

Testing the Odds of Inherent versus Observed Over-dispersion in Neural Spike Counts

Wahiba Taouali¹, Giacomo Benvenuti¹, Pascal Wallisch²,
Frédéric Chavane¹ and Laurent U. Perrinet¹

Institutions

¹ Institut de Neurosciences de la Timone, CNRS, Aix-Marseille Université, Marseille, France

² Center for Neural Science, New York University, New York, USA

Abbreviated title for the running head

Odds of Inherent versus Observed Over-dispersion

Corresponding author contact

Name : Wahiba Taouali

E-mail: wahiba.taouali@univ-amu.fr

Phone number : +33.662 866 453

Physical address: Institut de Neurosciences de la Timone (UMR 7289), CNRS, Aix-Marseille Université

27, Bd Jean Moulin, 13385 Marseille Cedex 05, France

BibTex entry

```
@article{Taouali15,  
  title = {Testing the Odds of Inherent versus Observed  
          Over-dispersion in Neural Spike Counts},  
  author = {Taouali, Wahiba and Benvenuti, Giacomo and Wallisch, Pascal and Chavane, Frédéric},  
  journal = {Journal of Neurophysiology},  
  year = {2015},  
  doi = {10.1152/jn.00194.2015},  
  issn = {1522-1598},  
  volume = {115},  
  number = {1},  
  pages = {434-444},
```

```

url = {http://dx.doi.org/10.1152/jn.00194.2015},
url = {http://invibe.net/LaurentPerrinet/Publications/Taouali15},
pmid = {26445864},
publisher = {American Physiological Society},
}

```

Abstract

The repeated presentation of an identical visual stimulus in the receptive field of a neuron may evoke different spiking patterns at each trial. Probabilistic methods are essential to understand the functional role of this variance within the neural activity. In that case, a Poisson process is the most common model of trial-to-trial variability. For a Poisson process, the variance of the spike count is constrained to be equal to the mean, irrespective of the duration of measurements. Numerous studies have shown that this relationship does not generally hold. Specifically, a majority of electrophysiological recordings show an “*over-dispersion*” effect: Responses that exhibit more inter-trial variability than expected from a Poisson process alone. A model that is particularly well suited to quantify over-dispersion is the Negative-Binomial distribution model. This model is well-studied and widely used but has only recently been applied to neuroscience. In this paper, we address three main issues. First, we describe how the Negative-Binomial distribution provides a model apt to account for overdispersed spike counts. Second, we quantify the significance of this model for any neurophysiological data by proposing a statistical test, which quantifies the odds that over-dispersion could be due to the limited number of repetitions (trials). We apply this test to three neurophysiological tests along the visual pathway. Finally, we compare the performance of this model to the Poisson model on a population decoding task. We show that the decoding accuracy is improved when accounting for over-dispersion, especially under the hypothesis of tuned over-dispersion.

Keywords : Spike counts, over-dispersion, Negative-Binomial distribution, Decoding, Tuning function.

Introduction

One of the most common ways to study the central nervous system is to use a stimulus-response paradigm. A system, which could be a neuron or a population of neurons, is provided with an input (a stimulus) and the resulting state change (response) of the system is quantified in order to better characterize functional associations. A main problem with this approach is that the system’s response to repeated presentations of an identical visual stimulus often exhibits high variability, particularly in the spike count statistics (see Figure ??). As a consequence, statistical methods are essential to characterize the information carried by neural populations.

The Fano factor [1] is commonly used to quantify the variability of spike counts [2–8]. It is defined by $FF(\Delta t) = \frac{\sigma^2(\Delta t)}{\mu(\Delta t)}$ where μ and σ^2 are the average and the variance over multiple trials of the number of spikes in some time window of length Δt , respectively. This variability is most often characterized by the Poisson spiking Model (PM). By the definition of this model, the spike count follows a homogeneous Poisson process with a rate parameter λ (the expected number of spikes that occur per unit of time). If k is the observed number of spikes fired by a neuron in response to a stimulus in a time window Δt , the probability of such an observation is given by a Poisson distribution of parameter $f = \lambda\Delta t$, which gives the expected number of spikes per sample in the considered time window:

$$P(k) = \frac{f^k e^{-f}}{k!} \quad (1)$$

It follows that $\mu = f$ and $\sigma^2 = f$ and thus that the Fano factor under the Poisson model has a value of 1, regardless of λ or Δt . Such a property is in accordance with a wide range of observations of variability in cortical responses to visual stimuli [6, 7, 9–11] showing variance-to-mean ratios of spike counts that equal or just exceed unity. However, for a large majority of reported evoked responses, the variability is higher

than expected from a Poisson model, especially in mammalian visual cortex [5, 12–15], whereas Sub-Poisson variability has been observed only under specific conditions [2].

In fact, the variability of neural responses depends on different parameters. Some studies reported that the Fano factor values can vary depending on the brain region [2, 16, 17]. Indeed, neural responses in higher cortical areas [18] or in other non-cortical sensory areas such as the retina [19] and the thalamus [2] show substantially lower variability. It has been also shown that some neural features such as the refractory period [2, 19] lead to a decrease of variability. Moreover, the nature of the stimulus itself could contribute to the variability of neuronal responses. It has been suggested that increasing the complexity of the stimuli by using natural-like statistics increases the reliability [20–22]. In addition, it has been reported that the variability of evoked responses is not stationary even during a trial time-course and decreases after the presentation onset of most sensory stimuli [23]. Finally, a recent study [24] on behaving animals explored the role of decision and other cognitive processes in shaping variability.

In this paper, we mainly explore over-dispersion, that is, cases where there is greater variance in the spike count than might be expected by the Poisson model. Our hypothesis is that the Poisson process only accounts for the intrinsic noise induced by the spiking mechanism itself. Indeed, under the Poisson process hypothesis, the trial-to-trial variability only results from the stochastic properties of the neuron [25–27]. Thus, the PM can not account for other significant sources of variability including hidden contextual variables such as attention, varying cortical states [10, 28–30], perceptual effects [31] and overt behaviors (such as the precise eye position) [32]. Thus, we consider an alternative compound distribution, called Negative-Binomial, that generalizes the Poisson distribution with a dispersion parameter that directly controls the ratio between the mean and variance. Such a doubly-stochastic model was recently proposed to model the variability of neural dynamics showing more complex variability than what is expected from a Poisson model [33–36]. Our aim is to further explore the potential suitability of this doubly-stochastic model to account for response variability in neurophysiological data.

The paper is organized as follows. First, we describe the Negative-Binomial encoding model (NBM) and two possible implementations of over-dispersion sources. Second, we propose a statistical test for evaluating whether the observed over-dispersion is significantly better described by the Negative-Binomial model than the Poisson model. To assess the generality of this model, we then apply this test to three data sets of spiking responses in different animals, different states and different visual stimuli: in the lateral geniculate nucleus (LGN) neurons of anaesthetized mice presented with drifting gratings, in the primary visual cortex (V1) of awake macaque monkeys viewing oriented moving bars and in the middle temporal cortex (MT) of anaesthetized macaque monkeys presented with drifting gratings. Finally, we compared the efficiency of PM and NBM models on a population decoder using the most dispersed data set (MT).

1 The Negative-Binomial model: encoding of inter-trial over-dispersion

Several models have been developed to account for over-dispersion such as generalized Poisson [37], zero-inflated Poisson [38–40] or quasi-Poisson [41, 42] models. However, only a few such as the Negative-Binomial model (NBM) are well known and widely available. The Negative-Binomial model has been studied and tested in different fields such as epidemiology [43], accident statistics [44], biology [45, 46], etc. However, it has only recently been applied to evoked neural responses [36, 47] where the variability was traditionally supposed to be well modeled by a Poisson model (PM).

A Negative-Binomial model depends on a distributional form equivalent to a compound stochastic process. It has recently been represented as a Polya-Gamma mixture of normals [48], but its standard representation is the Gamma-Poisson mixture. It corresponds to a doubly stochastic process where the Poisson process has a probability distribution \mathcal{P} with a parameter λ which is itself a random variable generated from a Gamma distribution $\mathcal{G}(\phi, \beta)$ with a shape parameter ϕ and a scale parameter β . When we consider a random variable x generated by this compound process (i.e. $x \sim \mathcal{P}(\lambda)$ with $\lambda \sim \mathcal{G}(\phi, \beta)$), the resulting distribution of this variable corresponds to a Negative-Binomial distribution (NB) which is usually defined by its shape parameter

ϕ and its scale parameter $p = \frac{1}{\beta+1}$:

$$x \sim \mathcal{NB}(\phi, p = \frac{1}{\beta+1}) \quad (2)$$

Using the Negative-Binomial model, the probability of getting k spikes from a given cell with parameters ϕ and p in response to a specific stimulus is given by:

$$P(k; \phi, p) = \frac{\Gamma(k + \phi) \cdot p^\phi \cdot (1 - p)^k}{\Gamma(\phi)k!} \quad (3)$$

We deduce the mean (μ), the variance (σ) and the Fano factor (FF) as:

$$\mu = \phi \cdot \beta, \quad \sigma^2 = \mu + \frac{\mu^2}{\phi} \quad \text{and} \quad FF = \frac{\sigma^2}{\mu} = 1 + \frac{\mu}{\phi} \quad (4)$$

This distribution approaches the Poisson probability of parameter μ for large values of ϕ , but allows to reach larger-than-Poisson variances for small ϕ values. We end up with a multivariate distribution that is - in practice - driven by two parameters: the mean spike count (μ) and the parameter (ϕ) that we call the “*inverse-dispersion parameter*”¹.

Under the NBM hypothesis, the distribution of inter-spike intervals (ISIs) is given by a Loamax distribution (equivalent to an exponential distribution with a random parameter generated from a Gamma distribution). The mean and the variance of such a distribution are given by the following expressions: mean(ISIs) = $\frac{\beta}{(\phi-1)}$ defined for $\phi > 1$, variance(ISIs) = $\frac{\beta^2\phi}{(\phi-1)^2(\phi-2)}$ defined for $\phi > 2$. The resulting squared coefficient of variation is given by :

$$CV^2 = \frac{\phi}{(\phi-2)} = 1 + \frac{1}{\frac{\phi}{2} - 1} \quad (5)$$

defined for $\phi > 2$. CV^2 is higher than 1, which implies that a Negative-Binomial process is more irregular than a Poisson process and approaches Poisson regularity for large values of ϕ . However, unlike under the Gamma distribution hypothesis made in [49], the Negative-binomial model does not allow us to systematically relate over-dispersion to the positive deviation from the null-hypothesis $FF = CV^2$. Thus, to better qualify this over-dispersion parameter, we need to focus on how to generate the Gamma input drive that feeds the Poisson generator.

Over-dispersion may result from extrinsic sensory noise: NBM as a mixture of Poisson and LogNormal distributions. Many studies assume a Log-Normal or a Gamma distribution when analyzing spike count distributions as it most often produces similar results [41, 50]. Indeed, a Log-Normal process might be a good approximative model for a Gamma process. The underlying spiking process could result from a simple encoding model: Let us assume that a noisy sensory input is transformed to a response by a linear correlation measure with a cell’s receptive field (for instance, in the early visual system, selective to a given orientation) and then through a non-linearity into a stochastic spike count (Figure ??). If we consider additive extrinsic Gaussian noise as the sensory noise and an exponential function as the non-linearity, the resulting input drive follows a Log-normal distribution that is closely fitted by a Gamma distribution. Using such a model, Figure ?? shows that over-dispersion could simply result from an additive normal sensory noise and the non-linearity of the spiking response.

However, even if log-Normal and Gamma distributions are often used to model the same phenomena, looking at their log functions - in general - shows clear differences in their respective skewness. Furthermore, there is no exact morphing between the two distributions that allows to link their respective parameters analytically. As such, this approximation does not allow a precise quantification of the relationship between the noise source and the over-dispersion parameter.

¹The inverse of the parameter ϕ , usually denoted by α , has been often called in the literature as the “*dispersion parameter*”. It has been occasionally named the “*over-dispersion parameter*” because it does not account for FF values smaller than one.

Over-dispersion may result from redundancy within a neural population: NBM as a mixture of Poisson and Exponential distributions. The Gamma distribution could also result from a mixture of distributions. Indeed, it is well known that the sum of n independent and identically distributed exponential random variables $\mathcal{E}(\frac{1}{\lambda})$ (the parameter of an exponential distribution is by definition the inverse of its mean, here called λ) follows a Gamma distribution $\mathcal{G}(n, \lambda)$, where n and λ are its shape and scale parameters respectively. As such, exponential distributions have previously been used as a simple model of variability in formal neural networks [51–53]. They correspond to a model of the distribution of time between two Poisson events. In addition, such a distribution maximizes the entropy of spike counts (knowing the average count) and thus maximizes the carried information [54].

From a biological point of view, to drive a Gamma input - with an inverse-dispersion parameter ϕ equal to n - to the output neuron considered as a Poisson generator, this neuron should receive as input the sum of n exponential variables having the same mean λ (see Figure ??). Thus, ϕ reflects the amount of convergence to a neuron which is - for instance - consistent with a model of the convergent connectivity from V1 to area V5/MT [55–57]. In addition, it was observed that the simulation of networks based on biologically realistic parameters [58] could lead to complex dynamics, showing in particular an excess of variability [59]. Thus, we could expect an over-dispersion in V5/MT resulting from dimensionality reduction (i.e. projecting inputs from a high dimensional space to a lower dimensional one [60]).

To summarize, we have focused on two main possible sources of over-dispersion in spike counts. The first, is an extrinsic source that accounts for sensory noise, conveyed along the thalamo-cortical pathway. The second is a source of noise that emerges from convergent inputs to a population, including feedforward, lateral and feedback. We have shown that the convergence of inputs between two data spaces may both result in an over-dispersion. In the next section, we will focus on quantitatively testing evidence for this over-dispersion.

2 The Fano Gamma test: How to quantify over-dispersion evidence?

The Fano factor test is misleading Most studies concerned with the characterization of spike count variability using different methods and formal tests start with the Fano factor (FF). However, most of the methodological studies measuring trial-to-trial variability in firing rates [8, 61, 62] do not take the estimation uncertainty introduced by the limited number of trials into account. Each trial may be considered as a realization of the random process and when accumulating evidence from N trials, the average and the variance converge with an expected error of $1/\sqrt{N}$. Instead, the computed mean and variance are often treated as real features of the distribution whereas they actually correspond to estimations given the (necessarily) limited set of measurements. Thus, considering only the FF value may induce errors in conclusions regarding the assumption of probability distribution. In particular, when the observed FF deviates from 1, is this the result of the limited number of trials or due to an over-dispersion? Or when a value is measured close to 1, can it be safely attributed to a Poisson model?

The chi-square test is problematic To evaluate how likely the observed data was generated by a Poisson model, the most common approach is to use a chi-square goodness of fit test (χ^2). This statistical tool can be used to estimate how closely an expected distribution matches an observed distribution of a discrete quantitative variable (having only finite possible values). However, it has been shown that small expected frequencies may result in a loss of power, that is, a tendency to not reject a false null hypothesis (or to reject a true null hypothesis) [63–65]. This is a potential problem when we deal with small sample sizes and its outcome also depends on how data is pooled into categories. Indeed, it often makes sense to pool the small frequencies together. Moreover, as this test is one-tailed, it does not distinguish evidence for over-dispersion from evidence for under-dispersion.

The Fano Gamma test To better quantify the over-dispersion while accounting for sample size and without dealing with pooling problems, we propose a statistical test based on an empirical sampling distribution of

the Fano factor under the Negative-Binomial assumption. Here, we opt for an analytic procedure that we call Fano Gamma test (*FG*). It allows for a simple computation (or tabulation) of probability bounds for specific distributions hypotheses. That is not the case for other existing procedures [49, 66, 67]. This procedure characterizes the probability distribution of the estimated FF. Thus, it allows one to compute FF bounds with respect to p-value precision limits. Starting from previous analyses [68–70] about the distribution of the Fisher dispersion index, we propose here that, for a Negative-Binomial distribution, the FF asymptotically follows a Gamma distribution with a shape parameter $\frac{n-1}{2}$ and a scale parameter $\frac{2(\frac{\mu}{\phi}+1)}{(n-1)}$:

$$FF_{Negative-Binomial} \sim \Gamma\left(\frac{n-1}{2}, \frac{2(\frac{\mu}{\phi}+1)}{(n-1)}\right) \quad (6)$$

where n is the number of spike count observations. We tested the accuracy of this approximation through simulation by comparing it to the empirical estimates of the true distribution of the Fano factor for any (μ, ϕ, n) triplet. We found that the p-value upper bounds resulting from the approximative formula are slightly over-conservative. Thus, it is worthwhile mentioning that using it to reject the NBM hypothesis is based on a strong assumption (i.e. a rejected realization could be a good one because its p-value is between the approximation and the empirical upper bound) but we do not detail this analysis here. For large values of ϕ , the Negative-Binomial distribution approaches a Poisson distribution and the previous formula gives:

$$FF_{Poisson} \sim \Gamma\left(\frac{n-1}{2}, \frac{2}{n-1}\right) \quad (7)$$

Eden & Kramer [70] studied the FF distribution under the Poisson distribution hypothesis, and have shown that the convergence to this asymptotic distribution is fast. Thus, this analysis provides a formal test for both hypotheses (PM and NBM) given the observed variability in the spiking.

We first compared the power of this test to the chi-square test for the general NBM hypothesis and for the specific PM hypothesis on surrogate data of spike counts. Figure ?? shows that the Fano Gamma test did better than a chi-square goodness of fit test at a significance level of 0.025. This latter test has a larger probability to make a type *II* error for both distributions (dashed red lines in Figure ??C and dashed blue lines in Figure ??D).

Over-dispersion evidence from neurophysiological data We applied these tests to evaluate evidence against ²the Poisson model on three data sets.

The first data set results from recording signals from mouse LGN neurons (72 cells) in response to moving gratings. The stimulus was presented several times in 12 different directions. The second data set comes from extra-cellular recordings in area V1 (67 cells) of two awake macaque monkeys in response to an oriented moving bar. The task was repeated several times for 12 different directions (orthogonal to the bar’s orientation). The third data set results from recording in area MT (40 cells) of three anaesthetized macaque monkeys in response to moving gratings. The stimulus was presented several times in 16 different directions. See the methods section for more details on the data collection.

We show the results in Figure ?. For the three data sets, we found that the Fano Gamma test gives lower evidence against the PM and the NBM compared to the FF and χ^2 tests (as expected given our discussion above). The Fano Gamma test shows, as do the other tests, that the evidence of over-dispersion is only significant in MT (with a rate of 64%) among cell/condition pairs, compared to LGN and V1 (with rates of 19% and 26%, respectively). Similarly, it shows that evidence against the NBM is significantly lower in the MT dataset (15%) compared to LGN and V1 data sets (52% and 41%, respectively).

²Indeed, an hypothesis test is a statistical analogy to proof by contradiction. Thus, we used the “evidence against” term that is the result of the statistical test and that means the rejection plausibility. Importantly, the evidence against a claim $\mathcal{H}0$ does not imply the evidence (plausibility) of the opposite claim $\mathcal{H}1$, i.e., this is not the same as a likelihood based test that allow explicitly to quantify the weight of evidence for one hypothesis.

For these particular datasets where the number of trials is limited (mostly less than 20 trials per condition), we have shown that in the LGN and V1 data sets, the high values of Fano factor observed could be simply related to the sampling of the stochastic process itself. Indeed, looking only at the FF values induces a false belief that these datasets exhibit a significant degree of over-dispersion, while the FG test reliably rejects this hypothesis. The χ^2 test shows slightly more evidence against the Poisson model than the FG test which could be explained by categorization issues. It shows also very strong evidence against the NBM model known to be suitable for over dispersed data. Finally, we have shown that the situation is different for the MT data set where we could indeed show evidence of a higher variability than expected from a Poisson model and less evidence against a Negative Binomial model. It is worth mentioning that our statistical test (as validated on surrogate data) allows us to make this conclusion, while other tests fail to do so. Since we mostly observe over-dispersion in the MT data set, we propose to compare the performance of an NBM-based to a PM-based direction decoder applied to this data set.

3 NBM-based direction decoder: what if the over-dispersion is tuned ?

To evaluate the gain in decoding when accounting for the over-dispersion using NBM compared to PM, we extended the classical probabilistic decoding approach proposed by Jazayeri & Movshon [71] and tested the population decoder on the MT data set (presenting the highest over-dispersion probability). Based on a Poisson model of trial-to-trial variability applied to single neurons and a model of tuning for a population coded variable (direction θ). This decoding approach allows us to infer the stimulus (the hidden variable θ) given a particular neuronal response (an observed spike raster Y).

The decoding algorithm consists of maximizing the posterior probability $P(\theta|Y)$ function of the estimated direction θ given a distribution hypothesis. Bayes' rule gives: $P(\theta|Y) = \frac{P(Y|\theta)P(\theta)}{P(Y)}$, where the evidence term $P(Y)$ is a normalization term independent of θ . So, if we assume that there is no prior knowledge on θ ($P(\theta)$ constant), maximizing the posterior under the PM hypothesis is equivalent to maximizing the following log-likelihood function $LP(\theta)$:

$$LP(\theta) = \sum_{i=1}^N k_i \log[f_i(\theta)] - \sum_{i=1}^N f_i(\theta) - \sum_{i=1}^N \log[k_i!], \quad (8)$$

where f_i is a function of the stimulus parameter θ and corresponds to the mean response of each cell i , k_i is the number of trials and N is the number of cells. Similarly, under the NBM hypothesis, we maximize the following log-likelihood function $LNB(\theta)$ in order to estimate θ , where $p_i(\theta) = \frac{\phi_i(\theta)}{\phi_i(\theta) + f_i(\theta)}$ and ϕ_i corresponds to the inverse-dispersion parameter of each cell i as a function of the stimulus parameter θ :

$$LNB(\theta) = \sum_{i=1}^N \log\left[\frac{\Gamma(k_i + \phi_i(\theta))}{\Gamma(\phi_i(\theta))k_i!}\right] + \sum_{i=1}^N \phi_i(\theta) \log[p_i(\theta)] + \sum_{i=1}^N k_i \log[1 - p_i(\theta)], \quad (9)$$

Given that objective, let us first define the tuning functions of the cells. These correspond to the mean spike count (the mean parameter of the Negative-Binomial model) with respect to direction. Thus, we defined for the MT data set a generative tuning function as von Mises functions peaking on the preferred direction. The parameters corresponding to the scaling, maximum and minimum values (κ_o , κ_d , R_0 , $R_0 + R_m$) are free to vary. The main tuning function is given by :

$$f_{MT}(\theta) = R_0 + R_m \cdot \frac{e^{\kappa_d \cdot \cos(\theta - \theta_0)} - e^{-\kappa_d}}{e^{\kappa_d} - e^{-\kappa_d}} \quad (10)$$

Figure ??B shows the estimated tunings of f for 16 different cells.

Then, we should consider the estimation of inverse-dispersion parameter ϕ . We tested two hypotheses. First, having no prior on the tuning of ϕ as a function of the stimuli direction, we supposed that this parameter

is stimulus independent (constant chosen as the average value of its estimates over the different conditions) and varying between cells. Second, we tested the hypothesis of a stimulus dependent inverse-dispersion parameter. We considered a von Mises tuning of ϕ centered on the preferred direction for each cell. Figure ??D shows the estimated tunings of ϕ for 16 different cells. Please note that our aim is not to have an optimal fit but to show that a simple tuning hypothesis of ϕ can improve the quality of decoding, even if the estimation of ϕ values is difficult (see discussion).

Finally, to perform population decoding, we used a leave-one-out (LOO) cross validation method. In this method, we randomly pick a single trial from the data set for each neuron and each condition. Then, we apply the decoding approach previously described: we infer an estimate of the encoded direction (stimulus direction) that maximizes the log-likelihood functions based on the Poisson model (eq. ??) or the Negative-Binomial model (eq. ??) using the spike count of each cell given by the picked trial (as k_i) and the estimated values of mean and inverse-dispersion using the remaining trials (as ϕ_i and f_i). Then, we compare the estimated direction angle to the stimulus direction for the PM model and the NBM model under tuned and un-tuned dispersion hypotheses.

The resulting decoding error values for both models applied to MT data show that the precision of decoding of the PM and the NBM under the non-tuned dispersion hypothesis are very similar (Figure ??A). However, we found that there is a significant gain in decoding when using the NBM under the tuned dispersion hypothesis (Figure ??C). To test if this is related to some data or estimation bias, we performed the same test on surrogate data. We generated Negative-Binomial data using the estimated parameter and applied the two decoding schemes. Similarly, the results of decoding on surrogate data show that the precision of decoding is not improved when using non-tuned dispersion (Figure ??B). However, when considering additive knowledge about over-dispersion tuning, there is a clear gain in accuracy and precision (mean and variance Figure ??D) which seems to follow decoding performance in real data.

Moreover, comparing the goodness of fit of these models to MT data favors strongly the tuned NBM. Indeed, we used the Bayesian information criterion (BIC) that penalizes models with additional parameters [72] to determine which model best explained our data. With the BIC measure, 37 (respectively 28) over the 40 cells were best fit with the tuned NBM (respectively non-tuned NBM) and the remaining were best fit with the Poisson model.

In summary, we have shown two main results in this section. First, the Poisson model performs as well as the Negative-Binomial model in the decoding task most of the time but not for good reasons as the data set is not well described by a Poisson variability model. Second, the Negative-Binomial model with tuned over-dispersion performs better and could be a generic model of neural computation.

4 Discussion

In this paper, we explored over-dispersion in neural spike counts using a doubly stochastic model, namely the Negative-Binomial model (NBM), that could disentangle different sources of over-dispersion. We have first explained that some commonly used procedures to quantify over-dispersion (the Fano factor and chi-square goodness of fit tests) could be problematic. Then, we have shown that the alternative analytical procedure that we propose (the Fano Gamma test) better characterizes over-dispersion in three different data sets (different animals, tasks, conditions and areas). Using the Fano Gamma test we show that there is weak evidence of over-dispersion in LGN and V1 data sets and low evidence against the NBM in MT data set, especially for high mean spiking values, a trend expected from [36].

In the last section, we compared the performances of the traditional Poisson model to the more recent Negative-Binomial model on a population decoding algorithm (direction decoding from spike recording in MT cells of an anesthetized macaque monkey). Under the Poisson hypothesis, a linear decoding model is optimal, as proposed by Jazayeri & Movshon [71]. We show here that this decoder is still efficient (mean decoding error < 10 degrees) even though there is strong evidence against the Poisson hypothesis. This is partly due to the tuning profile of the likelihood functions: the direction corresponding to the maximum

likelihood will converge toward the stimulus' direction even if the noise model is not optimal.

Intuitively, one could expect that using the compound model that accounts for the observed over-dispersion, one should obtain better decoding results. We first considered the hypothesis of non-tuned cells' dispersion (i.e. each cell dispersion parameter is constant). The Poisson model and The Negative-Binomial model with non-tuned dispersion show similar direction decoding results, not only on real data, but also on surrogate data. However, note that, similarly to what was reported before [36], we found that the likelihood values of the estimated directions are always higher for the Negative-Binomial based decoding compared to the Poisson based decoding (not shown), but with similar bandwidths. This leads to similar decoding accuracies, despite the fact that the NBM fits better to the data. Indeed, the spike counts corresponding to the response of the cell to its preferred direction are more likely to be generated by an NBM with higher inverse-dispersion values, i.e. Poisson-like. This could explain why PM performance could be comparable to NBM performance in decoding.

In addition, the estimation of the inverse-dispersion parameter is known to be difficult. Indeed, the estimation of this parameter has been widely studied in different fields [41, 73–75] and several estimation methods were provided such as the maximum likelihood method [76], or the method of moment and maximum extended quasi-likelihood method [77]. A wide range of studies have shown that the estimation of this parameter is very challenging and can be significantly biased, especially when dealing with small data sets. It is worthwhile to point out this problem here: Since neurophysiological experiments are costly and time consuming, we often deal with small data sets. Thus, having more knowledge and control about the dispersion tuning is crucial.

Therefore, we tested a simple possible form of dispersion tuning, that is, the tuning of the inverse-dispersion having similar preferred direction as the tuning of the spike count (using a von Mises function). We found that under the hypothesis of a stimulus-dependent inverse-dispersion, the NBM performs qualitatively better.

Moreover, such a tuning function could explain complex Fano factor tuning profiles that were recently reported [34]. Our results lead us to propose that the reported trial-by-trial variability among MT neurons - showing a directional tuning that is not trivially explained by firing rate variations alone - may simply be explained by a bell-shaped tuning of the over-dispersion (see Figure ??). However, contrary to what they expect, we did not observe a decrease of over-dispersion with respect to the spontaneous level. Indeed, applying the Fano-Gamma test to MT spontaneous spiking data (results not shown) yields lower evidence against a PM (47.5%) compared to that for evoked spiking (64.1%, Figure 5.D) but similar evidence against NBM (15%) compared to evoked spiking (16.3%, Figure 5.D). This is more in line with an increase of under-dispersion, but which still in accordance with the known fact that FF is decreased when a stimulus is applied.

Where could such a tuning of the over-dispersion come from? One first hypothesis could be that over-dispersion results from the sensory noise. Then, as suggested by Goris, Movshon & Simoncelli [36], over-dispersion could also reflect an excitability state of the cell. However, both of these hypotheses cannot account for stimulus dependent variability. Alternatively, we suppose that this property is better apprehended at the population scale. Going back to the encoding paradigm described in the first section (Figure ??), an alternative hypothesis would be that the over-dispersion arise from a stimulus-dependent degree of network convergence (within MT or from V1). Such a scheme could result from a center-surround connectivity profile in the direction domain: more excitatory input for a stimulus close to the preferred direction of the recipient cell.

This study calls for an extension over a wider range of experimental observations. First, we would like to extend this work to more datasets with more trials to explore if our predictions are still valid, both for teasing out sources of variability but also to evaluate population decoding. It would be mostly interesting to compare this extra variability along the sensory pathways to associative areas to confirm the tendency that we observed with a gradual increase of over-dispersion from the lower to associative areas. A second important perspective is to consider the role of experimental conditions. Indeed, it is likely that results would greatly vary depending on the arousal state of the animal, the nature of the stimulation or the behavioral task. In particular, behaviorally more relevant stimulations (such as the model-based synthesis of textures [78] may

lead to a more precise response [20].

5 Materials and Methods

In our analyses, we used three different data sets from neurophysiological recordings.

LGN data set This data set [79] contains the spiking responses of 72 LGN neurons in the anesthetized mouse to drifting grating (square wave (n=24) and sinusoidal (n=48)). Many of these neurons are already contained in the data set published in [80]. Contrarily to the two other data sets, it is an openly available data set.

V1 data set Experiments were conducted at the INT on two adult male rhesus macaque monkeys (*Macaca mulatta*). Experimental protocols have been approved by the Marseille Ethical Committee in Neuroscience (approval #A10/01/13, official national registration #71–French Ministry of Research). All procedures complied with the French and European regulations for animal research, as well as the guidelines from the Society for Neuroscience.

Surgical preparation : Monkeys were chronically implanted with a head-holder and a recording chamber located above the V1 and V2 cortical areas. In a second surgery, a search coil was inserted below the ocular sclera to record eye movements [81].

Behavioral task and training : Monkeys were trained to fixate - within a window of 1 to 2 degrees of diameter - on a red target presented at the center of the screen during the entire duration of the trial (1.5 s). The eye position was monitored by the scleral search coil technique [81, 82] and the animal's behavior controlled using the REX package [83].

Unit Recording : The two monkeys were chronically implanted with a head-holder and a recording chamber located above cortical area V1. A computer controlled microdrive (MT-EPS, Alpha Omega, Israel) was used to trans-durally insert a micro-electrode (FHC, 0.5-1.2 M Ω at 1 KHz) in the right hemisphere. Spikes were sorted online using a template matching algorithm (MSD, Alpha Omega).

Visual Stimulation : Visual stimulation protocols have been produced using in-house software (developed by Gérard Sadoc, Acquis1-Elphy, Biologic CNRS-UNIC/ANVAR). Stimuli were back-projected on a translucent screen covering 80°×80° of the monkey's visual field, at a distance of 1 m, using a retro-projector (resolution: 1280×1024 pixels at 60 Hz). The mean luminance of the motion stimulus was 22.2 cd/m² and the background was kept constant to about 2.24 cd/m². The display was gamma calibrated by means of a lookup table. Once a cell is isolated, a sparse noise (SN) stimulus was first used to quantitatively map out the RF (a grid of 10×10 squares with a side length of 0.6 degrees, 10.9 cd/m² darker or brighter than the background of 11.1 cd/m², ~15 trials). Once the RF was properly located and estimated, a direction tuning paradigm was launched. In this paradigm, a moving bar (0.5×4°) translated across the receptive field (the displacement covers 3° on either side of the RFc) in 12 possible directions (spaced by 30 degrees of polar angle, at a speed of 6.6°/s).

MT data set Experiments were conducted at NYU-CNS on 3 anesthetized, paralyzed, adult macaque monkeys (*Macaca nemestrina*) of either sex. All procedures were conducted in compliance with the National Institute of Health Guide for the Care and Use of Laboratory Animals, and with the approval of the New York University Animal Welfare Committee.

Surgical preparation : The standard procedures for the surgical preparation of animals and single-unit recordings have been reported in detail previously [84]. Briefly, experiments typically lasted 5 to 6 days, during which we maintained anesthesia with infusion of sufentanil citrate (6-30 $\mu\text{g kg}^{-1} \text{h}^{-1}$) and paralysis with infusion of vecuronium bromide (Norcuron; 0.1 mg kg⁻¹ h⁻¹) in isotonic dextrose-normosol solution. We monitored vital signs (heart rate, lung pressure, end-tidal pCO₂, EEG, body temperature, urine flow, and

osmolarity) and maintained them within appropriate physiological ranges. Pupils were dilated with topical atropine. The eyes were protected with gas-permeable contact lenses, and refracted with supplementary lenses chosen through direct ophthalmoscopy. At the conclusion of data collection, the animal was sacrificed with an overdose of sodium pentobarbital.

Unit Recording : Extracellular recordings were made with quartz-platinum-tungsten microelectrodes (Thomas Recording), advanced mechanically into the brain through a craniotomy and small durotomy. To record from MT we passed microelectrodes through a small durotomy centered roughly 16 mm lateral to the midline and 3 mm posterior to the lip of the lunate sulcus at an angle of 20 degrees from horizontal in a ventro-anterior direction. Area MT was identified by the brisk direction-selective responses of isolated neurons. We made recordings from every single unit with a spike waveform that rose sufficiently above noise to be isolated. Stimuli were presented in random order.

Visual Stimulation : We presented visual stimuli on a gamma-corrected CRT monitor (Eizo T966; mean luminance: 33 cd/m²) at a resolution of 1.280×960 with a refresh rate of 120 Hz. Stimuli were presented using Expo software (<http://corevision.cns.nyu.edu>) on an Apple Macintosh computer. For each isolated unit, we first determined its ocular dominance and occluded the non-preferred eye. We presented circularly windowed sinusoidal grating stimuli (16 directions around the clock, 22.5 degrees apart) to map each cell's receptive field, determined its preferred size and speed, and then measured selectivity for orientation or spatial frequency.

The time window considered to calculate the spike counts was the duration of stimulation for the grating stimuli (LGN and MT data sets) and 400 ms covering approximately the time when the bar stimulus crossed the cell's receptive field for V1 data set. All data were analyzed using Python and Matlab (The MathWorks Inc., Natick, MA, USA).

Acknowledgments

WT, FC and LP were supported by EC FP7-269921, "BrainScaleS", GB by FACETS ITN project (EU funding, grant number 237955), a 'Marie-Curie Initial Training Network' and PW was supported by F32-EY019833 from the NIH. We wish to thank Philippe Foundation for a travel grant support to F.C. The authors are indebted to T. Movshon and his laboratory for giving the opportunity to F.C. to perform MT experiments at CNS/NYU.

References

1. Fano, U. Ionization yield of radiations. II. The fluctuations of the number of ions. *Physical Review* **72**, 26 (1947).
2. Kara, P., Reinagel, P. & Reid, R. C. Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron* **27**, 635–646 (2000).
3. Alitto, H. J., Moore, B. D., Rathbun, D. L. & Usrey, M. M. A comparison of visual responses in the lateral geniculate nucleus of alert and anaesthetized macaque monkeys. *The Journal of physiology* **589**, 87–99 (2011).
4. Zhuang, J. *et al.* Layer 4 in primary visual cortex of the awake rabbit: contrasting properties of simple cells and putative feedforward inhibitory interneurons. *The Journal of Neuroscience* **33**, 11372–11389 (2013).
5. Tolhurst, D. J., Movshon, J. A. & Dean, A. F. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research* **23**, 775–785 (1983).
6. Softky, W. R. & Koch, C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *The Journal of Neuroscience* **13**, 334–350 (1993).

7. Shadlen, M. N. & Newsome, W. T. The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. *The Journal of Neuroscience* **18**, 3870–3896 (1998).
8. Koch, C. & Poggio, T. Predicting the visual world: silence is golden. *Nature Neuroscience* **2**, 9–10 (1999).
9. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience* **10**, 1157–1169 (1993).
10. Buracas, G. T., Zador, A. M., DeWeese, M. R. & Albright, T. D. Efficient Discrimination of Temporal Patterns by Motion-Sensitive Neurons in Primate Visual Cortex. *Neuron* **20**, 959–969 (1998).
11. McAdams, C. J. & Maunsell, J. H. R. Effects of Attention on the Reliability of Individual Neurons in Monkey Visual Cortex. *Neuron* **23**, 765–773 (1999).
12. Heggelund, P. & Albus, K. Response variability and orientation discrimination of single cells in striate cortex of cat. *Experimental Brain Research* **32**, 197–211 (1978).
13. Dean, A. F. The variability of discharge of simple cells in the cat striate cortex. *Experimental Brain Research* **44**, 437–440 (1981).
14. Schiller, P. H., Finlay, B. L. & Volman, S. F. Short-term response variability of monkey striate neurons. *Brain research* **105**, 347–349 (1976).
15. Vogels, R., Spileers, W. & Orban, G. A. The response variability of striate cortical neurons in the behaving monkey. *Experimental Brain Research* **77**, 432–436 (1989).
16. Gur, M. & Snodderly, D. M. High Response Reliability of Neurons in Primary Visual Cortex (V1) of Alert, Trained Monkeys. *Cerebral Cortex* **16**, 888–895 (2006).
17. Kayser, C., Logothetis, N. K. & Panzeri, S. Visual enhancement of the information representation in auditory cortex. *Current Biology* **20**, 19–24 (2010).
18. Maimon, G. & Assad, J. A. Beyond Poisson: Increased Spike-Time Regularity across Primate Parietal Cortex. *Neuron* **62**, 426–440 (2009).
19. Berry, M. J., Warland, D. K. & Meister, M. The structure and precision of retinal spiketrains. *PNAS* **94**, 5411–5416 (1997).
20. Baudot, P. *et al.* Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. *Frontiers in Neural Circuits* **7** (2013).
21. Haider, B. *et al.* Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. *Neuron* **65**, 107–121 (2010).
22. Borst, A. & Theunissen, F. E. Information theory and neural coding. *Nature Neuroscience* **2**, 947–957 (1999).
23. Churchland, M. M. *et al.* Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience* **13**, 369–378 (2010).
24. Nienborg, H., Cohen, M. R. & Cumming, B. G. Decision-related activity in sensory neurons: correlations among neurons and with behavior. *Annual review of neuroscience* **35**, 463–483 (2012).
25. Carandini, M. Amplification of Trial-to-Trial Response Variability by Neurons in Visual Cortex. *PLoS Biol* **2**, e264+ (2004).
26. Mainen, Z. F. & Sejnowski, T. J. Reliability of spike timing in neocortical neurons. *Science* **268**, 1503–1506 (1995).
27. Schneidman, E., Freedman, B. & Segev, I. Ion Channel Stochasticity May Be Critical in Determining the Reliability and Precision of Spike Timing. *Neural computation* (1998).

28. Arieli, A., Sterkin, A., Grinvald, A. & Aertsen, A. Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science* **273**, 1868–1871 (1996).
29. Tsodyks, M., Kenet, T., Grinvald, A. & Arieli, A. Linking Spontaneous Activity of Single Cortical Neurons and the Underlying Functional Architecture. *Science* **286**, 1943–1946 (1999).
30. Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A. & Arieli, A. Spontaneously emerging cortical representations of visual attributes. *Nature* **425**, 954–956 (2003).
31. Ress, D. & Heeger, D. J. Neuronal correlates of perception in early visual cortex. *Nature Neuroscience* **6**, 414–420 (2003).
32. Gur, M., Beylin, A. & Snodderly, D. M. Response variability of neurons in primary visual cortex (V1) of alert monkeys. *The Journal of Neuroscience* **17**, 2914–2920 (1997).
33. Churchland, A. K. *et al.* Variance as a signature of neural computations during decision making. *Neuron* **69**, 818–831 (2011).
34. Ponce-Alvarez, A., Thiele, A., Albright, T. D., Stoner, G. R. & Deco, G. Stimulus-dependent variability and noise correlations in cortical MT neurons. *PNAS* **110**, 13162–13167 (2013).
35. Ponce-Alvarez, A., Kilavik, B. E. & Riehle, A. Comparison of local measures of spike time irregularity and relating variability to firing rate in motor cortical neurons. *Journal of Computational Neuroscience* **29**, 351–365 (2010).
36. Goris, R. L. T., Movshon, J. A. & Simoncelli, E. P. Partitioning neuronal variability. *Nature Neuroscience* **17**, 858–865 (2014).
37. Consul, P. C. & Jain, G. C. A generalization of the Poisson distribution. *Technometrics* **15**, 791–799 (1973).
38. Lambert, D. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* **34**, 1–14 (1992).
39. Miaou, S. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis and Prevention* **26**, 471–482 (1994).
40. Shankar, V., Milton, J. & Mannering, F. Modeling accident frequencies as zero-altered probability processes: an empirical inquiry. *Accident Analysis and Prevention* **29**, 829–837 (1997).
41. McCullagh, P. & Nelder, J. A. *Generalized linear models* (London England Chapman and Hall 1983., 1989).
42. Wedderburn, R. W. M. Quasi likelihood functions, generalized linear models, and the GaussNewton method. *Biometrika* **61**, 439–447 (1974).
43. Byers, A. L., Allore, H., Gill, T. M. & Peduzzi, P. N. Application of negative binomial modeling for discrete outcomes: a case study in aging research. *Journal of clinical epidemiology* **56**, 559–564 (2003).
44. Poch, M. & Mannering, F. Negative binomial analysis of intersection accident frequencies. *Journal of Transportation Engineering* **122**, 105–113 (1996).
45. Anscombe, F. J. Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika*, 358–382 (1950).
46. Bliss, C. I. & Fisher, R. A. Fitting the negative binomial distribution to biological data. *Biometrics* **9**, 176–200 (1953).
47. Scott, J. & Pillow, J. W. Fully Bayesian inference for neural models with negative-binomial spiking. *Advances in neural information processing systems*, 1898–1906 (2012).
48. Polson, N. G., Scott, J. G., *et al.* On the half-Cauchy prior for a global scale parameter. *Bayesian Analysis* **7**, 887–902 (2012).

49. Nawrot, M. P. *et al.* Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods* **169**, 374–390 (2008).
50. Atkinson, A. C. Regression diagnostics, transformations and constructed variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1–36 (1982).
51. Treves, A. & Rolls, E. T. What determines the capacity of autoassociative memories in the brain ? *Network: Computation in Neural Systems* **2**, 371–397 (1991).
52. Levy, W. B. & Baxter, R. A. Energy efficient neural codes. *Neural Computation* **8**, 531–543 (1996).
53. Baddeley, R. An efficient code in V1? *Nature* **381**, 560–561 (1996).
54. Shannon, C. E. A Mathematical Theory of Communication. *The Bell System Technical Journal* **27**, 379–423, 623–656 (1948).
55. Simoncelli, E. P. & Heeger, D. J. A model of neuronal responses in visual area MT. *Vision Research* **38**, 743–761 (1998).
56. Rust, N. C., Mante, V., Simoncelli, E. P. & Movshon, J. A. How MT cells analyze the motion of visual patterns. *eng. Nature Neuroscience* **9**, 1421–1431 (2006).
57. Wang, H. X., Heeger, D. J. & Landy, M. S. Responses to second order texture modulations undergo surround suppression. *Vision Research* **62**, 192–200 (2012).
58. Voges, N. & Perrinet, L. Phase space analysis of networks based on biologically realistic parameters. *Journal of Physiology-Paris* **104**, 51–60 (2010).
59. Voges, N. & Perrinet, L. Complex dynamics in recurrent cortical networks based on spatially realistic connectivities. *Frontiers in Computational Neuroscience* **6** (2012).
60. Haykin, S. . *Neural Networks: A Comprehensive Foundation* (Prentice Hall, 1999).
61. Teich, M. C., Heneghan, C., Lowen, S. B., Ozaki, T. & Kaplan, E. Fractal character of the neural spike train in the visual system of the cat. *JOSA A* **14**, 529–546 (1997).
62. Dayan, P. & Abbott, L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* 1st (The MIT Press, 2001).
63. Cochran, W. G. Some methods for strengthening the common χ^2 tests. *Biometrics* **10**, 417–451 (1954).
64. Yates, D., Moore, D. S. & McCabe, G. P. *Practice Of Statistics Advanced Placement* 1st (W. H. Freeman, 1999).
65. Campbell, I. Chi squared and Fisher Irwin tests of two by two tables with small sample recommendations. *Statistics in Medicine* **26**, 3661–3675 (2007).
66. Amarasingham, A., Chen, T.-L. L., Geman, S., Harrison, M. T. & Sheinberg, D. L. Spike count reliability and the Poisson hypothesis. *The Journal of Neuroscience* **26**, 801–809 (2006).
67. Jackson, J. & Redish, A. D. Network dynamics of hippocampal cell-assemblies resemble multiple spatial maps within single tasks. *Hippocampus* **17**, 1209–1229 (2007).
68. Hoel, P. G. On Indices of Dispersion. *The Annals of Mathematical Statistics* **14**, 155–162 (1943).
69. Kathirgamatamby, N. Note on the Poisson Index of Dispersion. *Biometrika* **40** (1953).
70. Eden, U. T. & Kramer, M. A. Drawing inferences from Fano factor calculations. *Journal of Neuroscience Methods* **190**, 149–152 (2010).
71. Jazayeri, M. & Movshon, J. A. Optimal representation of sensory information by neural populations. *Nature Neuroscience* **9**, 690–696 (2006).
72. Box, G. E., Jenkins, G. M. & Reinsel, G. C. *Time Analysis, Forecasting and Control* 1994.

73. Gourieroux, C., Monfort, A. & Trognon, A. Pseudo maximum likelihood methods: Applications to Poisson models. *Econometrica: Journal of the Econometric Society*, 701–720 (1984).
74. Lawless, J. F. Negative binomial and mixed Poisson regression. *Canadian Journal of Statistics* **15**, 209–225 (1987).
75. Dai, H., Bao, Y. & Bao, M. Maximum likelihood estimate for the dispersion parameter of the negative binomial distribution. *Statistics and Probability Letters* **83**, 21–27 (2013).
76. Piegorsch, W. W. Maximum likelihood estimation for the negative binomial dispersion parameter. *Biometrics*, 863–867 (1990).
77. Clark, S. J. & Perry, J. N. Estimation of the negative binomial parameter κ by maximum quasi-likelihood. *Biometrics*, 309–316 (1989).
78. Sanz-Leon, P., Vanzetta, I., Masson, G. S. & Perrinet, L. U. Motion Clouds: model-based stimulus synthesis of natural-like random textures for the study of motion perception. *Journal of Neurophysiology* **107**, 3217–3226 (2012).
79. Scholl, B., Tan, A. Y. Y., Corey, J. & Priebe, N. J. *Spiking responses of LGN neurons in the mouse to drifting gratings* CRCNS.org. <http://dx.doi.org/10.6080/K05Q4T1N>. 2013.
80. Scholl, B., Tan, A. Y. Y., Corey, J. & Priebe, N. J. Emergence of orientation selectivity in the Mammalian visual pathway. *Journal of Neuroscience* **33**, 10616–10624 (2013).
81. Robinson, D. A. A method of measuring eye movement using a scleral search coil in a magnetic field. *Bio-medical Electronics, IEEE Transactions on* **10**, 137–145 (1963).
82. Collewijn, H., Van der Mark, F. & Jansen, T. C. Precise recording of human eye movements. *Vision Research* **15**, 447–IN5 (1975).
83. Hays, A. V. A UNIX-based multiple process system for real-time data acquisition and control. *WESCON Conf Proc* **2**, 1–10 (1982).
84. Cavanaugh, J. R., Bair, W. & Movshon, J. A. Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology* **88**, 2530–2546 (2002).

Figures

Figure 1: **Trial-to-trial variability of neural responses.** (Left) We show the spiking activity of a representative cell of the area MT of a macaque monkey (see Methods section) in response to oriented drifting gratings. Each subplot corresponds to a raster plot of the responses to 16 equally distributed motion directions. There is a high trial-to-trial variability in the spiking responses. (Right) To quantify this variability, we plot the directional tuning curves in response to the same 16 directions (mean spike count, black lines) along with the range of \pm the standard deviation (red lines) for a population of 16 cells from the same area. The subplot with a star corresponds to the representative cell on the left. The area corresponding to the mean \pm the standard deviation expected under Poisson distribution hypothesis ($\sigma = \sqrt{\mu}$) is plotted in blue. In this experiment, most of the cells show a higher variability than expected from the Poisson model.

Figure 2: Over-dispersion may result from extrinsic sensory noise: NBM as a mixture of Poisson and LogNormal distributions. We show here a simple linear / non-linear spiking model with three processing steps: (A) The correlation between an noisy Gabor image (input) and a population of Gabor filters at different orientations models the linear processing in the receptive field (RF) of a visual neuron. The green, black and red lines correspond respectively to one trial (simulation), the average and the average \pm the standard deviation (within 1000 trials). (B) This input is transformed using a static, exponential non-linearity multiplied by a scaling factor corresponding to the maximum spike count. The input drive resulting from a Gaussian input noise and the non-linearity follows a Log-normal distribution that is closely fitted by a Gamma distribution. (C) This drive is finally transformed into spike counts using a Poisson point process with the mean as the input drive. The resulting tuning curve (mean spike count) is plotted in black along with the range of \pm the standard deviation (red line). The area corresponding to the mean \pm the standard deviation expected given the Poisson model ($\sigma = \sqrt{\mu}$) is plotted in blue. The resulting spike counts variance is higher than what is expected by a Poisson model, as can be observed in neurophysiological data (see Figure ?? - Right).

Figure 3: Over-dispersion may result from redundancy within a neural population. NBM as mixture of Poisson and Exponential distributions. Let us consider, within a neural population, a subset of ϕ cells with a similar selectivity and thus driven by a similar correlation measure $\lambda = \mathcal{C}(\theta)$. This measures the correlation between the input Gabor and the Gabor representing the orientation θ characteristic of their receptive field (RF). Let us further assume that the filter cells exhibit an exponential distribution of spike counts $\mathcal{E}(\frac{1}{\lambda})$ with mean λ , thus maximizing the entropy of the input drive. Then, the pooling of the resulting spike counts results in an input drive $\mathcal{G}(\phi, \lambda)$ equal to the sum of exponential variables having the same mean equal to λ . It can be shown that the Poisson spiking process driven by this input drive is equivalently described by a Negative-Binomial probability distribution function with parameters λ (mean) and ϕ (inverse-dispersion) (see equation ??). Such a distribution describes observed trial-to-trial variability well and in this model, the parameter of dispersion is directly correlated with the redundancy, i.e., the convergence of similar input cells.

Figure 4: Evaluating the power of chi-square test (χ^2) test versus Fano Gamma test (FG) when testing evidence against Poisson (PM) model (in blue) and Negative-Binomial (NBM) model (in red). (Top) Two examples of spike count histograms resulting from artificially generated data (surrogate data) using the NBM (A) and the PM (B). These examples were chosen such that, when using the χ^2 test, the model used to synthesize the data is rejected. For each histogram, we plotted in blue the expected distribution from a PM and in red the expected distribution from a NBM. For these examples, the sample size is 50. (Bottom) χ^2 test (dashed lines) and FG test (solid lines) were applied to Negative-Binomial (C) and Poisson (D) surrogate data for different sample sizes with a significance level of 0.025. The mean is fixed to 10 spikes, the dispersion parameter is equal to 1 spike and the test was repeated 1000 times for each sample size. (E) and (F) show the results in (C) and (D), respectively, corresponding to small sample sizes. The rejection rate is higher for χ^2 test when the null hypothesis is the correct distribution (the model generating the data) and lower for the incorrect distribution than for the proposed Fano Gamma test. The results thus show this test is more powerful than the classically used χ^2 test.

Figure 5: Evaluating evidence against Poisson and Negative-Binomial models on visual neurophysiological data (A) The Fano Gamma test (FG) results : Evidence against the Poisson model. Each subplot shows the measures of variability for the different cells and conditions as points on a (mean, variance) plot. The straight line corresponds to $FF=1$ (Poisson), while dots correspond to observations. Blue dots correspond to observations with no evidence against Poisson, red dots correspond to no evidence against the Negative-Binomial model (NBM) and black dots correspond to strong evidence against the Poisson model (PM) on the top and against NBM on the bottom based on the empiric distribution of the Fano factor (FG test). The results show that there is more evidence of over-dispersion in MT and that the proportion of cells showing over-dispersion seems to increase along the visual hierarchy. (B) The percentage of observations with evidence against PM and NBM are given with respect to three criteria: the Fano factor test discards the observations with a Fano factor value higher or lower than its formal expected value (equal to one for the Poisson model), the χ^2 test and the FG test with a significance level of 0.025. The bar plots correspond to the percentage of observations with evidence against the NBM (left) and the PM (right). The tests were applied to 3 data sets (from the left to the right, respectively): LGN cells (mouse), V1 cells (awake macaque monkey) and MT cells (anaesthetized macaque monkeys). The results show that the Fano factor values are misleading when the number of trials is not taken into account. Looking only at the FF gives strong evidence against PM and NBM that is not confirmed by the other tests. The χ^2 test shows slightly more evidence against the Poisson model than the FG test which could be explained by categorization issues. It shows also very strong evidence against the NBM model known to be suitable for over dispersed data. In summary, the FG test gives more coherent results: weak evidence of over-dispersion in LGN and V1 data sets and low evidence against the NBM in MT data set.

Figure 6: Mean and dispersion tunings (Left) The box plots represent the distribution of cells' mean spike count (A) and the distribution of their inverse-dispersion values (C) function of the different directions centered around the preferred direction of each cell. The "whiskers" for each direction indicate the 25th and the 75th percentiles. The colored dashes (blue for mean and red for inverse dispersion) represent the median values. The inverse-dispersion parameter (C) shows a tuning less regular than the mean (A) but that could, as a first approximation, be fitted by a bell-shaped function. (Right) Tuning curves of the mean spike count (B, in blue) and the inverse-dispersion (D, in red) for 16 cells resulting from, respectively, the fit of a von Mises function (equivalent to a circular Gaussian) to each cell estimated parameters from their repeated responses to the 16 directions used for training (gray points). Each subplot corresponds to a cell. The tuning functions of the inverse-dispersion are supposed to be centered on the cell's preferred direction deduced from the mean fit.

Figure 7: Population decoding : Poisson model versus Negative-Binomial model. Variation of decoding error values using the Negative-Binomial model (NBM) function compared to that using the Poisson model (PM) both applied to real MT data (A and C), then to surrogate data generated using the parameters inferred from the real data (B and D). We tested two alternatives of the NBM : one with an over-dispersion parameter stimulus-independent (non-tuned dispersion) and the other fitting this parameter by a von Mises function of the direction centered on the preferred direction (tuned dispersion). Points correspond to the median error values and bars to 25%-75% confidence interval within 100 trials (horizontal for PM and vertical for NBM). The Poisson model and The Negative-Binomial model with non-tuned dispersion show similar results. However, the Negative-Binomial model with stimulus-dependent dispersion parameters shows better results compared to the two others. This suggests that tuned over-dispersion could play a role in the neural code.

Figure 8: **Complex tuning of Fano factor (FF) values for different motion directions.** (A) An example of a mean response tuning curve. (B) An example of an inverse-dispersion tuning curve. (A) and (B) were analytically generated by Von Mises (circular Gaussian) functions centered on a given preferred direction (PD). (C) Let us suppose that the spiking count of a cell is driven from a Negative-Binomial model characterized by mean values given in (A) and inverse dispersion values given in (B). The expected resulting Fano factor values are given by the relationship equation ($FF = 1 + \frac{f}{\phi}$). For example, the mean value (7.14) and the inverse dispersion value (21.47) corresponding to the PD result in an FF value of 1.33. The resulting FF profile could not be explained only by the mean response tuning as in a Poisson model.