

*The information  
content of neurons  
engaged in  
population temporal  
coding in early  
olfactory circuits*

---

## 4.1 Abstract

It has long been believed that the information content of single neurons in the first olfactory processing nucleus is very low, both because of the overlapping nature of the ensemble representation of odors and because of their low signal to noise ratio, which has been estimated to be as low as  $10^{-4}$  (Freeman, 1990). Here, I take the point of view of the organism and applied to olfaction for the first time an algorithm to identify the odor presented by observing spike trains of projection neurons in the antennal lobe of the locust. The information in one spike train of a single neuron sufficed to identify a stimulus among several presented on up to 95% of all trials. I characterize the timescale of optimal information extraction for two different neural codes: while information in single neurons is robust to a variety of readout temporal scales, including a rate code, the information rate across assemblies of neurons is significantly greater when taking temporal response patterns into account. PN assemblies are shown to be most informative when decoded with a time constant on the order of several hundred milliseconds. This is shown to be due to the burstiness of PN spike trains: The timescale at which PN response patterns are found to vary is on average significantly longer than that previously reported (Wehr and Laurent, 1996). I characterize the information present in assemblies of increasing numbers of neurons. Finally, I characterize the reliability and sparseness of the representation as a function of the timescale of the code with which it is read out.

---

## 4.2 Introduction

Perhaps the foremost task of the early olfactory system is to convey enough information about olfactory stimuli to be able to discriminate between odors with different behavioral relevance. In the olfactory bulb of vertebrates and the antennal lobe of insects, it has long been believed that information is coded by distributed assemblies of neurons. In addition, these neurons have been thought to be very noisy, with signal-to-noise ratios close to  $10^{-4}$  (Freeman, 1990). The information in single neurons has thus been considered insufficient to identify odors. The present work quantifies such information present in single projection neurons (PNs) in the antennal lobe of the locust.

## 4.3 Results

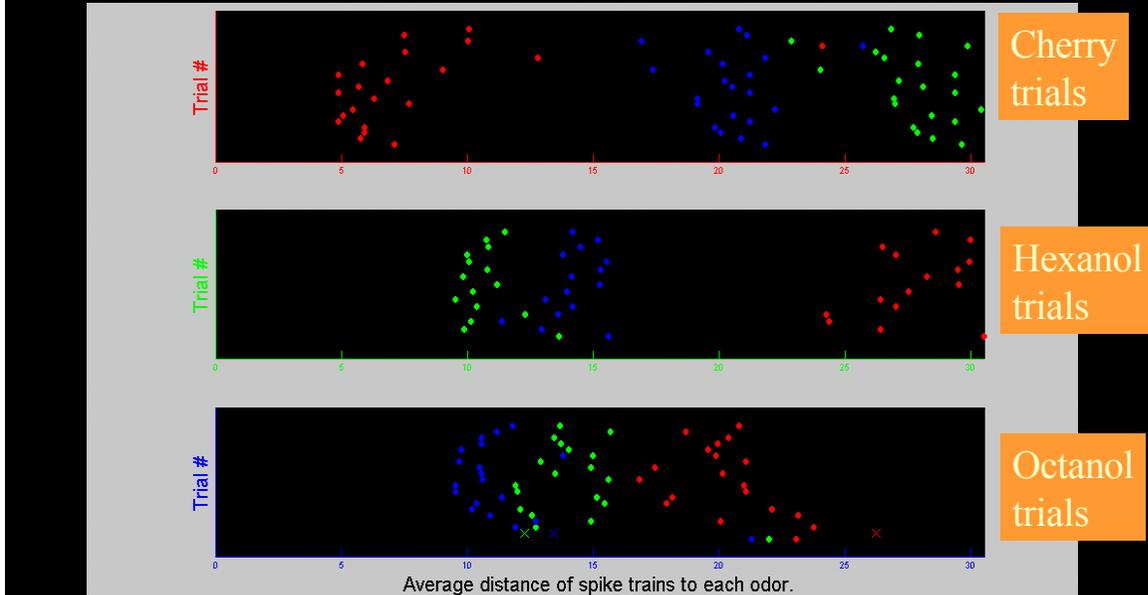
Single PNs allowed correct odor identification in 81% of trials on average when a single concentration of each odor was presented (chance = 50%) ( $n=12$  PNs) and above 95% for several PNs exposed to three odors (chance = 33%) (Fig. 4.1). When multiple concentrations of each odor were presented, the task of the recognition of odor identity was naturally made more difficult. Single PNs nevertheless exhibited enough information to recognize odor identity correctly in up to 50% of all trials on average (chance = 26%) ( $n=46$  PNs) (Fig. 4.2).

To find the duration at which any further information in spike trains becomes redundant with earlier information, I varied the length of the spike trains used for classification systematically. The information content of spike trains saturated at windows of 0.5–1 sec ( $n=13$  PNs, Fig. 4.4). The mean informational content of spike trains for the discrimination of odor quality peaked at odor onset and suffered a fall 500 msec after odor onset ( $n=13$  PNs, Fig. 4.5).

---

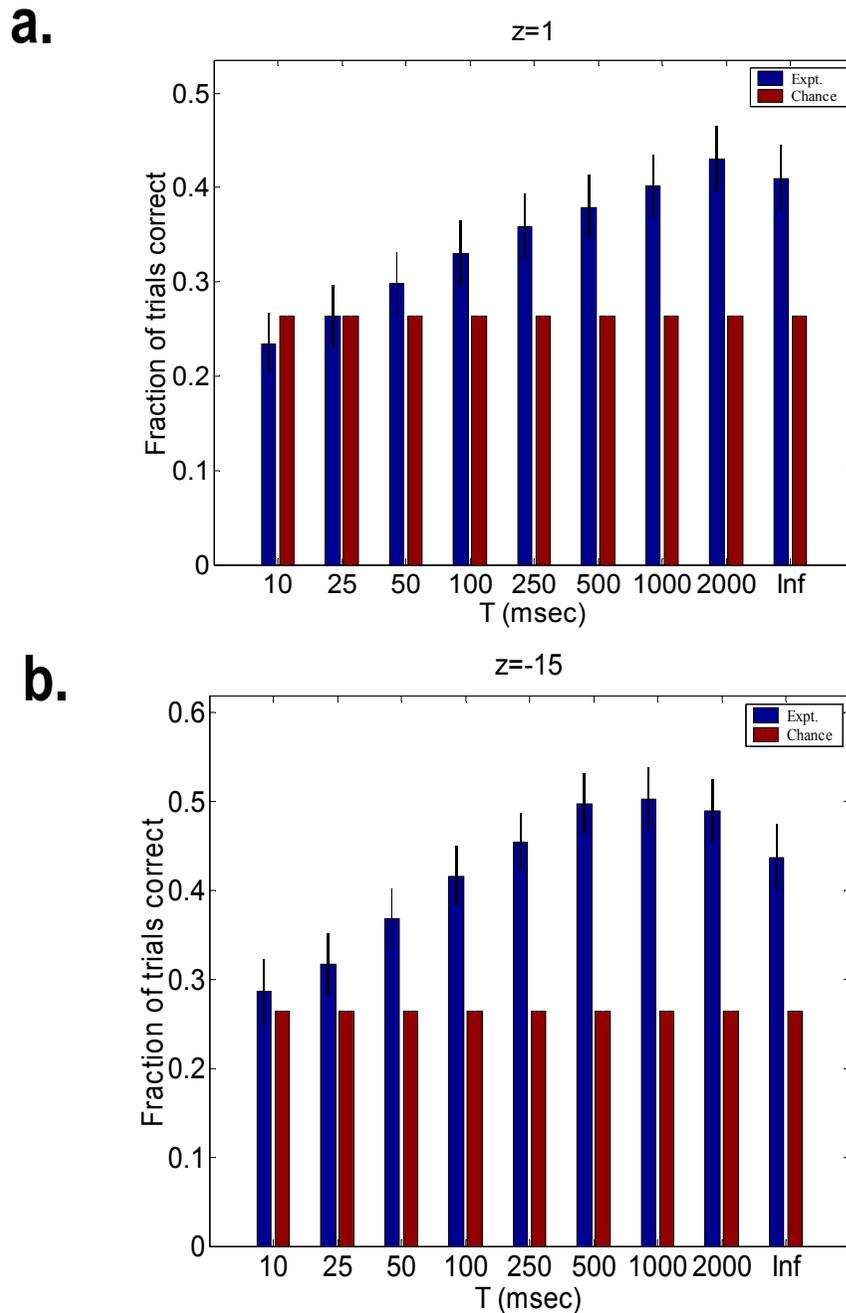
## Odors can be identified reliably based on the output of single neurons (97% correct below)

Distance to cherry trials. Distance to hexanol trials. Distance to octanol trials.

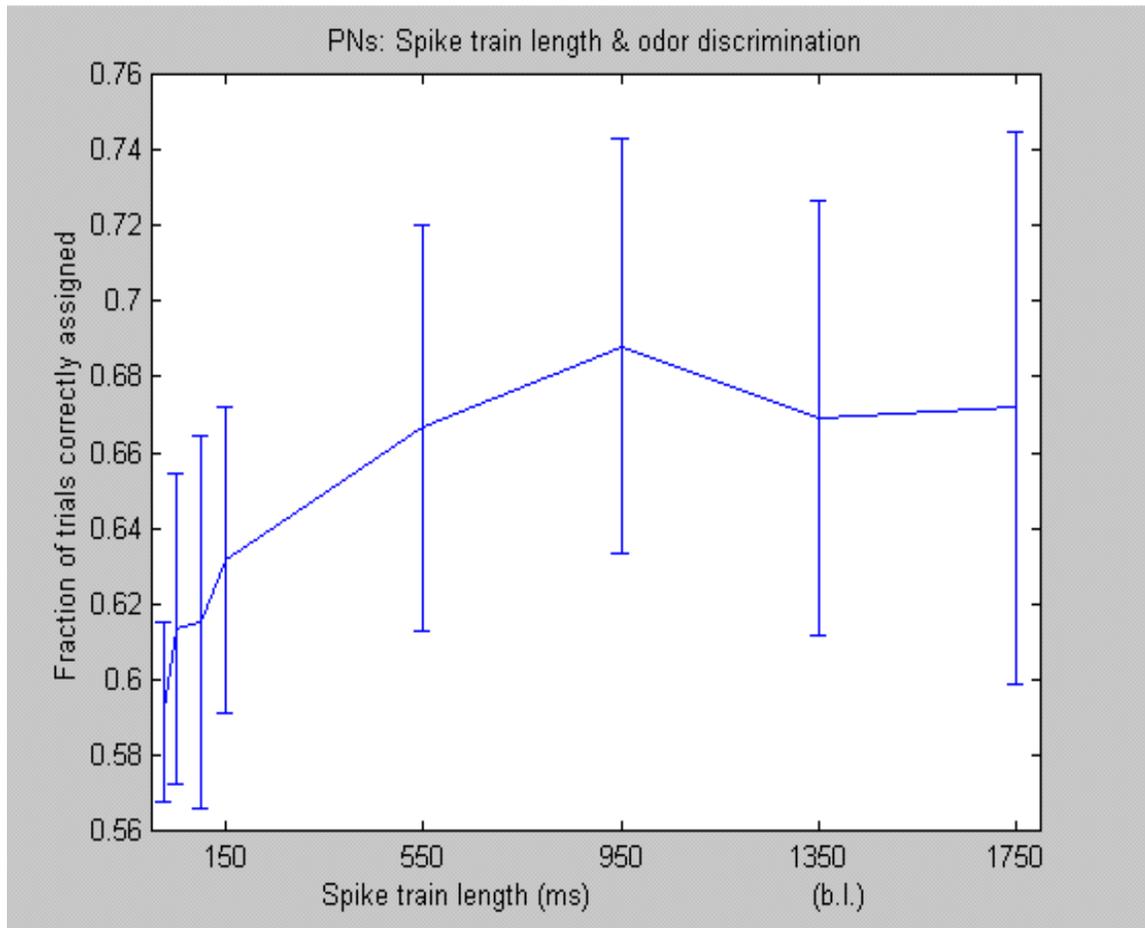


**Figure 4.1. Information in single PNs can reliably identify the odor presented.** Each plot shows all trials for one odor. Each line shows one trial; the first trial of the corresponding series is displayed at the bottom of each plot. The axis for each point represents the mean distance between a spike train and all (other) spike trains for a given odor (class). Trials whose distance to their own class is smaller than that for other classes are correctly classified; others are shown as crosses —only the second octanol trial is misclassified, as hexanol.

Although for individual PNs temporal information could yield significantly better classification than spike counts alone (Fig. 4.5), on average across all PNs tested, information in single neurons was barely greater using temporal information than using spike counts alone (Fig. 4.2). Furthermore, firing rate variations over longer timescales was less variable across trials and concentrations, and therefore classification based on spike count was better than that based on T values on the order of the length of an oscillation cycle (50 ms) (Fig. 4.2).

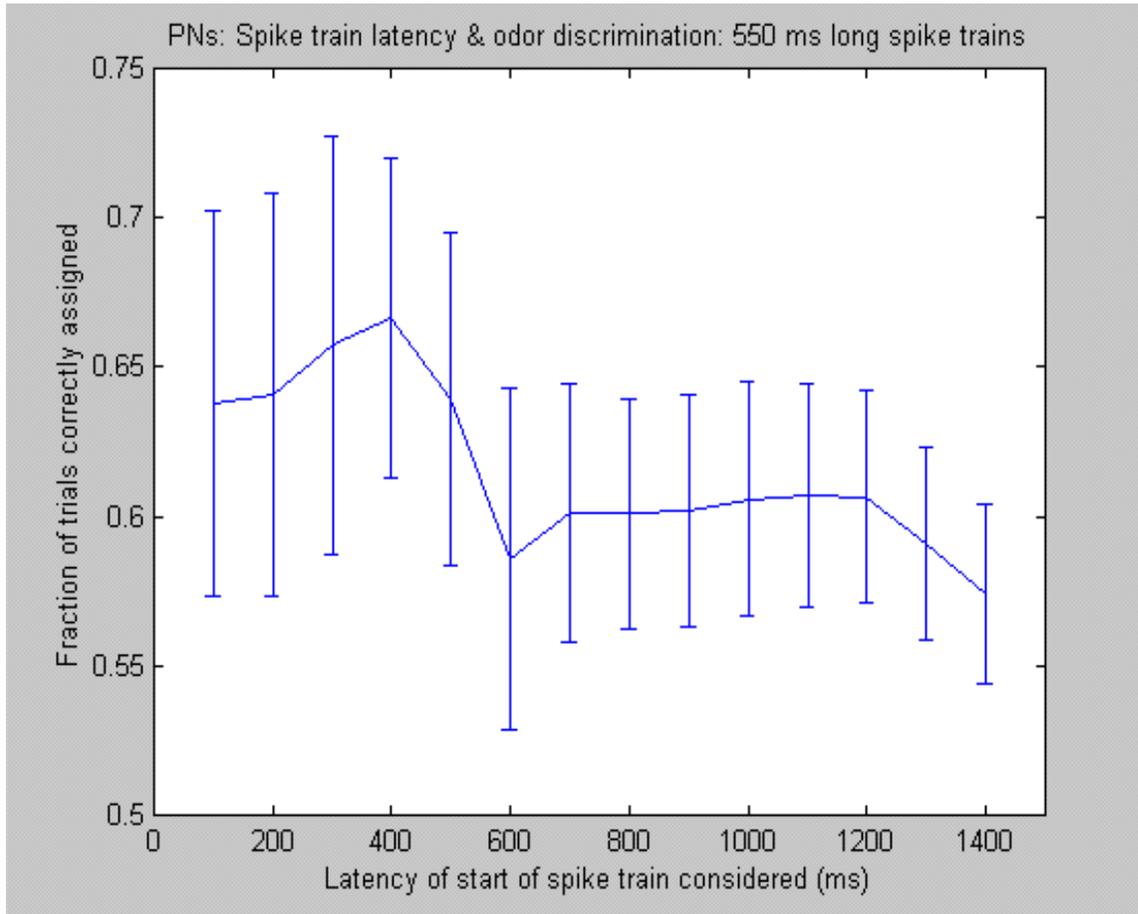


**Figure 4.2.** Fraction of trials for which the odor was correctly classified as a function of the timescale at which the spike trains of single PNs were decoded. Each trial is assigned to the class with the lowest arithmetic mean distance to it (a) or to the class with the lowest geometric mean distance to it, with exponent=-15 for the averaging (b) (blue bars,  $n=47$  PNs). The former classifies spike trains to the class closest on average over all trials for that odor; the latter amounts to classification into the class with the closest trials altogether. Each odor was presented for 1 sec at various concentrations from 2% to 100% of saturated vapor pressure. For a complete explanation of  $T$ , see methods. Red bars denote chance levels of classification.



**Figure 4.3. Odor discrimination as a function of the length of the spike trains used for classification.** The latency was kept at the one yielding best classification in the range starting from 1 second before the odor onset to 1 second after the odor onset. Discrimination was performed for a single concentration of different odors.

To test whether this was the consequence of using single cells for classification, I classified the spike trains of 19 PNs simultaneously recorded by Stijn Cassenaer in response to one concentration of 16 odorants. Classification was performed using two different neural codes that employ different ways to pool across neurons (see Reich et al., 2001). The first, which I call population code in keep-



**Figure 4.4.** Odor discrimination as a function of the latency of the start of the spike train considered. Discrimination was performed for a single concentration of different odors.

ing with Reich et al.'s nomenclature, simulates an integrate-and-fire downstream decoder by treating spikes from all neurons equally regardless of neuronal identity. One spike train was computed that aggregates the spikes across all PNs, and classification was performed based on distances between these aggregate spike trains. Using this algorithm, classification was significantly better using temporal information than using spike counts alone (Fig. 4.6), but was hardly better than using just a single cell. The second algorithm, called labeled line code, does not throw away information about neuronal identity of each spike. All neurons receive equal weight in the classification decision:

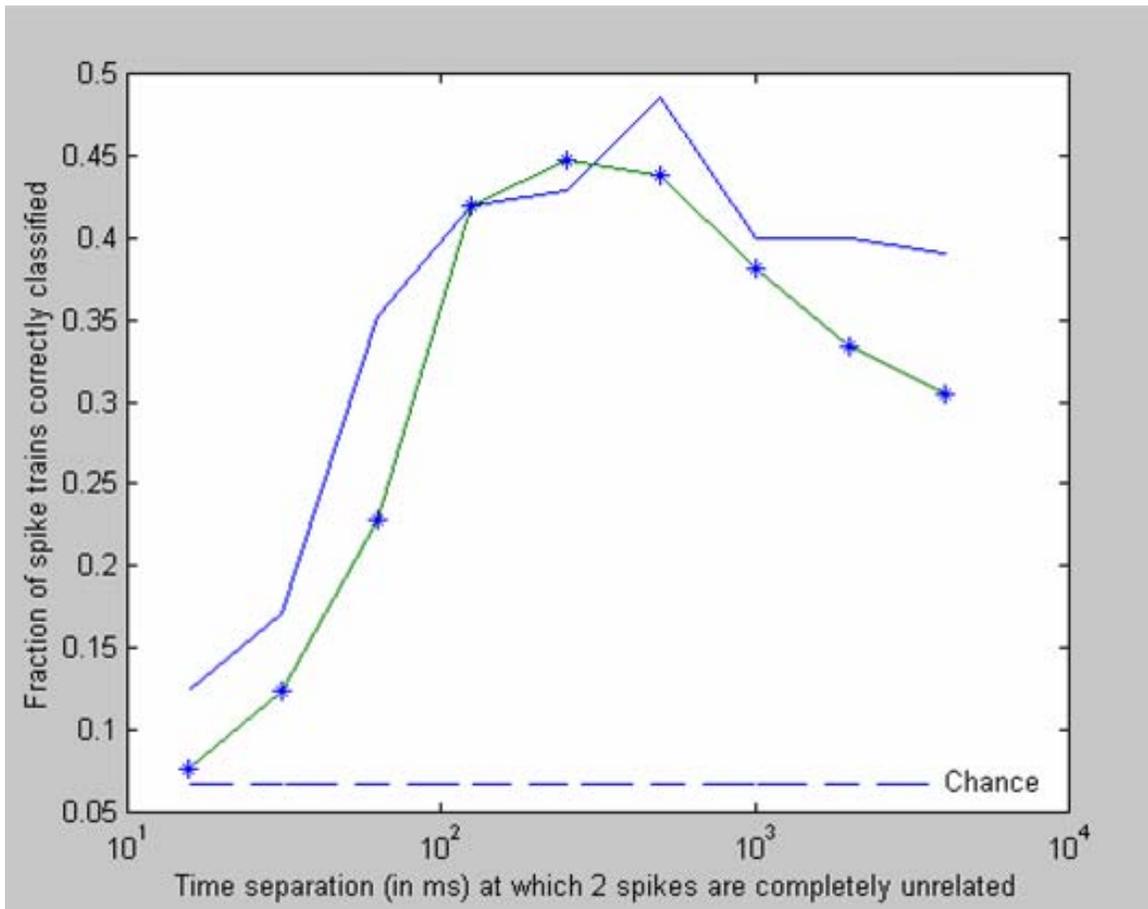


Figure 4.5. A single PN exhibits better classification using temporal information than using spike count alone ( $T$  tending to Infinity) when classification is to the closest class performing linear averaging across all trials (green) and geometric averaging with exponent = -15 (blue) (see Methods; see Fig. 4.2 for mean classification across 19 PNs).

the distance used for classification is the sum of the distances for individual cells. Because each PN carries different stimulus information, classification using the labeled line code was significantly better than that using the population code (Fig. 4.7). Classification using the labeled line code yielded 98% correct classification among the 16 odors presented, and presented almost perfect classification even when only using spike counts (Fig. 4.7).

To test whether this timescale independence of the labeled line code was due to a saturation in information due to the high ratio of # of neurons/ # of odors, I calculated classification rates as a

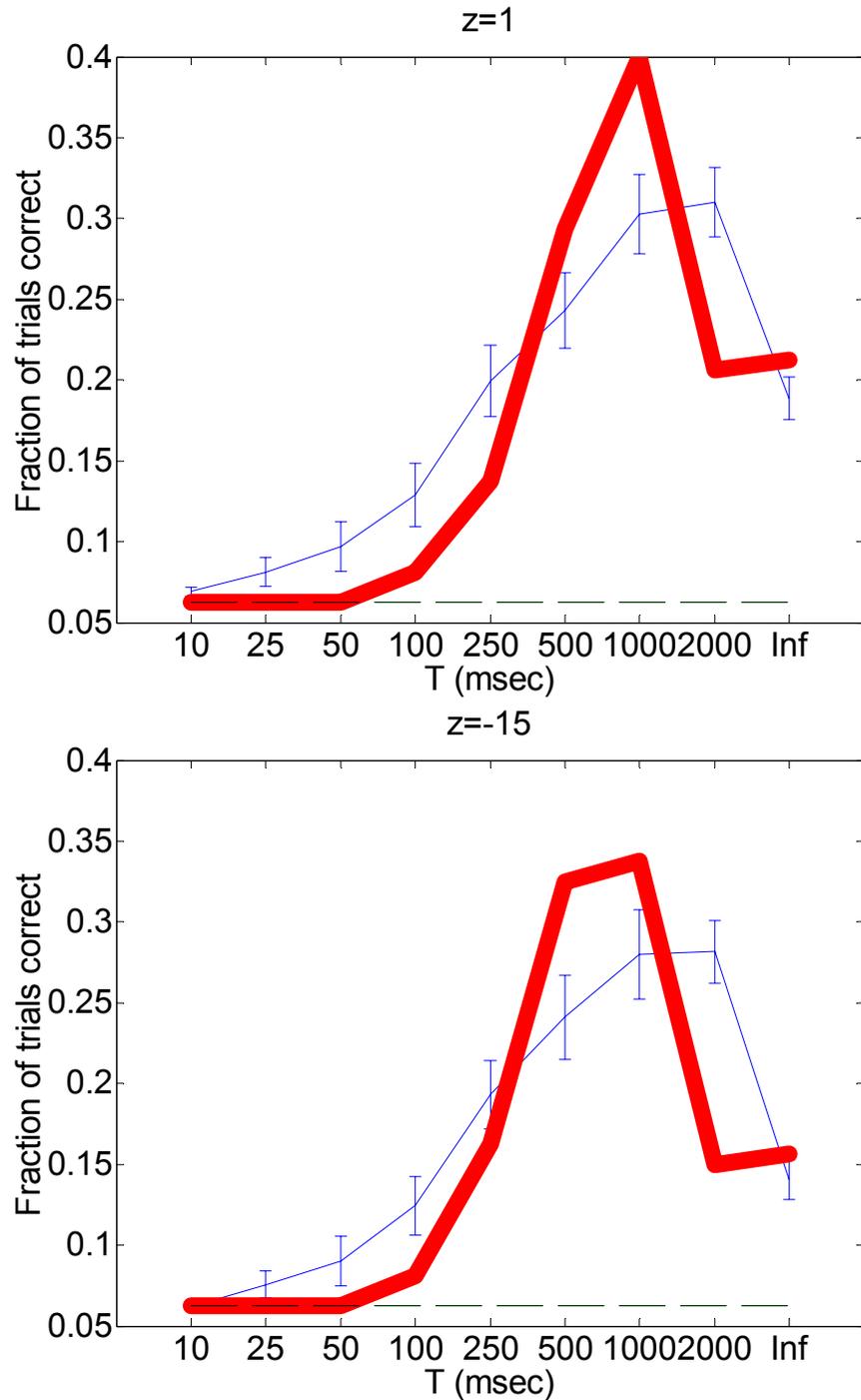


Figure 4.6. Odor discrimination by an assembly of 19 simultaneously recorded PNs among 16 odors using the population code (red, see text) and using single PNs (blue; means and S.E.M.), as a function of the timescale  $T$  of decoding (see Methods), for linear averaging across trials ( $z=1$ , above) and geometric averaging ( $z=-15$ , below). For each trial, 7.5 sec beginning with the onset of a 3 sec long odor pulse were used for discrimination. Chance levels are indicated by the black dashed line.

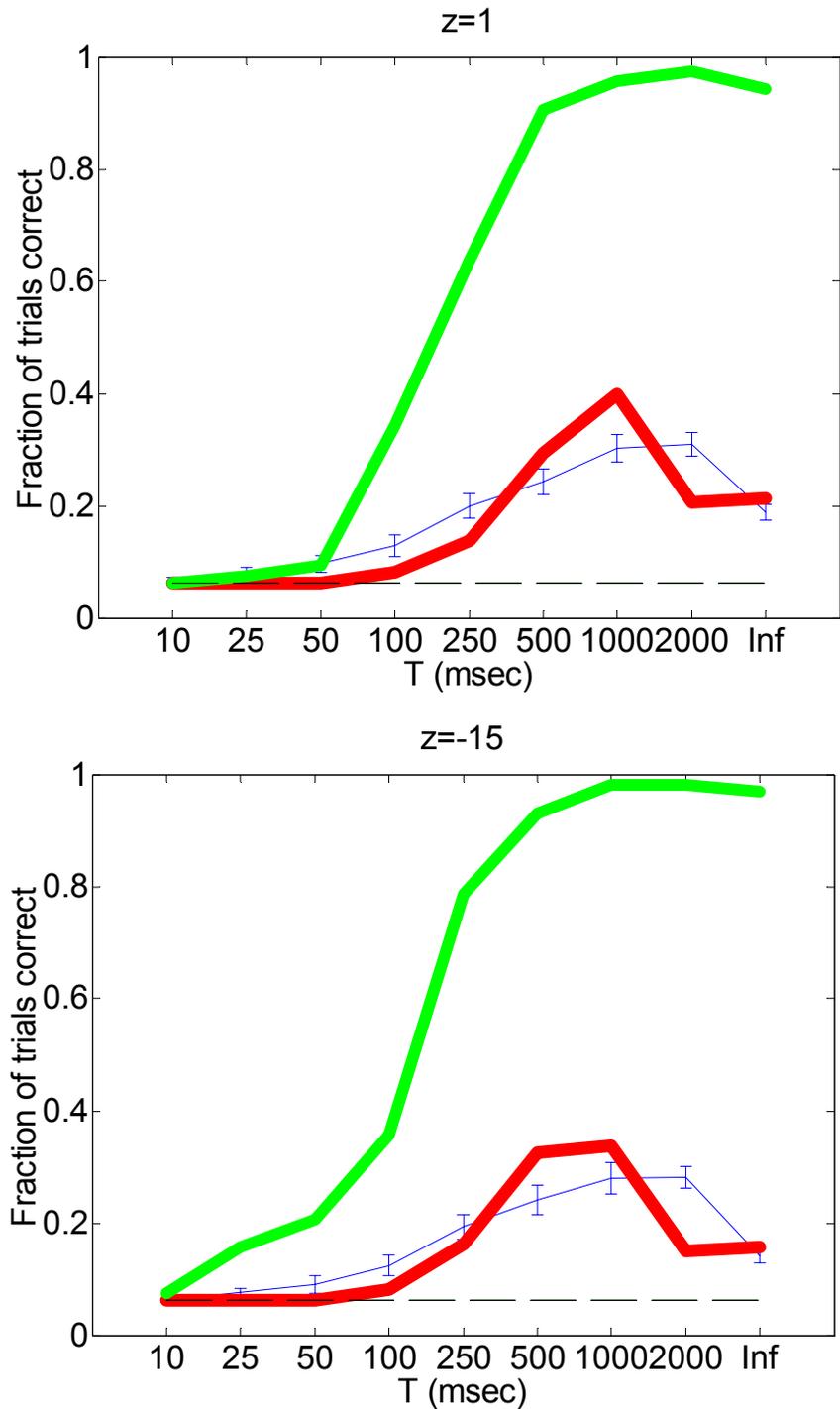


Figure 4.7. Odor discrimination by an assembly of 19 simultaneously recorded PNs among 16 odors using the labeled line code (green, see text), the population code (red) and single cells (blue), as a function of the timescale  $T$  of decoding (see Methods), for linear averaging across trials ( $z=1$ , above) and geometric averaging ( $z=-15$ , below). For each trial, 7.5 sec beginning with the onset of a 3 sec long odor pulse were used for discrimination. Chance classification levels are indicated by the black dashed line.

function of the number of neurons used in decoding for each decoding timescale used (Fig. 4.8). For each number of neurons, 40 randomly chosen subsets of the 19 neurons were used (unless the maximum number of combinations was less than 40, in which case all combinations were used). The same information is plotted in a different format in Fig. 4.9, which shows that indeed, for smaller PN assemblies for which the information content is not saturated, temporal information improves classification accuracy. The transmitted information per cell is approximately constant at 0.3 bits/neuron after the first neuron and up to within 0.5 bits of the total stimulus information available in our experiments (Fig. 4.18). Assuming this linearity holds throughout the entire antennal lobe when the stimulus set is sufficiently large, this yields a total bandwidth of 250 bits for the locust antennal lobe (830 neurons x 0.3 bits/neuron), which would allow the discrimination of  $10^{75}$  odors.

Classification rates, however, were maximal for timescales much larger than the timescale of the oscillations which were previously hypothesized to form the basis for a temporal code based on the fact that some PNs exhibit different firing probabilities for successive cycles (Wehr and Laurent, 1996). This suggested that PN responses might exhibit significant correlations on timescales smaller than 1 second. To test that, I calculated the matrix of conditional probabilities

$$P(x,y) = p(\# \text{ of spikes in cycle } N=y \mid \# \text{ of spikes in cycle } N-1=x)$$

over all trials of all concentrations of all odors presented to 46 PNs (data collected by the author; Fig. 4.10) and over all trials of all odors at one concentration presented to 12 PNs (data collected by Katrina MacLeod; Fig. 4.11). These matrices showed a strong correlation between the number of spikes in successive non-overlapping 50 msec windows. This correlation was present both following odor presentations ( $r=0.67$ ,  $p \ll 10^{-6}$ , Spearman ranksum correlation test, 3 sec period following odor stimulation, Figs. 4.10-4.11) and during the 1 sec period preceding odor presentation ( $r=0.64$ ,  $p \ll 10^{-6}$ , Spearman ranksum correlation test, Fig. 4.12). The effect is very significant: over all trials

---

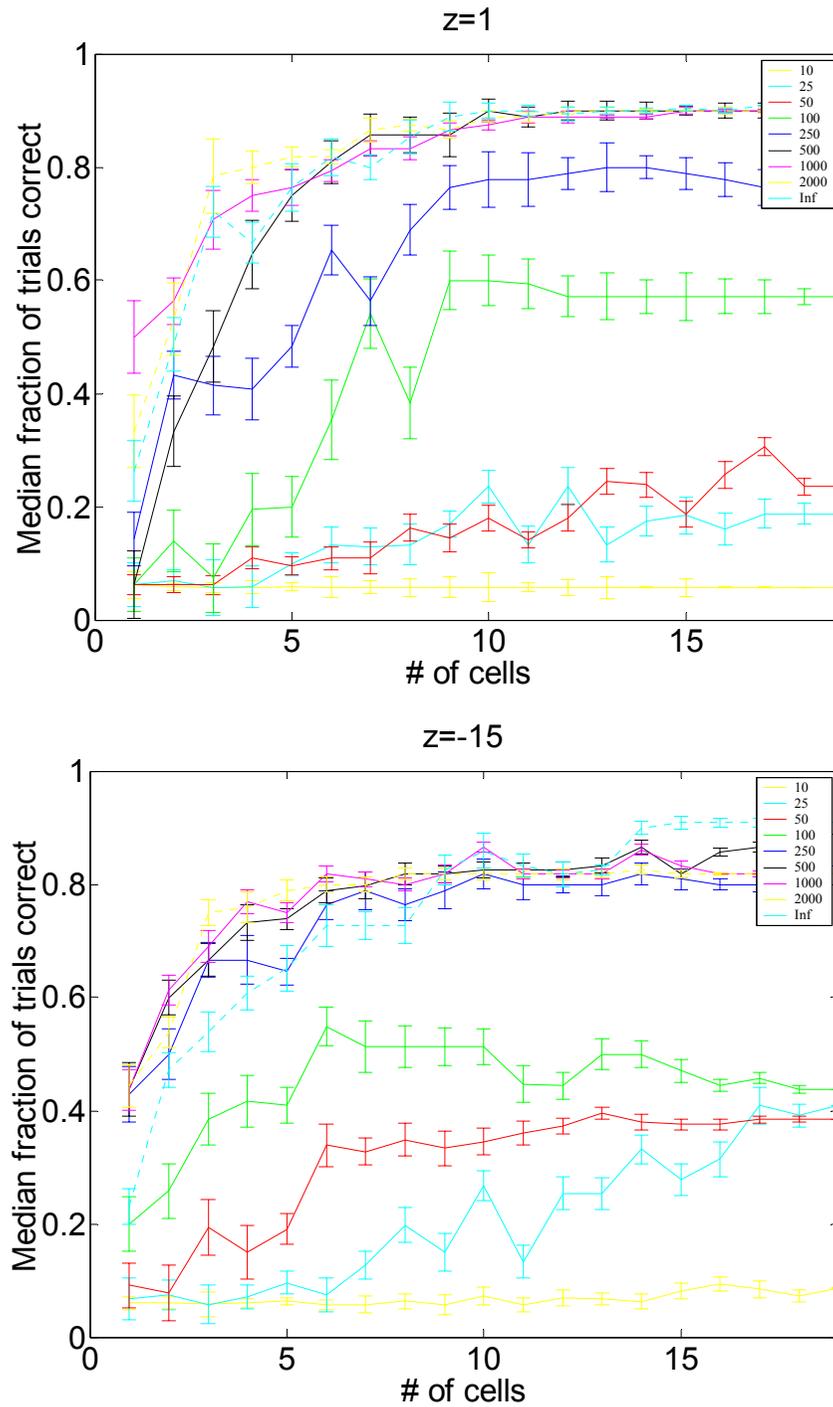
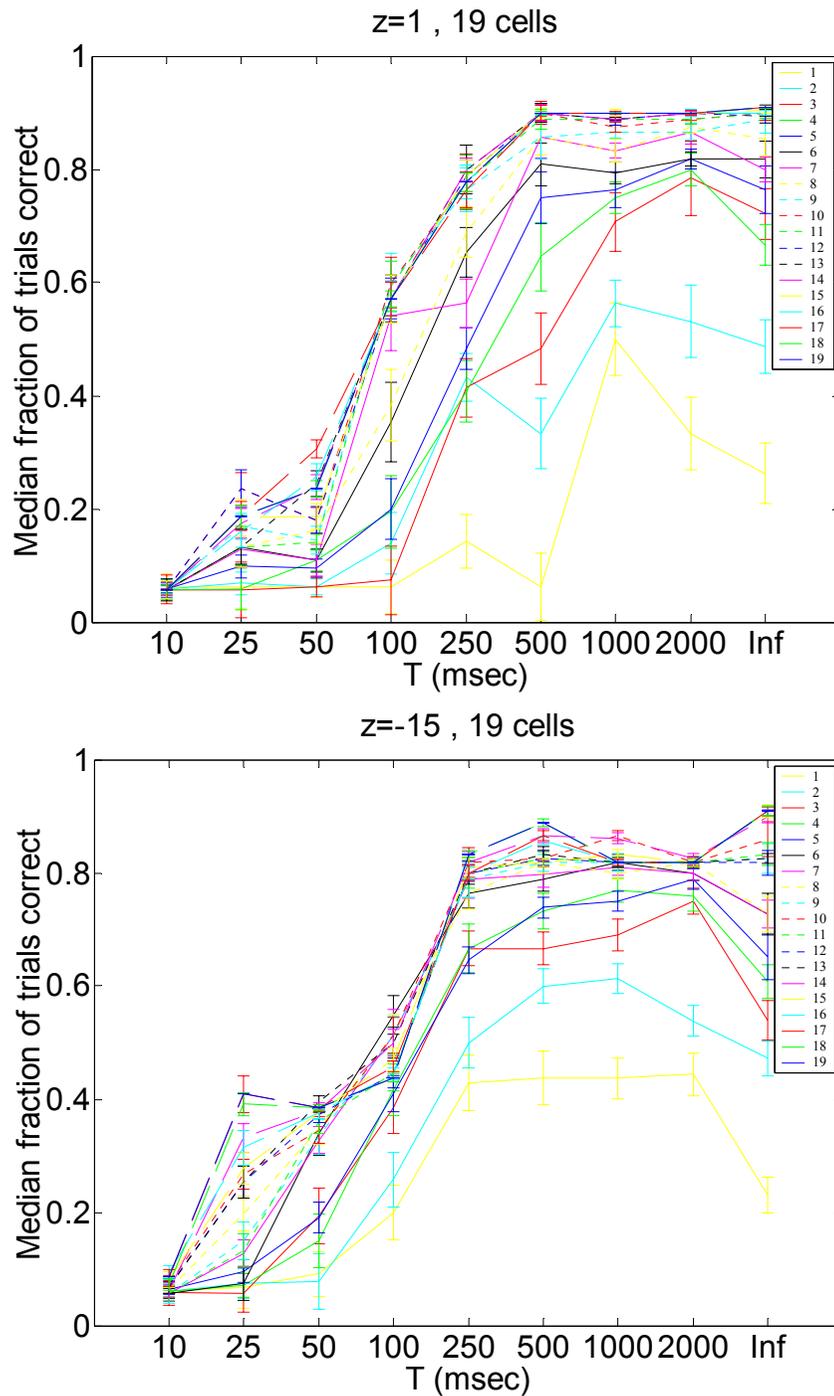
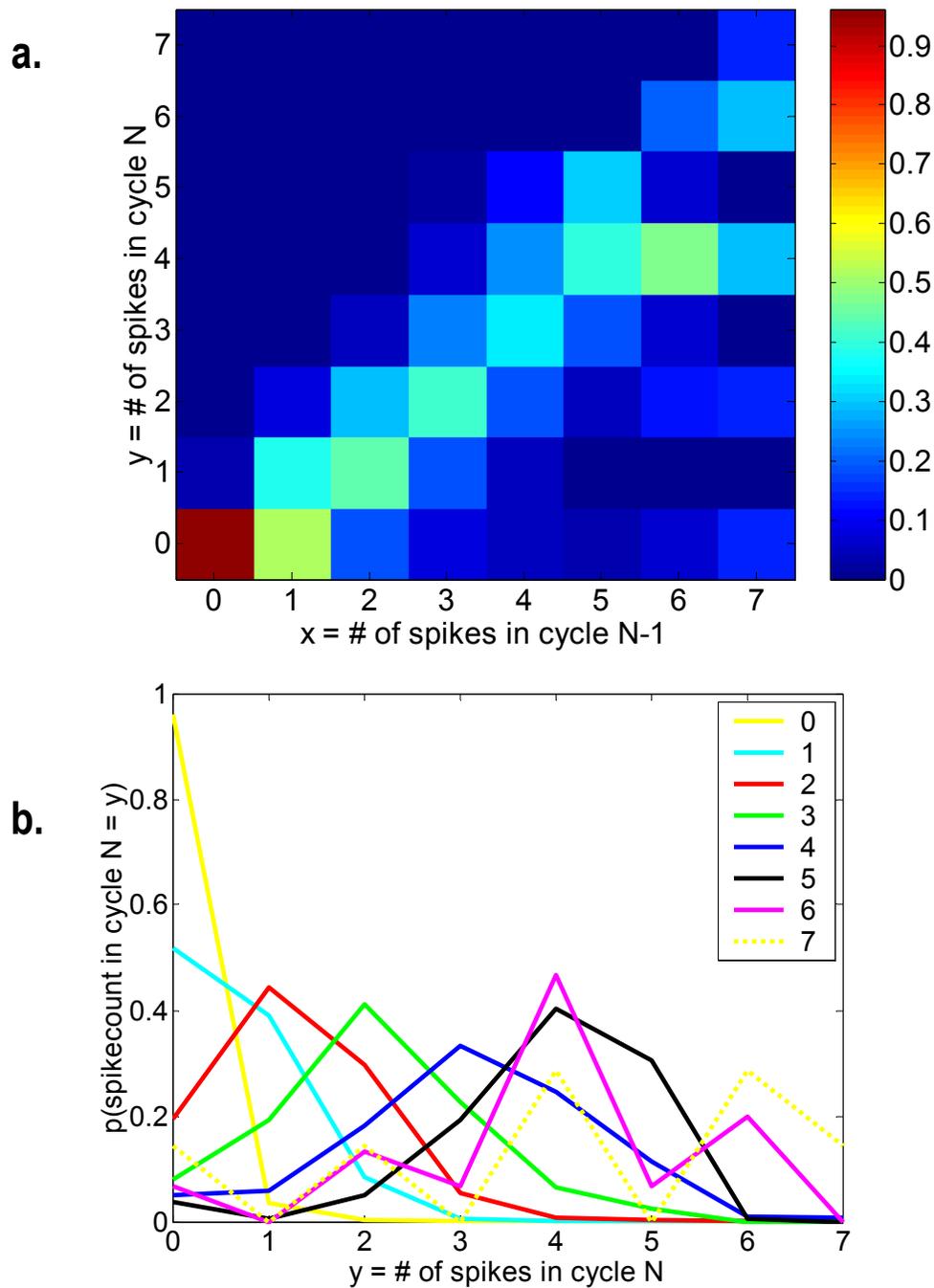


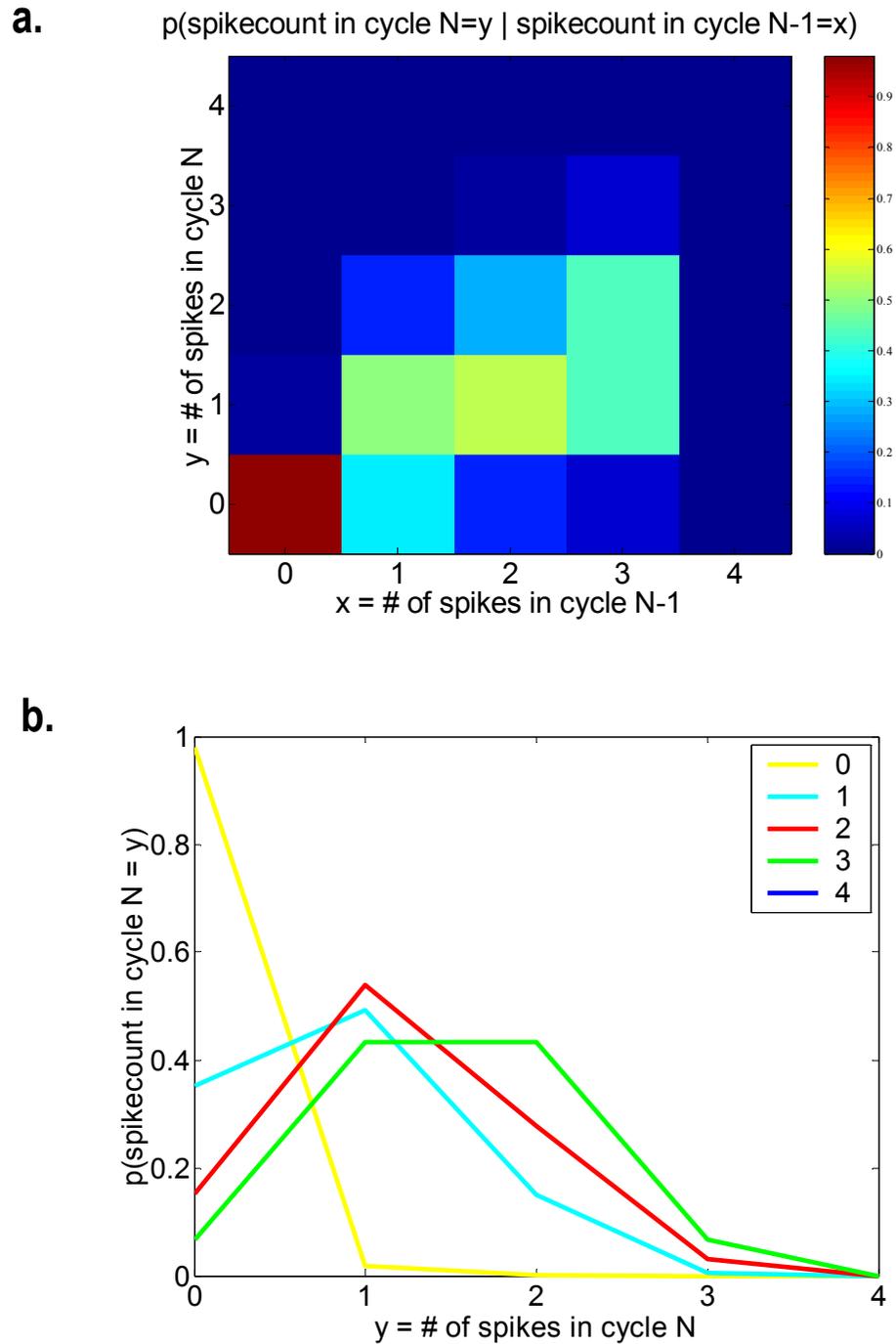
Figure 4.8. Odor discrimination for the labeled line code as a function of the size of the cell assembly used for decoding. Each line represents a different timescale used for decoding (T values, inset).



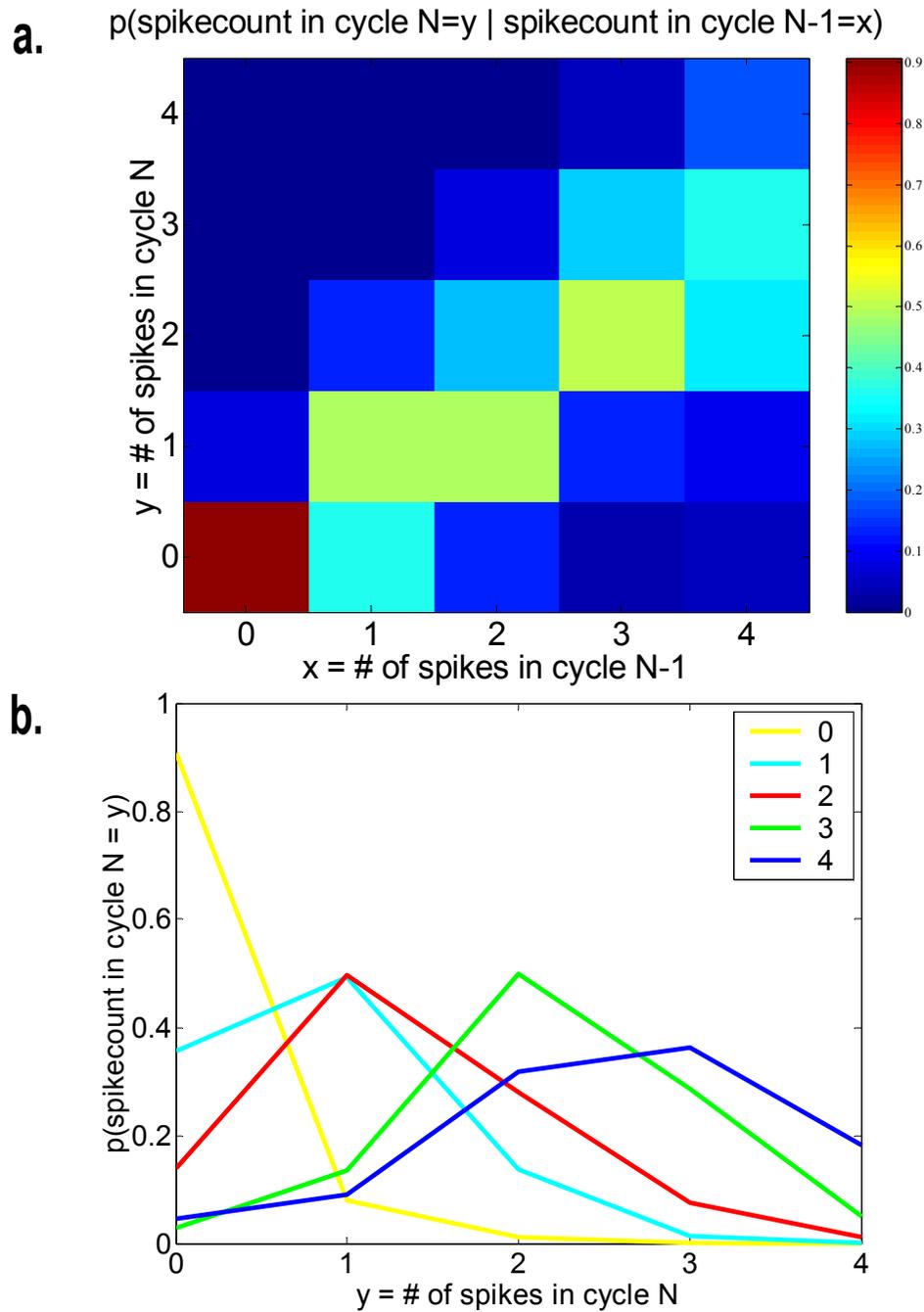
**Figure 4.9.** Odor discrimination for the labeled line code as a function of the timescale used for decoding ( $T$ ). Each line represents a different size of the cell assembly used for decoding (key, inset).



**Figure 4.10. PNs are bursty:** The probability of firing of a PN in any given cycle is significantly enhanced if the PN has fired in previous cycles, and the number of spikes in successive cycles is significantly correlated. **a)**  $P(x,y) = p(\# \text{ of spikes in cycle } N=x \mid \# \text{ of spikes in cycle } N-1=y)$ . Probabilities calculated over several concentrations of more than 100 PN-odor pairs of 46 PNs. **b)** The same data plotted as line plots of the probability distributions of the # of spikes in cycle N-1. Each curve represents a different # of spikes in cycle N (see legend). Note the shift rightward in the curves as the # of spikes in cycle N increases.



**Figure 4.11. PN odor responses are bursty.** These plots were computed for a set of 12 PNs different from those in Fig. 4.10 for a single concentration of each odor (data courtesy of Katrina MacLeod). See Fig. 4.10 legend for details.



**Figure 4.12.** PNs' basal activity is bursty: The probability of firing of a PN in any given cycle during basal activity between odor stimulations is significantly enhanced if the PN has fired in previous cycles, and the number of spikes in successive cycles is significantly correlated. **a)**  $P(x,y) = p(\# \text{ of spikes in cycle } N=y \mid \# \text{ of spikes in cycle } N-1=x)$ . **b)** The same data plotted as line plots of the probability distributions of the # of spikes in cycle  $N-1$ . Each curve represents a different # of spikes in cycle  $N$  (see legend). Note the shift rightward in the curves as the # of spikes in cycle  $N$  increases. These plots were calculated for the same cells as Fig. 4.11.

for all odors for the 46 PNs in Fig. 4.10 and during the 1-second period of odor stimulation, the probability of encountering a spike in any 50 msec long time window was 0.04 if there had been no spikes in the preceding 50 msec window, but jumped to 0.61 if there had been one or more spikes in the preceding window (see Table 1).

**TABLE 1. Conditional probabilities of PN spike counts in 50 msec windows during 1 sec period of odor stimulation**

X	P(0 spikes   X spikes in previous 50-msec)	P(1 spike   X spikes in previous 50-msec)	P(2 spikes   X spikes in previous 50-msec)	P(3 spikes   X spikes in previous 50-msec)	P(4 spikes   X spikes in previous 50-msec)
0	0.9641	0.0313	0.0041	0.0005	0.0000
1	0.4661	0.4056	0.1174	0.0102	0.0008
2	0.1866	0.4359	0.3092	0.0623	0.0060
3	0.0987	0.1947	0.4409	0.2211	0.0445
4	0.0625	0.0625	0.2768	0.3482	0.2500
Row probabilities may not add up to 1 due to windows with 5+ spikes.					Calculated over all trials for all odors for 46 PNs.

To confirm this correlation and measure its timescale, I calculated spike-triggered firing rate averages for 119,351 spikes in 10,730 trials in 377 cell/odor/conc datasets for 46 PNs. Beyond a refractory period, PNs exhibited a large positive autocorrelation with a time constant of several hundred milliseconds both during odor responses (Fig. 4.13a) and baseline firing (Fig. 4.13b).

In the locust, the output of PN assemblies is decoded by Kenyon cells (KCs) in the mushroom bodies<sup>1</sup>. The decoding scheme for individual KCs is quite different from the classification algorithms

1. As well as by Lateral Horn inhibitory neurons.

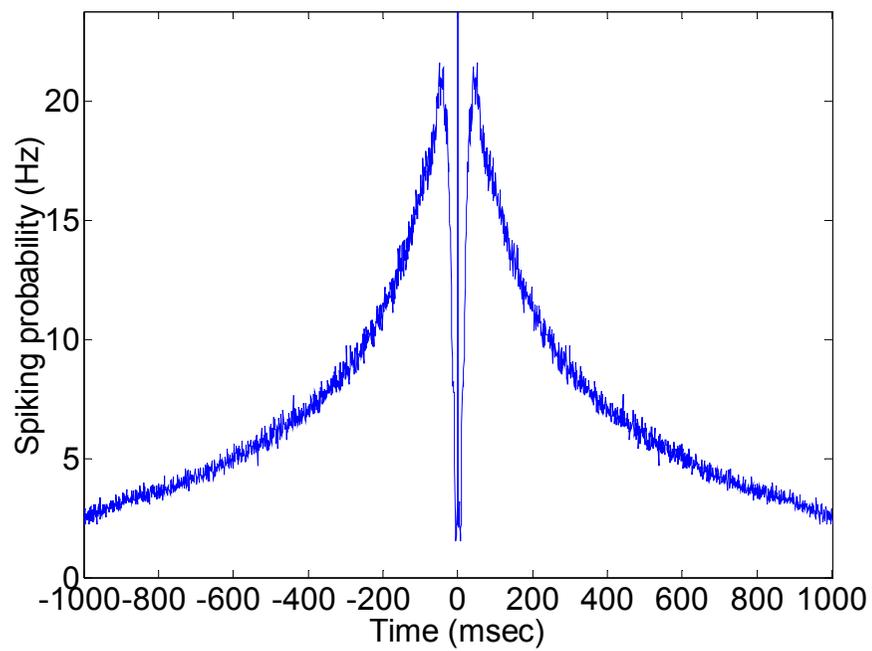
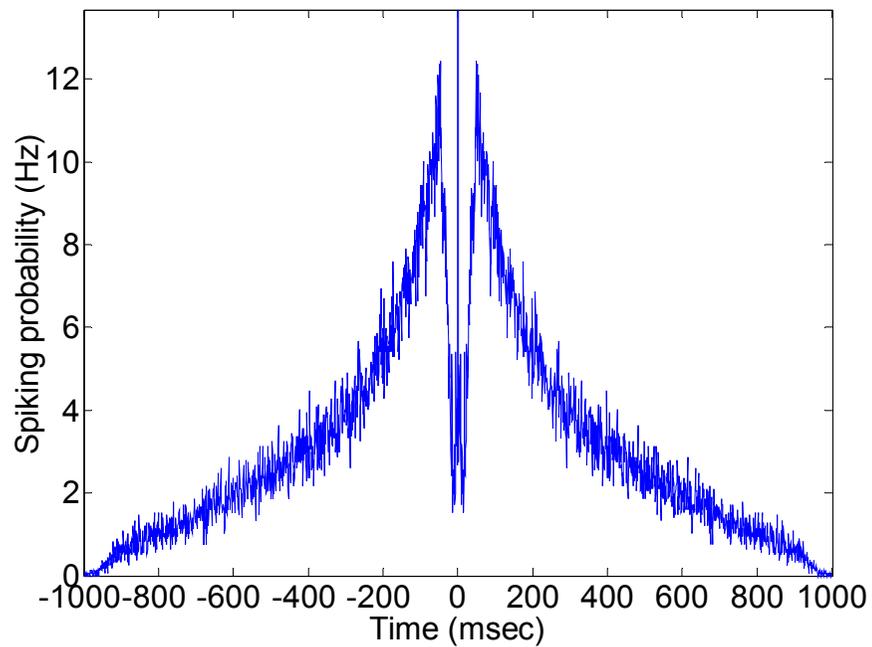
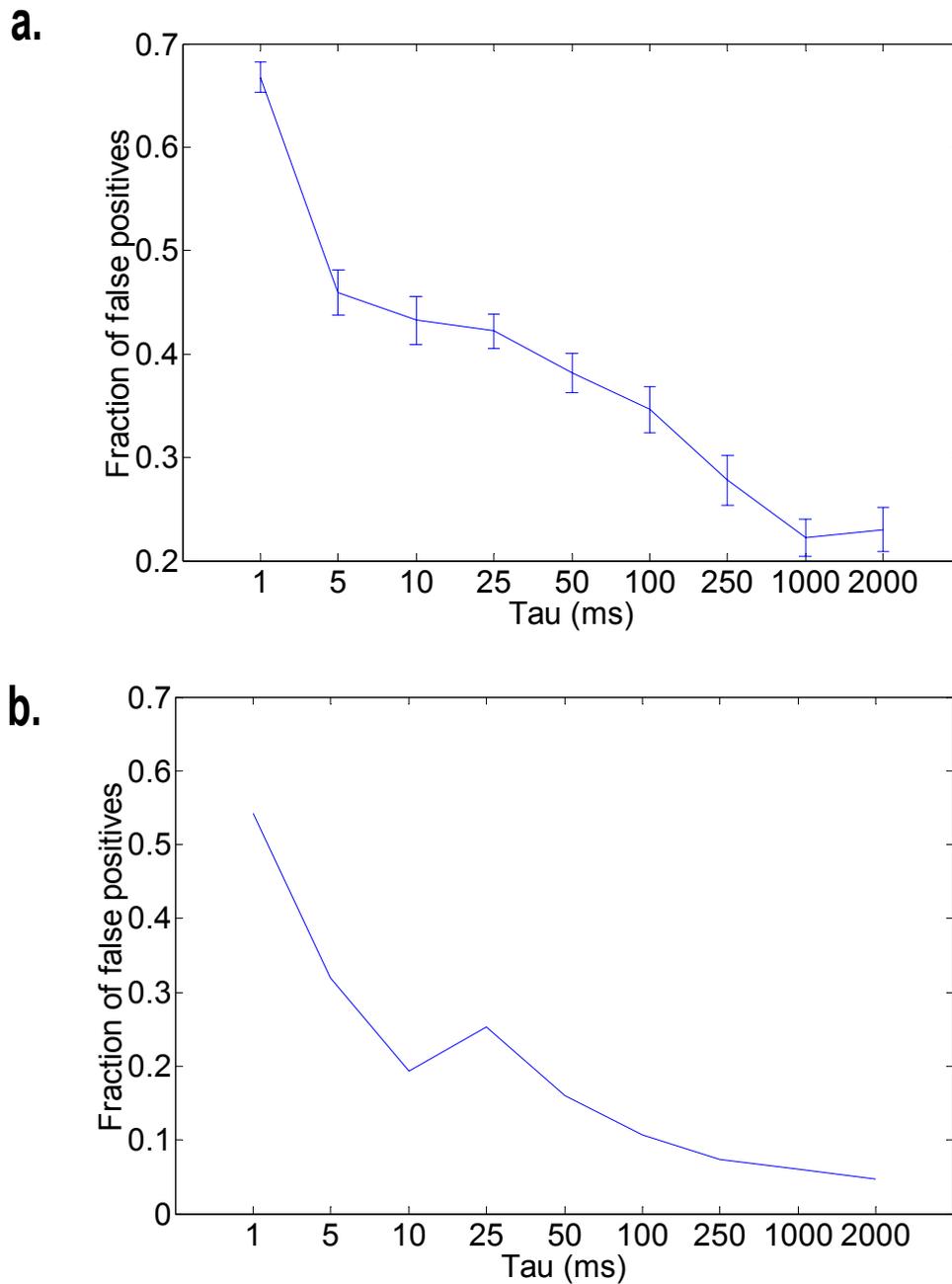
**a.****b.**

Figure 4.13. PNs exhibit a positive autocorrelation with a time constant of several hundred milliseconds both during odor responses (a) and baseline firing (b).

employed above: KCs exhibit sparse representation of odors, often responding to only one odor. It is thus reasonable to expect that the properties of the decoding algorithms are likely to minimize false positives and false negatives in a representation where each KC codes only for the odor that excites it the most, rather than maximizing odor discrimination across the entire spectrum of odors. To study the effect of the timescale of decoding on the reliability of the encoding of single odors by KCs, I created simple model KCs that smoothed each PN's response by convolving it with a Gaussian of standard deviation  $\tau$  and then integrated the smoothed inputs of 10 PNs together additively. The number 10 was chosen because combinatorial arguments and existing data on PN-KC connectivity suggest that any one KC integrates inputs from a number of PNs that is less than 20, and that it fires upon activation of a subset of these probably not exceeding 10 (Bäcker and Laurent, unpublished results). On any given trial, the model KC fired if and only if its PN inputs exceeded a threshold value at any point in time. No dynamics were considered for the KC. I then calculated, for each of 16 odors, the threshold value such that all trials of that odor elicited a response from the model KC. The odor which yielded the highest such threshold was selected as the model KC's preferred odor, and the KC's threshold was set to the corresponding threshold value. Then, the proportion of trials for non-preferred odors which elicited a response from the model KC was calculated, and called the proportion of false positives. If the representation is sparse and reliable, this proportion should be low; otherwise, it will be higher. The proportion of false positives was then computed as a function of the timescale  $\tau$  at which the PN inputs were smoothed. As was the case with odor discriminability using Victor and Purpura's algorithm above, odor discriminability with this sparse coding scheme also was optimal for timescales on the order of 1-2 sec: the proportion of false positives decreased with increasing  $\tau$  (Fig. 4.14). Sensitivity to variations in PN spike trains on a timescale of 1-2 oscillation cycles did not contribute to enhance the reliability and sparseness of the representation.

---



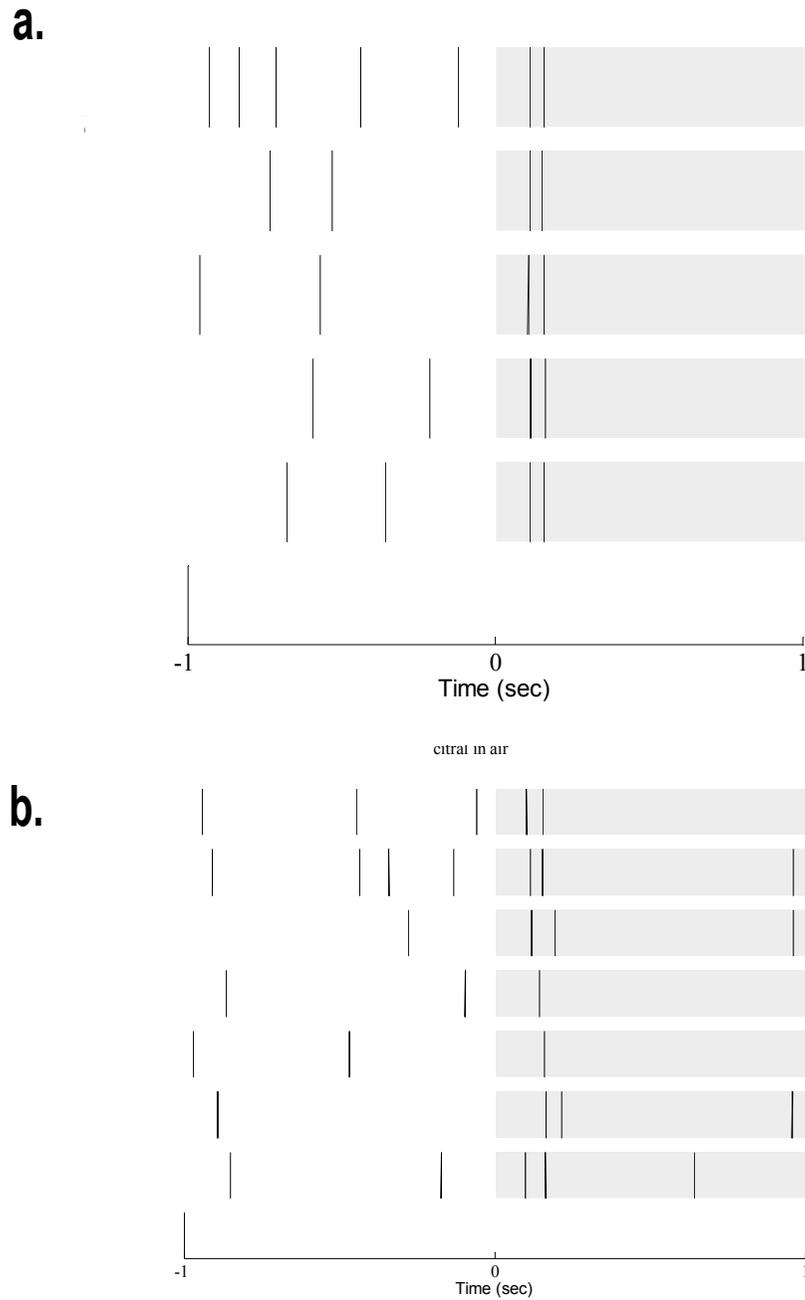
**Figure 4.14.** The fraction of false positive responses of a model Kenyon cell coding for 1 odor out of 16 and with its threshold set to have no false negatives decreases as the timescale ( $\tau$ ) of the decoding algorithm is increased (see Text for details). **a)** Mean fraction of trials which yield false positives over 20 different PN assemblies. **b)** Minimum fraction of trials which yield false positives over 20 different PN assemblies, i.e., fraction of false positive trials for the most discriminating set of PNs.

## 4.4 Discussion

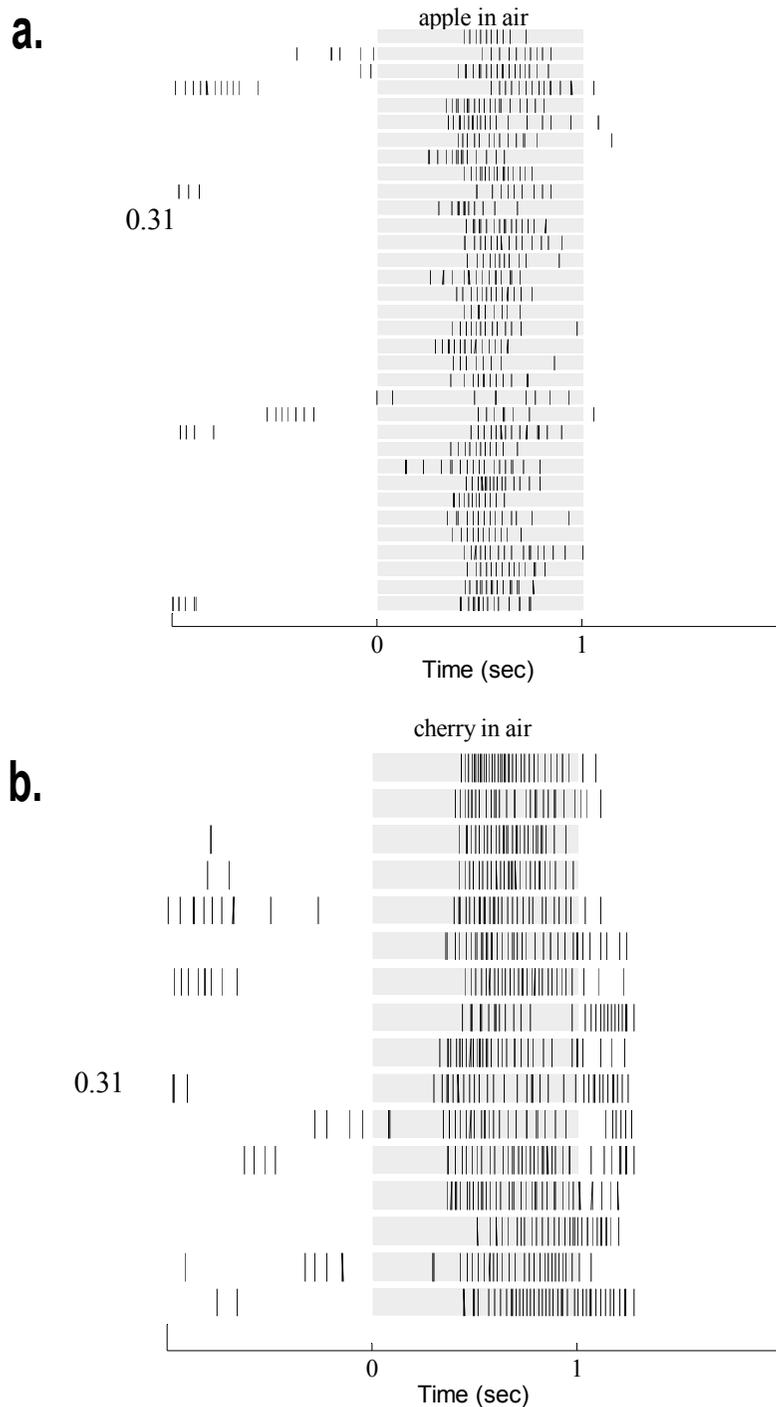
In summary, I have shown that odor information in single PNs allows classification among several odors significantly above chance levels. This information peaks at odor onset and decays about 500 ms later. The information content of single PN spike trains saturates at 500-1000 ms. Different PNs carry non-redundant information: correct classification rates were significantly higher if I kept track of neuronal identity than if responses were summed over the PN assembly. Increasing the size of the PN assembly increased correct classification rates, saturating at about 10 PNs. No further improvement was seen by increasing the size of the PN assemblies from 10 to 19. The optimal timescale for decoding proved to be of the order of 1-2 seconds, yielding significantly higher classification rates than the timescale of the oscillation cycles of 50 milliseconds. This was true both for single PNs and PN assemblies, and both for discrimination among all odors presented and a sparse coding scheme in which each decoder encoded a single odor. The preferred timescale of decoding can be explained by the observation that PN responses, both during odor presentation and between odor stimuli, are highly correlated on the timescale of several hundred milliseconds.

This correlation over timescales of hundreds of milliseconds is in sharp contrast with the results of Wehr and Laurent (1996), who reported on 4 PNs some of whose firing probabilities changed abruptly from one oscillation cycle to the next for any given odor. Although I have observed such PNs with highly precise and fast-varying firing rates in my data as well (Fig. 4.15), they constitute a minority of all PNs recorded, as evidenced by the analysis across 77 PNs presented here. More representative of the majority is a bursty PN (Fig. 4.16). The differences cannot be explained by a differential recording bias in favor of bursty PNs with high firing rates on my part, because the same degree of correlation in PN responses was observed not only in the recordings of Katrina MacLeod, but also in extracellular tetrode recordings in which the position of the electrodes is not manipulated

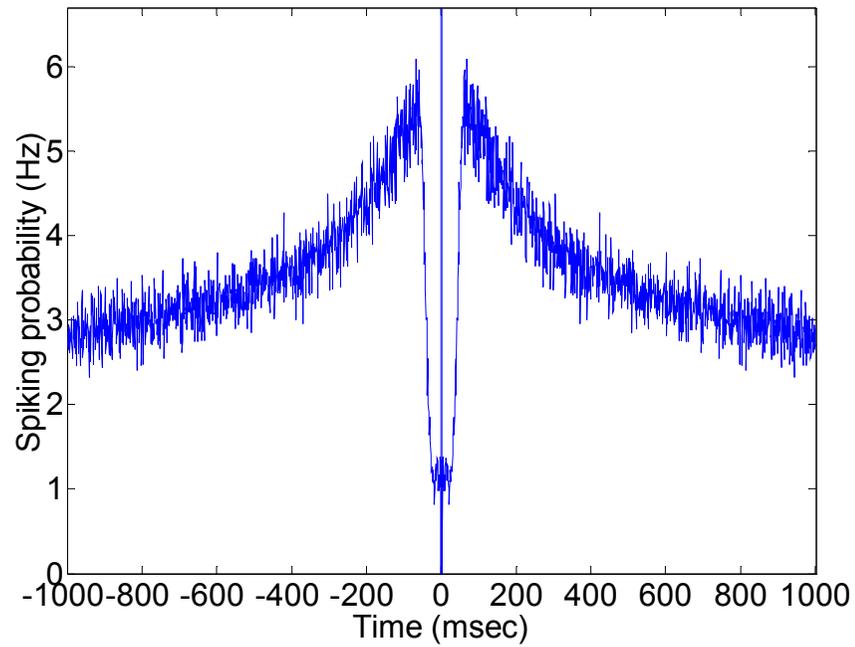
---



**Figure 4.15. The response to cherry (a) and citral (b) of a PN with precise and brief response patterns, as described by Wehr and Laurent (1996). Responses such as these are found in a minority of PNs. They are typically characterized by short latencies, low intertrial variability and short duration. These probably constitute a sub-type of PNs, since their responses to all odors all typically fall in the same category of precise, brief, early responses. The odors were presented from  $t=0$  to 1 s.**



**Figure 4.16.** The response to apple (a) and cherry (b) of a PN in another locust than that shown in Fig. 4.15. This PN, more representative of the majority of PN recordings than that in Fig. 4.15, responds to odors with bursts of spikes lasting several hundred milliseconds. This type of neuron is responsible for the highly significant slow autocorrelation observed across all 77 PNs analyzed above.



**Figure 4.17. Spike-triggered average firing probability for 19 PNs simultaneously recorded with 2 tetrodes (data courtesy of Stijn Cassenaer).** As with intracellular recordings (Fig. 4.13), a positive autocorrelation on a timescale of several hundred milliseconds is observed.

for individual cells (Fig. 4.17).

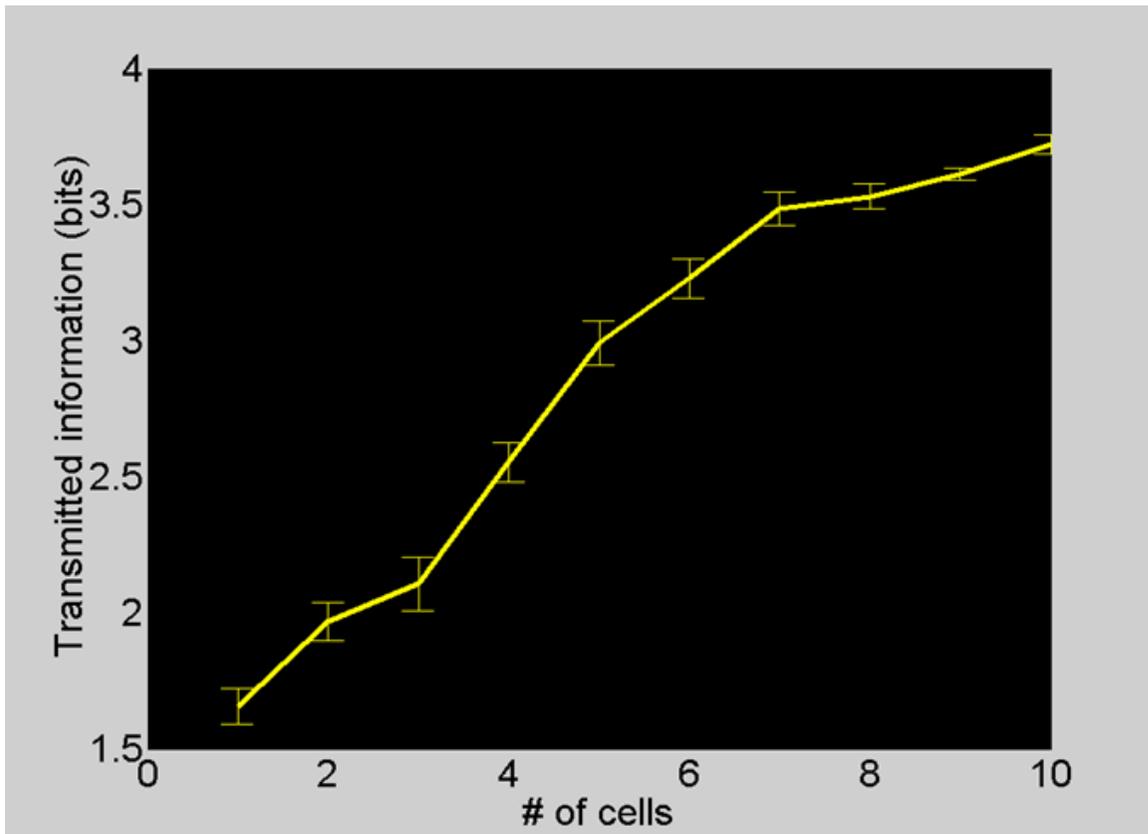


Figure 4.18. Transmitted information as a function of the number of projection neurons used for odor identification.

The work of Stopfer and colleagues (1997) and Chapter 5 of this thesis (part of which has been published as MacLeod et al., 1998) has shown that synchronization on a timescale much smaller than that seen here to be optimal for decoding is required for fine olfactory discrimination and the read-out of PN assemblies by downstream neurons. The functional advantage conferred by such selectivity remains unknown, and will be addressed in the first part of the next chapter. It is possible that neuronal biophysics makes it impossible to integrate over timescales of several hundred milliseconds, making the ideal decoding algorithm biologically implausible. Alternatively, it is possible that the analysis of larger numbers of simultaneously recorded PNs, or the analysis of decoding algo-

rhythms that approximate Kenyon cells more closely than those used herein, will reveal a role for fast timescales in the decoding of PN assemblies. Finally, it is possible that fast timescales are useful in the encoding of the rapidly varying signals in natural dynamic odor plumes rather than the more uniform odor pulses used in the experiments in this thesis and in the previous work of the laboratory.

## **4.5 Applications**

The method applied here to classify spike trains into the stimulus classes most likely to have given rise to them, and slight variations on it, have been widely applied (MacLeod, 1999; Stopfer and Laurent, 1999; Friedrich and Laurent, 2001; Bhazenov et al., 2001, Ch. 8, this thesis) since the publication of parts of this work (MacLeod et al., 1998).

## **4.6 Methods**

### **Surgery, Odor delivery and Electrophysiology**

See Chapters 7-9, Methods.

### **Clustering analysis**

The clustering analysis is based on the cost-based metric methods (Victor and Purpura, 1997) according to which a 'distance' is computed between spike trains. This distance is defined as the cost paid to transform one spike train into the other using three elementary steps: insertion; deletion

---

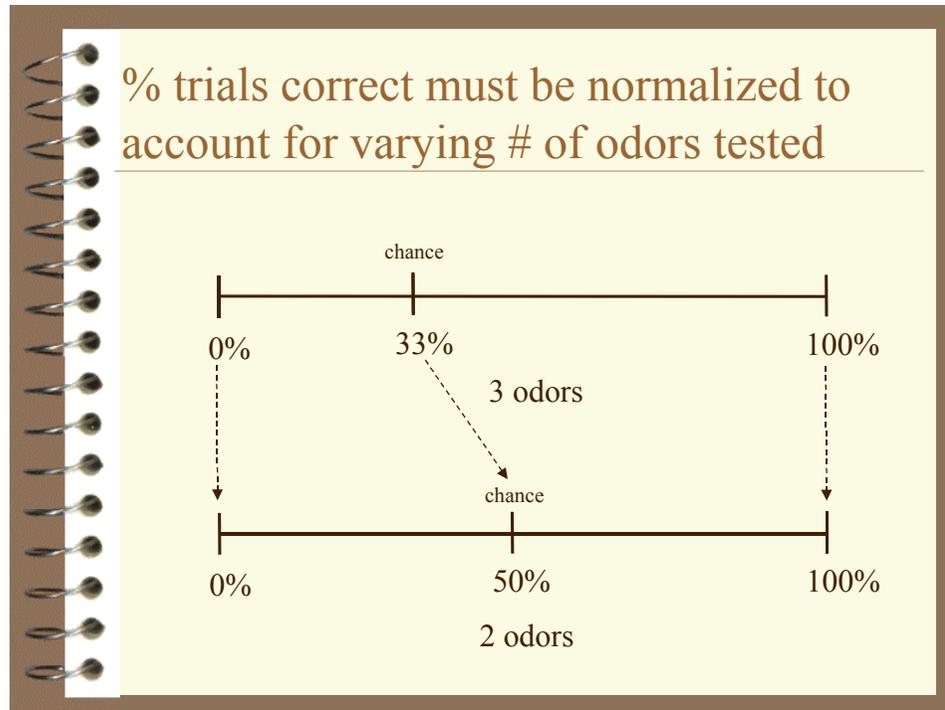


Figure 4.19. Normalization of fraction correct to a two-alternative-forced-choice scenario.

of a spike (each at a cost of 1); and displacement of a spike by 1 ms (cost of  $2/T$  for each displacement, where  $T$  is the maximum separation in ms allowed between the spike time in one train and that in the other). I used a range for  $T$  between 16 and 4,000, with  $T = 150$  providing the best classification overall. Results were not greatly different for  $16 \leq T \leq 1000$ . Classification was carried out using two methods: In one, the mean distance between a spike train and all spike trains of a stimulus class was an arithmetic mean (all points equal); in the second, the mean ( $M$ ) was geometric, with the exponent set to -15 (less weight to outliers):

$$M = \sum_i d_i^{-15}$$

Percent correct results from choosing among all odors without restriction to pairwise assignments in those neurons that responded to more than two odors. For each neuron  $i$ , chance level is thus  $1/m_i$ , where  $m_i$  is the number of odors to which neuron  $i$  responded. The effective number of odors for the mean percentage correct was calculated as  $\langle 1/m_i \rangle^{-1}$ . When mean rates are shown as 50%, mean percentage correct was then normalized to a two-alternative-forced-choice scenario by dividing the difference between the percentage correct observed and the chance level given the effective number of odors, by 0.5, i.e., the maximum improvement above chance possible after normalization. For the lag and duration analyses, in order to obtain a distribution of percentages for all datasets to be able to compare across lag or duration values, percentages of trials correct were normalized for each dataset using a piecewise linear transformation between the space for the actual number of odors for each dataset and a normalized space with 2 odors, such that 0, 100% and chance levels were fixed points and mapped to the corresponding point in the other space, and all other points underwent a linear mapping using the closest fixed points as anchors (Fig. 4.19).

For the estimation of discrimination as a function of latency and spike train duration, this method was then applied to successive sliding windows of the spike trains, varying the duration and lag of the windows with respect to the time of stimulus onset (Fig. 4.20).

---

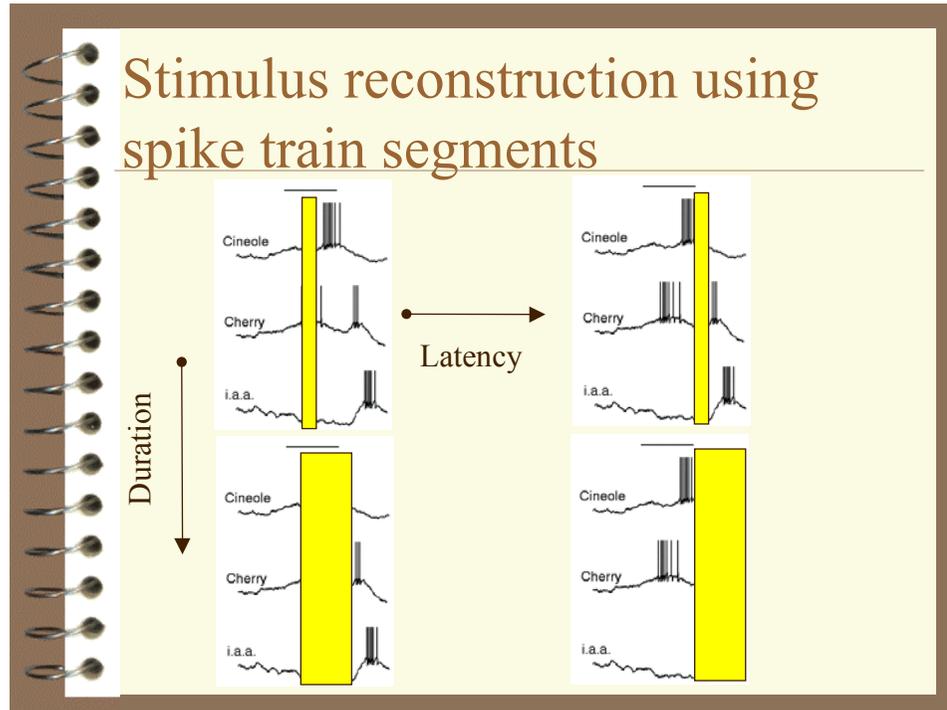


Figure 4.20. Schematic diagram of method used to estimate odor discrimination as a function of latency and spike train duration.