Biol Cybern (2006) 94: 180–191
DOI 10.1007/s00422-005-0039-3

ORIGINAL PAPER

Biological Cybernetics

Benoit Gaillard · Hilary Buxton · Jianfeng Feng

Population approach to a neural discrimination task

Received: 24 May 2004 / Accepted: 1 November 2005 / Published online: 6 December 2005

Abstract This article gives insights into the possible neuronal processes involved in visual discrimination. We study the performance of a spiking network of Integrate-and-Fire (IF) neurons when performing a benchmark discrimination task. The task we adopted consists of determining the direction of moving dots in a noisy context using similar stimuli to those in the experiments of Newsome and colleagues . We present a neural model that performs the discrimination involved in this task. By varying the synaptic parameters of the IF neurons, we illustrate the counter-intuitive importance of the second-order statistics (input noise) in improving the discrimination accuracy of the model. We show that measuring the Firing Rate (FR) over a population enables the model to discriminate in realistic times, and even surprisingly significantly increases its discrimination accuracy over the single neuron case, despite the faster processing. We also show that increasing the input noise increases the discrimination accuracy but only at the expense of the speed at which we can read out the FR.

1 Introduction

Discriminating between inputs is a basic task for the visual system because it underlies recognition, decision, reasoning and action. Experiments with random dot stimuli are classical ways to carry out a discrimination task. In such experiments, the task is to find the global direction of motion in a visual display consisting of random dots, a certain percentage of which are moving coherently in one direction or its opposite. Thus the difficulty of the task depends on the number of dots moving coherently. Newsome, Shadlen and Coworkers (1992, 1994, 1996, 2001) have experimented on this discrimination process in macaque monkeys since the macaque visual system is one of the best understood and also currently one of the best animal models we have for the human visual system. The monkeys had to watch the random dots stimulus and then make a deliberative eye movement towards one of two targets that indicated the direction of motion in this display. The experimenters studied the Firing Rates (FRs) of neurons situated in the lateral intraparietal area (LIP) of the cortex, and in the extrastriate visual cortex (areas MT and MST), of monkeys that were performing the simple discrimination task.

The FRs of the neurons in areas MT and MST accurately follow the directional information of the display, and the expected performance of an ideal observer discriminating the direction according to those FRs matches the psychological performance of the monkeys. Furthermore, the FR of the LIP neuron allows the experimenter to guess the action the monkey will chose, even when the decision is in error. This is characteristic of a motor control neuron, but the FR of the LIP neuron, for a given decision, also depended on the actual global direction of the dots. These neurons are part of the pathway between perception and decision, without being pure sensory or motor neuron, they illustrate a typical sensorimotor decision process. Recently, a wealth of experimental findings and model predictions suggest that what happens here is a process of accumulation of evidence, whereby a competition is induced between pools of neurons, and when the difference of activity reaches a threshold, then the decision is taken. We have modelled a building block of this more general decision-making system. We model neurons of area MT that, solely influenced by the stimulus, provide the evidence that informs the competition between the pools of neurons in area LIP. This role of MT neurons has been experimentally demonstrated by Parker and Newsome (1998),Britten et al. (1996), and Britten (2003). The discrimination accuracy of the MT neuron is a fundamental parameter for the dynamics of the decision making based on the evidence they provide. The time needed to read out the output of these neurons is also a critical parameter for decision-making models.

B. Gaillard (✉) · H. Buxton
Department of Informatics, Sussex University, Brighton BN1 9QH, UK
E-mail: bg22@sussex.ac.uk; hilaryb@sussex.ac.uk

J. Feng
Department of Computer Science, Warwick University, Coventry CV4 7AL, UK
E-mail: jianfeng.feng@warwick.ac.uk

We have implemented a detailed model of the MT neurons that provide the evidence for the decision making of monkeys during a two choice discrimination task, and studied how its performance varies with model parameters. We show that increasing r, the ratio between inhibitory and excitatory inputs to the neuron, significantly increases the discrimination accuracy of the model, although it also increases the postsynaptic input noise, traditionally considered harmful. Using an Integrate-and-Fire (IF) model, and modelling the production of each action potential, enables us to study timing issues inherent in the overall decision-making process: How long do we need to read out the FR of the model reliably enough to reach an acceptable discrimination accuracy?

Hence, as well as measuring the FR of one single neuron, we measured it over a population of neurons, which is much more biologically plausible. It enables us to model discrimination within a realistic time scale, and is consistent with the studies of Newsome and coworkers "Perceptual judgments of motion direction are thought to depend on the pooled activity of at least 100 neurons" (Parker and Newsome 1998). The population consists of identical neurons that have strictly feedforward connections. Furthermore, we discovered that the discrimination accuracy of the population model is surprisingly better than that of the single neuron, even if we use the same number of spikes to measure the FRs.

In Sect. 2, we describe the benchmark visual dicrimination task on which we apply our model. Then, in Sect. 3 we describe in detail the methods that we used: we explain the assumptions made for the IF model, the structure of the population of neurons, and the parameters that we control in order to monitor the input noise. In Sect. 4, we present the results and their interpretation, and in Sect. 5 we relate our contribution to current research on visual decision making.

2 The discrimination task

The discrimination task is the one that is used as a benchmark task for decision theory by many authors (e.g., Wang 2002; Shadlen and Gold 2004) and that has been set up by Newsome and coworkers (e.g. 1996). In this set of experiments, the monkeys had to watch a display of dots, a certain percentage of them moving consistently in one direction or its opposite, and the rest of the dots appearing at random places on the screen as a perturbing noisy context. Then they had to signify, by an eye movement, which was the consistent direction. Within this framework, the task could be made more or less difficult by modifying the percentage of coherently moving dots.

Since we focus on the discrimination accuracy of the IF model, our connections are strictly feedforward and we do not take into account past information. We model only one key stage of the decision making, the discrimination of stimuli on which the accumulation of evidence is based.

We assume that the discriminating neuron receives p synaptic input composed of an actual signal perturbed by a noise. If a percentage n_c/p of dots moves coherently in one direction, the same percentage of the synapses will receive coherent input. Furthermore, we assume that the spike trains arriving at those synapses are correlated. The rest of the $p - n_c$ synapses will receive randomly distributed inputs. The synaptic inputs are modelled as Poisson processes. This is a very simple model of the input to MT neurons, but it has been shown with experimental evidence that motion detectors in area V1 and in area MT/MST are constituted of columns of neurons as expressed in Tovée (1996), and models have been proposed for this organization by Heeger et al. (1995); Simoncelli and Heeger (1998).

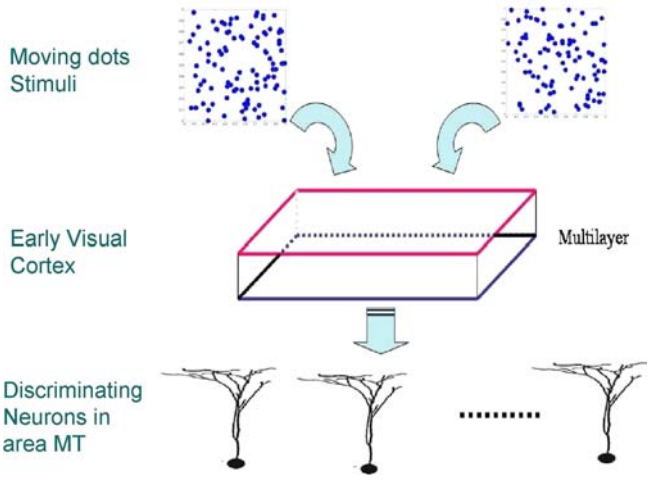

Fig. 1 In article we use a schematic plot of the model. The multilayer part represents the early visual cortex, from the retina to MT. It is treated as a "black box" that gives a preprocessed input to the represented discriminating neurons of area MT

The outputs of the discriminating neuron are spike trains whose FRs are related to the input of the movement. In order to evaluate the discrimination accuracy, an ideal observer uses the following rule: if the FR is larger (or smaller) than a criterion, it represents the decision that the dots move up (or down). The criterion is determined statistically, after measuring, over hundred trials, the two first movement of the distribution of the FRs (One FR per trial). See the Appendix for more details. Because of the noise in the stimulus, and because of the synaptic input noise, this decision can be erroneous, for example, the FR can be bigger than the criterion when the movement is downwards. Of course, the clearer the stimulus, the more widely separated the efferent spike trains, and thus the fewer errors the model makes.

3 The neuron model

3.1 The Integrate-and-Fire model

The discriminating neuron model used here is the classical IF model, as used in Gerstner and Kistler (2002), Tuckwell (1988) and Feng (2001). The dynamics of the membrane potential below threshold are defined as follows:

$$\mathrm{d}V = -L(V - V_{\mathrm{rest}})\mathrm{d}t + \mathrm{d}I_{\mathrm{syn}}(t)$$

where L is the decay coefficient and the synaptic input is

$$I_{\text{syn}}(t) = a\sum_{i=1}^{p} E_i(t) - b\sum_{j=1}^{q} I_j(t) \tag{3.1}$$

with $E_i(t)$ and $I_j(t)$ being Poisson process with rates $\lambda_{i,E}$ and $\lambda_{i,I}$, a and b being the positive magnitudes of the EPSP and IPSP, and p and q the numbers of excitatory and inhibitory synapses.

As seen in the previous subsection, we modelled each neuron to have 100 synapses. Each synapses receives a Poisson process with a rate λ_i that is proportionally linked to the direction of one of the 100 moving dots on the screen, and normalized between 0 and 100 Hz. For example, if one dot moves rightwards, the rate of the Poisson input to the corresponding synapse will be 0, if it moves downwards, the rate will be 25 Hz, leftwards, 50 Hz and if the dot moves upwards, the rate of the Poisson process to the corresponding synapse will be 75 Hz. So for n_c dots that move coherently, n_c synapses receive coherent inputs, namely $\lambda = 25\,Hz$ for a downwards movement, and $\lambda = 75 Hz$ for an upwards movement. Of course, one neuron in area MT has many more synapses than 100, and more than one synapse receives information concerning one moving dot. However the IF model that we use enables us to group together synapses that receive similar input and consider them as one. The n_c synapses that receive coherent inputs are correlated by a constant c, that reflects the correlation of activity of different synapses as studied in Feng (2001) and Zohary et al. (1994).

The main parameter that we study is the ratio between inhibitory and excitatory synaptic inputs, we call this ratio r: $r = \frac{q}{p}$, from Eq. (3.1). Then we simplify the equation by changing the notation and letting $\lambda_{i,I} = r\lambda_{i,E}$, and $p = q$. This simplification is justified by the fact that the summation of two Poisson processes with a rate $m1$ and $m2$, respectively, will be a Poisson process with a rate $m1 + m2$. So, the summation of $q = p$ Poisson processes with rate $r\lambda_{i,E}$ will be a Poisson process with a rate $pr\lambda_{i,E}$. This quantity of inhibitory postsynaptic input is exactly the same as the one we would obtain with the original notation. This simplification is commonly used, e.g. by Feng (2003). Then, using the diffusion approximation, as is common in the literature (Tuckwell 1988; Ricciardi and Sato 1990) and assuming that $a = b$, we reach the simplified following description of the dynamics of our discriminating neuron:

$$\mathrm{d}V = -\frac{V\mathrm{d}t}{\gamma} + \mu dt + \mathcal{N}\sigma\sqrt{\mathrm{d}t}$$

where:

$$\mu = a\sum_{j=1}^{p}(1-r)\lambda_j;$$

$$\sigma^2 = a^2\left[\sum_{j=1}^{p}(1+r)\lambda_j + \sum_{i,j=1,i\neq j}^{n_c} c(1+r)\sqrt{\lambda_i\lambda_j}\right]$$

with V being the membrane potential. In our studies:

- The ratio between inhibitory inputs and excitatory inputs: r is variable.
- The number of synapses (corresponding to the number of dots in an experiment): $p = 100$.
- $\lambda(j)$ is the rate of the Poisson process incoming to the jth synapse. It is proportional direction of the jth dot. $0\,Hz < \lambda < 100\,Hz$, the coherent signals being $\lambda = 25\,Hz$ or $\lambda = 75\,Hz$.
- The time decay parameter: $\gamma = 20\,ms$.
- The time step for the integration: $\mathrm{d}t = 0.01\,ms$.
- N_{spikes} is the number of output spikes that we use to measure the FR.
- The correlation coefficient between coherent motion: $c = 0.1$.
- The number of coherent inputs: $n_c \leq p$ is variable. Coherent inputs are dots that move consistently in one direction.
- The resting membrane potential: $V_{rest} = 0\,mV$.
- The threshold membrane potential: $V_{threshold} = 20\,mV$.
- $\mathcal{N}$ is a normally distributed random variable (mean 0, variance 1); in the formal IF model, $\mathcal{N}\sqrt{\mathrm{d}t}$ is the standard Brownian motion.

3.2 The ratio between inhibitory and excitatory inputs (r) augments the input noise

We can interpret the equation of the dynamics of the membrane potential of the IF model (3.1) as a leaky neuron ($-\frac{V\mathrm{d}t}{\gamma}$) whose synapses receive a Poisson input of rate μ ($+\mu\mathrm{d}t$), perturbed by a synaptic stochastic noise ($\mathcal{N}\sigma\sqrt{\mathrm{d}t}$). In effect, $\mathcal{N}$ is a normally distributed random variable. Since this stochastic perturbation is proportional to $(1 + r)$ and the mean is proportional to to $(1 - r)$, the stochastic effect of the synapses must increase with r. We have measured this input noise increase within the model, and presented the results in the curves represented in Fig. 2.

As explained in Feng (2001), an increase in the coefficient of variability in the input will increase the coefficient of variability of the efferent spike train of the neuron. This increased variability affects the width of the histograms that represent a set of measures of the output FR. Thus, intuitively, it should be more difficult to discriminate between two inputs from their efferent FRs. However, detailed investigations of the spiking cells have shown that the temporal jitter of the output spike is linked to the temporal jitter of the input synapses by the following equation: $\sigma_{out} = k\sigma_{in}$, with $k < 1$ (Koch 1999). That means that the variability of the input is significantly reduced by the neuronal computation. What is more is that Feng and his colleagues (Feng et al. 2003) have formally proved that the discrimination is easier when the coherent inputs (those upon which we discriminate) are correlated. In this article, we use a correlation coefficient of $c = 0.1$ for synapses that receive the coherent input. This is justified by the organization of area MT: Neurons detecting the same direction will be grouped together, as shown by Liu and Newsome (2003). This means that they will fire together

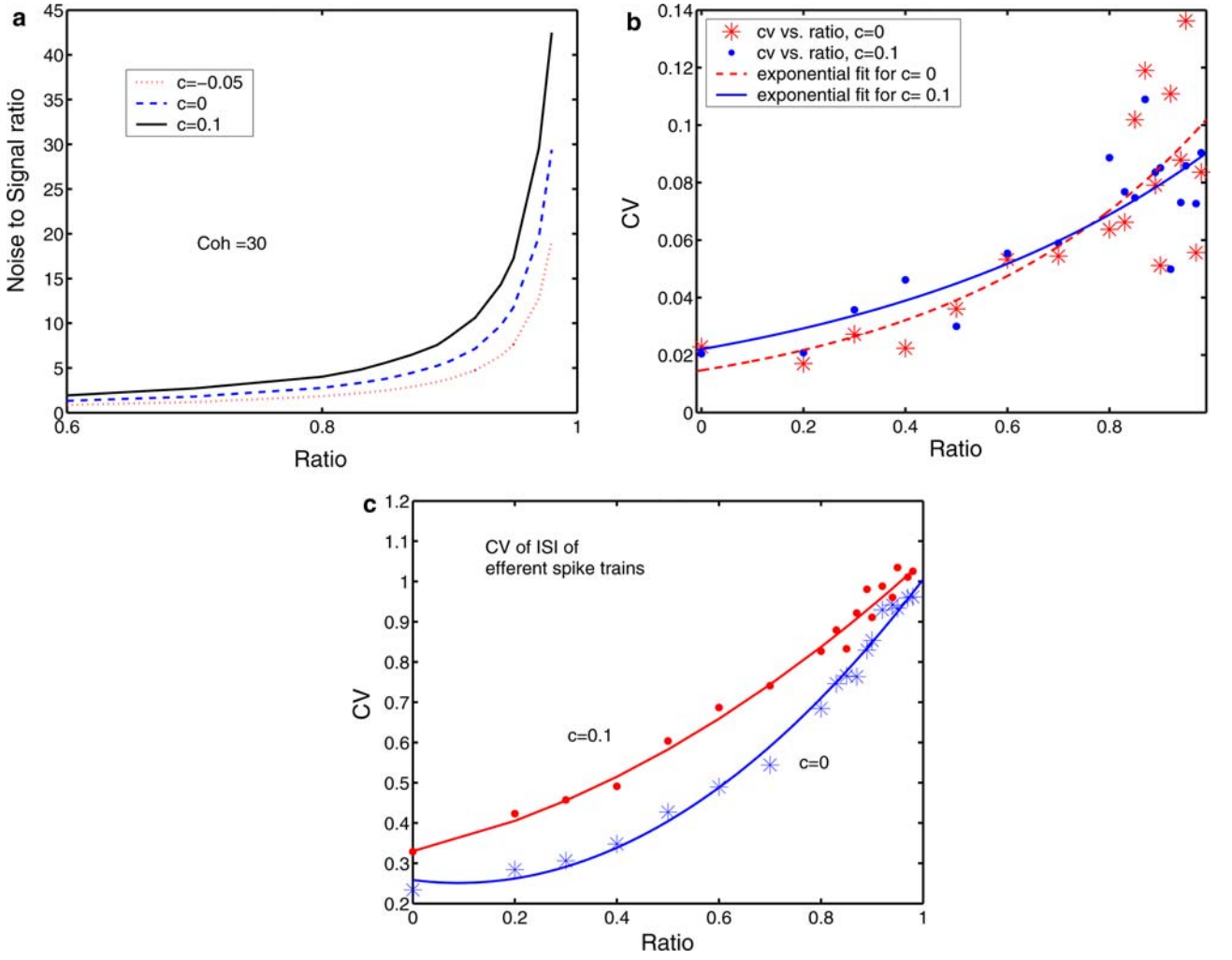

Fig. 2 In (**a**) we show the noise to signal ratio ($\frac{\sigma}{\mu}$, in the model) of the postsynaptic input to the neuron as a function of r. This ratio, that measures the importance of the second-order statistics in the synaptic input, increases dramatically when we increase the number of inhibitory synapses (i.e. r), and with synaptic correlation. The data was generated, for each value of r, by generating the stimulus 100 times, measuring the parameters σ and μ of the postsynaptic input, and taking the mean value of ($\frac{\sigma}{\mu}$ over these 100 repetitions. On curves (**b**) and (**c**), however, we see that the CV of the neuron's output is hardly affected by the synaptic correlation, only by r. Curve (B) shows the CV, that is, dispersion, of the FR(variance divided by mean of the sample of FRs measured on 100 experiments), with and without synaptic correlation ($c = 0$ or c=0.1). To measure the FR, we measured the time to 100 spikes, reversed it and normalized it. The curves are exponential fits with two parameters. The standard errors of the regression (root mean squared errors of the fits) are, respectively, RMSE $= 0.022$ ($c = 0$) and RMSE $= 0.014$ ($c = 0.1$). Curve (**c**) shows the CV of the Inter-Spike Interval over the hundred spikes used to measure the FR, with and without synaptic correlation ($c = 0$ or c=0.1).The curves are quadratic fits. The standard errors of the regression (root mean squared errors of the fits) are, respectively, RMSE $= 0.012$ ($c = 0$) and RMSE $= 0.009$ ($c = 0.1$).

as explained, for example, in Sheth et al. (1996). Furthermore, it has been shown (Zohary et al. 1994) that in area V5 of the visual cortex of the monkeys, the level of correlation is 0.1 and, although it is weak, has a significant impact on the global behavior.

3.3 Population coding

It has been shown by Feng et al. (2003) that measuring the FR over only ten spikes significantly reduces the discrimination accuracy of the neuronal system. However, especially in the case of a ratio r close to one, the FR becomes very small, because the generation of a spike becomes an exceptional event, entirely dependent on the random variable $\mathcal{N}$. This leads to unrealistic processing times, thus the FR approach is said to be much too slow and consequently, many people argue that it cannot be what is really used in the brain. Waiting for one neuron to emit a hundred spikes before taking a visual decision would indeed require unrealistic processing time, since we know that the visual system processes even complex information very fast, as argued, for example, in VanRullen and Thorpe (2000). They show that even for discriminating between very high-level categories of objects,

"we only have time for one spike per neuron," which leads them to reject the FR approach (VanRullen and Thorpe 2001).

In order to overcome this daunting flaw, instead of discriminating with the measured FR of one neuron (Feng and Liu 2002; Feng et al. 2003), we can discriminate with the measured FR of a population of neurons. The basic advantage is that each neuron only needs to fire a few spikes, which allows us to work with biologically realistic times. We can choose to measure this FR within a time window, and take into account only the spikes produced before 100 ms, for example, or we can decide to measure the FR on the first hundred or thousand spikes. The longer (or the more spikes), the more reliable the statistics are.

The population of neurons is organized very simply. We simulate a number n of strictly feedforward neurons (100, in this article), that is, they are not-laterally connected and there is no feedback. They are all the same as the neuron that we describe in 3.1, they all receive the same input. In order to measure the FR of the population, we measure the time t needed for the population to emit a number s of spikes (typically 100, in this article) and we divide this number s by n and t: $FR = \frac{s}{n \times t}$

4 Results

4.1 Input noise improves discrimination accuracy

We use the total probability of misclassification (TPM) as a measure of the performance of the system. The TPM is the number of errors divided by the total number of trials, and in fact corresponds to the degree of overlapping between the distributions of the outputs of the system over 100 trials, for each type of stimulus. We explain in greater detail, in the Appendix, how we calculate it.

We performed simulations of the discrimination task with one neuron, measuring the FR with its first hundred spikes. Measuring the FR over the first hundred spikes allows us not only to measure the discriminative performance of the neuron, but also to compare the time needed to compute the FR with the time of the population, as seen in the following subsection. We carried out experiments to measure the performance on the whole range of ratios between inhibitory and excitatory input, from $r = 0$ to $r \approx 1$, along the range of input coherence, from 0% of the dots moving coherently to 60%. (For larger coherences, the TPM is too small to be compared between the various ratios). When the correlation coefficient is $c = 0.1$ and the ratio r between inhibitory and excitatory inputs increases, the postsynaptic input noise increases, as seen in Sect. 3.2, but the discrimination improves. Figure 3 illustrates this result and its reasons: in the histogram representations of the efferent FRs for a simple discrimination task, the first-order difference (mean) decreases less than the second-order difference (width) between the two histograms. When the ratio r goes from $r = 0.98$ to $r = 0.7$, we measured that, whilst the difference between the mean FR of the outputs corresponding to upwards and downwards stimuli

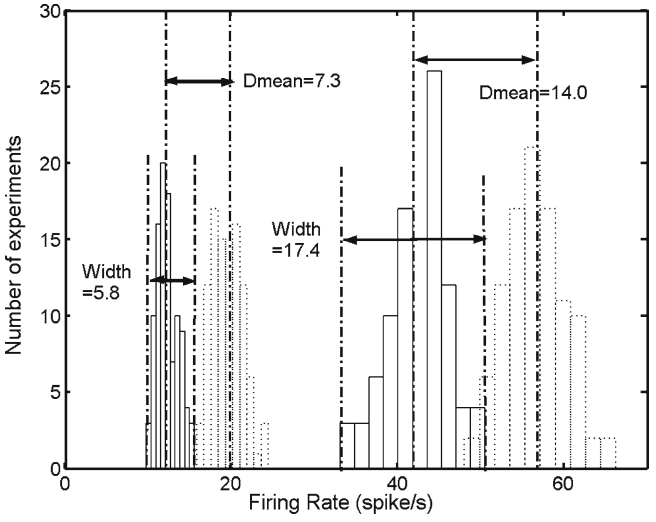

Fig. 3 Illustration of the improvement of the discrimination accuracy with r. Each histogram is a graphical representation of the distribution of the efferent FRs over a hundred repeats of the same experiment. From *left to right*: $r = 0.98$, *downward*; $r = 0.98$, *upward*; $r = 0.7$, *downward*; $r = 0.7$, *upward*. The coherence of the kinematogram was 15. We see that the difference between the means of two *histograms* decreases much less than the width of each histogram, when r increases. Whilst the difference between the means of upwards and downwards input is approximately halved ($\frac{D^{0.7}_{\text{mean}}}{D^{0.98}_{\text{mean}}} = 1.9$), the difference between the width of the *histograms* is almost divided by 3: in the downwards case, ($\frac{D^{0.7}_{\text{width}}}{D^{0.98}_{\text{width}}} = 3.02$). That leads to the separability of the output (the histograms do not overlap) when $r = 0.98$, whilst the responses are not separable (*histograms* do *overlap*) when $r = 0.7$. This performance improvement is quantified by measuring the total probability of misclassification (TPM): $\text{TPM}_{r=0.98} = 0.01$ whilst $\text{TPM}_{r=0.7} = 0.02$

is approximately doubled ($\frac{D^{0.7}_{\text{mean}}}{D^{0.98}_{\text{mean}}} = 1.92$), the width of the histograms is almost multiplied by a factor 3 (In the downwards case, $\frac{D^{0.7}_{\text{width}}}{D^{0.98}_{\text{width}}} = 3.02$). This visual results translates into the quantitative measure of the TPM: $\text{TPM}_{r=0.98} = 0.01$ whilst $\text{TPM}_{r=0.7} = 0.02$. Each histogram represents the distribution of the FRs of the neurons for a hundred experiments repeated with the same stimulus direction. Note that the mean value of the FR for simulation (corresponding to each histogram) has little information relevance for the discrimination task, as it is just an offset, and the first effect of adding inhibitory inputs is just a reduction in this FR. As we intuitively guess, inhibitory inputs will reduce the rate of the incoming signal, thus reduce the output FR (except in some extremal cases, as described in Feng and Wei 2001). What is important for discrimination is to have a large gap between the means relative to the width of each histogram. This is illustrated in Fig. 3. With $r = 1$, the efferent FR depends solely on the synaptic noise. Because of the correlation between synapses, the strength of this synaptic input noise (σ) is more dependent on the coherent stimuli than the deterministic input drive (μ). Furthermore, the stronger the input noise, the larger the efferent FR (This is a consequence of the diffusion process that governs the dynamics

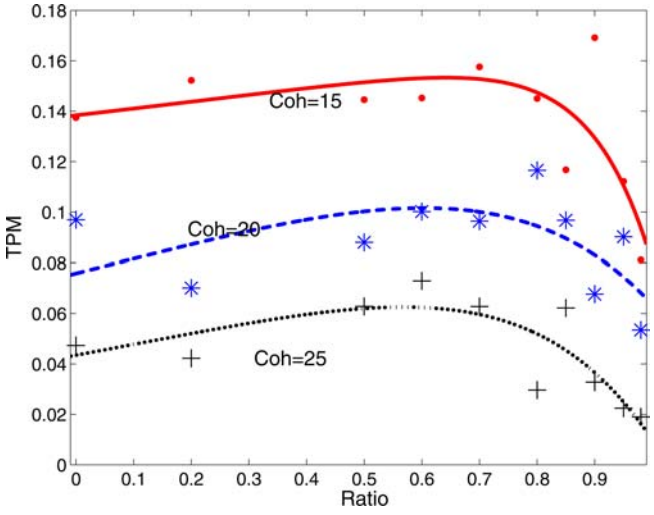

Fig. 4 Three curves that represent the error measured by TPM as a function of r, for three coherence values, 15% (*dot symbols*), 20% (*star symbols*), 25% (*plus symbols*). Each single symbol represents one measure of the TPM for the corresponding parameters (ratio and coherence) (see Appendix for more details on how we calculate it). The curves are exponential fits of the data points, with four parameters. The standard errors of the regression (root mean squared errors of the fits) are respectively RMSE $= 0.02$, RMSE $= 0.019$ and RMSE $= 0.013$. We see that the TPM drops significantly for ratios close to 1, $r >= 0.9$. We also see that it is naturally easier to discriminate when the input is more coherent

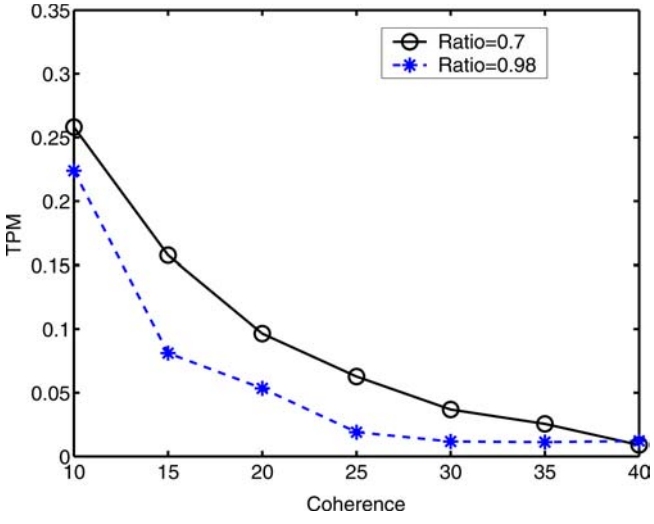

Fig. 5 Illustration of the decrease of TPM with increasing coherence, and the influence of the ratio, r, on this decrease. The curves were obtained by measuring the TPM, as explained in the Appendix, for each value of the ratio r and of the coherence. This TPM is represented by the *symbols*, and the *curves* join the data points with interpolated *straight lines*

of the membrane potential). Consequently, the discrimination power of the neuron is increased by the increase of r. Theoretically, we can understand this when we know that adding inhibitory input to a leaky IF cell "has a subtractive rather than a divisive effect on FRs" (Koch 1999): Plotting the curves of efferent FR as a function of the input rate, the adding of inhibitory input shifts the curves to the right with very little effect on the slopes. This means that the relative distance between the FR generated by two different stimuli is preserved, even if the overall output mean and variance is reduced, which is exactly what is needed in order to improve the discrimination accuracy, and is illustrated in Fig. 3.

These results confirm that the task is obviously easier when more dots move coherently, and confirm the counter-intuitive result that increasing the noise to signal ratio of the postsynaptic input, by increasing r, increases the discrimination accuracy of the neuron. Feng and Liu (2002) and Feng et al. (2003) have analytically proved the simple fact that for $r = 1$ the discrimination is better than for $r = 0$. Here we have a more detailed description of how the probability of error of the single neuron is not monotonically decreasing, but reaches a peak around a ratio equal to 0.7, as we can see in Fig. 4. We also show the importance of the separability of the input to the discrimination performance in Fig. 5.

4.2 The population firing rate is more accurate

We compared the performance of the two models, the single neuron and the population one, over the range of ratios and coherence. We see, in Figs. 6 and 7, that the discrimination of the population code is much more accurate than that of the single neuron's FR, for the same number of spikes.

The better performance of the population can be explained as follows: the efferent FR of the population is not exactly the same as that of a single neuron because the assumption of ergodicity does not hold. In the population approach, we use the first hundred spikes of a hundred neurons to measure the FR, which means that we use on average one spike per neuron, long Inter-Spike Interval (ISIs) are unlikely to be produced. On average, the population will produce a hundred spikes before one neuron has generated a spike with an ISI longer than twice the mean ISI. Thus, across all the experiments, we will not obtain exceptionally low FRs, which are generated by exceptionally long ISIs in the corresponding efferent spike train. That means that the distribution of FRs, represented as an histogram, such as in Fig. 3 or Fig. 11, has its left part cut. Consequently, the histograms overlap less, which ultimately means that the TPM is lower (see the Appendix for more details on the TPM). Graphically, it means that the output FR histograms are better separated than what we see in Fig. 3, and consequently the discrimination accuracy of the model is increased.

4.3 Time considerations

4.3.1 Time to hundred spikes

In this part, we want to compare the time it takes for the model to generate 100 spikes, which is a number that enables us to reliably estimate the FR. This approach enables us to compare the two approaches with an objective criterion. Figure 8 shows us the main improvement of using a population code:

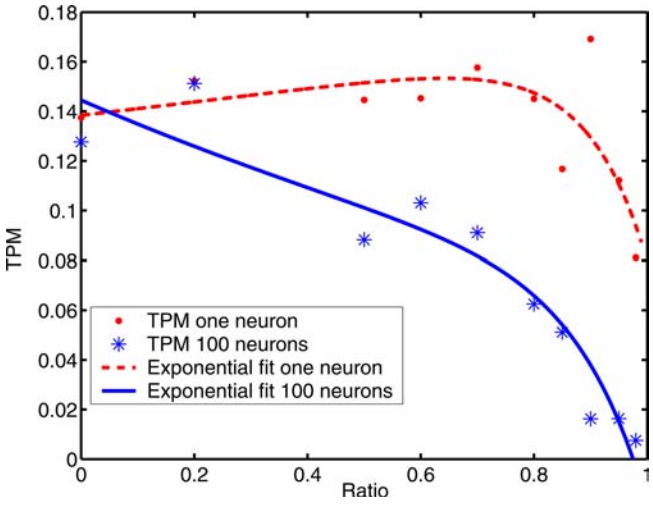

Fig. 6 The coherence of the stimulus was 15%. The continuous line represents the TPM versus r, for a population of a hundred neurons, using a hundred spikes. The *dashed line* represents the TPM with the same setup, but measuring the FR from one neuron. Both of them are exponential fitting with four parameters of the data represented with *dots* (single neuron) and *stars* (population).The standard errors of the regression (root mean squared errors of the fits) are, respectively, RMSE $= 0.018$ and RMSE $= 0.019$. The results obtained from the population of neurons are clearly better, almost one order of magnitude, when r is close to 1

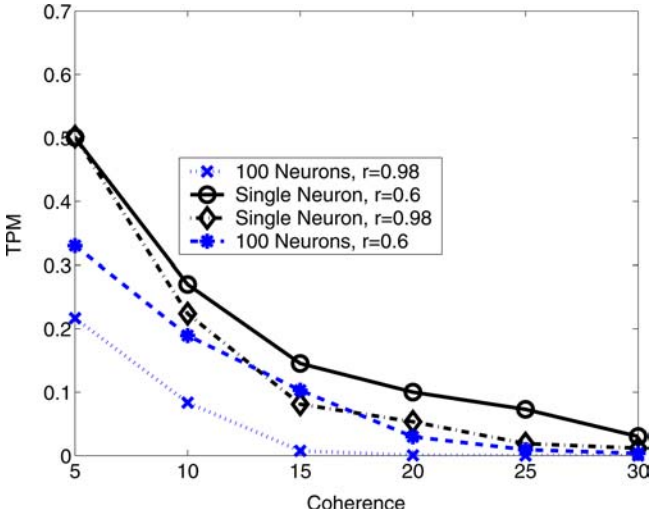

Fig. 7 Comparison of the discrimination performance with a single neuron and a hundred neurons using the FR measured on the first hundred spikes. The results are presented for two different values of r. The curves were obtained by measuring the TPM, as explained in the Appendix, for each value of the ratio r and of the coherence. This TPM is represented by the symbols, and the curves join the data points with interpolated straight lines. It is striking that the single neuron is much less accurate

it is much quicker. In this figure, we see a dramatic increase of the time when r tends to one. This is related to the fact that, with perfectly balanced inputs, there is no deterministic synaptic input. The only post-synaptic input is created by synaptic noise, which means that reaching the threshold potential to generate a spike is exceptional.

The time to hundred spikes decreases with the coherence of the stimuli: the more dots move coherently, the larger the FR. This is due to the correlated synapses: if more dots move coherently, the positive correlation of the coherent synapses will lead to a generally larger input noise.

Obviously, the population code is a much faster code: in the case of an almost balanced ratio r, the single neuron needs around ten seconds to generate a hundred spikes, for coherence values typically between 10 and 20%. Furthermore, the time to hundred spikes of the population code, with a ratio r close to 1, only goes up to the order of one second with a hundred neurons, and 100 ms with a thousand neurons, without any loss of accuracy. With the increase of the population size, the time to generate a hundred spikes significantly decreases. Using the first hundred spikes of a thousand neurons, we would use 0.1 spike per neuron, on average.

4.3.2 TPM as a function of the size of the time window: A trade-off between accuracy and time

We know that, for most of the biological systems the absolute performance must take into account not only the accuracy at realizing the task, but also the time spent to achieve it (see, e.g. Ratcliff et al. 1999 or Usher and McClelland 2001). Hence, processing time is an important criterion for comparing the various models. Thus, to put the TPM in perspective, we have to measure the evolution of the quantity of errors with the size of the time window during which we collect spikes.

We can see, in Fig. 9, the performance of the model for various values of r, as a function of the time, within a one second time window. It shows clearly that we reach much better results with $r \approx 1$ than for $r = 0$, but that we reach acceptable performances much faster with $r \approx 0$. Bearing in mind that significantly increasing the size of the population would allow us to reach such results in a significantly shorter time, these high performances seem relevant to the natural time constraints.

4.3.3 Time to a good discrimination

Because of the Gaussian approximation of the distribution of the efferent FR over 100 repeats, the TPM will never be exactly zero. However, we can consider that a good discrimination performance is attained when TPM $= 0.1$, and that an acceptable discrimination performance is reached for TPM $=$ 0.2. As seen in the previous subsection, the TPM of the model depends on the size of the time windows during which we collect the spikes to accurately evaluate the FR of population. This is a typical case of speed-accuracy trade-off.

We have measured $TPM_r(t)$, as in Fig. 9, for the whole range of ratios, and evaluated t_r such that $TPM_r(t_r) = 0.1$ or 0.2, respectively. We have performed this analysis for a coherence equal to 15%. Then we have fitted these data points with an exponential model with four parameters. The data clearly has an exponential behavior, and we had to use four

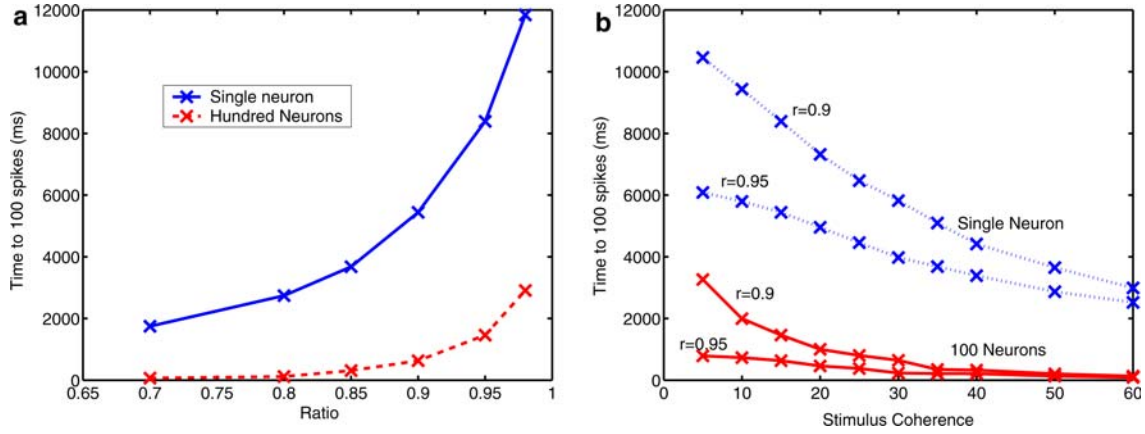

Fig. 8 **a** Time to get a hundred spikes versus r, with a population of a hundred neurons and with a single neuron. Coherence=15. We see that this time increases dramatically when r tends to one. As predicted, the population of a hundred neurons generates hundred spikes more quickly, between five and ten times quicker. **b** Time to get a hundred spikes versus the coherence of the stimulus, with $r = 0.9$ and $r = 0.95$. The *dotted lines* represent the results for a population of a single neuron, the *solid lines*, one hundred

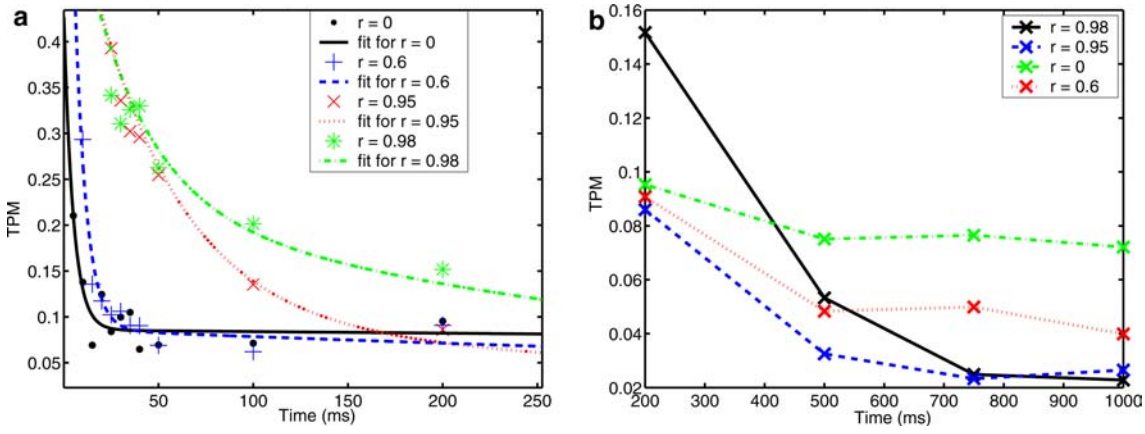

Fig. 9 Comparison of the evolution of the TPM with the size of the time window, for $r = 0.98$, $r = 0.95$, $r = 0.6$ and $r = 0$. **a** Zoom on the behavior for shorter time windows ($t < 250\,ms$). These curves show that the TPM decreases much faster with the time when r is low. The curves are exponential fits to the data with 4 parameters. The root mean squared errors of the fits are, respectively, RMSE $= 0.052$, RMSE $= 0.044$, RMSE $= 0.018$ and RMSE $= 0.022$. **b** Comparison of the evolution of the TPM for long-time windows, reaching to one second, for $r = 0.98$, $r = 0.95$, $r = 0.6$ and $r = 0$. The data points are obtained by measuring the TPM, as explained in the Appendix. We see that for longer time windows, the "slower" strategy that consists in using $r \approx 1$ outperforms the fast, small ratio approaches

parameters to obtain satisfactory standard errors between the exponential fit and the data. Thus we obtained numerical evaluations of the time to a good TPM (TPM $= 0.1$) as a function of r ($T_1(r)$), and of the time to an acceptable TPM (TPM $= 0.2$) as a function of r ($T_2(r)$):

$$T_1(r) = a \times e^{(b.r)} + c \times e^{(d.r)}$$

$$T_2(r) = f \times e^{(g.r)} + h \times e^{(k.r)}$$

where $a = 27.94$, $b = 0.6279$, $c = 1.903e-008$, $d = 23.8$, $f = 3.349e+006$, $g = 6.936$, $h = -3.349e+006$, $k = 6.936$.

We can see an illustration of $T_1(r)$ and $T_2(r)$ in Fig. 10. Figure 10 also shows the data points that were used to evaluate the fitting.

5 Discussion

5.1 Summary of the results

As a first result, we have shown that the discrimination performance increases with the ratio r between inhibitory and excitatory inputs, given that we use the same number of spikes to measure the efferent FR. This is counter-intuitive because, on the other hand, r increases the post-synaptic input noise.

We obtained the surprising result that the population code, without lateral connections is more accurate than the single neuron when we use the same the number of spikes to evaluate the FR. Leaving aside any time considerations, measuring the FR on the first hundred spikes of the population enables much more accurate discriminations than measuring the FR

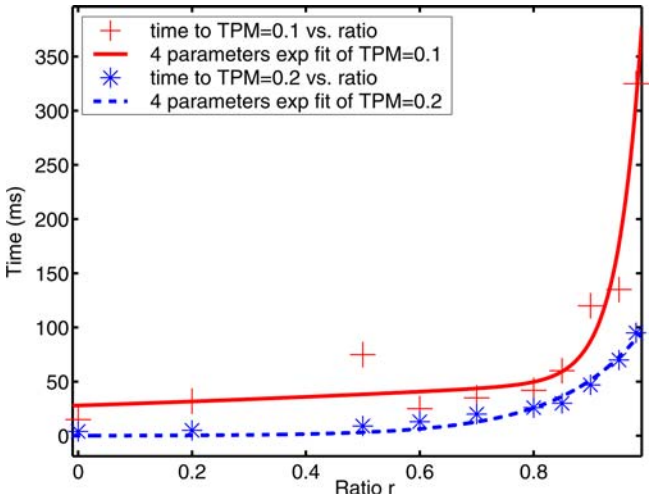

Fig. 10 Illustration of the numerical estimation of the time to reach an acceptable discrimination performance. The coherence of the input is 15%. The data was obtained using the evaluations of the function representing TPM versus time. For each value of r we found t such that $\mathrm{TPM}(t) = 0.1$ or $\mathrm{TPM}(t) = 0.2$. The curves are exponential fits to the data with four parameters. The root mean squared errors of the fits are $\mathrm{RMSE} = 28.74$ (TPM $= 0.1$) and $\mathrm{RMSE} = 6.602$ (TPM $= 0.2$)

on the first hundred spikes of one neuron. This is due to the fact that we rule out the long ISIs.

Although increasing the ratio r increases the performance, it also increases the processing time significantly. The study of the discrimination performance as a function of the time window during which we collect spikes shows that the probability of misclassification decreases much faster for the smaller ratios. We numerically evaluated an expression of the time needed to reach an acceptable TPM as a function of the ratio.

Furthermore, we have seen that only ratios close to one can reach a level of performance unreachable by the FR of a population with exclusively excitatory synapses. These very high performances are reached at the cost of a very long discrimination process.

5.2 Comparison with other work

Experimental evidence for our model predictions Romo et al. (2003) show that positive correlation is not always harmful to discrimination accuracy, as it can successfully remove the noise from the input. In their article, they show that this happen when pools of sensory neurons have opposite slopes characterizing their activity versus intensity of stimulus functions.

Their research experimentally confirms what our model predicts with numerical simulation. In our model, increasing the intensity of the inhibition (i.e. the ratio r) increases the discrimination accuracy when the correlation is positive (here, $c = 0.1$), but is useless if $c = 0$ (unpublished simulation results). The activity (FR) of the inhibitory input is modelled to be the same as the activity (FR) of the excitatory input, but with the opposite effect on the synapse: thus the post-synaptic effect is opposite. When the postsynaptic input contributions are plotted against the coherence, the inhibitory and excitatory contributions have opposite slopes, and are positively correlated. In both their research and ours, this leads to an improved discrimination accuracy. These situations are particular cases of noise substraction by negative correlation. This phenomenon has been neglected but is now being studied by Durrant et al. (2005).

Building block of larger decision models Our model is consistent with larger models that simulate decision processes. For example, in the work of Wang (2002), decision is modelled by a recurrent network with an inhibitory feedback loop. This neural network, after a few time steps, diverges to the decision of the system. In the work of Rao (2004) the decision is made by a Bayesian network. Both of these models are consistent with the idea of the accumulation of evidence, as a decision variable that triggers the expression of a decision when a threshold is reached. This more general idea has been studied by Gold and Shadlen (2003) and Shadlen and Gold (2004). More precisely, Gold and Shadlen (2003) propose a model for how a decision variable, such as the ones we can find in area LIP, accumulate motion information that is represented in area MT and MST. Our model gives detailed insights into how the very general motion information represented in MT or MST can be coded by IF neurons into a signal that enables discrimination relevant to the task. Experimental results (Ditterich et al. 2003) as well as psychological work (Ratcliff and Rouder 1998) support this idea. This decision process can be made faster or slower, according to the parameters of the network. In particular, Ratcliff proposes that the speed of the accumulation variable, the "drift rate" varies with the difficulty of stimuli, and varies as well with various trials of the same difficulty. This drift rate variation is used to explain the variations of the reaction time (RT) and the accuracy of the decision. It is straightforward to imagine how the drift rate depends on the information given by our model: it will be lower if the TPM is high, because the FR that would generate errors would have a negative effect on the accumulation process.

Speed-accuracy trade-off Our model enables better discrimination at the cost of longer processing times. This trade-off between accuracy and speed of processing is a typical pattern of decision making. The time needed to reach an efferent spike train that reflects a reasonable decision goes from a few dozen of milliseconds to hundred of milliseconds. This large discrepancy is consistent with experimental results studied by Shadlen and Gold (2004), where they show that the RT of a monkey, in this precise discrimination task, can go up to 800 ms, according to the required reliability and to the difficulty of the task. Of course we only model one neuronal layer, hence we cannot precisely compare our results to the RT, but the magnitude is similar. Furthermore, the RT depends on the speed at which the evidence is expressed.

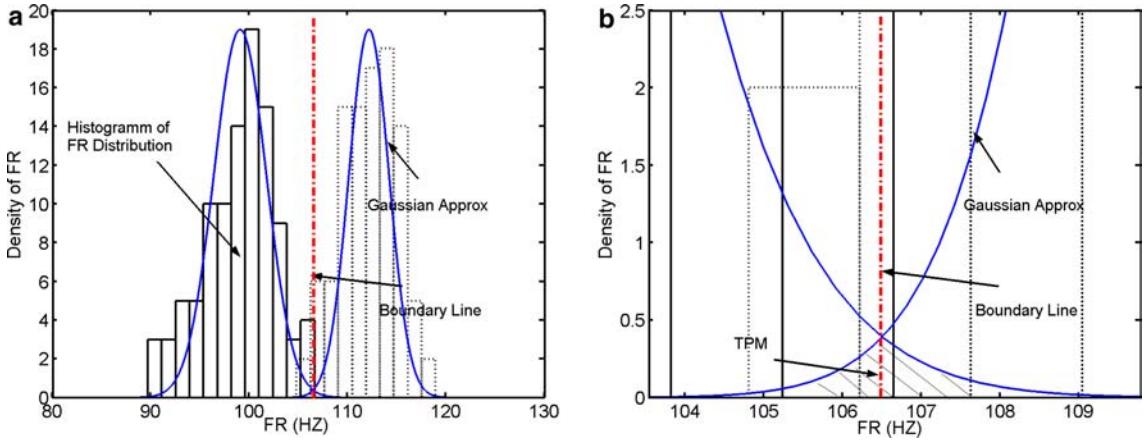

Fig. 11 Graphical illustration of how the TPM is calculated from the boundary line between two Gaussian distributions of FR. (**b**) is a detailed zoom of (**a**). The *histograms* represent the distribution of FR measured for the hundred repeats of the same experiments, upwards and downwards. Experiments made with a coherence of 22%, a single neuron, FR measured on hundred spikes, r=0. We see that the Gaussians fit the distribution of FR reasonably well

5.3 Future work

The methods used here to interpret the efferent spike train by comparing the FR to a decision boundary line is not a likely brain process and has yet to be modelled, to have a plausible model of overall decision making.

Our model is "static" in so far as it does not depend on the past, and it models only one stage of the perceptual decision. Even though the visual process does not necessarily occur in continuous time (explained in VanRullen and Koch 2003), we need to devise a model that keeps past information to accumulate. This accumulation of evidence can be implemented by a larger dynamical model of competing units, whose divergence indicates the decision. The instantaneous evidence provided by our MT model can be what is accumulated by such a models. Such larger dynamical models have already been presented by Lee et al. (1999) and by Wang (2002), and these fit the biological data well as seen in Heekeren et al. (2004), Sugrue et al. (2005), and Rorie and Newsome (2005). We too are studying the dynamics of decisions made by competing pools of IF neurons, but with a special emphasis on the role of second-order statistics such as the variance of the low-level background activity of the brain. The stimulus related evidence is provided by a slightly adapted version of the model that we present in this article. If we increase the viewing time, the discrimination neurons that we modelled here provide a FR that varies with time and that is used as evidence in the decision making. This is beyond the scope of this article but is the object of current work (Gaillard et al. 2005).

In order to have a realistic input from the primary visual cortex that projects to the MT discriminating neurons of our model, we can incorporate models presented by Heeger et al. (1995) and Simoncelli and Heeger (1998) for velocity encoding. Their model justifies the fact that the synapses that receive coherent movements are correlated: movement detectors are grouped in columns, thus the neurons detecting the upward movement are close to each other, thus they will probably send correlated signals.

5.4 Conclusion

Even if our model is still far from realistically simulating the biological process, it is a good tool to study the effects of various neuronal parameters on psychophysical performance. As pointed out by Feng and his colleagues (Feng et al. 2003), the improvement with the ratio between inhibitory and excitatory inputs is independent of the set of parameter values, thus it is transposable to different neuronal processes. As discussed, the model is consistent with several biologically related models and with biological data. Thus, it is a very good basis for building for a more general decision-making model that can be related to biophysical measures such as error rates or RT. That is the direction of our future research. Furthermore, it could be a good source of inspiration for experimental biologists to make hypotheses on the dynamics and numbers of real neurons involved in the discrimination to decision pathways, which are fundamental to any purposeful activity.

Appendix: measuring the performance: the total probability of misclassification (TPM)

Firing rate distribution During the experiments, we present the same setup to the discrimination model one hundred times: a certain percentage of dots move downward (or upward), and the rest is a random perturbation. Each time, we measure the FR of the model. Then, we do the same for the opposite motion, with the same coherence. For each new display, the random perturbation is recreated. So we obtain two distributions of FR, that are obviously not deterministic because of the variation in the stimuli (randomly created perturbation), and because of the synaptic noise. We know that the output of an IF neuron has distribution statistics of a renewal process, its rate being the FR. The FR distributions can then, according to the renewal theorem (Feller 1971;

Tuckwell 1988), be modelled as a Gaussian distribution:

$$G(x) = \frac{1}{2\pi\sigma^2} \exp^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where σ^2 is the variance of the sample of FRs and μ its mean. The absolute value of the term that multiplies the exponential to determine the size of the area under the Gaussian can be an arbitrary constant, since we calculate the probability of misclassification as a fraction. This is illustrated in the shape of the histogram of the distribution of the output FR, as in Fig. 11.

Discrimination boundary Given the moving dot input, the model gives us a FR from which we are able to tell the direction of the motion. The criterion here will be the comparison of this FR to a value, the "discrimination boundary". If the FR is greater than this value, then the movement is upwards, otherwise downwards. This boundary value is the point where the two Gaussian curves, which model the upwards and downwards motion distributions of FR, overlap. This is illustrated in Fig. 11, and its value can be formally determined.

Total probability of misclassification (TPM) The TPM is the number of misclassifications divided by the total number of classifications. Using the Gaussian model for the output FR, these numbers will be the areas under the curves. If we classify an upward pattern of input as downward, it means that the FR is smaller than the boundary value, even if it has been generated by an upwards motion . This is the area of right-hand Gaussian that falls on the left of the boundary. Consequently, the number of errors will be the area under the right-hand Gaussian on the left of the boundary, plus the area under the left-hand Gaussian on the right of the boundary. The total number of classifications is given by the sum of the two areas under the whole right-hand and left-hand Gaussian curves.

$$\mathrm{TPM} = \frac{\int_{-\infty}^{\beta} G_{\mathrm{up}}(x)\mathrm{d}x + \int_{\beta}^{\infty} G_{\mathrm{down}}(x)\mathrm{d}x}{\int_{-\infty}^{\infty} G_{\mathrm{up}}(x) + G_{\mathrm{down}}(x)\mathrm{d}x}$$

where β is the abscissa of the boundary line, G_{up} and G_{down} are respectively the Gaussian functions that model the upward and downward movements.

References

Britten K (2003) The middle temporal area:Motion processing and the link to perception: In: Chalupa WJ (ed) The visual neurosciences, MIT Press, Boston

Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. Visual Neurosci 13:87–100

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion : a comparison of neuronal and psychophysical performance. J Neurosci 12:2331–2355

Deng Y, Williams P, Liu F, Feng JF (2003) Discriminating between different input signals via single neuron activity. J Phys A: Math Gen 36(50):12379–12398

Ditterich J, Mazurek M, Shadlen MN (2003) Microstimulation of visual cortex affects the speed of perceptual decisions. Nature Neurosci 6(8):891–898

Durrant S, Kendrick K, Feng JF (2005) Algorithms for exploiting negative correlation. Biosystems (accepted)

Feller W (1971) An introduction to probability theory and its applications, vol. 2, 2nd edition. Wiley, New York

Feng JF (2001) Is the integrate-and-fire model good enough? - a review. Neural Netw 14:955–975

Feng JF (2003) Effects of correlated and synchronized stochastic inputs to leaky integrator neuronal model. J Theor Biol 222:151162

Feng JF, Liu F (2002) A modelling study on discrimination tasks. Biosystems 67:67–73

Feng JF, Wei G (2001) Increasing inhibitory inputs increases neuronal firing rate. J Phys A: Math Gen 34:7493–7510

Gaillard B, Buxton H, Feng JF (2005) Neuronal model of decision making. In: Feng J, Qian M, Jost J (eds) Networks: from biology to theory. Springer, Berlin Heidelberg New York (in preparation)

Gerstner W, Kistler W (2002) Spiking neuron models single neurons, populations plasticity. Cambridge University Press, Cambridge

Gold JI, Shadlen MN (2003) The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. J Neurosci 23:632–651

Heeger DJ, Simoncelli EP, Carandini M, Movshon JA (1995) Computational models of cortical visual processing. In: Proc national academy of science, vol. 93, pp 623–627

Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. Nature 431:859–862

Koch C (1999) Biophysics of computation: information processing in single neurons. Oxford University Press, Oxford

Lee DK, Itti L, Koch C, Braun J (1999) Attention activates winner-take-all competition amongst visual filters. Nature Neurosci 2: 375–381

Liu J, Newsome WT (2003) Functional organisation of speed tuned neurons in visual area MT. J Neurophysiol 89:246–256

Parker AJ, Newsome WT (1998) Sense and the single neuron: Probing the physiology of perception. Annu Rev Neurosci 21:227–277

Rao RPN (2004) Bayesian computation in recurrent neural circuits. Neural Comput 16(1):1–38

Ratcliff R, Rouder JN (1998) Modeling response times for two-choice decisions. Psychol Sci 9:347–356

Ratcliff R, Van Zandt T, Mc Koon G (1999) Connectionnist and diffusion models of reaction time. Psychol Rev 106:261–300

Ricciardi LM, Sato S (1990) Diffusion process and first-passage-times problems. In: Ricciardi LM (ed) Lectures in applied mathematics and informatics. Manchester University Press, Manchester

Romo R, Hernandez A, Zainos A, Salinas E (2003) Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. Neuron 38:649–657

Rorie AE, Newsome WT (2005) A general mechanism for decision-making in the human brain? Trends Cogn Sci 9 (advanced online publication)

Shadlen MN, Gold JI (2004) The neurophysiology of decision making as a window on cognition. In: Gazzaniga MN (ed) The cognitive neurosciences, 3rd edn. MIT Press, Cambridge

Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J Neurophysiol 86:1916–1935

Shadlen MN, Newsome WT (1996) Motion perception: seeing and deciding. PNAS 93:628–633

Sheth BR, Sharma J, Rao SC, Sur M (1996) Orientation maps of subjective contours in visual cortex. Sci 274:2110-2115

Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. Vision Res 38:743–761

Sugrue LP, Corrado GS, Newsome WT (2005) Choosing the greater of two goods: neural currencies for valuation and decision making. Nature Rev Neurosci 6(5):363–75

Tovée Martin J (1996) An introduction to the visual system. Cambridge University Press, Cambridge

Tuckwell HC (1988) Introduction to theoretical neurobiology, vol 2. Cambridge University Press, Cambridge
Usher M, McClelland J (2001) On the time course of perceptual choice: the leaky competing accumulator model. Psychol Rev 108:550–592
VanRullen R, Koch C (2003) Is perception discrete or continuous? Trends Cognit Sci 7(5):207–213
VanRullen R, Thorpe S (2000) Is it a bird is it a plane? Ultra-rapid visual categorization of natural and artifactual categories. Perception 30(6):655–668
VanRullen R, Thorpe S (2001) Rate coding vs temporal order coding: what the retinal ganglion cells tell the visual cortex. Neural Comput 13:1255–1283
Wang XJ (2002) Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. Neuron 36:955–968
Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. Nature 370:140–143