

Honors Independent Study Paper

Examining Neural Coding and Jitter

Methods

An Introduction

Paper By: Matthew Kulec

Worked with: Professor Asohan Amarasingham and Anthony Alston

December 21, 2018

1 Neural Coding with Techniques and Theory

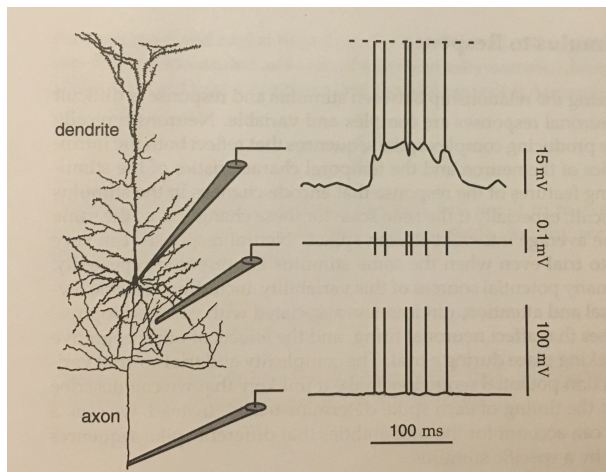


Figure 1: Image courtesy of Dayan and Abbott, 2001

Neuroscience is very well defined as a “fill in the blanks” kind of science. Indeed, much is unknown about the brain and assumptions are made to simplify analysis. It is best to examine the underlying biological process of a neuron to explain why spike timing is the most useful information that can be extracted. A neuron operates by sending electrical impulses along its dendrites and axons until it reaches its axon terminal where between the presynaptic neuron and the postsynaptic neuron an exchange of neurotransmitters takes place (Dayan and Abbott, 2001). These neurotransmitters cause ion channels to open in the postsynaptic neuron and the process repeats itself, with a rising potential difference that if reached above a threshold level, triggers a spike to be generated (Dayan and Abbott, 2001). The most important thing of interest to a neuroscientist are the spikes themselves, rather than the measure of a neuron’s potential. The image above depicts electrodes (which are incorporated to gain data on a neuron’s firing) being inserted at various points along the neuron. The electrode at the axon has an associated graph where spikes, or sharp jumps in the potential (action potential), are clearly visible as opposed to the subthreshold potential. Essentially, the subthreshold potential is severely attenuated along the neuron and what is left are the spikes

themselves (Dayan and Abbott, 2001). Only spikes propagate regeneratively down axons (Dayan and Abbott, 2001) which makes them a suitable object of research. Although there are differences in action potentials when it comes to duration, amplitude, and shape, they are generally ignored and the duration is treated as if it is instantaneous (in practice this figure is about 1 ms) (Dayan and Abbott, 2001). A collection of spikes (whether binary or not) forms a spike train, or spike sequence.

A very important assumption made about spike times in research is whether to consider the independent spike code or the correlation code, and by “code” it is meant neural code. In other words, do individual spikes encode independently of each other or dependently (referenced respectively) (Dayan and Abbott, 2001). A remark is that the timing of these spikes may be governed by some nonconstant rate function $\lambda(\tau)$ at which spikes occur. If the spikes are indeed independent, then to specify matters more succinctly, the interspike intervals, or the waiting times between two consecutive neurons, are independent of one another because all the information is encapsulated in the timing of these spikes (Gallager, 2011). Under this assumption, these interspike intervals all follow an exponential distribution with probability $\exp(-\int_0^T \lambda(\tau) d\tau)$ (Gallager, 2011) where T here represents the random variable for the length of the interspike interval. The following section is part of a full proof (proof is original but incorporates elements from Gallager, 2011).

Some additional remarks to make is that a spike train $0 \leq \xi_1 \leq \xi_2 \leq \xi_3 \leq \dots \xi_i \leq \dots \leq \Xi$ of period Ξ is, by definition, an arrival process because the times ξ_i are themselves random variables in an increasing sequence (Gallager, 2011). Also, an arrival process is best complemented with a counting process $N(y)$ for $y > 0$ which denotes the count of arrivals (spikes) by time y (Gallager, 2011). Let's represent T_i as the random variable for the length of an interspike interval between ξ_i

and ξ_{i-1} . To illustrate,

$$\sum_{i=1}^k T_i = \xi_k$$

Focusing on some time y_0 and assume there is a “wiggle room” α on the time domain such that

$0 \leq y_0 + \alpha \leq \xi_1$ or equivalently $0 \leq y_0 + \alpha \leq T_1$. Then,

$$\begin{aligned} \Pr\{T_1 > y_0 + \alpha | N(y_0) = 0\} &= \Pr\{T_1 > y_0 + \alpha | T_1 > \alpha\} = \frac{\Pr\{T_1 > y_0 + \alpha \cap T_1 > y_0\}}{\Pr\{T_1 > y_0\}} \\ &= \frac{\Pr\{T_1 > y_0 + \alpha\}}{\Pr\{T_1 > y_0\}} = \frac{\exp\left(-\int_0^{y_0+\alpha} \lambda(\tau) d\tau\right)}{\exp\left(-\int_0^{y_0} \lambda(\tau) d\tau\right)} = \exp\left(-\int_{y_0}^{y_0+\alpha} \lambda(\tau) d\tau\right) \end{aligned}$$

The first equality follows because if $T_1 > \alpha$ then certainly by time α , we have yet to await a spike;

they correspond to the same event. The third equality results because $\{T_1 > y_0\} \subseteq \{T_1 > y_0 + \alpha\}$.

The fourth equality stems from the fact that there are $n = 0$ arrivals in $N(y) = n$ where $N(y) = n$ is a Poisson random variable with probability:

$$\Pr\{N(y) = n\} = \frac{1}{n!} \left(\int_0^y \lambda(\tau) d\tau \right)^n \exp\left(-\int_0^y \lambda(\tau) d\tau\right)$$

Now for any T_{i+1} where $\xi_i \leq y_0 + \alpha \leq \xi_{i+1}$, it can be argued that:

$$\begin{aligned} \Pr\{T_{i+1} > y_0 + \alpha - \xi_i | N(y_0) = i, \xi_i\} &= \Pr\{T_{i+1} > y_0 + \alpha - \xi_i | T_{i+1} > y_0 - \xi_i, \xi_i\} \\ &= \Pr\{T_{i+1} > y_0 + \alpha - \xi_i | T_{i+1} > y_0 - \xi_i\} = \frac{\Pr\{T_{i+1} > y_0 + \alpha - \xi_i \cap T_{i+1} > y_0 - \xi_i\}}{\Pr\{T_{i+1} > y_0 - \xi_i\}} \\ &= \frac{\Pr\{T_{i+1} > y_0 + \alpha - \xi_i\}}{\Pr\{T_{i+1} > y_0 - \xi_i\}} = \frac{\exp\left(-\int_{\xi_i}^{y_0+\alpha} \lambda(\tau) d\tau\right)}{\exp\left(-\int_{\xi_i}^{y_0} \lambda(\tau) d\tau\right)} = \exp\left(-\int_{y_0}^{y_0+\alpha} \lambda(\tau) d\tau\right) \end{aligned}$$

Therefore by induction, any interspike interval T_i shares the same probability. The only difference there is, is due to the rate function $\lambda(t)$, but this relates to them being independent of one another because this derivation holds for any y_0 , even when $y_0 = \xi_i$ and because of equality four. So if the interspike intervals are independent, then their probability distribution must be the same. Equality one holds because if by time y_0 , there are i arrivals, then surely T_{i+1} in length must be greater than

$y_0 - \xi_i$. The ξ_i arrival must exist when you have the event $\{T_{i+1} > y_0 - \xi_i\}$ so equality two holds. There are no arrivals in an interspike interval $(\xi_i, y_0 + \alpha]$ and $(\xi_i, y_0]$ satisfying equality five. A final remark is that because these probabilities are the same, the memoryless property is used (Gallager, 2011). That is, given that an arrival has not yet happened after some time, the remaining waiting time is said to have no “memory” of its previous waiting (Gallager, 2011); a shift in time doesn’t change the probability. To make this intuitively sound, picture a store owner waiting for his next customer. If this process is indeed memoryless, him waiting some time s puts him in no better position than expecting the next customer at time 0. Interestingly, it turns out that the only continuous random variable that is memoryless turns out to be the exponential (Gallager, 2011). An arrival process whose distribution is the one just derived is called a *non-homogeneous Poisson process*. To conclude, the probability distribution for each interspike interval need not be the same if the correlation code is assumed because then there is extra information that must be accounted for. Most research revolves around the independent spike coding assumption because it is more robust and convenient (Dayan and Abbott, 2001). In reality, it is found that the information carried by correlations between two or more spikes is rarely larger than 10% of the information carried by the spikes when considered independently (Dayan and Abbott, 2001).

To obtain simulated spike times for theoretical purposes given some rate function $\lambda(t)$, this is similar to simulating coin flips on a computer. Since flipping coins is essentially a Bernoulli procedure, the expected number of heads in a coin flip is $1/2$. Therefore, we can use the algorithm: if $1/2 \geq U_{[0,1]}$ then heads. Here, $U_{[0,1]}$ is a sample from the uniform distribution on the interval $[0, 1]$. Using a similar approach, we must find the expected number of arrivals in a time bin of infinitesimal length δ . This is because a Poisson process is best described as a “limiting version” of the Bernoulli procedure. The probability distribution for the number of “successes” for some finite number

of M independent Bernoulli experiments is the binomial distribution. As $M \rightarrow \infty$ and $p \rightarrow 0$, the distribution becomes Poisson. The number of arrivals $N(y_0)$ by time y_0 follows a Poisson distribution (Gallager, 2011). Therefore,

$$\Pr\{N(y_0 + \delta) - N(y_0) = 0\} = \exp(-\lambda(t)\delta) = 1 - \lambda(t)\delta + \frac{\lambda^2(t)}{2}\delta^2 + \dots = 1 - \lambda(t)\delta + o(\delta)$$

$$\Pr\{N(y_0 + \delta) - N(y_0) = 1\} = \lambda(t)\delta \exp(-\lambda(t)\delta) = \lambda(t)\delta - \lambda^2(t)\delta^2 + \frac{\lambda^3(t)}{2}\delta^3 + \dots = \lambda(t)\delta + o(\delta)$$

$$\Pr\{N(y_0 + \delta) - N(y_0) \geq 2\} = o(\delta)$$

Where the $o(\cdot)$ notation represents the part of the expansion when $\lim_{\delta \rightarrow 0} \frac{o(\delta)}{\delta} = 0$. As $\delta \rightarrow 0$, these terms virtually disappear when compared to δ . We find that the expectation is:

$$\mathbb{E}[N(y_0 + \delta) - N(y_0)] = \sum_{n \geq 0} n \Pr\{N(y_0 + \delta) - N(y_0) = n\} = \lambda(t)\delta$$

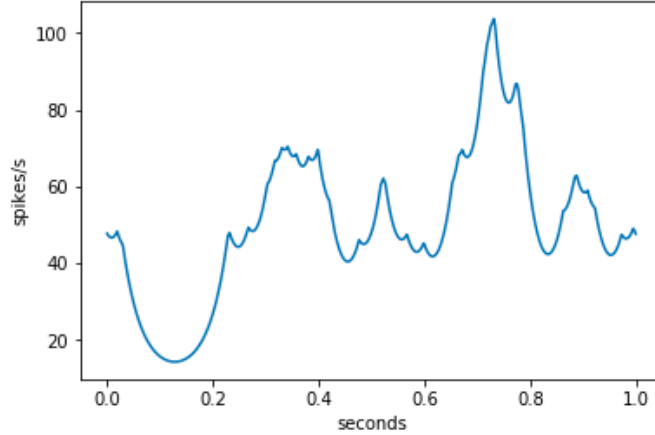
So, the algorithm for generating spike times will be: if $\lambda(t)\delta \geq U_{[0,1]}$, then a spike occurs.

An example of a theoretical rate function (Amarasingham et al., 2012) is as follows:

$$f_k(t) = \left[10 + \sum_{j=1}^{40} \sum_{l=-\infty}^{\infty} \frac{1}{2\sigma/\sqrt{2}} \exp\left(-\frac{|t+l-\mu_{j,k}|}{\sigma/\sqrt{2}}\right) \right] \mathbb{I}\{t \in [0, 1)\}$$

Where k indicates the trial number, $\mu_{1,k}, \dots, \mu_{40,k}$ that are IID samples from the uniform distribution on the interval $(0, 1)$. Because of the $\mu_{j,k}$ parameter, this Poisson rate function is stochastic and thus a *Cox process* (Amarasingham et al., 2012). This parameter is important because neurons are known to exhibit bursting, a phenomena where neurons fire rapid spikes for some time period. Bursting can be accounted for by allowing $\lambda(t)$ to fluctuate in order to represent the high firing rate that is associated with a burst (Dayan and Abbott, 2001) and that is what these samples are responsible for. In bursting, there is a dependence between spikes because given that the neuron is experiencing bursting, the chance of another spike occurring given that one already happened is high. Since $\lambda(t)\delta$ is a probability, then $\lambda(t)$ must be the probability density. Below is a graph of

the function over a single trial (any k) over the interval $[0, 1]$ because it is zero otherwise due to indicator function. Notice how the peaks are sometimes “rough,” for lack of a better word. This depicts how model incorporates bursting because we can pick adjacent δ ’s under the part of the graph such as $[0.35, 0.4]$ and evaluate their $\lambda(t^*)$ at some t^* within their δ ’s, and we find that these probabilities are relatively high and comparable.



Another assumption (now comparing the neurons themselves rather than their spikes) is the independent neuron coding hypothesis. It states that when considering a population of neurons, they are statistically independent (Dayan and Abbott, 2001). That is, there is no shared information between 2 or more neurons. Of course, they can act dependently, in which case the hypothesis fails. It turns out that synchrony, short for (nearly) synchronous firing, between two or more neurons can be a suitable measure for testing the merit of this hypothesis (Dayan and Abbott, 2001). A function that measures the number of synchronies, or spike pairs, is known as a *cross-correlation histogram* (Amarasingham et al., 2012) and is given by:

$$c(\tau) = \sum_{(l,k) \in N_1 \times N_2} \mathbb{I}\{|l - k - \tau| \leq 0.001\} = \sum_{(l,k) \in N_1 \times N_2} \mathbb{I}\{|l - (k + \tau)| \leq 0.001\}$$

Where $(l, k) \in N_1 \times N_2$ represents all the possible pairs of spike times in the cartesian product of sets of spikes (spike trains) for 2 neurons. The second equality is to emphasize a shift, or lag τ ,

for one of the spike times. So, essentially, the summand adds one whenever the delay between one spike time and another shifted spike time is within 0.001 seconds. Recall that a spike's duration is roughly 1 ms (Dayan and Abbott, 2001). Therefore, allowing at most 1 ms of overlap is a reasonable figure to work with.

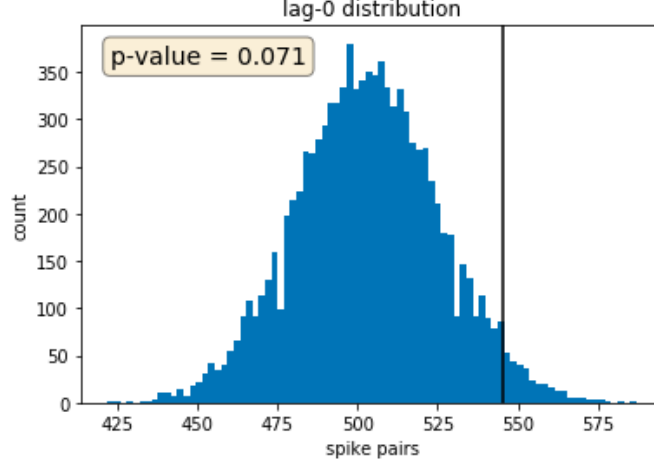
We can test the independent neuron coding assumption for two neurons with a hypothesis test. We make an inference about our data, or the null hypothesis H_0 while H_1 offers an alternate hypothesis in case our assumptions are proven wrong (Pishro-Nik, 2014). Hypothesis tests involve functions that take in the data which are called test statistics. The only possibilities for error are when we reject H_0 , but in reality it is the correct hypothesis, or if we accept H_0 when it is indeed false. The former is called a *type 1 error* while the latter is called a *type 2 error* (Pishro-Nik, 2014). The probability for a type 1 error is at most α . This is called the *significance level*, and it sets a line between rejection of H_0 and a failure to reject H_0 (Pishro-Nik, 2014). Canonically speaking, $\alpha = 0.05$, though in some cases, a lower value may be desired. It turns out that there is a tradeoff between the two errors. There is also a lower bound for the value of α and this is the p-value (Pishro-Nik, 2014). P-values are usually depicted graphically as a vertical line placed on some value of the domain where a gaussian resides. P-values at the tail end of the normal distribution signify a rejection of the null hypothesis H_0 because then the probability $\Pr\{\text{type 1 error}\} = \Pr\{\text{type 2 error}\}^c = \alpha$ is very small. In other words, due to the trade off, the probability that we get fooled and accept H_0 when it is indeed false is high. Likewise when α is large (usually when $\alpha > 0.05$), the probability that we reject H_0 when it is indeed true is high, so we shouldn't reject it.

To test for independence between two theoretical spike trains, it can be proposed that H_0 be the null hypothesis that the spike trains are independent, while H_1 be otherwise (Amarasingham

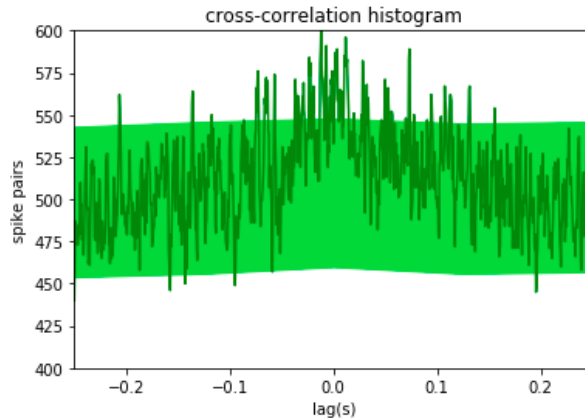
et al., 2012). We can generate spike times with the Poisson Process spike generation algorithm from $k = 100$ concatenated rate functions (both previously mentioned) so that $\Xi = 100$ seconds (Amarasingham et al., 2012). We do procedure twice so that we yield N_1 and N_2 . The cross-correlation histogram will serve as the test statistic. In addition, create $M = 9999$ surrogate spike trains where you pick N_2 , say, and shuffle the spike times grouped by their trial ($[l]$ is an easy way to identify what trial spike time l in N_2 belongs to because each trial is only 1 second long) so that the end result are M of these permuted versions of N_2 (Amarasingham et al., 2012). Since cross-correlations care about order (an unintended shape will happen if we ignore, otherwise), the integer components for each element in each permuted version of N_2 must be dropped and the same integer components in the original N_2 are attached element-wise. Compute the cross correlation at $\tau = 0$ for every pair $(N_1, N_2), (N_1, N_2^{(1)}), \dots, (N_1, N_2^i), \dots, (N_1, N_2^M)$ where i is a marker for every permuted version of N_2 . The choice for a p-value is then a valid one (Amarasingham, 2012):

$$\text{p-value} = \frac{1}{M+1} \sum_{m=0}^M \mathbb{I}\{c_m(0) \geq c_0(0)\}$$

The subscript m in c_m denotes every cross-correlation function. If $m = 0$, then this is the original cross correlation function. Whereas if $1 \leq m \leq M$, then this is the cross correlation of the original, uncopied N_1 and some permuted version of N_2 . What this equation is evaluating is the proportion of times synchrony by chance is greater than the observed synchrony (Amarasingham et al., 2012). Intuitively speaking, if it is greater than for a great proportion of times ($> \alpha$), then we should expect independence since there is nothing special between our two “neurons.” On the other hand, if the proportion is reasonably small ($\leq \alpha$), then due to a higher than expected synchrony there must be something special between our two “neurons” (or formally speaking, a dependence). Below is a histogram for the values of the M cross-correlation histograms and the corresponding p-value.



Note that the choice of α will dictate whether we reject H_0 or fail to reject it. If we use $\alpha = 0.05$, then with a p-value of 0.071 such that $0.071 > 0.05$ signifies a failure to reject H_0 , or that the 2 “neurons” with spike trains N_1 and N_2 are independent. However, if we chose $\alpha = 0.10$ so that $0.10 > 0.071$, then this is a rejection of H_0 by that standard, and so our 2 “neurons” are dependent. We may also desire a plot of $c(\tau)$ with *pointwise acceptance bands* $[A(\tau), B(\tau)]$ (Amarasingham et al., 2012) such that $A(\tau)$ (lower bound) is the 0.025Mth element in a sorted (by value) list of the cross correlation histogram functions $c_0(\tau), c_1(\tau), \dots, c_M(\tau)$. Likewise, $B(\tau)$ (upper bound) represents the 0.975Mth element of the same sorted list. The reason for 0.975 and 0.025 is because $0.975 - 0.025 = 0.95$. In other words, there is a 95% *confidence interval*. Below is a plot of $c_0(\tau)$ for $\tau \in [-0.25, 0.25]$ along with these acceptance bands.



An open question in neuroscience is at what scale, or *temporal resolution*, must spike times be observed to extract the most useful information (Dayan and Abbott, 2001)? What makes this problem really hard is the fact that it is an extension of the question: are spikes themselves independent? (i.e. the independent spike coding hypothesis) (Dayan and Abbott, 2001). This is because temporal structure includes information about the dynamics of an observed stimulus and the nature of neural encoding (Dayan and Abbott, 2001). Suppose that we are not interested in whether 2 neurons are independent or dependent, rather if there is structure at some chosen scale. The interval jitter method (Amarasingham et al., 2012) allows us to test the null hypothesis H_0 that there is no temporal structure at a scale finer than Δ ms. This parameter Δ is coming from the procedure used to generate jittered versions (instead of permuted ones) of a spike train:

$$\left(\xi_1^{(m)}, \dots, \xi_i^{(m)}\right) = \text{sort} \left(\Delta \lfloor \xi_1 / \Delta \rfloor + \Delta U_{[0,1],1}^{(m)}, \dots, \Delta \lfloor \xi_i / \Delta \rfloor + \Delta U_{[0,1],i}^{(m)} \right)$$

Where ξ_g such that $1 \leq g \leq i$ represents a spike time in a train of length i . The (m) indicates the m th jittered version where once again $1 \leq m \leq M$. So what this is saying, is, bin the interval $[0, \Xi]$ so that bins are obtained of some length Δ . $\Delta \lfloor \xi_i / \Delta \rfloor$ moves the the i th spike in a spike train (of some neuron) to the start of its bin of length Δ (Amarasingham et al., 2012), while $\Delta U_{[0,1],i}^{(m)}$ “jitters,” or moves the spike within that bin uniformly. The ordering of the spikes is kept the same, but now spikes experience a “noise.” Once these jittered versions are created, the rest of the hypothesis testing follows the same procedure as the permutation test described previously. To state the null hypothesis H_0 more rigorously, would be saying that (Amarasingham et al., 2012)

$$\Pr \left\{ \boldsymbol{\xi} = \{\xi_1, \dots, \xi_i\} | N(\boldsymbol{\xi}) = \mathbf{n}, |N| = \frac{\Xi}{\Delta} \right\} = \frac{\mathbb{I}\{N(\{\xi_1, \dots, \xi_i\}) = \mathbf{n}\}}{\sum_{\mathbf{y} \subseteq \{0,1\}^i} \mathbb{I}\{N(\mathbf{y}) = \mathbf{n}\}}$$

That is, given some set of binary spikes $\boldsymbol{\xi} = \{\xi_1, \dots, \xi_i\}$ with length i , partition the set into Ξ/Δ pieces (number of jitter intervals) so that the vectorized function $N(\cdot)$ returns the set \mathbf{n} , in which each element is the count for the number of spikes (1’s) of each partitioned piece (thus, the

cardinality of N denoted by $|N|$ must be equal to Ξ/Δ). The denominator is the cardinality of the sample space: the number of all subsets \mathbf{y} of $\{0,1\}^i$ (where $|\{0,1\}^i| = 2^i$) such that $N(\mathbf{y}) = \mathbf{n}$. To summarize, as long as there are n_1 spikes in bin 1, \dots , $n_{\Xi/\Delta}$ spikes in bin Ξ/Δ , the placement of these spikes doesn't matter (i.e. follows a uniform distribution) (Amarasingham et al., 2012). This attests to there being no temporal structure at a scale finer than Δ ms. Another way to see if indeed there is no finer structure than Δ ms is if the synchrony count is less once interval jitter has been used for some Δ .

Lastly, a generalization of the interval jitter method is the pattern jitter method (Amarasingham et al., 2012). This method is excellent for controlling for phenomena as bursts because now bundles of spikes are treated as single entities and they can be interval jittered. In other words, the local statistics are preserved (Amarasingham et al., 2012). What constitutes such an “entity” or *pattern* is given by the following rules. Suppose you have a spike train $\xi_1 \leq \dots \leq \xi_i$ such that the entire spike train fits in the time interval $[0, \Xi]$. Let R be the history parameter. In a burst, R may be the distance that those bursted spikes are at no distance greater than R from each other if they are indeed part of that burst, or equivalently the pattern. A bundle of spike times ξ is a pattern iff $\xi_j - \xi_{j-1} > R$, $\xi_{k+1} - \xi_k > R$, and $\xi_{l+1} - \xi_l \leq R$ where $j \leq l < k$ (Amarasingham et al., 2012). In words this means that a pattern contains spikes that are no more than a distance R from each other, while the entire pattern is separated from other patterns by a distance greater than R . The null hypothesis H_0 then transforms into the following statement: the conditional distribution of ξ given that these rules are satisfied is uniform, similar to that of the interval jitter null hypothesis (Amarasingham et al., 2012).

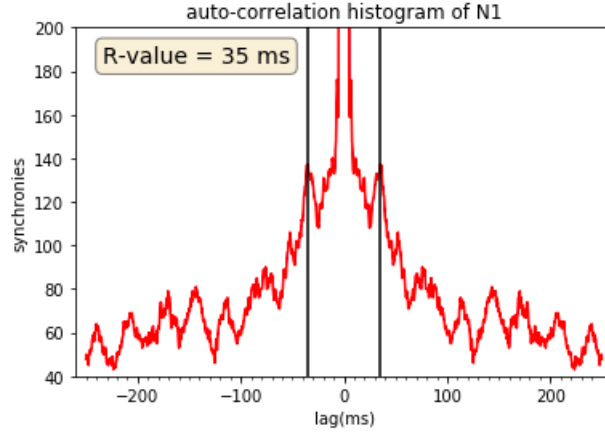
2 Pattern Jitter Modeling: An Application

A demonstration of pattern jitter hypothesis testing can be shown. But first, the spike train data that will be utilized comes from an important paper. To provide context, it is important to state yet another series of related challenges facing the neuroscience community. Especially in the visual cortex, there are a type of cells called *simple cells* and they each have connections to parts of your vision, called a *receptive fields* (Farah, 2000). When presented with a certain stimulus, these simple cells respond with a high firing rate if the edges in their receptive field are located and oriented at a particular angle (Farah, 2000). Each simple cell “prefers” a particular orientation and location (Farah, 2000). The obvious question then remains: how does the brain assemble all these simple cell constituents to recognize an image? Understanding image recognition in the biological sense can help shed light on image recognition in machine learning such as the problem of *invariance versus selectivity* (Geman, 2006). Classification tasks done by a machine excel at recognizing dark handwritten numbers against a white background. However, the background need not always be that uninteresting. Suppose you have a picture with a bird on a tree. The tree has flowers whose colors are close to that of the bird’s feathers. The goal is have a model that is robust to coloring, texturing and contrast (Geman, 2006). For the features that visually characterize the bird, the model should be *selective* to, whereas the features that don’t visually characterize the bird, but the surrounding environment, the model should be *invariant* to (Geman, 2006). In the paper that uses the data of interest to us (Von der Heydt et al., 2015), it is theorized that a type of cell called a *grouping cell* facilitates input from many simple cells (Von der Heydt et al., 2015). Neurons that belong to the same grouping cell are called *consistent* in the sense that they share a *common border-ownership* (Von der Heydt et al., 2015). What this means is similar to the orientation selectivity of simple cells just described. Parts of a figure such as the top of a bird’s

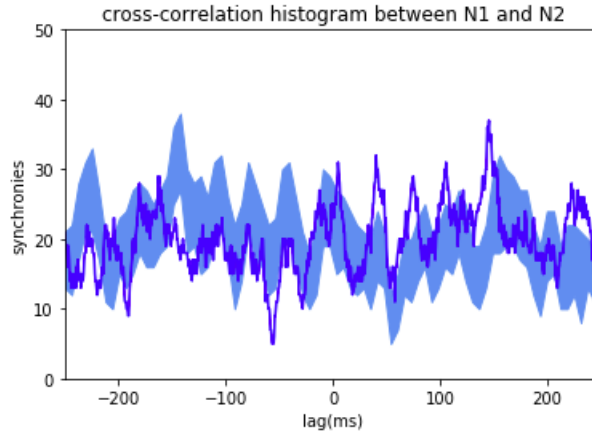
head may appear closer than the tree background, which is occluded. At that point, the border is said to be owned by bird (Von der Heydt et al., 2013). Its wing may also be brought forward, in which case now the edge of that wing may appear closer and thus dominates that border (Von der Heydt et al., 2013). Evidence for such grouping cells can be proven by increased synchrony of neurons that share a common border ownership (Von der Heydt et al., 2015). For our purposes here, which is to propose a simple demonstration, it is fruitful to omit the exact details of the experiment. A general overview is that they presented images to monkeys. Images that had a common figure were classified as the *bound* conditions whereas images with different figures were classified as *unbound* (Von der Heydt et al., 2015). They also control for attention, but they found “no evidence in [their] data for a role of synchrony in coding attention” (Von der Heydt et al., 2015). An important finding was that “consistent pairs showed significant synchrony with binding even for intervals as narrow as 10 ms” (Von der Heydt et al., 2015). Here, “intervals” refers to the fact that they used the interval jitter method for $\Delta = 10\text{ms}$. The question is, would accounting for bursts (by using pattern jitter) affect the synchrony count? We would expect a change in synchrony for $\Delta < 10\text{ ms}$ if there is finer temporal structure. If the null hypothesis H_0 is true, then there is no finer temporal structure and thus no probable change in synchrony. On the other hand, we can say that a part of the “significant synchrony” may be attributed to bursting.

To find a suitable choice for the parameter R , we can construct an autocorrelation function and see at what point $\tau \neq 0$ in the lag domain does a noticeable peak appear. Notice that if we only pick one spike train and compare it with itself, so that N_1 and N_2 are the same spike train, we essentially obtain an autocorrelation. To show just an example, the temporal recordings from two neurons (with spike trains N_1 and N_2) in the consistent, bound and attend condition were selected. Preprocessing the data (provided by Von der Heydt et al, 2015) so that the trial onset times Ω_j at

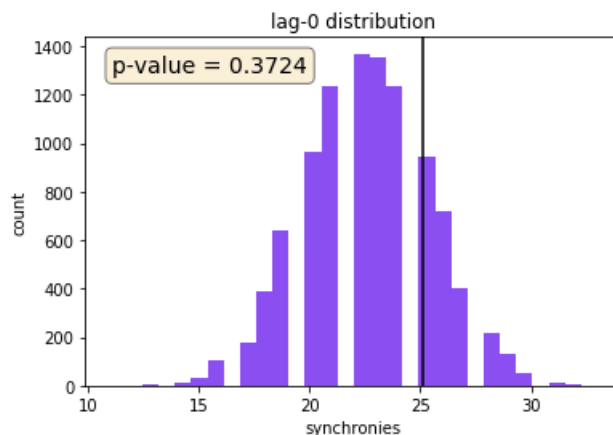
trial j were accounted for in a way that the filtered times were in the range $[\Omega_j + 400, \Omega_j + 1050]$ ms (the 400-1050 time range is when the images were presented), then concatenating over a total of 30 trials left a spike train of increasing sequence such that $\max_i \xi_i \approx 19.5$ seconds and $\min_i \xi_i \approx 0$ seconds. Below is the graph of the autocorrelation with dark lines cutting at the noticeable peaks such that $\tau \neq 0$. The chosen R value is then 35 ms.



This is a plot of the cross correlation histogram with pattern jitter derived pointwise acceptance bands. The value of $R = 35$ ms and $\Delta = 10$ ms was used to create $M = 9999$ pattern jittered versions of N_1 thanks to a special algorithm (Harrison, 2009). These pattern jittered versions was then incorporated in the acceptance bands and for computing the lag-0 distribution.



This is the plot of the lag-0 distribution.



We see that no matter if we choose $\alpha = 0.10$ or $\alpha = 0.05$, the p-value is greater which means that we fail to reject the null hypothesis H_0 that there is no temporal structure at a scale finer than $\Delta = 10$ ms. This means that even for smaller Δ we shouldn't expect smaller counts of synchrony. Of course, the data that was used for this “mini-experiment” isn't large enough which poses problems due to possible noise, but nonetheless the shows what can be accomplished with jitter methods like pattern jitter.