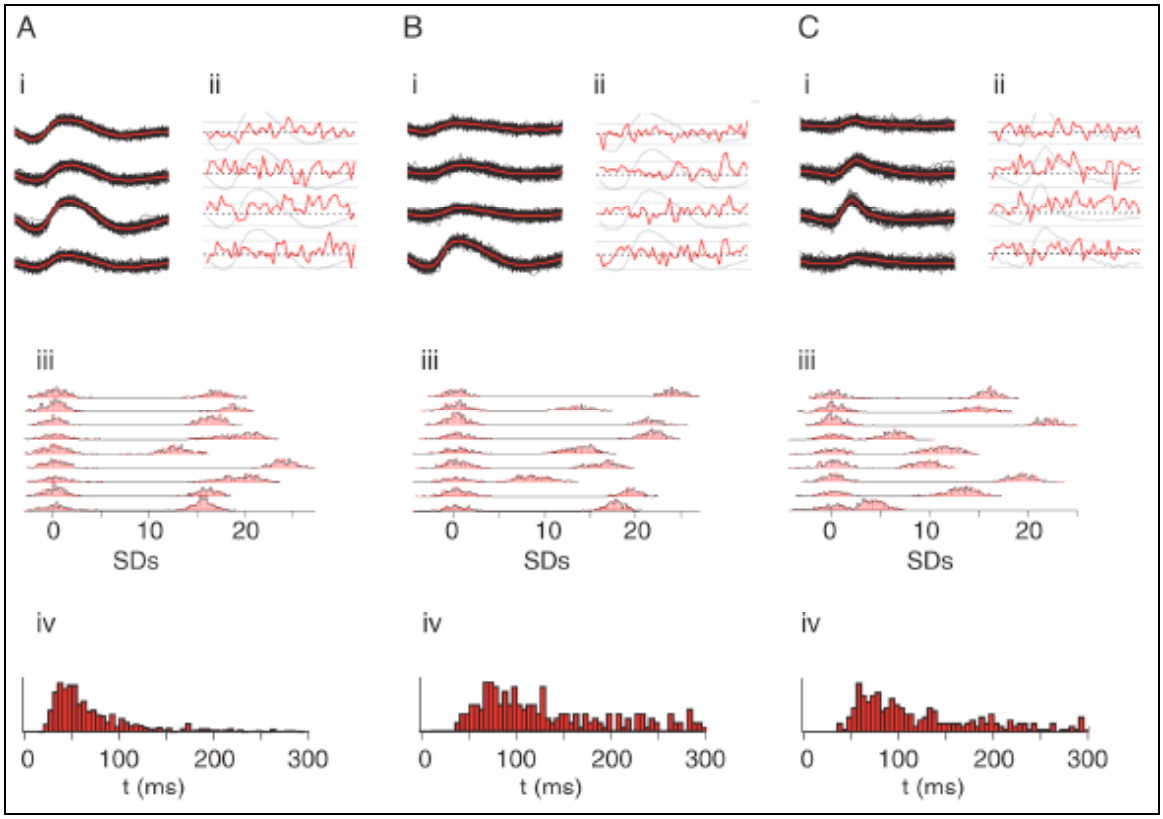
Supplementary Information

SI 1: *Spike clustering and PN unit isolation*

Because the only antennal lobe neurons that produce sodium action potentials are projection neurons (in locust, antennal lobe local neurons do not produce sodium action potentials, Laurent and Davidowitz, 1994), all spikes recorded with tetrodes could be attributed to PNs without ambiguity.

SI Figure 1:



This figure illustrates spike clustering from tetrode recordings of PN activity. We provide examples of spike clusters from categories 1, 2 and 4 of cluster isolation quality (SI Fig1**A, B, C** respectively). The three clusters are taken from the same data-set and have comparable signal-to-noise and variability properties. Each example belongs to the lower range of each category (*i.e.*, these examples are

among the worst ones). For each cluster (A-C): **i**: Overlay and average (red) of the raw waveforms (3 ms at 15 kHz sampling, or 45 samples) on each of the 4 tetrode sites. **ii**: SD test (Pouzat et al., 2002), showing the SD of each waveform (red) within its predicted 0.95 confidence interval (dotted lines) of the mean (stippled line). **iii**: Projection test (Pouzat et al., 2002), plotting the distance (in units of SD) between a tested cluster (left-hand-side distribution) and every other recorded cluster (right-hand-side distributions), projected on the axis connecting their centers. Mean cluster separation becomes clearly worse from A to C. **iv**: ISI test (Pouzat et a., 2002) showing, for each cluster, the inter-spike interval (ISI) distribution, with a characteristic minimal interval due to refractoriness. To minimize risks of misclassification, we used data only from PN spike clusters from the first two quality categories (examples in A and B), as described in methods. The example in A falls just within the limits of the highest separation category (no event of any other cluster falls within 7 SDs of the cluster center); the example in B falls just within the limits of the second category (no other event within 5 SDs of the cluster center); the example in C is within the 4th category (less than 5 SD minimal distance between cluster centers). See methods for details.

SI 2: ***Structure and odor-evoked activity of the locust olfactory system***

Recent work in locust (Perez Orive et al., 2002; 2004; Stopfer et al., 2003; Mazor and Laurent, 2005) showed that the representations of odors are dramatically sparsened between the first and second olfactory relays—the antennal lobe and mushroom body, respectively. In the antennal lobe (AL, Figure 1A), the insect analog of the vertebrate olfactory bulb, odors are represented by distributed patterns of projection neuron (PN) activity. There are 830 PNs in each antennal lobe. When tested over 1 second-long odor pulses, individual PNs respond on average to about half of all odors presented, with odor- and neuron-specific firing patterns (Mazor and Laurent, 2005; Perez-Orive et al., 2002;

Laurent, 2002; Laurent and Davidowitz, 1994; Laurent, 1996). At baseline, individual PNs fire spontaneously at rates between 2.5 and 4 spikes/s (Perez-Orive et al., 2002; Mazor and Laurent, 2005). By contrast, in the mushroom body (Figure 1A), odors are represented by small subsets of highly selective Kenyon cells (KCs) in a large population of 50,000 neurons (Perez-Orive et al., 2002; Perez-Orive et al., 2004; Stopfer et al., 2003). Individual KCs respond equally specifically to mixtures (*e.g.*, a fruit or flower blend) or to single-molecule odors; these spiking responses occur on a background of very little spontaneous firing (Laurent and Naraghi, 1994; Perez Orive et al., 2002; Mazor and Laurent, 2005). KC responses are associative—their high specificity matches only that of combinations of PNs—and often concentration invariant (Stopfer et al., 2003).

PNs are the sole source of olfactory inputs to KCs. Previous work has identified several mechanisms contributing to KC specificity: for example, upon the presentation of an odor to the animal, PN responses phase-lock to a common 20-30 Hz oscillation (Laurent and Davidowitz, 1994); this synchronization is exploited for coincidence detection by their targets (Perez-Orive et al., 2002 and 2004). Because PN responses evolve over successive cycles of this periodic population output (Wehr and Laurent, 1996; Laurent et al., 1996; Mazor and Laurent, 2005) and because PNs generally fire 0 or 1 spike per cycle, the oscillation cycle is the time unit over which PN output is most appropriately characterized (Perez-Orive et al., 2002; Laurent, 2002; Mazor and Laurent, 2005). Coincidence detection by KCs enables piecewise decoding of PN activity vectors—the sets of PNs active during each odor-evoked oscillation cycles (Perez-Orive et al., 2002; Perez-Orive et al., 2004; Laurent and Naraghi, 1994; Wehr and Laurent, 1996); each KC fires if it encounters an “appropriate” instantaneous PN activity vector within one oscillation cycle. KCs may thus be thought of as binary classifiers of PN activity vectors (Laurent, 2002).

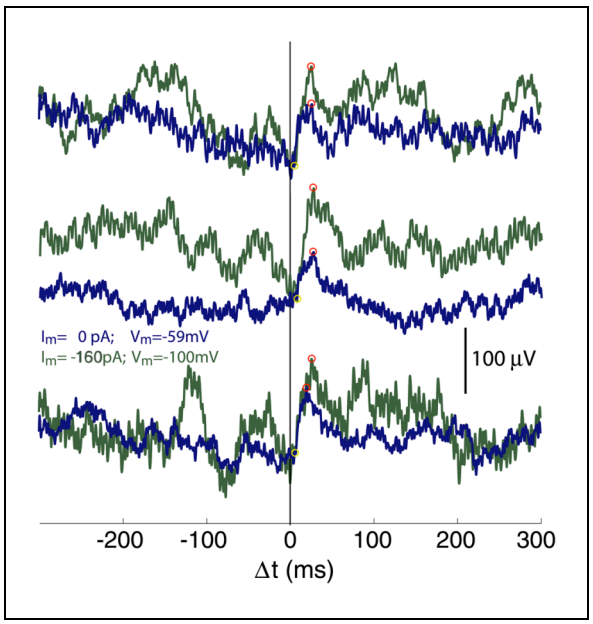
SI 3: *Technical issues*

a. Do the STAs represent EPSPs?

While the feed-forward circuits we studied are relatively simple, estimating the statistics of connectivity between two neuron populations in the brain is difficult (Chadderton et al., 2004; Franks and Isaacson, 2006). First, determining what constitutes a *bona fide* EPSP from a spike-triggered average waveform can be ambiguous, because average EPSP shapes are affected by noise (and thus, by event numbers) and by spike discharge statistics. We used three different methods to detect putative EPSPs, of which one only was based on visual inspection. All three yielded the same results. We also controlled for the effect of PN spike autocorrelations and found that the shape of the averaged KC EPSP was always consistent with the expected effect of PN discharge statistics. Second, high firing correlations between putative presynaptic neurons could influence estimates of connectivity. This is unlikely here: PN-PN spike time correlations calculated over 75 pairs during baseline were flat at all relevant time intervals. We also sampled more than 100 PN-PN pairs with simultaneous tetrode and intracellular recordings (data not shown): not a single direct connection was ever detected between PNs (Ron Jortner, Ofer Mazor and Gilles Laurent, in preparation). Third, experimental design could have led to an undesired bias in our sample; this is unlikely because all putative connections were detected *post hoc*, only after the PN spikes had been clustered and separated. Also, the mean KC EPSP amplitude was so small that single events could never be detected on line from within the noise of our intracellular recordings. The only selection criteria we applied to our data were based on recording quality, and were applied prior to any assessment of connectivity (see methods). This led to the elimination of over half of our initial data-set of paired recordings; loosening our selection criteria for analysis (for example by including PNs with less well separated waveform clusters) led to nearly identical results on

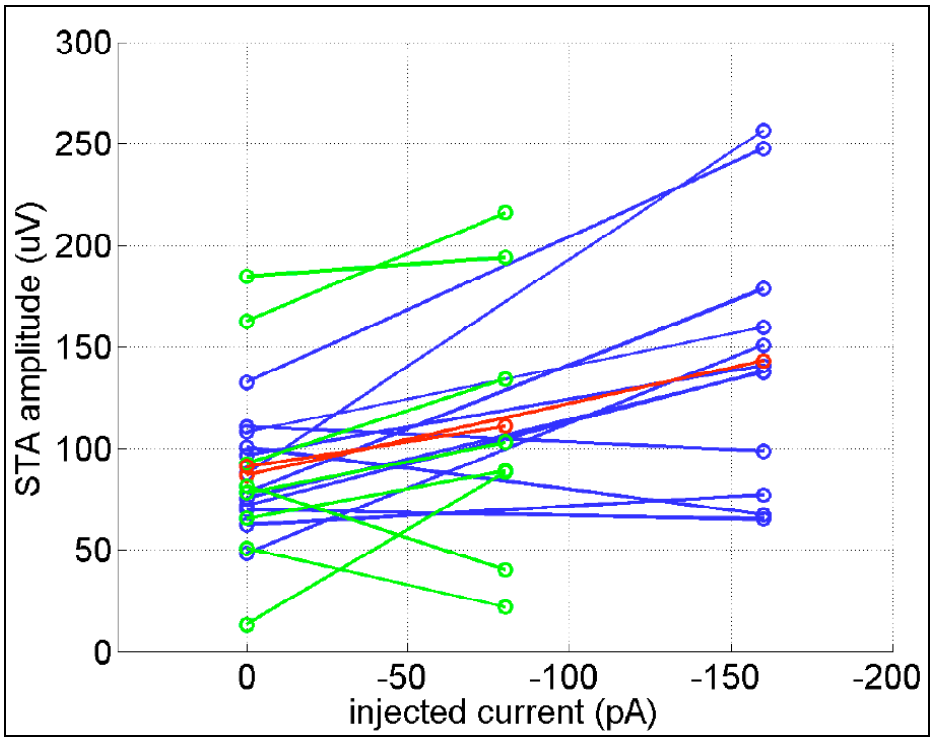
connection statistics. Fourth, PNs (tetrodes) and KCs (intracellular) were selected randomly and in all accessible regions of the antennal lobe and mushroom body, consistent with the widespread PN axon collateral projections in the mushroom body in this and other large insects (Ernst and Boeck, 1983). Together, these observations increase the confidence in our estimate that PN-KC mean connectivity is around 50 ± 13% ($p<0.05$). This estimate is consistent with three further observations: that there always existed physical overlap between any labeled KC's dendritic tree and PN's axon collateral arbor; that the modulations of membrane potential recorded in any single KC were highly correlated with those of a simultaneously recorded local field potential, reflecting the activity of all PNs; that most KCs tested with more than one PN were found to be connected (as assessed by the STAs) with half of those PNs. We also note that EPSP amplitude was a function of KC membrane polarization: three examples are given below (SI Fig 2A), showing EPSPs at resting potential (blue) and at hyperpolarized potentials (green). Group data are plotted in SI Fig 2B for two values of direct hyperpolarizing currents. Means over experiments are shown in red.

SI Figure 2A:



These experiments confirm that PN-evoked EPSPs are chemically mediated and consistent with direct PN input.

SI Fig 2B:



b. Firing threshold estimates

Estimating the firing threshold of KCs (in mV and in numbers *f* of PN inputs) is more difficult. First, because their thresholds are high, KCs fire rarely (Perez-Orive et al., 2002); and opportunities for appropriate measurements are scarce. Second, because they were small, EPSP amplitudes were measured on averages of hundreds of sweeps. If PN-KC synaptic transmission failure rates were high, non-failure events could be significantly larger than the population mean (86 ± 44 μV). High failure rates seem unlikely, however, because we found no evidence for homosynaptic facilitation, normally correlated with low release-probability (Katz and Miledi, 1968; Dobrunz and Stevens, 1997). More real is the possibility of non-linear dendritic summation. Indeed, we know that KC dendrites

summate supra-linearly in a voltage-dependent manner (Perez-Orive et al., 2002; 2004; Laurent and Naraghi, 1994); a realistic estimate of the number of simultaneous PN inputs required to bring a KC to threshold would thus be lower. We also know that PN spike times during individual odor-evoked oscillation cycles are distributed (±10 ms) (Laurent and Davidowitz, 1994; Wehr and Laurent, 1996) and that the decay time constant of KC EPSPs decreases to just a few ms when KCs are sufficiently depolarized (Perez-Orive et al., 2002; Laurent and Naraghi, 1994): jitter in PN spike times and sharpening of KC EPSPs together lead to sub-linear summation. While acting as a selective filter for tightly synchronized PN inputs, these nonlinearities act on summation in opposite ways; the extent to which one might dominate the other is so far unknown. With these observations, our initial estimate of the KCs' firing thresholds was $f \approx 100$ concurrent PN inputs per oscillation cycle (*i.e.*, about a quarter of their mean inputs). This range for f is commensurate with the mean number of PN spikes produced per oscillation cycle (100-150) during an odor stimulus (Mazor and Laurent, 2005) but, as described in the discussion section, probably unrealistically high.

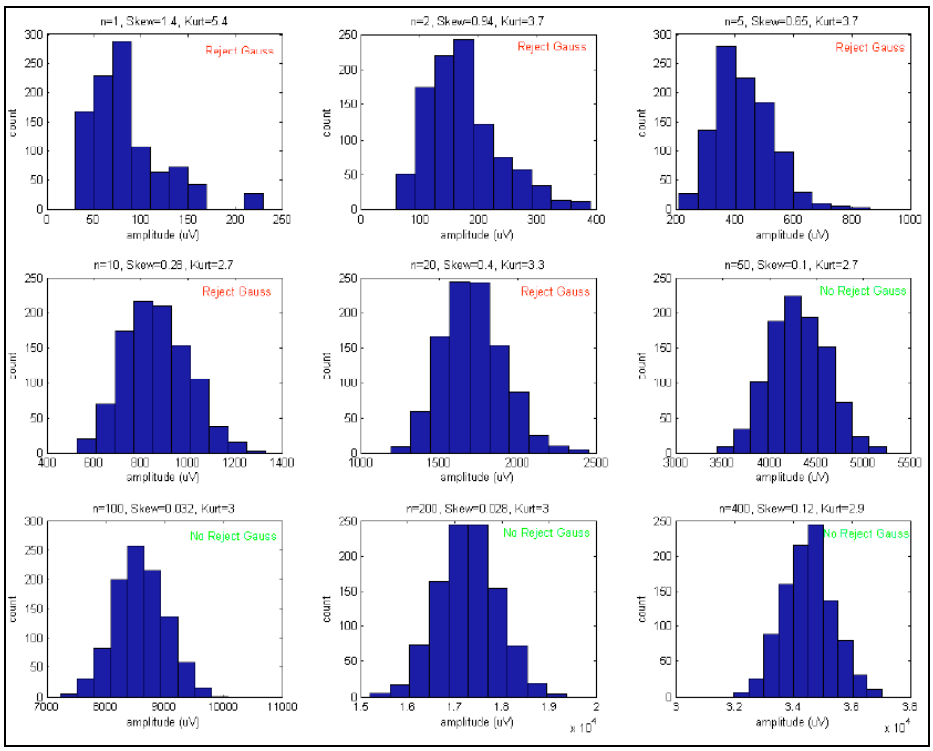
c. Importance of EPSPs' amplitude distribution for threshold estimate.

Our data reveal a distribution of EPSP amplitudes; we estimated the firing threshold f (in numbers of concurrent PN spikes) by dividing the KC spike threshold by the mean amplitude of the PN-evoked EPSPs. Is, however, the EPSP amplitude distribution Gaussian? And if not, would our estimate of f be different if we took the shape of the distribution into account?

We first note that the EPSP amplitude distribution was indeed skewed (3^{rd} and 4^{th} moments of the amplitude distribution: 1.4 and 5, respectively). Under such conditions, the *Central Limit Theorem* says that if a set of n independent random variables come from a distribution with a finite variance, then their sum will be approximately normally distributed and increasingly so as n approaches infinity.

We thus ran simulations, by drawing sets of *n* random events (EPSPs) randomly from the experimental (skewed) EPSP amplitude distribution (where *n*=1, 2, 5, 10, 20, 50, 100, 200 and 400) and measured their sum, 1,000 times over. As predicted, the distributions of those summed EPSPs became more Normal as *n* increased (SI Figure 3); for $n \geq 50$,the Gaussian hypothesis could no longer be rejected (Lilliefors test, $p<0.05$).

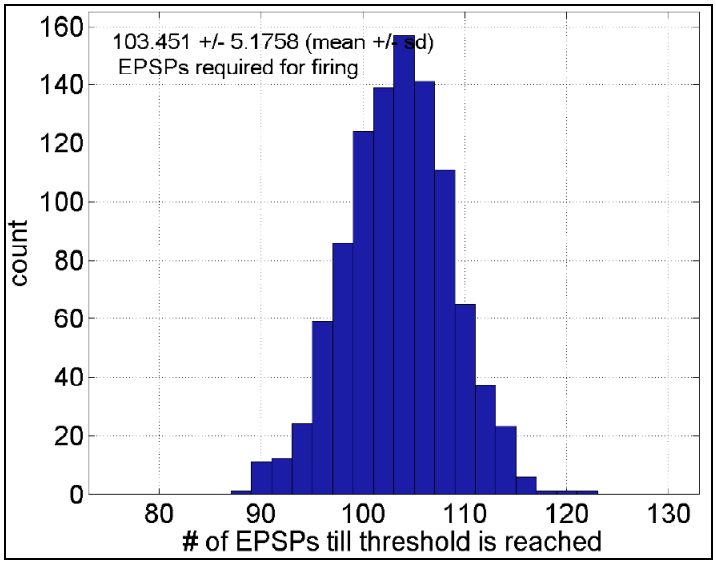
SI Figure 3:



We then ran a set of simulations in which we drew, one by one, EPSPs out of the skewed, experimental distribution (with mean amplitude = 86μV) and summed them until spike threshold (8.9mV) was crossed. This operation was repeated 1,000 times. The frequency distribution of *f* (number of simultaneous spikes, or EPSPs) is shown in SI Figure 4. In these simulations, the number *f* of inputs required to cross the threshold was 103.4±5 (mean ± SD, 1000 draws), in excellent agreement with the predicted value: 103.1±4.4 (assuming Gaussian

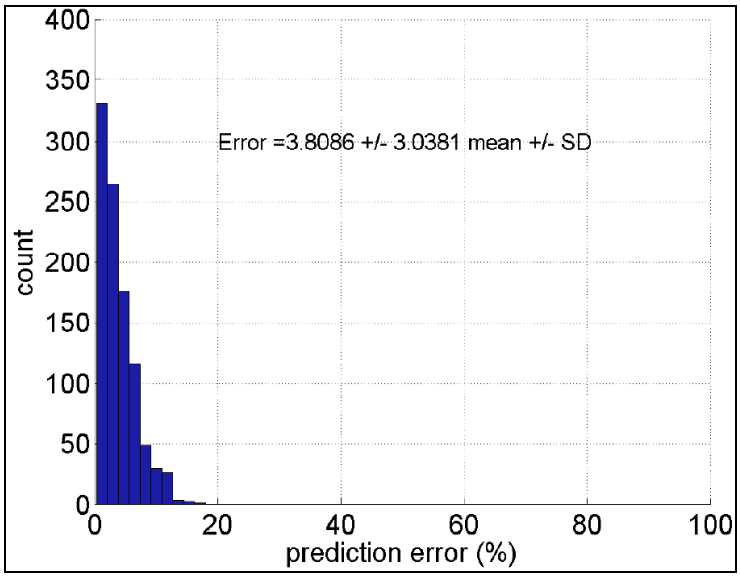
distribution of the sum).

SI Figure 4:



We plot below the distribution of RMS error (in %) between simulated and predicted values for *f*. The error is less than 4% on average (SI Fig 5).

SI Figure 5:



In conclusion, our estimates of *f* are only minimally influenced by the fact

that the EPSPs are drawn from a distribution that is not Gaussian, in great part because thresholds are high (>50).

SI 4: *Expressing KC response probability p given antennal lobe state (a) and KC firing threshold (f)*

Method 1.

Let us assume some network state of the antennal lobe, with *a* active PNs out of *n.* A hypothetical Kenyon cell samples *m* out of the *n* PNs. We will express the probability that this network state drives the KC across its firing threshold, *f* (expressed in numbers of PN inputs). If $f > a$, the threshold can never be crossed; we will therefore assume $f \leq a$.

We begin by calculating the probability *p* that **exactly** *f* active PNs fall within the *m* sampled by the KC. The total number of ways to pick *m* neurons out of *n* is

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} \qquad (1).$$

The number of combinations where **exactly** *f* active cells are in the group sampled by the KC is

$$\binom{a}{f}\binom{n-a}{m-f}$$

Dividing this number by the total number of combinations, we get the probability

$$p(pick \cdot exactly \cdot f) = \binom{n}{m}^{-1}\binom{a}{f}\binom{n-a}{m-f} \qquad (2).$$

Similarly, the number of combinations where **at least** *f* active PNs $(f \leq a)$ fall within the *m* sampled by the KC is

$$\sum_{i=f}^{a} \binom{a}{i}\binom{n-a}{m-i},$$

Again dividing by the total number of combinations, the probability is

$$p(pick \cdot at \cdot least \cdot f) = \underline{\binom{n}{m}^{-1} \sum_{i=f}^{a} \binom{a}{i}\binom{n-a}{m-i}} \qquad (3).$$

Method 2.

A different way to approach this problem is as follows. We choose one KC, hard-wired to *m* PNs (out of *n*), and draw an antennal lobe network state with *a* active PNs (among *n*); what is the probability *p* that at least *f* of the active PNs fall within the *m* sampled by the KC?

The total number of ways to pick *a* active PNs out of *n* is

$$\binom{n}{a} = \frac{n!}{a!(n-a)!} \qquad (4).$$

The number of combinations where *f* active PNs are in the group sampled by the KC is

$$\binom{m}{f}\binom{n-m}{a-f}.$$

Thus, the probability is:

$$p(pick \cdot exactly \cdot f) = \binom{n}{a}^{-1}\binom{m}{f}\binom{n-m}{a-f} \qquad (5).$$

Similarly, the number of combinations where **at least** *f* active PNs *(f ≤ a)* fall within the *m* sampled by the KC is

$$\sum_{i=f}^{a}\binom{m}{i}\binom{n-m}{a-i},$$

Dividing this number by the total number of combinations, the probability is:

$$p(pick \cdot at \cdot least \cdot f) = \binom{n}{a}^{-1} \sum_{i=f}^{a} \binom{m}{i}\binom{n-m}{a-i} \qquad (6).$$

Equivalence.

Equations (3) and (6) are equivalent:

$$\underline{\binom{n}{a}^{-1} \sum_{i=f}^{a} \binom{m}{i}\binom{n-m}{a-i}} = \frac{a!(n-a)!}{n!} \sum_{i=f}^{a} \frac{m!(n-m)!}{i!(m-i)!(a-i)!(n-m-a+i)!}$$

$$= \frac{m!(n-m)!}{n!} \sum_{i=f}^{a} \frac{a!(n-a)!}{i!(a-i)!(m-i)!(n-m-a+i)!}$$

$$= \binom{n}{m}^{-1} \sum_{i=f}^{a} \binom{a}{i}\binom{n-a}{m-i}.$$