

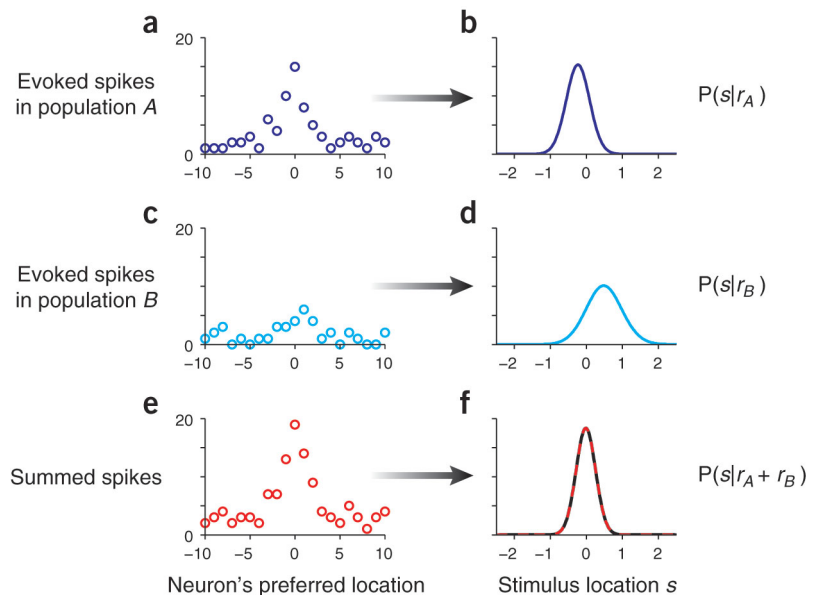
# Noisy neurons can certainly compute

Emilio Salinas

**How do neurons combine separate pieces of information that are only partially reliable? Surprisingly, their noise properties may simplify the underlying computations while allowing them to maintain optimal performance.**

Uncertainty is encountered in nearly every aspect of life. For instance, in a murder trial, the key problem is determining whether particular events happened or not, and the jury must reach a conclusion by weighing multiple pieces of evidence—testimonies, physical exhibits, prior histories—each of which is only partially reliable. However, the pooling of multiple sources of partially reliable information is fundamental even for the simplest perceptual processes: when I'm trying to find my cat, who often escapes into the back yard, a slight movement of shadows and a weak sound of trodden leaves may indicate that he is hiding somewhere behind the bushes. If combined properly, the two cues may provide a better estimate of the cat's position than either of them alone, but as in the case of the trial, the two signals may not be equally informative or equally reliable. Thus, neural circuits must be specially adapted to deal with uncertainty at multiple levels, and to weigh different signals according to their corresponding reliabilities. Now, new theoretical results by Ma and Beck in Alex Pouget's laboratory, in collaboration with Peter Latham, provide an important and surprising clue to this problem<sup>1</sup>.

Mathematically, the rules for calculating optimal probabilistic estimates are known as Bayesian inference, and psychophysical experiments show that, under various laboratory conditions, human performance in simple tasks that require the combination of uncertain cues is essentially as good as it can be given the available information<sup>2–4</sup>. Such optimal performance indicates that the



**Figure 1** Optimal pooling of unreliable signals is simplified by Poisson noise. (a) Each point represents the number of spikes produced by a model neuron in population A in response to the presentation of a stimulus. The set of 21 responses is denoted as  $r_A$ . (b) A value along the x-axis indicates a possible stimulus location. The corresponding value on the curve is the probability that, having observed the responses in a, the stimulus location was s. (c,d) As in a and b, but for population B, which has a smaller overall response amplitude. This causes the probability density  $P(s|r_A)$  to be wider than  $P(s|r_B)$ . (e) Summed spikes from populations A and B. (f) The red trace is the probability density for stimulus location s given the summed responses. The black trace is the product of the curves in b and d, normalized, which is the optimal way to take into account both  $r_A$  and  $r_B$ . The two traces are identical only when response noise is Poisson-like<sup>1</sup>. The stimulus that evoked the responses in this example was located at 0.

brain must implement some form of Bayesian inference, but so far, how this happens has remained unclear. The new paper by Ma and colleagues<sup>1</sup> shows that the variability of cortical neurons has a form that greatly simplifies one of the critical operations in this process, the pooling of individual cues.

To appreciate these results, first note that cortical neurons are highly variable, in the sense that responses evoked by identical stimuli typically change dramatically from

one presentation to the next, approximately following Poisson statistics<sup>5–7</sup>. So, if the average response of a neuron in a given condition is 10 spikes, its variance is also roughly equal to 10 spikes. To put this in perspective, radioactive decay, which is the prototypical example of a random process, closely follows Poisson statistics. Such high variability could conceivably serve some important functional purpose, such as speeding up the responses of a neuronal

Emilio Salinas is in the Department of Neurobiology and Anatomy, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157-1010, USA.  
e-mail: esalinas@wfuwmc.edu

population<sup>8</sup>, or making a network more robust to synaptic loss<sup>9</sup>, but it is commonly viewed as a nuisance that mainly reduces the computational accuracy of neural circuits. Nuisance or not, it turns out that constructing model neurons with realistic variability is rather difficult<sup>10–12</sup>, so such ‘neuronal noise’ has for some time been puzzling. In a somewhat ironic turn of events, Ma *et al.* point out that neurons can compute probabilities efficiently precisely because their variability is Poisson-like.

The advantage of Poisson noise is illustrated in **Figure 1** using two populations of model neurons, *A* and *B*. Both populations consist of 21 neurons whose firing rates encode the value of a stimulus, *s*. For population *A*, this stimulus may correspond, for instance, to the position of my cat’s shadow along the hedge, whereas for *B* it may correspond to the position of the sound that he makes while stepping on leaves, measured along the same dimension. As is often the case with real neurons, each model neuron fires maximally when the stimulus originates from a particular point, the neuron’s preferred location. Here, the auditory cue is assumed to be less accurate than the visual one. In the model, accuracy is controlled by the maximum spike rate in a population, so the maximum spike rate of *B* is smaller than that of *A*. The goal of the example is to compare how accurately we can infer the cat’s location when the two groups of neurons are analyzed independently (**Fig. 1a–d**) versus when they are pooled optimally (**Fig. 1e,f**), following the Bayesian rules.

Having obtained the two sets of neuronal responses, the next step is to estimate the cat’s position from each of them independently. How this is done is not really important now—it may be done through a dedicated neural circuit or through some other downstream process. What matters is that there is a separate step, the Bayesian decoder, that can generate an estimate for a single population. In fact, what the decoder produces is a probability function describing how likely it is for the cat to be found at each point along the hedge. The probability functions derived from the activities in *A* and *B* are  $P(s|r_A)$  and  $P(s|r_B)$ , respectively (**Fig. 1b,d**). The former is narrower than the latter because the *A* population is more reliable, so its determination of *s* is more localized, that is, less uncertain.

Up to this point, the form of the noise does not make any difference. However, the

next crucial step is to consider the *A* and *B* activities together to produce a single, most accurate estimate of *s*. According to Bayesian inference, the recipe for this is to multiply the two individual probability functions. Although this is not particularly complicated mathematically, it is not at all obvious how neurons could do it. The problem is that to perform the calculation in the straightforward way, the brain would need to process each population’s activity using the decoder, maintain an explicit representation of the individual probability functions, and then somehow multiply them point by point. Furthermore, if the two sets of responses were combined incorrectly, a significant loss of information could ensue, making the estimate far worse. This is where the new results come to the rescue.

The main observation made by Ma *et al.* is that if the *A* and *B* activities are just added, neuron by neuron, the combined activity that results can be processed by the Bayesian decoder in the same way as either of the original two, but with the advantage that the result will be as accurate as possible, as if the two probability functions had been multiplied (**Fig. 1e,f**). Mathematically, the statement is  $P(s|r_A + r_B) = c P(s|r_A) P(s|r_B)$ , where *c* is a normalization constant. This is a great simplification, regardless of what the decoder actually is, but it is valid only when the neuronal noise is Poisson-like.

In this example, there are many simplifying assumptions. However, what sells the new results is their generality (see the supplementary material linked to ref. 1). When the two populations have different numbers of neurons with different types of response functions (tuning curves) that are distributed irregularly along the stimulus range, the proper combination of signals is just slightly more complicated—instead of simply adding the activities neuron by neuron, they need to be added or subtracted in various proportions. Thus, in the general case, Poisson-like noise makes the optimal pooling operations linear, which is basically as simple as they could possibly be.

What Ma *et al.* have done is interpret the variability of neurons in a novel way, such that neuronal variability is used to represent uncertainty. The idea that populations of neurons can represent both the average value of a quantity and its uncertainty is not new<sup>13</sup>. The new insight is that the particular form of noise that most cortical neurons display, Poisson noise, simplifies tremendously

the pooling of partially reliable pieces of information while adhering to the strict rules of optimal probabilistic inference.

This will definitely not be the end of the story, though—neuronal noise and the neural implementation of Bayesian computations are issues of great interest. First, the theory makes specific predictions that must be checked experimentally. The most direct confirmation would be to find that responses to separate, uncertain cues are indeed combined linearly during optimal performance of a cue-integration task. This would require some clever task designing and sophisticated neuronal recording, but should be possible. Also, not all cortical neurons are expected to be concerned with probabilistic representations, so specific correlations may exist between differences in function and differences in noise properties. Second, the theory covers only one step in the chain of events between sensory activation and a motor reaction, but other contextual factors, such as behavioral goals or prior expectations, might also influence how neuronal signals are integrated. Finally, other schemes for computing probability or likelihood functions have been proposed<sup>14</sup>, and there may be more alternatives.

Nevertheless, the possibility that neuronal noise is essential for implementing a Bayesian calculus is extremely exciting, because it is consistent with the assumption that the brain must perform very near an optimal point, limited by fundamental constraints such as energetic costs or neuronal numbers, not by sloppiness in its computing elements.

1. Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. *Nat. Neurosci.* **9**, 1432–1438 (2006).
2. Ernst, M.O. & Banks, M.S. *Nature* **415**, 429–433 (2002).
3. Kording, K.P. & Wolpert, D.M. *Nature* **427**, 244–247 (2004).
4. Miyazaki, M., Yamamoto, S., Uchida, S. & Kitazawa, S. *Nat. Neurosci.* **9**, 875–877 (2006).
5. Dean, A. *Exp. Brain Res.* **44**, 437–440 (1981).
6. Holt, G.R., Softky, W.R., Koch, C. & Douglas, R.J. *J. Neurophysiol.* **75**, 1806–1814 (1996).
7. Compté, A. *et al. J. Neurophysiol.* **90**, 3441–3454 (2003).
8. Van Vreeswijk, C. & Sompolinsky, H. *Science* **274**, 1724–1726 (1996).
9. Basalyga, G. & Salinas, E. *Neural Comput.* **18**, 1349–1379 (2006).
10. Softky, W.R. & Koch, C. *J. Neurosci.* **13**, 334–350 (1993).
11. Shadlen, M.N. & Newsome, W.T. *Curr. Opin. Neurobiol.* **4**, 569–579 (1994).
12. Stevens, C.F. & Zador, A.M. *Nat. Neurosci.* **1**, 210–217 (1998).
13. Zemel, R.S., Dayan, P. & Pouget, A. *Neural Comput.* **10**, 403–430 (1998).
14. Jazayeri, M. & Movshon, J.A. *Nat. Neurosci.* **9**, 690–696 (2006).