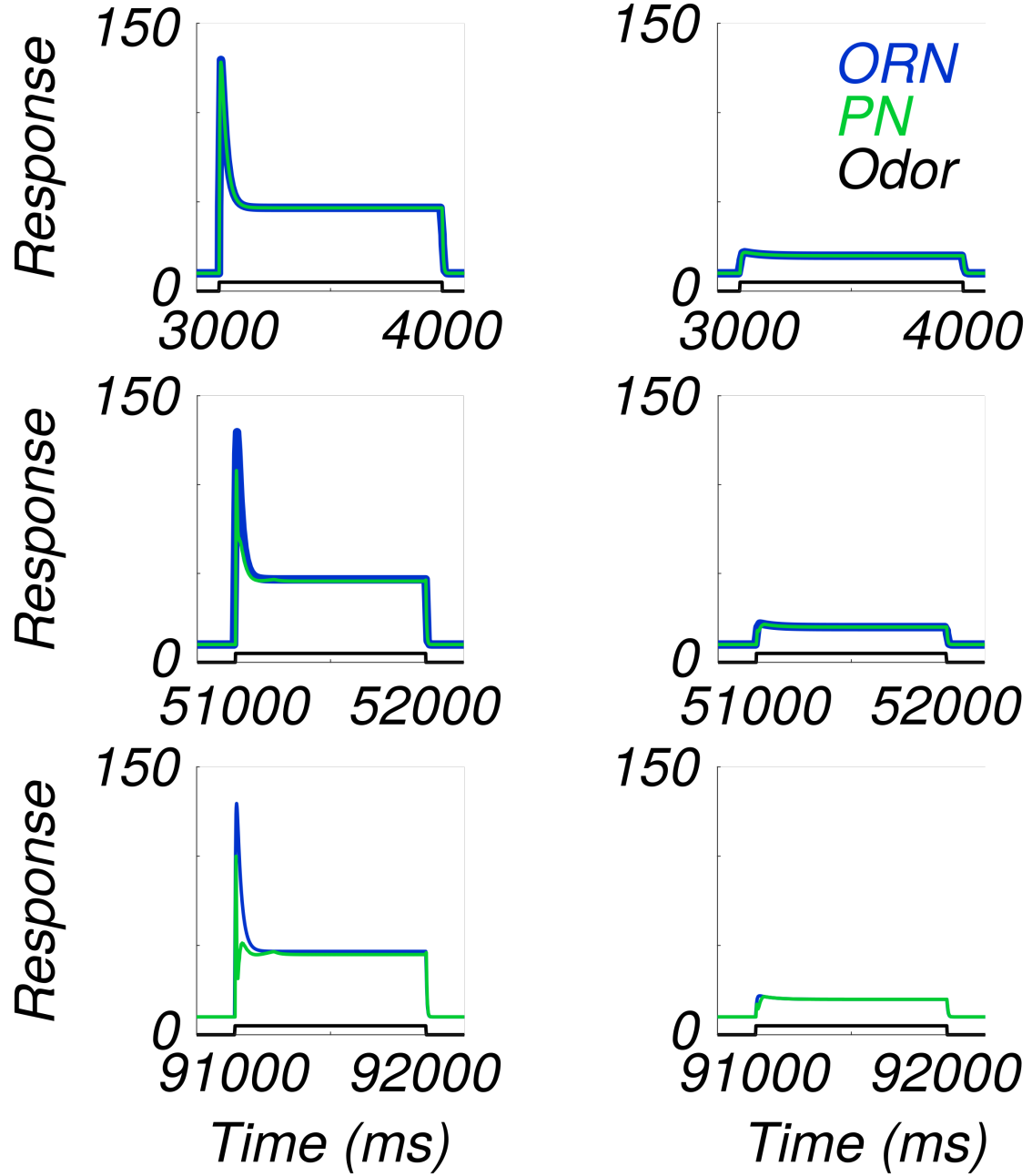
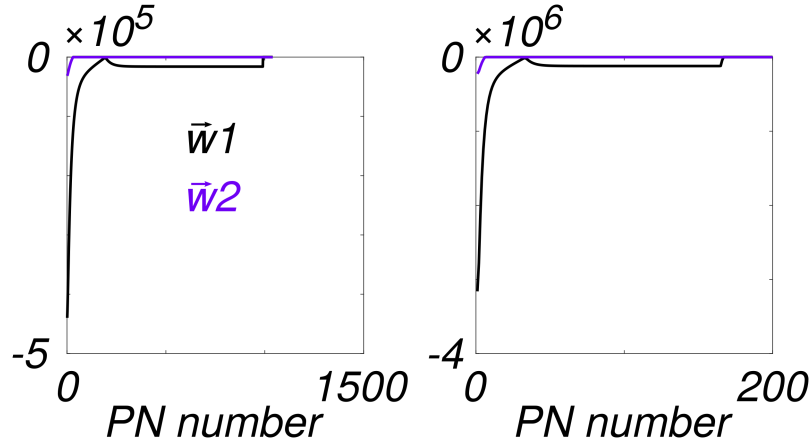


Figure 13: **LN weight vectors after a 120s simulation (29 odor presentations)** The left panel depicts weights using the delta function LN basis set; the right panel depicts weights using the gaussian LN basis set. The black trace depicts the left column of Figures 11 and 12, while the purple trace depicts the ORN from the right column (with the weaker ORN onset transient).



7.1 Problems and inaccuracies

In the creation of such a simplistic model, a number of tradeoffs had to be made. Chief among them is the unrealistic nature of the inhibition - in our model, we had a scalar adjust inhibition to a level that was "just-right", strong enough to measurably affect the signal, while not reducing the signal to zero. This was very much a hand-tuned quality. The values used for β were chosen to be in this ideal window. A model where inhibition was either inherently balanced with the strength of excitation, or at least a physiological justification for the mechanism of inhibition would be better moving forward. Additionally, inhibition in the ORN-PN synapse seems to be presynaptic and divisive[7; 8]. Our model unrealistically represented ORN-PN inhibition as both postsynaptic and subtractive. The scale of the weights also varied across five or six orders of magnitude, which is again unrealistic, but may be solved by a switch to a model with divisive inhibition.

References

- [1] Curtis C Bell. An Efference Copy Which is Modified by Reafferent Input. *Science*, 214(4519):450–453, 1981.
- [2] Curtis C Bell, Angel Caputi, Kirsty Grant, and Jacques Serrier. Storage of a sensory

- pattern by anti-Hebbian synaptic plasticity in an electric fish. *Proc. Natl. Acad. Sci. U. S. A.*, 90(May):4650–4654, 1993. doi: 10.1073/pnas.90.10.4650.
- [3] Curtis C Bell, Victor Z Han, Yoshiko Sugawara, and Kirsty Grant. Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature*, 405(June), 2000. doi: 10.1038/nature01070.1.
- [4] Ann Kennedy, Greg Wayne, Patrick Kaifosh, Karina Alviña, L F Abbott, and Nathaniel B Sawtell. A temporal basis for predicting the sensory consequences of motor commands in an electric fish. *Nat. Neurosci.*, 17(3):416–22, 2014. doi: 10.1038/nn.3650.
- [5] Katherine I Nagel and Rachel I Wilson. Biophysical mechanisms underlying olfactory receptor neuron dynamics. *Nat. Neurosci.*, 14(2):208–216, 2011. doi: 10.1038/nn.2725.
- [6] Katherine I Nagel and Rachel I Wilson. Mechanisms Underlying Population Response Dynamics in Inhibitory Interneurons of the Drosophila Antennal Lobe. *J. Neurosci.*, 36(15):4325–4338, 2016. doi: 10.1523/JNEUROSCI.3887-15.2016.
- [7] Katherine I Nagel, Elizabeth J Hong, and Rachel I Wilson. Synaptic and circuit mechanisms promoting broadband transmission of olfactory stimulus dynamics. *Nat. Neurosci.*, 18(1):56–65, 2015. doi: 10.1038/nn.3895.
- [8] Shawn R Olsen, Vikas Bhandawat, and Rachel I Wilson. Divisive normalization in olfactory population codes. *Neuron*, 66(2):287–299, 2010. doi: 10.1016/j.neuron.2010.04.009.
- [9] Rachel I Wilson. Early olfactory processing in Drosophila: mechanisms and principles. *Annu. Rev. Neurosci.*, 36:217–41, 2013. doi: 10.1146/annurev-neuro-062111-150533.