Replica-free evaluation of the neuronal population information with mixed continuous and discrete stimuli: From the linear to the asymptotic regime

Valeria Del Prete

Department of Mathematics, King's College London, Strand WC2R2LS, London, United Kingdom
(Received 23 January 2003; published 11 September 2003)

Recent studies have explored theoretically the ability of populations of neurons to carry information about a set of stimuli, both in the case of purely discrete or purely continuous stimuli, and in the case of multidimensional continuous angular and discrete correlates, in the presence of additional quenched disorder in the distribution. An analytical expression for the mutual information has been obtained in the limit of large noise by means of the replica trick. Here, we show that the same results can actually be obtained in most cases without the use of replicas, by means of a much simpler expansion of the logarithm. Fitting the theoretical model to real neuronal data, we show that the introduction of correlations in the quenched disorder improves the fit, suggesting a possible role of signal correlations—actually detected in real data—in a redundant code. We show that even in the more difficult analysis of the asymptotic regime, an explicit expression for the mutual information can be obtained without resorting to the replica trick despite the presence of quenched disorder, both with a Gaussian and with a more realistic thresholded-Gaussian model. When the stimuli are mixed continuous and discrete, we find that with both models the information seem to grow logarithmically to infinity with the number of neurons and with the inverse of the noise, even though the exact general dependence cannot be derived explicitly for the thresholded-Gaussian model. In the large noise limit, lower values of information were obtained with the thresholded-Gaussian model, for a fixed value of the noise and of the population size. On the contrary, in the asymptotic regime, with very low values of the noise, a lower information value is obtained with the Gaussian model.

I. INTRODUCTION

The mutual information, extensively used in the theory of communication [1,2], has been more recently proposed as a measure of the coding capacity of real neurons in the brain (see, for example, Refs. [3–5], for a general overview). Information estimates, both from real data and in pure theoretical modeling, ideally quantify how efficiently an external observer might discriminate between several correlates of behavior on the basis of the firing of single or multiple cells.

Several theoretical studies have explored the ability of one population of neurons to encode external stimuli, relevant to behavior [6–10]; others have tried to assess how efficiently the information is transmitted across several layers of a network, which may represent distinct stages of processing in some brain area [11–13].

In most cited works, the replica trick [14] has been successfully used in order to derive an explicit expression for the mutual information. As we will show in detail in the following section, from the formula of the information replicas do appear as a natural methodological choice, due to the presence of the logarithm of a sum of conditional probabilities depending on some quenched parameters. Yet in the cited works no attempt has been done to verify whether the same results can be obtained without resorting to replicas, even in the cases [6,7,10] where the evaluation could be carried out without any additional assumption of replica symmetry.

Moreover, an exact estimate of the mutual information regardless of the population size N and of the noise σ is often unachievable, so that an analytical expression can be provided only in some limit cases. It might well be that restricting oneself to these cases makes the use of replicas redundant or at least an alternative choice to other methods.

In particular, Refs. [6,7] have used replicas to study the initial linear rise of the information, characterized by small population sizes and large noise in the firing distributions of the neurons; this limit would roughly correspond to a high temperature regime for a physical system such as a spin glass. It is reasonable to think that this limit can be treated and solved without replicas, since it is known that annealed and quenched averages coincide in the high temperature regime.

Here, we first reconsider the analysis performed in Refs. [6,7]; we show that, in the limit when the noise σ is large and the population size N is small, the same analytical expressions for the information can be obtained without the use of the replica trick, by means of a simple Taylor expansion of the logarithm, regardless of the nature of the stimulus whether purely discrete or mixed continuous and discrete, and both with a Gaussian and with a more realistic thresholded-Gaussian firing distribution.

In the particular case of mixed continuous angular and discrete stimuli [7], the distribution had been parametrized in order to model the firing of neurons recorded from the motor cortex of monkeys performing arm movements, categorized according to a direction and a "type" [15,16]. Restricted to this dataset, correlations in the preferred direction of a given unit across different movement types were actually observed, but the impact of such correlations on the information content was not quantified. Thus, here we investigate theoretically whether correlations introduced in the quenched parameters characterizing the distribution can improve the fit of

real information curves provided by the model. This would suggest that such correlations are information bearing, or better, depress the information, leading to a redundant code.

We move then to the limit of large population sizes and small noise. An attempt to study this regime in the presence of purely discrete stimuli by means of replicas was unsuccessful in Ref. [6].

Here, we show that even in the asymptotic regime, in the case of purely discrete stimuli, an analytical expression for the mutual information can be provided without the use of the replica trick.

Another replica-free approach to this limit was proposed in Ref. [9], applicable both to the case of continuous and discrete stimuli, for a generic firing distribution, provided that it can be factorized into single neuron probability density functions. No additional quenched disorder was assumed in the distribution. Here, we try and apply this method to our particular model and we find the assumption under which we retrieve our original approximation.

Finally, in Ref. [7] it has been shown that, when limited to the initial linear regime, both the Gaussian and the thresholded-Gaussian model provide the same analytical expression for the mutual information, except for renormalization of a noise parameter. In particular, lower values of the information were obtained with the thresholded-Gaussian model. We investigate this issue in the asymptotic regime comparing the leading term of the information for both models.

II. POPULATION INFORMATION IN THE INITIAL LINEAR REGIME

A. Coding of purely discrete and mixed continuous and discrete stimuli in a Gaussian approximation

The firing of neurons emerging from the analysis of real data is characterized by strong irregularities and by a wide variability. The choice of a Gaussian model as a possible firing rate distribution might therefore seem unrealistic and unjustified. Yet with a large sample of data, it is likely that most irregularities in the distribution average out; their presence if often due to a too poor sampling, which in turn biases information estimates, so that smoothing with a Gaussian or other kernels has become a standard procedure in data analysis (see Ref. [4] for a review of several regularizing procedures). The advantage in using a Gaussian approximation is easier mathematical analysis, which allows for the derivation of an explicit expression for the mutual information [6–8,10]. Moreover, at least in the regime of the initial information rise, the use of a more realistic model leads to the same mathematical expression for the mutual information, except for a renormalization of the noise [7]. This last issue will be discussed in more detail in the following section.

Let us consider a population of N independent cells which fire to a set of p discrete stimuli, parametrized by a discrete variable s, according to a Gaussian distribution

$$p(\{\eta_i\}|s) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp-[(\eta_i - \eta_i^s)^2/2\sigma^2], \quad (1)$$

where η_i is the firing rate of the ith input neuron, while η_i^s is its mean rate in response to stimulus s.

The mutual information between the neuronal firing rates $\{\eta_i\}$ and the stimuli s reads

$$I(\{\eta_i\},s) = \sum_s p(s) \int \prod_i d\eta_i p(\{\eta_i\}|s) \log_2 \frac{p(\{\eta_i\}|s)}{p(\{\eta_i\})}. \quad (2)$$

Since $p(\{\eta_i\})$ can be written as $\Sigma_s p(s)p(\{\eta_i\}|s)$, it is easy to show that the mutual information can be expressed as the difference between the entropy of the firing rates $H(\{\eta_i\})$ and the equivocation $\langle H(\{\eta_i\}|s)\rangle_s$:

$$I(\{\eta_i\},s) = H(\{\eta_i\}) - \langle H(\{\eta_i\}|s)\rangle_s, \quad (3)$$

with

$$\langle H(\{\eta_i\}|s)\rangle_s = -\sum_s p(s) \int \prod_i d\eta_i p(\{\eta_i\}|s)$$
$$\times \log_2 p(\{\eta_i\}|s), \quad (4)$$

$$H(\{\eta_i\}) = -\sum_s p(s) \int \prod_i d\eta_i p(\{\eta_i\}|s) \log_2$$
$$\times \left[\sum_{s'} p(s')p(\{\eta_i\}|s') \right]. \quad (5)$$

The variables $\{\eta_i\}$ in $p(\{\eta_i\}|s')$ are *quenched*: the sum on the stimuli s' should be performed, and the logarithm taken, for any fixed configuration $\{\eta_i\}$, before integrating on $\{\eta_i\}$. The replica trick, devised to perform averages of the partition function across quenched disorder in spin glasses [14], seems to apply also to this case. Yet, contrary to what is found in the theory of spin glasses, where the connectivities vary on a much longer time scale with respect to the spins and therefore they are quenched, here the presence of quenched disorder does not reflect any real distinction between two separate time scales. In fact, the same sum appears outside the logarithm, and if one were able to explicitly derive $p(\{\eta_i\})$ from $p(\{\eta_i\}|s')$ there would be no need for replicas to evaluate $H(\{\eta_i\})$.

In the specific case of distribution (1), $p(\{\eta_i\})$ has a functional dependence on the configuration of the average rates $\{\eta_i^s\}$ and it cannot be explicitly derived except for some trivial cases, such as

$$\eta_i^s = \eta_i^0 \quad \forall \ s, \quad (6)$$

where the information is obviously zero, since $p(\{\eta_i\}|s)$ does not depend on s anymore; or the opposite noiseless limit, where the cells fire at each stimulus s always with a pattern $\{\eta_i^s\}$ and the configurations $\{\eta_i^s\}$ across the stimuli do not overlap. In this case, when the stimuli are equally likely, so that $p(s) = 1/p$, one has

$$p(\{\eta_i\}|s) = \delta(\{\eta_i\} - \{\eta_i^s\}), \qquad (7)$$

$$p(\{\eta_i\}) = 1/p \sum_{s'} \delta(\{\eta_i\} - \{\eta_i^{s'}\}), \qquad (8)$$

and since the average configurations $\{\eta_i^s\}$ do not overlap, it is easy to see that the mutual information reaches the upper bound of $\log_2 p$.

In a more realistic context, the average firing rates $\{\eta_i^s\}$ are not kept fixed, reflecting the strong variability of the neural activity detected in real data. Therefore, in order to obtain an information estimate independent of a particular configuration of the selectivities, the variables $\{\eta_i^s\}$ are considered quenched and the information must finally be averaged across the distribution of $\{\eta_i^s\}$:

$$I(\{\eta_i^s\}) \rightarrow \langle I(\{\eta_i^s\}) \rangle_\eta, \qquad (9)$$

$$\langle F(\{\eta_i^s\}) \rangle_\eta = \int d\{\eta_i^s\} \varrho(\{\eta_i^s\}) F(\{\eta_i^s\}). \qquad (10)$$

$I(\{\eta_i^s\})$ is the mutual information between the neuronal firing rates $\{\eta_i\}$ and the stimuli s evaluated according to Eq. (2), for a particular configuration of the mean rates $\{\eta_i^s\}$.

This approach has been followed in Refs. [6,7], where the replica trick has been used to perform the analytical evaluation.

Let us consider the case where quenched disorder is uncorrelated and identically distributed across units and across the p discrete correlates

$$\varrho(\{\eta_s^i\}) = \prod_{i,s} \varrho(\eta_s^i) = [\varrho(\eta)]^{Np}. \qquad (11)$$

As already shown in Refs. [6,7], it is easy to prove that for a population of independent units, the equivocation $\langle H(\{\eta_i\}|s) \rangle_s$ is additive.

By $I(\{\eta_i\},s)$ and $H(\{\eta_i\})$ and $\langle H(\{\eta_i\}|s) \rangle_s$ in the following, I will implicitly mean the corresponding quenched averaged quantities. Inserting Eq. (1) in Eq. (4) one obtains

$$\langle H(\{\eta_i\}|s) \rangle_s = \frac{N}{2\ln 2}(1 + \ln 2\pi\sigma^2). \qquad (12)$$

We turn now to the more difficult evaluation of the rate entropy. Inserting Eq. (1) in Eq. (5) and using the equivalence

$$\ln\left[\sum_{s'} p(s') \prod_i \frac{e^{-(\eta_i - \eta_i^{s'})^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}\right]$$

$$= -\frac{N}{2}\ln 2\pi\sigma^2 - \sum_i \frac{\eta_i^2}{2\sigma^2}$$

$$+ \ln\left[\sum_{s'} p(s') \prod_i e^{[2\eta_i \eta_i^{s'} - (\eta_i^{s'})^2]/2\sigma^2}\right], \qquad (13)$$

one can integrate out the second term on the right-hand side in Eq. (13); the result, added to the first term, simplifies with the equivocation, Eq. (12); rearranging all the terms, the mutual information can be written in the following form:

$$I(\{\eta_i\},s) = -\left\langle \frac{1}{\ln 2} \sum_s p(s) \int \prod_i d\eta_i \prod_i e^{-(\eta_i - \eta_i^s)^2/2\sigma^2} \right.$$

$$\left. \times \ln\left[\sum_{s'} p(s') \prod_i e^{-[(\eta_i^s)^2 + (\eta_i^{s'})^2 - 2\eta_i \eta_i^{s'}]/2\sigma^2}\right] \right\rangle_\eta . \qquad (14)$$

Due to the presence of the sums and of the quenched disorder under the logarithm, an analytical expression of the mutual information cannot be obtained in the general case; yet the evaluation can be performed in some limit cases. We focus here on the initial regime, where the number of cells is not large compared to the noise. The asymptotic regime for large population sizes will be discussed later on.

As it has been shown in Ref. [6], one way to get rid of the logarithm and to perform the quenched averages and the sums in Eq. (14) is by means of the replica trick [14]; yet, in the limit when the noise σ is very large and the population size N is not large, a straightforward and natural approach consists of performing a simple Taylor expansion of the exponentials under the logarithm

$$\ln\left[\sum_{s'} p(s') \prod_i e^{-[(\eta_i^s)^2 + (\eta_i^{s'})^2 - 2\eta_i \eta_i^{s'}]/2\sigma^2}\right]$$

$$\simeq \ln\left[1 - \sum_{s'} p(s')\left(\sum_i \frac{(\eta_i^s)^2}{2\sigma^2} + \sum_i \frac{(\eta_i^{s'})^2}{2\sigma^2}\right.\right.$$

$$\left.\left. - \sum_i \frac{\eta_i \eta_i^{s'}}{\sigma^2} - \frac{1}{2}\sum_{ij} \frac{\eta_i \eta_j \eta_i^{s'} \eta_j^{s'}}{\sigma^4}\right)\right]$$

$$\simeq - \sum_{s'} p(s')\left(\sum_i \frac{(\eta_i^s)^2}{2\sigma^2} + \sum_i \frac{(\eta_i^{s'})^2}{2\sigma^2}\right.$$

$$\left. - \sum_i \frac{\eta_i \eta_i^{s'}}{\sigma^2} - \frac{1}{2}\sum_{ij} \frac{\eta_i \eta_j \eta_i^{s'} \eta_j^{s'}}{\sigma^4}\right)$$

$$- \frac{1}{2}\sum_{s'} p(s')\sum_{s''} p(s'')\sum_{ij} \frac{\eta_i \eta_j \eta_i^{s'} \eta_j^{s''}}{\sigma^4}. \qquad (15)$$

The terms of the order of $1/\sigma^4$ must be kept because it can be shown that after integration on $\{\eta_i\}$ they will actually give a contribution of the of order of $1/\sigma^2$ to the mutual information.

Inserting expansion (15) in Eq. (14) and performing the integration on $\{\eta_i\}$, one obtains

$$I(\{\eta_i\},s) = \frac{1}{\ln 2}\frac{1}{2\sigma^2}\left\langle \sum_i\left[\sum_s p(s)(\eta_i^s)^2\right.\right.$$

$$\left.\left. - \sum_s \sum_{s'} p(s)p(s')\eta_i^s \eta_i^{s'}\right]\right\rangle_\eta$$

$$= \frac{1}{\ln 2}\frac{N}{2\sigma^2}\sum_s \sum_{s'} p(s)p(s')[\langle(\eta^s)^2\rangle_\eta$$

$$- \langle \eta^s \eta^{s'} \rangle_\eta], \qquad (16)$$

where we have used the fact that quenched disorder is uncorrelated and identically distributed across stimuli and neurons.

The averages across quenched disorder and across s, s' can be performed distinguishing between the cases $s = s'$, $s \neq s'$. The final result for the mutual information up to order N/σ^2 reads

$$I(\{\eta_i\}, s) = \frac{1}{\ln 2} \frac{N}{2\sigma^2} \frac{p-1}{p} \sigma_\eta^2, \quad \sigma_\eta^2 = \langle \eta^2 \rangle_\eta - \langle \eta \rangle_\eta^2.$$

$$(17)$$

The same result has been obtained in Ref. [6] by means of the replica trick. We have checked that the agreement between the two approaches is found also at higher orders in N/σ^2. Yet the derivation via the replica trick is clearly longer and more complicated, and *a priori* less controllable than the simple Taylor expansion used here to derive the same results.

The interest in the coding of purely discrete stimuli rises naturally from the need to provide a theoretical framework allowing a direct quantitative comparison with the results of real experiments. In fact, in a typical experimental protocol neural activity is recorded from some areas in the brain, while the subject (human or animal) is presented a discrete number of stimuli, or it is trained to perform a discrete number of tasks.

Yet natural stimuli are multidimensional and some of the dimensions can vary in a continuous domain. For example, a visual stimulus can be parametrized through its color (varying within a discrete set of possible choices) and its orientation (represented by continuous angle). It is therefore a primary theoretical interest to extend our results to the case where the stimulus may be multidimensional and the dimensions may be discrete and continuous.

In Refs. [7,8,16] the coding of movements categorized according to their direction (continuous dimension) and their *type* (discrete dimension) has been studied via direct information estimates from real data and pure theoretical modeling. In particular, in Ref. [7] the information between the neuronal firing rates and the movements has been evaluated in the limit of large noise and finite population size, in the presence of quenched disorder and resorting to the replica trick.

In analogy to the model studied in Ref. [7], let us consider a population of N neurons firing independently of one another to an external stimulus parameterized by an angle ϑ and a discrete variable s, according to a Gaussian distribution

$$p(\{\eta_i\}|\vartheta, s) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \{[\eta_i - \tilde{\eta}_i(\vartheta, s)]^2/2\sigma^2\}.$$

$$(18)$$

Such as in Eq. (1), η_i is the firing rate of the ith neuron; $\tilde{\eta}_j(\vartheta, s)$ is its average firing rate corresponding to the stimulus (ϑ, s),

$$\tilde{\eta}_i(\vartheta, s) = \varepsilon_s^i \bar{\eta}_i(\vartheta) + (1 - \varepsilon_s^i) \eta^f,$$

$$(19)$$

$$\bar{\eta}_i(\vartheta - \vartheta_{i,s}^0) = \eta^0 \cos^{2m}\left(\frac{\vartheta - \vartheta_{i,s}^0}{2}\right),$$

$$(20)$$

where ε_s^i and $\vartheta_{i,s}^0$ are sources of quenched disorder, distributed, respectively, between 0 and 1 and between 0 and 2π, and m is a positive integer; a justification of this choice on the basis of tuning curves actually observed in real data is provided in Ref. [7]. It is assumed that quenched disorder is uncorrelated and identically distributed across neurons and stimuli

$$\varrho(\{\varepsilon_s^i\}) = \prod_{i,s} \varrho(\varepsilon_s^i) = [\varrho(\varepsilon)]^{Np},$$

$$(21)$$

$$\varrho(\{\vartheta_{i,s}^0\}) = [\varrho(\vartheta^0)]^{Np} = \frac{1}{(2\pi)^{Np}}.$$

$$(22)$$

Equation (19) states that for each discrete correlate s each neuron i fires at an average rate modulating with ϑ around the preferred direction $\vartheta_{i,s}^0$ with an amplitude ε_s^i; alternatively, the average rate is fixed to a value η^f, independently of ϑ, with amplitude $1 - \varepsilon_s^i$. In Ref. [7] it has been shown that a similar choice for the average rate can effectively reproduce the main features of real neurons directional tuning curves.

The basic definitions (2),(4), and (5) as well as the initial treatment can be easily generalized to the case of structured stimuli via the replacements

$$\sum_s p(s) \rightarrow \sum_s p(s) \int d\vartheta p(\vartheta),$$

$$\eta_i^s \rightarrow \tilde{\eta}_i(\vartheta, s),$$

$$\langle \cdots \rangle_\eta \rightarrow \langle \cdots \rangle_{\varepsilon, \vartheta^0}.$$

$$(23)$$

It is easy to show that the mutual information can be expressed in a form analogous to Eq. (14):

$$I(\{\eta_i\}, \vartheta \otimes s)$$

$$= -\left\langle \frac{1}{\ln 2} \sum_s p(s) \int d\vartheta p(\vartheta) \int \prod_i d\eta_i \right.$$

$$\times \prod_i e^{-[\eta_i - \widetilde{\eta_i}(\vartheta, s)]^2/2\sigma^2} \ln\left[\sum_{s'} p(s') \int d\vartheta'\right.$$

$$\left. \times \prod_i e^{-\{[\tilde{\eta}_i(\vartheta, s)]^2 + [\tilde{\eta}_i(\vartheta', s')]^2 - 2\eta_i \tilde{\eta}_i(\vartheta', s')\}/2\sigma^2}\right]\right\rangle_{\varepsilon, \vartheta^0}.$$

$$(24)$$

We use again an expansion of the logarithm similar to Eq. (15); it is then easy to derive the analogous of Eq. (16):

$$I(\{\eta_i\},\vartheta\otimes s)\simeq\frac{1}{\ln 2}\frac{N}{2\sigma^2}\sum_s\sum_{s'}p(s)p(s')\int d\vartheta$$

$$\times\int d\vartheta'p(\vartheta)p(\vartheta')[\langle[\tilde{\eta}(\vartheta,s)]^2\rangle_{\varepsilon,\vartheta^0}$$

$$-\langle\tilde{\eta}(\vartheta,s)\tilde{\eta}(\vartheta',s')\rangle_{\varepsilon,\vartheta^0}],$$

From Eqs. (19)–(22), it is easy to verify that

$$\sum_s\int d\vartheta p(s)p(\vartheta)\langle[\tilde{\eta}(\vartheta,s)]^2\rangle_{\varepsilon,\vartheta^0}$$

$$=(\eta^0)^2[(A_2+\alpha^2-2\alpha A_1)\langle\varepsilon^2\rangle_\varepsilon+\alpha^2$$

$$+2\alpha(A_1-\alpha)\langle\varepsilon\rangle_\varepsilon], \tag{25}$$

$$\sum_s\sum_{s'}p(s)p(s')\int d\vartheta d\vartheta'p(\vartheta)p(\vartheta')$$

$$\times\langle\tilde{\eta}(\vartheta,s)\tilde{\eta}(\vartheta',s')\rangle_{\varepsilon,\vartheta^0}$$

$$=(\eta^0)^2\left[(A_1-\alpha)^2\left(\frac{p-1}{p}\langle\varepsilon\rangle_\varepsilon^2+\frac{1}{p}\langle\varepsilon^2\rangle_\varepsilon\right)\right.$$

$$\left.+\alpha^2+2\alpha(A_1-\alpha)\langle\varepsilon\rangle_\varepsilon\right], \tag{26}$$

$$A_1=\frac{1}{2^{2m}}\binom{2m}{m},\quad A_2=\frac{1}{2^{4m}}\binom{4m}{2m},\quad \alpha=\frac{\eta^f}{\eta^0}. \tag{27}$$

Inserting Eqs. (25) and (26) in Eq. (25), one obtains the final expression for the mutual information up to order N/σ^2:

$$I(\{\eta_i\},\vartheta\otimes s)\simeq\frac{1}{\ln 2}\frac{N(\eta^0)^2}{4\sigma^2}\left[\frac{p-1}{p}2(\alpha-A_1)^2\sigma_\varepsilon^2\right.$$

$$\left.+2[A_2-(A_1)^2]\langle\varepsilon^2\rangle_\varepsilon\right]. \tag{28}$$

In Ref. [7] it has been shown that the same expression for the information in the linear approximation is obtained in the case where the preferred directions do not modulate with the discrete stimuli: $\vartheta_{i,s}^0=\vartheta_i^0\ \forall\ s,i$. It can be easily proved that a different contribution to the information would derive in either case from the term $\langle\tilde{\eta}(\vartheta,s)\tilde{\eta}(\vartheta',s')\rangle_{\varepsilon,\vartheta^0}$; yet the differing term becomes zero when averaged across ϑ, ϑ'.

Setting $\vartheta_{i,s}^0=\vartheta_i^0\ \forall\ s$ corresponds to correlating the signal that each neuron carries about different stimuli s. Intuitively, since no difference in the preferred orientation can be detected any more while looking at distinct correlates s, one would expect an information loss. Such a loss is indeed present, as revealed from a detailed evaluation of the quadratic contribution in the population size N. We do not report the calculation, which is a trivial application of the perturbative theory very much similar to the one performed for the

linear approximation in N, and which consists in retaining all the terms of the order of $N^2(\eta^0)^4/\sigma^4$ out of the expansion of the logarithm.

The final expression, obtained adding the second-order contributions to the linear term, Eq. (28), reads in either case

$$I_2\simeq\frac{1}{\ln 2}\frac{N(\eta^0)^2}{4\sigma^2}\left[\frac{p-1}{p}2(\alpha-A_1)^2\sigma_\varepsilon^2+2[A_2-(A_1)^2]\langle\varepsilon^2\rangle_\varepsilon\right]$$

$$-\frac{1}{\ln 2}\frac{N^2(\eta^0)^4}{2(4\sigma^2)^2}\left\{\frac{p-1}{p^2}2[2(\alpha-A_1)^2\lambda_1]^2\right.$$

$$\left.+\frac{(\lambda_2)^2}{p}\left[\left(\frac{1}{2^{2m-1}}\right)^{4m-1}\sum_{\nu=0}\left[\binom{2m}{\nu}\right]^4\right]\right\}, \tag{29}$$

$$I_2^{corr}\simeq\frac{1}{\ln 2}\frac{N(\eta^0)^2}{4\sigma^2}\left[\frac{p-1}{p}2(\alpha-A_1)^2\sigma_\varepsilon^2+2[A_2-(A_1)^2]\right.$$

$$\left.\times\langle\varepsilon^2\rangle_\varepsilon\right]-\frac{1}{\ln 2}\frac{N^2(\eta^0)^4}{2(4\sigma^2)^2}\left\{\frac{p-1}{p^2}2[2(\alpha-A_1)^2\lambda_1]^2\right.$$

$$+\left[\frac{p-1}{p}(\lambda_1-\lambda_2)^2+\frac{(\lambda_2)^2}{p}\right]$$

$$\left.\times\left[\left(\frac{1}{2^{2m-1}}\right)^{4m-1}\sum_{\nu=0}\left[\binom{2m}{\nu}\right]^4\right]\right\}, \tag{30}$$

where by I_2^{corr} we mean the information in the correlated case $\vartheta_{i,s}^0=\vartheta_i^0\ \forall\ s,i$.

The same expression as in Eq. (30) has been obtained in Ref. [7] by means of the replica trick.

Figure 1, on the left, shows an example of neuron recorded in the supplementary motor area (SMA), whose preferred direction does not modulate with the discrete dimension (reproduced from Ref. [7]). Restrictedly to this dataset, such neurons were statistically dominant, even though the significance of such observation should be quantified by means of the analysis of other samples of cells.

On the right, we show the theoretical curves in the quadratic approximation for the uncorrelated and correlated cases [Eqs. (29) and (30), respectively] corresponding to the best linear fit [Eq. (28)] of the information as estimated from a population of SMA cells. To obtain such an estimate, for each value of N, we extracted random subsamples of cells out of the total population and we used a decoding procedure, finally averaging results across subsamples; details of the data analysis are given in Ref. [16]. The visual comparison shows that the introduction of correlations in the preferred directions improves the fit. Even far from proving that this precise type of signal correlation is the actual mechanism used by SMA cells, this result suggests that real cells transmit information firing in a correlated way.

B. Coding of mixed continuous and discrete stimuli with a thresholded-Gaussian model

Till now we have examined the information carried about stimuli characterized by discrete or mixed continuous and discrete dimensions, assuming that the firing of different cells is independent across cells and Gaussian distributed.
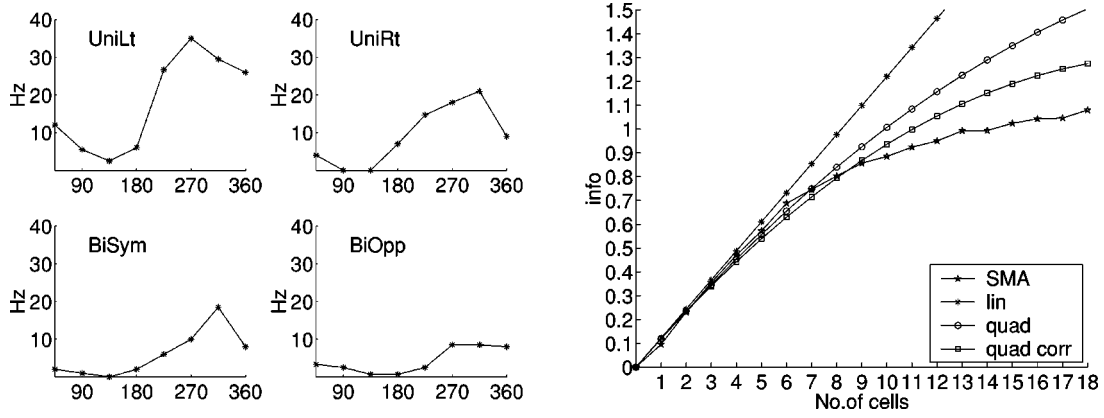
FIG. 1. Left: directional tuning for a cell recorded in the right supplementary motor area of a monkey performing four different types of arm movement. UniLt=unimanual left, UniRt=unimanual right, BiSym=bimanual symmetric, BiOpp=bimanual opposite. On the x axis angles are in degrees; reproduced from Ref. [7]. Right: comparison between the theoretical information curves corresponding to Eqs. (28) (lin in the legend), (29) (quad), (30) (quad corr), and the information estimated from a sample of cells recorded in the right supplementary motor area [16]; $m=1$, $p=2$, the distribution $\varrho(\varepsilon)$ in Eqs. (21) is just equal to 1/3 for each of the three allowed values of ε 0, 1/2, 1; $(\eta^0/2\sigma)^2 = 0.64$.

Yet, as already remarked, both assumptions provide a rough approximation of the firing distribution of real neurons. The most unjustified one seems to be the Gaussian assumption, since it implies that also negative rates have a nonzero probability of occurrence; *a priori* the information rise might result more or less seriously distorted.

This question has been investigated in Ref. [7], where a more realistic model has been proposed, truncating the Gaussian distribution and adding a δ peak in zero.

Let us consider once again a population of independent units firing to mixed continuous and discrete stimuli $\vartheta \otimes s$, where the single neuron distribution is written as follows:

$$P(\eta_i | \vartheta, s) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \{ [\eta_i - \tilde{\eta}_i(\vartheta,s)]^2 / 2\sigma^2 \} \Theta(\eta_i)$$

$$+ 2\{ 1 - \mathrm{erf}[\tilde{\eta}_i(\vartheta,s)/\sigma] \} \delta(\eta_i) \Theta(-\eta_i), \tag{31}$$

$\Theta(x)$ is the Heaviside step function and $\tilde{\eta}_i(\vartheta,s)$ has already been defined in Eq. (19). $\mathrm{erf}(x)$ is the error function:

$$\mathrm{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} dt\, e^{-t^2/2}. \tag{32}$$

In Ref. [7] the mutual information $I(\{\eta_i\}, \vartheta \otimes s)$ has been evaluated by means of the replica trick, in the limit of large noise σ. The interest in this limit arises since the larger σ is, the larger is the Gaussian weight assigned as a whole to negative rates; a consequence might be a larger distortion in the information values.

We show here how the same results can be obtained without the use of the replica trick.

As usual, the information can be expressed as the difference between the equivocation and the output entropy, analogously to Eqs. (4) and (5) and considering replacements (23). In Ref. [7] it is shown that the equivocation can be calculated quite easily as a sum of single neuron terms; assuming as usual that quenched disorder in uncorrelated and identically distributed across stimuli and neurons, according to Eqs. (21) and (22), one obtains

$$\langle H(\{\eta_i\} | \vartheta, s) \rangle_{\vartheta,s} = \frac{N}{2\ln 2} \left\{ [1 + \ln(2\pi\sigma^2)] \langle \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma] \rangle_{\varepsilon,\vartheta^0} - \left\langle \frac{\tilde{\eta}(\vartheta,s)}{\sqrt{2\pi}\sigma} e^{-[\tilde{\eta}(\vartheta,s)]^2/2\sigma^2} \right\rangle_{\varepsilon,\vartheta^0} \right. $$

$$\left. + 2\langle [1 - \mathrm{erf}\{\tilde{\eta}(\vartheta,s)/\sigma\}] \rangle_{\varepsilon,\vartheta^0} \ln\frac{\epsilon}{2} - 2\langle [1 - \mathrm{erf}\{\tilde{\eta}(\vartheta,s)/\sigma\}] \ln[1 - \mathrm{erf}\{\tilde{\eta}(\vartheta,s)/\sigma\}] \rangle_{\varepsilon,\vartheta^0} \right\}, \tag{33}$$

where we have used the representation of the δ function

$$\int_{-\infty}^{+\infty} dx\, \delta(x) F(x) = \lim_{\epsilon \to 0} \int_{-\epsilon/2}^{\epsilon/2} dx\, \frac{1}{\epsilon} F(x). \tag{34}$$

The equivocation diverges when $\epsilon \to 0$, but, as we will show later on, this divergence is canceled exactly by a corresponding term in the entropy of the responses, yielding a finite result for the mutual information.

The average across quenched disorder can be performed in the limit of large σ and expanding the error functions in Eq. (33) in powers of $1/\sigma$. First, we evaluate the entropy of the responses, since we will show that there is a partial cancellation of terms.

Considering Eq. (5) for the entropy of the responses and replacements (23), it is easy to show that in the case of distribution (31), one obtains

$$
H(\{\eta_i\}) = -\left\langle \sum_s p(s) \int d\vartheta p(\vartheta) \int \prod_i d\eta_i \prod_i \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp{-\{[\eta_i - \tilde{\eta}_i(\vartheta,s)]^2/2\sigma^2\}\Theta(\eta_i)} \right.\right.
$$

$$
+ (2\{1 - \mathrm{erf}[\tilde{\eta}_i(\vartheta,s)/\sigma]\}\delta(\eta_i)\Theta(-\eta_i)]\log_2\left[\sum_{s'} \int d\vartheta' p(\vartheta',s') \prod_i \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp \right.\right.
$$

$$
\left.\left.\left. -\{[\eta_i - \tilde{\eta}_i(\vartheta',s')]^2/2\sigma^2\}\Theta(\eta_i) + 2\{1 - \mathrm{erf}[\tilde{\eta}_i(\vartheta',s')/\sigma]\}\delta(\eta_i)\Theta(-\eta_i) \right] \right] \right\rangle_{\varepsilon,\vartheta^0} . \tag{35}
$$

Developing the products on the neuron index i and taking into account the symmetry in the distribution of different units, Eq. (35) can be rewritten as follows:

$$
H(\{\eta_i\}) = -\left\langle \sum_s p(s) \int d\vartheta p(\vartheta) \sum_{k=0}^{N} \binom{N}{k} \int_0^\infty \prod_{i=1}^k d\eta_i \int_{-\infty}^0 \prod_{i=k+1}^N d\eta_i \prod_{i=1}^k \frac{1}{\sqrt{2\pi\sigma^2}} \right.
$$

$$
\times \exp{-\{[\eta_i - \tilde{\eta}_i(\vartheta,s)]^2/2\sigma^2\}2^{N-k}\prod_{i=k+1}^N \{1 - \mathrm{erf}[\tilde{\eta}_i(\vartheta,s)/\sigma]\}\delta(\eta_i)\log_2\left[\sum_{s'} \int d\vartheta' p(\vartheta',s') \right.}
$$

$$
\left.\left. \times \left(\prod_{i=1}^k \frac{1}{\sqrt{2\pi\sigma^2}} \exp{-\{[\eta_i - \tilde{\eta}_i(\vartheta',s')]^2/2\sigma^2\}2^{N-k}\prod_{i=k+1}^N \{1 - \mathrm{erf}[\tilde{\eta}_i(\vartheta',s')/\sigma]\}\delta(\eta_i)} \right) \right] \right\rangle_{\varepsilon,\vartheta^0} , \tag{36}
$$

where we have used the following conventions:

$$
\prod_{i=1}^k x_i = 1 \quad \text{if} \quad k = 0, \tag{37}
$$

$$
\prod_{i=k+1}^N x_i = 1 \quad \text{if} \quad k = N. \tag{38}
$$

Extracting from the logarithm all the factors which do not depend on ϑ, s and integrating them on $\{\eta_i\}$, it is easy to show that the entropy of the responses can be expressed as follows:

$$
H(\{\eta_i\}) = \sum_s \int d\vartheta p(\vartheta,s) \left[\sum_{k=1}^N \binom{N}{k} \frac{k}{2\ln 2}(1 + \ln 2\pi\sigma^2)\langle \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^k \langle 1 - \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k} \right.
$$

$$
+ \sum_{k=0}^{N-1} \binom{N}{k}\frac{N-k}{\ln 2}\ln(\epsilon/2)\langle \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^k \langle 1 - \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k} + \sum_{k=1}^N \binom{N}{k}\frac{k}{2\ln 2}
$$

$$
\times \langle \tilde{\eta}^2(\vartheta,s)/\sigma^2 \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}\langle \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{k-1}\langle 1 - \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k} + \sum_{k=1}^N \binom{N}{k}\frac{k}{2\ln 2}
$$

$$
\left. \times \langle \tilde{\eta}(\vartheta,s)/\sqrt{2\pi}\sigma e^{-\tilde{\eta}^2(\vartheta,s)/2\sigma^2}\rangle_{\varepsilon,\vartheta^0}\langle \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{k-1}\langle 1 - \mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k} \right]
$$

$$
- \left\langle \sum_s \int d\vartheta p(\vartheta,s)\sum_{k=0}^N \binom{N}{k}\int_0^\infty \prod_{i=1}^k d\eta_i \prod_{i=1}^k \frac{1}{\sqrt{2\pi\sigma^2}}\exp{-\{[\eta_i - \tilde{\eta}_i(\vartheta,s)]^2/2\sigma^2\}\prod_{i=k+1}^N \{1 - \mathrm{erf}[\tilde{\eta}_i(\vartheta,s)/\sigma]\}} \right.
$$

$$
\left. \times \log_2\left[\sum_{s'} \int d\vartheta' p(\vartheta',s')\left(\prod_{i=1}^k \exp[2\eta_i\tilde{\eta}_i(\vartheta',s') - \tilde{\eta}_i^2(\vartheta',s')/2\sigma^2]\prod_{i=k+1}^N \{1 - \mathrm{erf}[\tilde{\eta}_i(\vartheta',s')/\sigma]\} \right) \right] \right\rangle_{\varepsilon,\vartheta^0} , \tag{39}
$$

where we have used equality (34) and we have assumed that quenched disorder is uncorrelated across units and stimuli.

Subtracting the equivocation, Eq. (33) from the entropy of the responses, Eq. (39), it is easy to see that after summation on k, the first two terms in Eq. (39) simplify with analogous terms in Eq. (33), so that the logarithmic divergence for $\epsilon \to 0$ cancel out.

Finally, the mutual information can be rewritten as follows:

$$
I(\{\eta_i\}, \vartheta \otimes s) = \sum_s \int d\vartheta p(\vartheta, s)\left[\frac{N}{\ln 2}\langle[1-\mathrm{erf}\{\tilde{\eta}(\vartheta, s)/\sigma\}]\ln[1-\mathrm{erf}\{\tilde{\eta}(\vartheta, s)/\sigma\}]\rangle_{\varepsilon, \vartheta^0} + \frac{N}{\ln 2}\left\langle\frac{\tilde{\eta}(\vartheta, s)}{\sqrt{2\pi}\sigma}e^{-[\tilde{\eta}(\vartheta, s)]^2/2\sigma^2}\right\rangle_{\varepsilon, \vartheta^0}\right.
$$

$$
+\frac{N}{2\ln 2}\left\langle\frac{\tilde{\eta}^2(\vartheta, s)}{\sigma^2}\mathrm{erf}[\tilde{\eta}(\vartheta, s)/\sigma]\right\rangle_{\varepsilon, \vartheta^0}\Bigg] - \left\langle\sum_s\int d\vartheta p(\vartheta, s)\sum_{k=0}^{N}\binom{N}{k}\int_0^\infty\prod_{i=1}^{k}d\eta_i\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\right.
$$

$$
-\{[\eta_i - \tilde{\eta}_i(\vartheta, s)]^2/2\sigma^2\}\prod_{i=k+1}^{N}\{1-\mathrm{erf}[\tilde{\eta}_i(\vartheta, s)/\sigma]\}\log_2\left[\sum_{s'}p(s')\int d\vartheta' p(\vartheta')\left(\prod_{i=1}^{k}\exp[2\eta_i\tilde{\eta}_i(\vartheta', s')\right.\right.
$$

$$
\left.\left.\left.-\tilde{\eta}_i^2(\vartheta', s')/2\sigma^2]\prod_{i=k+1}^{N}\{1-\mathrm{erf}[\tilde{\eta}_i(\vartheta', s')/\sigma]\}\right)\right]\right\rangle_{\varepsilon, \vartheta^0}. \tag{40}
$$

Equation (40) constitutes the final expression for the mutual information in the general case. To proceed with the analytical evaluation one must now resort to some approximation. As suggested in Ref. [7] in the limit when the noise σ is large, one can expand the error functions in Eq. (40):

$$
\mathrm{erf}(x) \simeq \frac{1}{2} + \frac{1}{\sqrt{2\pi}}x + o(x^2). \tag{41}
$$

A Taylor expansion can be performed also for the exponentials under the logarithm. One has

$$
\log_2\left[\sum_{s'}\int d\vartheta' p(\vartheta', s')\left(\prod_{i=1}^{k}\exp[2\eta_i\tilde{\eta}_i(\vartheta', s')\right.\right.
$$

$$
\left.\left.-\tilde{\eta}_i(\vartheta', s')^2/2\sigma^2]\prod_{i=k+1}^{N}\{1-\mathrm{erf}[\tilde{\eta}_i(\vartheta', s')/\sigma]\}\right)\right]
$$

$$
\simeq \log_2\left[\sum_{s'}p(s')\int d\vartheta' p(\vartheta')\left(1+\sum_{i=1}^{k}\frac{\eta_i\tilde{\eta}_i(\vartheta', s')}{\sigma^2}\right.\right.
$$

$$
-\sum_{i=1}^{k}\frac{\tilde{\eta}_i^2(\vartheta', s')}{2\sigma^2}+\frac{1}{2}\sum_{i,j=1}^{k}\frac{\eta_i\eta_j\tilde{\eta}_i(\vartheta', s')\tilde{\eta}_j(\vartheta', s')}{\sigma^4}\right)
$$

$$
\times\frac{1}{2^{N-k}}\left(1-\sum_{i=k+1}^{N}\sqrt{\frac{2}{\pi}}\frac{\tilde{\eta}_i(\vartheta', s')}{\sigma}\right.
$$

$$
\left.\left.+\frac{1}{\pi}\sum_{i\neq j}\frac{\tilde{\eta}_i(\vartheta', s')\tilde{\eta}_j(\vartheta', s')}{\sigma^2}\right)\right]. \tag{42}
$$

Developing all the products one can expand the logarithm again in powers of $1/\sigma$, very much similarly to what has been shown in Eq. (15). The result is then integrated on $\{\eta_i\}$

and the binomial sums in Eq. (40) can be performed. Details about the evaluation are given in Appendix B.

Finally, the mutual information can be written as follows:

$$
I(\{\eta_i\}, \vartheta \otimes s) \simeq \frac{1}{\ln 2}\frac{N}{2\sigma^2}\left(\frac{1}{2}+\frac{1}{\pi}\right)\sum_s\sum_{s'}p(s)p(s')
$$

$$
\times\int d\vartheta\int d\vartheta' p(\vartheta)p(\vartheta')[\langle[\tilde{\eta}(\vartheta, s)]^2\rangle_{\varepsilon, \vartheta^0}
$$

$$
-\langle\tilde{\eta}(\vartheta, s)\tilde{\eta}(\vartheta', s')\rangle_{\varepsilon, \vartheta^0}]
$$

$$
= \frac{1}{\ln 2}\left(\frac{1}{2}+\frac{1}{\pi}\right)\frac{N(\eta^0)^2}{4\sigma^2}\left[\frac{p-1}{p}2(\alpha-A_1)^2\sigma_\varepsilon^2\right.
$$

$$
\left.+2[A_2-(A_1)^2]\langle\varepsilon^2\rangle_\varepsilon\right], \tag{43}
$$

where we have used Eqs. (25) and (26) and A_1, A_2, and α have been defined in Eq. (27).

This result equals the expression obtained using replicas in Ref. [7], showing that even with a more complicated distribution, other than the simple Gaussian model, the evaluation can be carried out via a simple Taylor expansion of the logarithm, and no significant advantage derives from the use of the replica trick in the limit case of large noise.

Comparing Eqs. (43) and (28) one can notice that limitedly to the case of large noise, the effect of thresholding the Gaussian distribution with respect to the information is merely a renormalization of the noise for a factor of $1/\sqrt{1/2+1/\pi}$. In Ref. [7] it has been shown that this renormalization effect holds at higher orders in $1/\sigma^2$.

III. POPULATION INFORMATION IN THE ASYMPTOTIC REGIME OF LARGE N AND SMALL σ

We turn now to the analysis of the asymptotic regime in the information curve, for a large number of neurons. A first

attempt to solve this limit in the case of independent Gaussian units and discrete stimuli, as in Eq. (1) has been done in Ref. [6] by means of the replica trick. Yet, probably due to some too strong approximation in summing on replicas, the final analytical expression was incorrect, according to the authors.

The asymptotic behavior was then studied in Ref. [9] distinctly in the case of discrete and continuous stimuli, and for a generic distribution of independent units, yet in the absence of additional quenched disorder. We try here to go further and study the case where quenched disorder is present, as in distribution (1), and mixed continuous and discrete dimensions characterize simultaneously the stimulus structure. We focus first on the simpler case of purely discrete stimuli. In this context, we compare an approach which is equivalent in the nature of the approximation to the one presented in Ref. [6], yet replica-free, to the approach presented in Ref. [9]. We show in detail under which approximation the two approaches provide the same result.

A. Coding of discrete stimuli in a Gaussian approximation

Let us reconsider Eq. (14). When σ becomes very small the probability density $p(\eta_i|s)$ can be approximated by a δ function:

$$\frac{e^{-(\eta_i - \eta_i^s)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \to \delta(\eta_i - \eta_i^s). \tag{44}$$

This approximation, which corresponds to freezing the quenched disorder represented by the variables $\{\eta_i\}$ under the logarithm, has been used in Ref. [6] after getting rid of the logarithm itself by means of the replica trick. Under this approximation, the integration on $\{\eta_i\}$ can be performed and the mutual information can be rewritten as follows:

$$I(\{\eta_i\},s) \simeq -\frac{1}{\ln 2}\left\langle \sum_s p(s)\ln\left[\sum_{s'} p(s') \right. \right.$$

$$\left. \left. \times \prod_i e^{-(\eta_i^s - \eta_i^{s'})^2/2\sigma^2}\right]\right\rangle_\eta$$

$$= \log_2 p - \frac{1}{\ln 2}\left\langle \sum_s p(s) \right.$$

$$\left. \times \ln\left[1 + \sum_{s'\neq s}\prod_i e^{-(\eta_i^s - \eta_i^{s'})^2/2\sigma^2}\right]\right\rangle_\eta, \tag{45}$$

where upper bound $\log_2 p$ derives from the term with $s=s'$ in the sum on s' and we have used $p(s)=\text{const}=1/p$.

Since σ is small one can expand the logarithm

$$\ln\left[1 + \sum_{s'\neq s}\prod_i e^{-(\eta_i^s - \eta_i^{s'})^2/2\sigma^2}\right]$$

$$\simeq \sum_{s'\neq s}\prod_i e^{-(\eta_i^s - \eta_i^{s'})^2/2\sigma^2} + \frac{1}{2}$$

$$\times \sum_{s'\neq s}\sum_{s''\neq s}\prod_{ij} e^{-(\eta_i^s - \eta_i^{s'})^2/2\sigma^2}e^{-(\eta_j^s - \eta_j^{s''})^2/2\sigma^2} + \cdots. \tag{46}$$
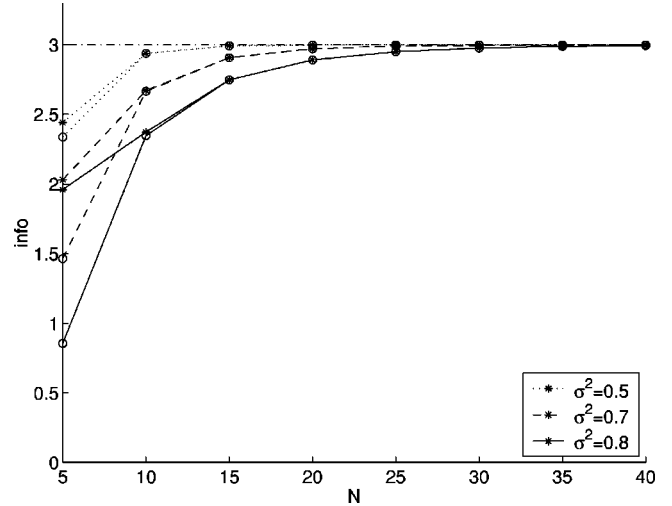

FIG. 2. Mutual information as in Eq. (47), as a function of the population size N. Different curves correspond to different values of the noise σ; circles are for the information with only the leading term of the order of σ^N in Eq. (47), while stars are for the full information with also the correction term of the order of σ^{2N}

In the Appendix, we show that inserting this expansion in Eq. (45), one can perform the quenched averages and derive an explicit expression for the mutual information. The final result reads

$$I(\{\eta_i\},s) \simeq \log_2 p\left[1 - \frac{p-1}{\log_2 p \ln 2}[S_1(\sqrt{2\pi}\sigma I_1)^N \right.$$

$$\left. -(p-2)S_2(2\pi\sigma^2 I_2)^N]\right], \tag{47}$$

where we have considered the leading term of the order of σ and the first correction of the order of σ^2, and one has

$$S_1 = \sum_{k=1}^\infty \frac{(-1)^{k+1}}{k^{N/2+1}}, \quad S_2 = \sum_{k=1}^\infty \frac{(-1)^{k+1}}{k^{N/2}},$$

$$I_1 = \int d\eta\varrho^2(\eta), \quad I_2 = \int d\eta\varrho^3(\eta). \tag{48}$$

When the noise goes to zero and the population size is large, the information reaches the upper bound of $\log_2 p$.

Figure 2 shows the mutual information according to Eq. (47) as a function of the population size and for different values of the noise σ. Circles and stars are, respectively, for the full mutual information with both the leading and the correction terms in Eq. (47) and with only the leading term of the order of σ^N. As it is evident from the plot, the larger σ, the slower the approach to the ceiling and larger the weight acquired by the correction term of the order of σ^{2N}.

An alternative replica-free method to study the asymptotic information regime has been proposed in Ref. [9]. In principle, the method looks quite efficient and, moreover, it can be applied both to continuous and to discrete stimuli.

Yet no additional quenched disorder affected the distributions considered in Ref. [9]. We try now to apply the method to our particular coding scheme.

Let us reconsider Eq. (2) for the mutual information. With a change of variables, it can be rewritten as

$$I(\{\eta_i\},s) = -\sum_s p(s) \int \prod_{s''\neq s} dX_{s''} \langle p(\{X_{s''}\}) \rangle_\eta \log_2$$

$$\times \left[\sum_{s'} p(s') \exp(X_{s'}) \right], \qquad (49)$$

$$X_{s'} = \frac{1}{N} \ln\left(\frac{p(\{\eta_i\}|s')}{p(\{\eta_i\}|s)} \right)$$

$$= \frac{1}{N} \sum_i (\eta_i^s - \eta_i^{s'})(\eta_i^s + \eta_i^{s'} - 2\eta_i)/2\sigma^2, \qquad (50)$$

$$p(\{X_{s'}\}) = \int d\{\eta_i\} p(\{\eta_i\}|s) \prod_{s'\neq s} \delta(X_{s'} - X_{s'}(\{\eta_i\},s,s')), \qquad (51)$$

where in deriving the explicit expression for $X_{s'}$ we have used distribution (1). This change of variable allows to move the quenched disorder from inside the logarithm to the distribution $p(\{X_{s'}\})$.

Using the integral representation for the δ function and integrating on $\{\eta_i\}$, Eq. (51) can be rewritten as follows:

$$p(\{X_{s'}\}) = \int \prod_{s'\neq s} \frac{dY_{s'}}{2\pi} \exp\left(-\iota Y_{s'}\left[X_{s'} + \frac{1}{N} \sum_i (\eta_i' \right.\right.$$

$$\left.\left. - \eta_i^s)^2/2\sigma^2 \right] \right) \prod_{s'\neq s} \prod_{s''\neq s} \exp\left(-\frac{Y_{s'}}{N} \frac{Y_{s''}}{N} \sum_i \right.$$

$$\times (\eta_i^{s'} - \eta_i^s)(\eta_i^{s''} - \eta_i^s)/2\sigma^2 \right). \qquad (52)$$

We notice that when N grows large and σ becomes small, the argument in the second exponential is of the order of $1/N$ with respect to the first one: a relatively flat Gaussian is multiplied times a strongly oscillating periodic function. Thus, we put the second function equal to 1 obtaining

$$p(\{X_{s'}\}) = \prod_{s'\neq s} \delta\left(X_{s'} + \frac{1}{N} \sum_i (\eta_i' - \eta_i^s)^2/2\sigma^2 \right). \qquad (53)$$

We notice that

$$\frac{1}{N} \sum_i (\eta_i' - \eta_i^s)^2/2\sigma^2$$

$$= -\int d\{\eta_i\} p(\{\eta_i\}|s) X_{s'}(\{\eta_i\},s,s')$$

$$= -\int d\{\eta_i\} p(\{\eta_i\}|s) \frac{1}{N} \ln\left(\frac{p(\{\eta_i\}|s')}{p(\{\eta_i\}|s)} \right)$$

$$= -\frac{1}{N} KL(s'||s), \qquad (54)$$

where $KL(s'||s)$ is the Kullback-Leibler divergence between the distributions $p(\{\eta_i\}|s')$ and $p(\{\eta_i\}|s)$. Thus, under this approximation, the distribution $p(\{X_{s'}\})$ factorizes into a product of δ-functions centered on the mean value of $X_{s'}$ $\forall$ s'.

Inserting Eq. (53) in the expression for the mutual information, Eq. (49), it is easy to show that integrating on $\{X_{s'}\}$ one reobtains Eq. (45). Thus, approximating the variables $\{X_{s'}\}$ with the mean of the distribution, as in Eq. (53), corresponds to the δ-function approximation, Eq. (44).

A more accurate estimate might be obtained calculating the correction given by the second exponential in Eq. (52) that we had previously neglected. Yet, integration on $\{Y_{s'}\}$ would lead to the introduction of matrices which depend on the quenched disorder $\{\eta_i^s\}$, which therefore must be averaged out first, in order to derive $\langle p(\{X_{s''}\}) \rangle_\eta$. The average might be performed specifying the distribution of the quenched disorder $\varrho(\eta)$, but even in this case, integrating out the variables $\{\eta_i^s\}$ introduces a nontrivial dependence on the variables $\{Y_{s'}\}$ which in turn can be integrated out by means of further ansatz. Details about the evaluation of the correction and of the results will be published elsewhere [17].

B. Coding of mixed discrete and continuous stimuli: Gaussian vs thresholded-Gaussian model

As we have shown in the preceding section, in the large σ limit both the technical evaluation and the final expression for the mutual information are formally the same whether the stimuli are discrete or continuous. On the other hand, in the asymptotic regime of small σ a qualitative difference exists between the case where the stimuli are discrete or continuous, in that in the former case the information is bounded by the entropy of the stimulus set, while in the latter the information grows to infinity as the noise goes to zero, or if the number of neurons grows to infinity for a finite noise. Here, we calculate the expression of this asymptotic growth, which corresponds to the upper bound in the case of purely discrete stimuli. We compare the two cases of Gaussian and thresholded-Gaussian models, in order to assess whether, as in the large σ regime, a plain relationship such as a renormalization of the noise links the two expressions.

Let us reconsider distribution (18). The expression for the mutual information is given by Eq. (24) that we recall

$I(\{\eta_i\}, \vartheta \otimes s)$

$$= -\left\langle \frac{1}{\ln 2} \sum_s p(s) \int d\vartheta p(\vartheta) \int \prod_i d\eta_i \right.$$

$$\times \prod_i e^{-[\eta_i - \tilde{\eta}_i(\vartheta,s)]^2/2\sigma^2} \ln\left[\sum_{s'} p(s') \int d\vartheta' \right.$$

$$\left.\times \prod_i e^{-\{[\tilde{\eta}_i(\vartheta,s)]^2 + [\tilde{\eta}_i(\vartheta',s')]^2 - 2\eta_i\tilde{\eta}_i(\vartheta',s')\}/2\sigma^2}\right]\right\rangle_{\varepsilon, \vartheta^0}.$$

(55)

In analogy with approximation (44), in the limit when σ becomes very small, we use

$$\frac{e^{-[\eta_i - \tilde{\eta}_i(\vartheta,s)]^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \to \delta(\eta_i - \tilde{\eta}_i(\vartheta,s)).$$

(56)

Under this aproximation, we obtain in analogy to Eq. (45)

$$I(\{\eta_i\}, \vartheta \otimes s) \simeq -\frac{1}{\ln 2}\left\langle \sum_s p(s) \int d\vartheta p(\vartheta)\ln \right.$$

$$\times\left[\sum_{s'} p(s') \int d\vartheta' p(\vartheta') \right.$$

$$\left.\times \prod_i e^{-[\tilde{\eta}_i(\vartheta',s') - \tilde{\eta}_i(\vartheta,s)]^2/2\sigma^2}\right]\right\rangle_{\vartheta^0, \varepsilon}.$$

(57)

In the preceding section, we have seen that in the case of p purely discrete stimuli, the upper bound $\log_2 p$ was given by the term with $s = s'$ under the logarithm. Therefore, in the case of mixed continuous and discrete stimuli, we expect the same term, after integration on ϑ', to give the logarithm of a coefficient depending on the ratio between N and σ^2, which grows to infinity when $\sigma \to 0$. Let us extract and calculate the term with $s = s'$ out of Eq. (57):

$$\int d\vartheta' p(\vartheta') p(s) \prod_i e^{-[\tilde{\eta}_i(\vartheta',s) - \tilde{\eta}_i(\vartheta,s)]^2/2\sigma^2}. \quad (58)$$

It is clear that when σ becomes small the major contribution to the integral comes from the values $\vartheta' \simeq \vartheta$. Therefore, we expand the difference $\tilde{\eta}_i(\vartheta',s) - \tilde{\eta}_i(\vartheta,s)$,

$$\tilde{\eta}_i(\vartheta',s) - \tilde{\eta}_i(\vartheta,s) \simeq \left.\frac{\partial\tilde{\eta}_i(\vartheta',s)}{\partial\vartheta'}\right|_{\vartheta'=\vartheta} (\vartheta' - \vartheta)$$

$$= \frac{\eta^0}{2}\varepsilon_s^i \sin(\vartheta - \vartheta_{i,s}^0)(\vartheta' - \vartheta),$$

(59)

where we have explicitly used expressions (19) and (20).

In the limit when $\sigma \to 0$ the resulting Gaussian distribution is approximated with a δ function and the integral on $d\vartheta'$ can be performed. Extracting this contribution out of the logarithm, the information can be rewritten as follows:

$$I(\{\eta_i\}, \vartheta \otimes s) \simeq \log_2\left[\frac{p\sqrt{\pi}}{\sqrt{2}} \frac{\sqrt{N}\eta^0}{\sigma} \right] + \sum_s \int d\vartheta p(s,\vartheta)$$

$$\times\left\langle \frac{1}{2}\log_2\left[\frac{1}{N} \sum_i (\varepsilon_s^i)^2\sin^2(\vartheta - \vartheta_{i,s}^0) \right]\right\rangle_{\vartheta^0, \varepsilon}$$

$$-\frac{1}{\ln 2}\left\langle \sum_s p(s) \int d\vartheta p(\vartheta)\ln\left[1 \right.\right.$$

$$+\frac{\eta^0}{\sqrt{8\pi\sigma^2}}\sqrt{\sum_{i=1}^N (\varepsilon_s^i)^2\sin^2(\vartheta - \vartheta_{i,s}^0)} \sum_{s'\neq s}$$

$$\left.\left.\times \int d\vartheta' \prod_i e^{-[\tilde{\eta}_i(\vartheta',s') - \tilde{\eta}_i(\vartheta,s)]^2/2\sigma^2}\right]\right\rangle_{\eta}.$$

(60)

The quenched average in the second term can be performed in the thermodynamic limit letting the average pass the logarithm in a mean field approximation. The third term behaves like $\log_2[1+\Delta]$, where Δ vanishes like $\sqrt{N}/\sigma \, \exp(-N/\sigma^2)$ when $\sigma \to 0$. The information in the leading term can be finally expressed in the following form:

$$I(\{\eta_i\}, \vartheta \otimes s) \simeq \log_2\left(\frac{p\sqrt{\pi}}{\sqrt{2}} \frac{\sqrt{N}\eta^0}{\sigma} \sqrt{\frac{\langle\varepsilon^2\rangle_\varepsilon}{2}} \right). \quad (61)$$

When the number of neurons grows to infinity and/or the noise tends to zero, the information grows logarithmically to infinity. Notice that the case of purely continuous stimuli can be retrieved by taking $p = 1$: not surprisingly, the continuous dimension plays a major role in determining the asymptotic growth of the information to infinity, with a relatively mild modulation according to the number p of discrete correlates.

We turn now to the case of the thresholded-Gaussian model, Eq. (31). As we have shown previously, the mutual information can be expressed as in Eq. (40). In the limit of small σ, we apply approximation (56). After a rearrangement of the terms, we obtain

$$
I(\{\eta_i\},\vartheta\otimes s) \simeq \sum_s \int d\vartheta\, p(\vartheta,s) \Bigg[\frac{N}{\ln 2} \langle [1-\mathrm{erf}\{\tilde\eta(\vartheta,s)/\sigma\}] \ln[1-\mathrm{erf}\{\tilde\eta(\vartheta,s)/\sigma\}] \rangle_{\varepsilon,\vartheta^0}
$$

$$
+ \frac{N}{\ln 2} \left\langle \frac{\tilde\eta(\vartheta,s)}{\sqrt{2\pi}\sigma} e^{-[\tilde\eta(\vartheta,s)]^2/2\sigma^2} \right\rangle_{\varepsilon,\vartheta^0} + \frac{N}{2\ln 2} \left\langle \frac{\tilde\eta^2(\vartheta,s)}{\sigma^2} \mathrm{erf}[\tilde\eta(\vartheta,s)/\sigma] \right\rangle_{\varepsilon,\vartheta^0} \Bigg]
$$

$$
- \Bigg\langle \sum_s \int d\vartheta\, p(\vartheta,s) \sum_{k=0}^{N} \binom{N}{k} \prod_{i=k+1}^{N} \{1-\mathrm{erf}[\tilde\eta_i(\vartheta,s)/\sigma]\} \Bigg\{ \sum_{i=1}^{k} \frac{k}{\ln 2} \frac{\tilde\eta_i^2(\vartheta,s)}{\sigma^2}
$$

$$
+ \log_2\Bigg[\sum_{s'} p(s') \int d\vartheta'\, p(\vartheta') \Bigg(\prod_{i=1}^{k} e^{-[\{\tilde\eta_i(\vartheta',s')-\tilde\eta_i(\vartheta,s)\}^2/2\sigma^2]} \prod_{i=k+1}^{N} \{1-\mathrm{erf}[\tilde\eta_i(\vartheta',s')/\sigma]\} \Bigg) \Bigg] \Bigg\} \Bigg\rangle_{\varepsilon,\vartheta^0}.
$$

$$(62)$$

Out of this expression, we aim at keeping the leading terms, which diverge when σ goes to zero and determine the asymptotic behavior. In the first approximation, we will neglect terms going to zero with σ, which would play a role in the first correction to the asymptotic value. As already done in the case of the of the Gaussian model, we consider the term with $s'=s$ in the discrete sum under the logarithm. Looking at the function to be integrated on ϑ', it is easy to see that the product of the k exponentials is maximal for values of ϑ' close to ϑ, while each one of the other $N-k$ factors containing error functions is maximal for cell and stimulus specific values of ϑ', namely, the ones corresponding to the smallest values of $\tilde\eta_i(\vartheta',s)$. Thus, the main contribution for the integral comes from the values of ϑ' close to ϑ and we can repeat the procedure as in Eqs. (59) and (60), obtaining

$$
I(\{\eta_i\},\vartheta\otimes s) \simeq \sum_s \int d\vartheta\, p(\vartheta,s) \Bigg[\frac{N}{\ln 2} \langle [1-\mathrm{erf}\{\tilde\eta(\vartheta,s)/\sigma\}] \ln[1-\mathrm{erf}\{\tilde\eta(\vartheta,s)/\sigma\}] \rangle_{\varepsilon,\vartheta^0} + \frac{N}{\ln 2} \left\langle \frac{\tilde\eta(\vartheta,s)}{\sqrt{2\pi}\sigma} e^{-[\tilde\eta(\vartheta,s)]^2/2\sigma^2} \right\rangle_{\varepsilon,\vartheta^0}
$$

$$
+ \frac{N}{2\ln 2} \left\langle \frac{\tilde\eta^2(\vartheta,s)}{\sigma^2} \mathrm{erf}[\tilde\eta(\vartheta,s)/\sigma] \right\rangle_{\varepsilon,\vartheta^0} \Bigg] - \Bigg\langle \sum_s \int d\vartheta\, p(\vartheta,s) \sum_{k=0}^{N} \binom{N}{k} \prod_{i=k+1}^{N} \{1-\mathrm{erf}[\tilde\eta_i(\vartheta,s)/\sigma]\}
$$

$$
\times \Bigg\{ \sum_{i=1}^{k} \frac{k}{\ln 2} \frac{\tilde\eta_i^2(\vartheta,s)}{\sigma^2} - \log_2\Bigg[\frac{p\sqrt{\pi}}{\sqrt{2}} \frac{\sqrt{k}\eta^0}{\sigma} \Bigg] - \frac{1}{2}\log_2\Bigg[\frac{1}{k} \sum_{i=1}^{k} (\varepsilon_s^i)^2 \sin^2(\vartheta-\vartheta_{i,s}^0) \Bigg]
$$

$$
+ \sum_{i=k+1}^{N} \log_2\{1-\mathrm{erf}[\tilde\eta_i(\vartheta,s)/\sigma]\} + \log_2\Bigg[1 + \frac{\eta^0}{\sqrt{8\pi\sigma^2}} \sqrt{\sum_{i=1}^{k} (\varepsilon_s^i)^2 \sin^2(\vartheta-\vartheta_{i,s}^0)} \Bigg]
$$

$$
\times \sum_{s'\neq s} \int d\vartheta' \Bigg(\prod_{i=1}^{k} e^{-[\{\tilde\eta_i(\vartheta',s')-\tilde\eta_i(\vartheta,s)\}^2/2\sigma^2]} \prod_{i=k+1}^{N} \{1-\mathrm{erf}[\tilde\eta_i(\vartheta',s')/\sigma]\} \Bigg) \Bigg\} \Bigg\rangle_{\varepsilon,\vartheta^0}.
$$

$$(63)$$

Performing the sums on k and taking the limit of small σ, it is easy to show that all terms simplify or vanish with higher orders in σ, except for the second and third term in the sum on k. The third term can be evaluated in a mean field approximation passing the quenched average under the logarithm: since N is very large, k is also large for most values. The leading asymptotic term finally reads

$$
I(\{\eta_i\},\vartheta\otimes s) \simeq \log_2\left(\frac{p\sqrt{\pi}}{\sqrt{2}} \frac{\sqrt{N}\eta^0}{\sigma} \sqrt{\frac{\langle\varepsilon^2\rangle_\varepsilon}{2}} \right) + \frac{1}{2} \sum_s \int d\vartheta\, p(\vartheta,s)
$$

$$
\times \sum_{k=1}^{N-1} \binom{N}{k} \langle 1-\mathrm{erf}[\tilde\eta_i(\vartheta,s)/\sigma] \rangle_{\varepsilon,\vartheta^0}^{N-k} \log_2\left(\frac{p\sqrt{\pi}}{\sqrt{2}} \frac{\sqrt{k}\eta^0}{\sigma} \sqrt{\frac{\langle\varepsilon^2\rangle_\varepsilon}{2}} \right).
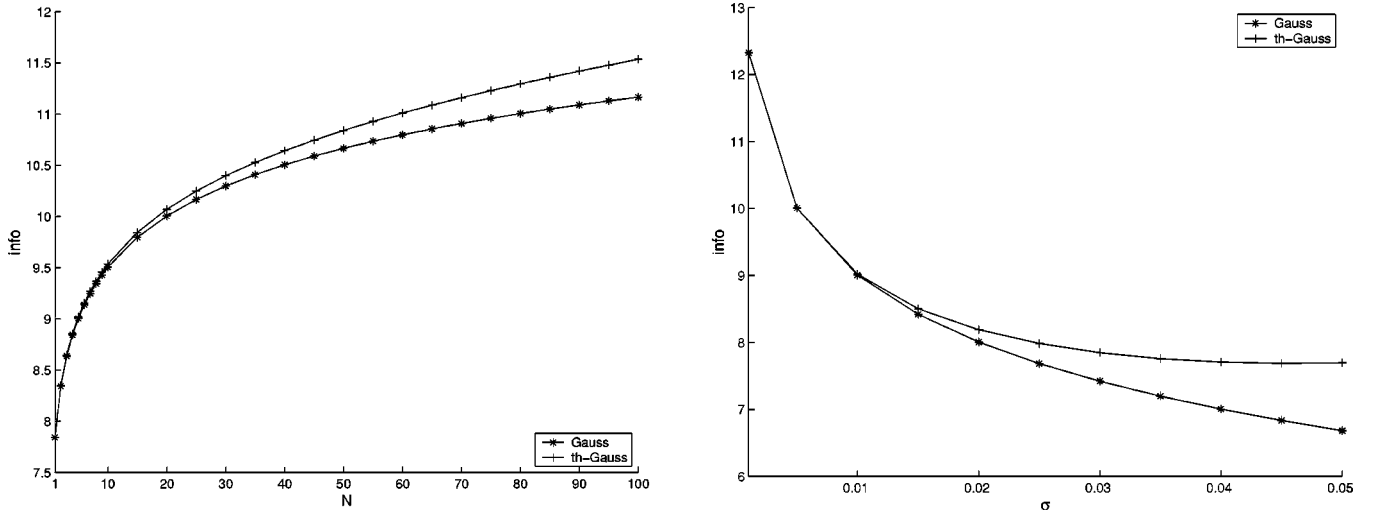$$

$$(64)$$

FIG. 3. Information for the Gaussian vs thresholded-Gaussian model with mixed continuous and discrete stimuli, according to Eqs. (61) and (64). $\eta^0 = 1$, $p = 4$, $\alpha = 0.2$, ε_s^i can take values 3/10, 6/10, and 9/10 with equal probability. Left: asymptotic behavior as a function of the population size N for $\sigma = 0.01$. Right: as a function of the noise σ for a population size of $N = 5$ cells.

Since the first term is exactly the information for the Gaussian model, Eq. (61), we see that already at the leading term, the rise to infinity with N is slightly higher for the thresholded Gaussian model. It is quite difficult to extract analytically the exact dependence on N and σ from Eq. (64). We have evaluated the sum on k numerically by means of a MATLAB code.

Figure 3(a) shows the mutual information for both models as a function of the population size. It must be said that while Eq. (61) is valid for generic values of the parameters, provided the noise is small, more restrictive assumptions underlie the derivation of Eq. (64). In particular, one has to exclude the values for which the tuning curve $\tilde{\eta}_i(\vartheta, s)$ is identically zero [namely, $\alpha = 0$ and $\varepsilon_s^i = 1$ with a finite weight in Eq. (19)]; the reason is that for very small values of the noise, the weight of the δ peak in distribution (31) becomes proportional to $\delta(\tilde{\eta}_i(\vartheta, s))$, and several terms in Eq. (63) cannot be neglected any more.

Moreover, for any fixed value of N there is an upper bound on the value of the noise, beyond which approximation (61) must be integrated by the neglected terms. This can be seen by direct analytical evaluation of each term.

As an example, we show in Fig. 3(b) this effect for a population of five cells: for values of σ close to 0.05, the decrease in the information for the thresholded Gaussian model starts slowing down, and for higher values of the noise, the information would even increase.

IV. DISCUSSION

We have presented a detailed analysis of the mutual information carried by one population of independent units about a set of stimuli, examining both the case where the stimuli are purely discrete and the case where they are characterized by an additional continuous angular dimension. In fact, even though in real experiments performed on trained animals the stimuli always vary in a discrete set, since even continuous dimensions are sampled on a finite number of points, in nature the real brain must learn how to discriminate highly dimensional stimuli or behavioral correlates, whose dimensions may equally be continuous or discrete. For our specific model, we have been inspired by the data recorded in the motor cortex of monkeys performing arm movements which might be parameterized according to a (continuous) direction and to a (discrete) "type" [15,16], We have focused on two possible limits, namely, the limit of finite population size and large noise, which corresponds to the initial information rise, and the asymptotic regime of large numbers of neurons and small noise. The limit of large noise has been recently studied in Ref. [7] by means of the replica trick. Here, we have shown that regardless of the structure of the stimulus whether continuous or discrete, the same results can be obtained without resorting to the replica trick, by a mere expansion of the logarithm. Moreover, we have shown that correlations introduced in the preferred orientations of each neuron, across different values of the discrete parameter, can increase the redundancy, depressing the information. This issue is biologically relevant since in several cortical areas neurons show a tuning for the direction and, in particular, in the dataset analyzed in Ref. [16] such correlations were indeed observed.

Modifying accordingly the theoretical model did improve the fit of real information curves, suggesting that the correlations detected in the data are information bearing—here, in a negative sense, they depress the information content.

We have been able to study analytically the asymptotic approach to the upper information bound with purely discrete stimuli, always without replicas, calculating both the leading term of the order of σ^N and the correction of the order of σ^{2N}. We have shown how to retrieve our results

using a different approach [9], always replica-free. This approach does allows to go beyond our original approximation. Yet, in the presence of additional quenched disorder as in the case of our specific model, further assumptions are necessary to proceed with the analytical evaluation of the correction. A careful analysis is still in progress and will be presented elsewhere [17].

Finally, we have evaluated the asymptotic information value in the presence of mixed continuous and discrete stimuli, which grows to infinity when the noise goes to zero and/or the number of neurons becomes large. We have found that the information grows to infinity logarithmically with N and with the inverse of the noise for the Gaussian model, while the exact dependence for the thresholded-Gaussian model is more difficult to detect. Under certain conditions and for very low values of the noise, the asymptotic information is higher for the thresholded-Gaussian model than for the pure Gaussian.

This result is quite interesting *per se*, but we refrain from any speculation leaving its interpretation to a more careful evaluation of the information. In particular, it will be interesting to evaluate the impact of the corrections neglected in Eq. (63) [17].

It must be said that the validity of our results is not checked here by means of numerical simulations. In a previous work [8], we have shown and discussed in detail that, since information is a very sensitive measure to limited sampling, in our specific case of a continuous rate model with additional quenched parameters, the numerical evaluation results extremely hard, especially in the most interesting limit of large population sizes. The simulations presented in Ref. [8] could be carried out using a decoding procedure which is meant to reduce the bias. Even so, the agreement with the analytical results was found only for a population size of maximum two cells, the curve deviating due to the distortion caused by decoding for larger population sizes. We are currently working in order to improve the numerical techniques and obtain a better check of our analytical results for most of the parameter space. Nonetheless, we think that the widely established difficulty in getting accurate numerical information estimates with models of this type makes our analytical efforts and the results presented here even more remarkable.

APPENDIX A: THE SMALL σ LIMIT IN PRESENCE OF PURELY DISCRETE STIMULI

Let us reconsider Eq. (45). Inserting expansion (46), one obtains

$$I(\{\eta_i\},s) \simeq \log_2 p - \frac{1}{\ln 2} \left\langle \sum_s p(s) \right.$$

$$\left. \times \left[\sum_{s' \neq s} \prod_i e^{-(\eta_i^s - \eta_i^{s'})^2/2\sigma^2} \right] \right\rangle_\eta$$

$$\simeq \log_2 p - \frac{1}{\ln 2} \left(\sum_s p(s) \sum_{k=1}^{\infty} \sum_{s_1 \neq s} \cdots \sum_{s_k \neq s} \right.$$

$$\left. \times \left\langle \exp\left[-\sum_{l=1}^{k} (\eta^s - \eta^{s_l})^2/2\sigma^2 \right] \right\rangle_\eta^N \right), \quad \text{(A1)}$$

where we have used the fact that quenched disorder is uncorrelated and identically distributed across neurons and stimuli.

When σ becomes very small one can use the following approximation:

$$e^{-(\eta^s - \eta^{s_1})^2/2\sigma^2} \rightarrow \sqrt{2\pi\sigma^2}\,\delta(\eta^s - \eta^{s_1}). \quad \text{(A2)}$$

Let us reconsider the term with $k=1$ in Eq. (A1):

$$\sum_{s_1 \neq s} \langle e^{-(\eta^s - \eta^{s_1})^2/2\sigma^2} \rangle_\eta^N$$

$$\rightarrow \sum_{s_1 \neq s} (\sqrt{2\pi}\sigma)^N \left[\int d\eta^s d\eta^{s_1} \varrho(s)\varrho(s_1)\delta(\eta^s - \eta^{s_1}) \right]^N$$

$$= (p-1)(\sqrt{2\pi}\sigma)^N \left[\int d\eta \varrho^2(\eta) \right]^N. \quad \text{(A3)}$$

Therefore, this term gives a contribution of the order of σ^N.

It is easy to check that at each order k, the term with $s_1 = s_2, \ldots, = s_k \neq s$ gives a contribution of the order of σ^N: in fact, while for a generic choice of $s_1, s_2, \ldots, s_k$ out of the p correlates one has finally several δ functions each of which, according to Eq. (A2), carries a factor of σ, when all the stimuli are equal only one δ function remains and the result is of the order of σ^N. Therefore, one has to sum all the contributions to calculate the exact coefficient determining the asymptotic approach of the information to the upper bound. For a generic order k in the expansion of the logarithm, one has

$$\frac{(-1)^{k+1}}{k} \sum_{s_1 = s_2, \ldots, = s_k \neq s} \left\langle \exp\left[-\sum_{l=1}^{k} (\eta^s - \eta^{s_l})^2/2\sigma^2 \right] \right\rangle_\eta^N$$

$$\rightarrow \frac{(-1)^{k+1}}{k} \sum_{s_1 \neq s} \langle e^{-k(\eta^s - \eta^{s_1})^2/2\sigma^2} \rangle_\eta^N \simeq \frac{(-1)^{k+1}}{k^{N/2+1}} (p-1)$$

$$\times (\sqrt{2\pi}\sigma)^N \left[\int d\eta \varrho^2(\eta) \right]^N. \quad \text{(A4)}$$

It is now clear that to obtain the contribution at all orders k, one must sum the series

$$S_1 = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^{N/2+1}}, \tag{A5}$$

and the final contribution to the mutual information up to order σ^N can be expressed as follows:

$$(p-1)S_1(\sqrt{2\pi}\sigma I_1)^N, \quad I_1 = \int d\eta \varrho^2(\eta). \tag{A6}$$

Inserting Eq. (A6) in Eq. (A1), one obtains the expression of the asymptotic approach to the ceiling up to order σ^N:

$$I(\{\eta_i\},s) \simeq \log_2 p \left[1 - \frac{p-1}{\log_2 p \ln 2} [S_1(\sqrt{2\pi}\sigma I_1)^N] \right]. \tag{A7}$$

This result can be easily extended to higher powers of σ. The first correction of the order of σ^{2N} to the leading term can be calculated through the same technique. Since at the kth order in the expansion of the logarithm the factor σ^N was obtained considering only the configuration where all the stimuli $s_1, \ldots, s_k$ are equal, it is clear that each configuration where all the stimuli except one are equal will generate a factor of σ^{2N}; in fact, if, say, l stimuli are different from one another among k, one will have to introduce a δ function for each of the l exponentials, according to Eq. (A2). Let us see in detail the kth order contribution, assuming, for example, that all stimuli are equal except s_1:

$$\frac{(-1)^{k+1}}{k} \sum_{s_1 \neq s} \sum_{s_2 = s_3, \ldots, = s_k \neq s, s_1}$$

$$\times \left\langle \exp\left[-\sum_{l=1}^{k} (\eta^s - \eta^{s_l})^2 / 2\sigma^2 \right] \right\rangle_\eta^N$$

$$\rightarrow \frac{(-1)^{k+1}}{k} \sum_{s_1 \neq s} \sum_{s_2 \neq s_1, s} \langle e^{-(\eta^s - \eta^{s_1})^2 / 2\sigma^2}$$

$$\times e^{-(k-1)(\eta^s - \eta^{s_2})^2 / 2\sigma^2} \rangle_\eta^N. \tag{A8}$$

A δ function is introduced according to Eq. (A2) for each of the two exponentials in Eq. (A8); since one has k possible choices for the stimulus which is different from the other $k-1$, the final result must be multiplied times a factor of k more; finally, one has

$$\frac{(-1)^{k+1}}{(k-1)^{N/2}} (p-1)(p-2)(2\pi\sigma^2)^N \left[\int d\eta \varrho^3(\eta) \right]^N. \tag{A9}$$

Summing all terms at any order k, the total contribution order σ^{2N} to the mutual information can be written as follows:

$$(p-1)(p-2)S_2(2\pi\sigma^2 I_2)^N, \quad S_2 = \sum_{k=1}^{\infty} \frac{(-1)^k}{k^{N/2}},$$

$$I_2 = \int d\eta \varrho^3(\eta). \tag{A10}$$

Inserting this result in Eq. (A7), one obtains the final expression for the mutual information up to order σ^{2N}, Eq. (47).

APPENDIX B: LARGE σ LIMIT FOR THE THRESHOLDED-GAUSSIAN MODEL

Let us reconsider Eqs. (42). Developing all the products one can expand the logarithm in powers of $1/\sigma$:

$$\log_2 \left[\sum_{s'} p(s') \int d\vartheta' p(\vartheta') \left(1 + \sum_{i=1}^{k} \frac{\eta_i \tilde{\eta}_i(\vartheta', s')}{\sigma^2} - \sum_{i=1}^{k} \frac{\tilde{\eta}_i^2(\vartheta', s')}{2\sigma^2} + \frac{1}{2} \sum_{i,j=1}^{k} \frac{\eta_i \eta_j \tilde{\eta}_i(\vartheta', s') \tilde{\eta}_j(\vartheta', s')}{\sigma^4} \right) \right.$$

$$\left. \times \frac{1}{2^{N-k}} \left(1 - \sum_{i=k+1}^{N} \sqrt{\frac{2}{\pi}} \frac{\tilde{\eta}_i(\vartheta', s')}{\sigma} + \frac{1}{\pi} \sum_{i \neq j} \frac{\tilde{\eta}_i(\vartheta', s') \tilde{\eta}_j(\vartheta', s')}{\sigma^2} \right) \right]$$

$$\simeq -(N-k) + \frac{1}{\ln 2} \ln \left[1 + \sum_{s'} \int d\vartheta' p(\vartheta', s') \left(\sum_{i=1}^{k} \frac{\eta_i \tilde{\eta}_i(\vartheta', s')}{\sigma^2} - \sum_{i=1}^{k} \frac{\tilde{\eta}_i^2(\vartheta', s')}{2\sigma^2} - \sum_{i=k+1}^{N} \sqrt{\frac{2}{\pi}} \frac{\tilde{\eta}_i(\vartheta', s')}{\sigma} \right. \right.$$

$$\left. \left. + \frac{1}{2} \sum_{i,j=1}^{k} \frac{\eta_i \eta_j \tilde{\eta}_i(\vartheta', s') \tilde{\eta}_j(\vartheta', s')}{\sigma^4} + \frac{1}{\pi} \sum_{i \neq j} \frac{\tilde{\eta}_i(\vartheta', s') \tilde{\eta}_j(\vartheta', s')}{\sigma^2} - \sum_{i=1}^{k} \sum_{j=k+1}^{N} \sqrt{\frac{2}{\pi}} \frac{\eta_i \tilde{\eta}_i(\vartheta', s') \tilde{\eta}_j(\vartheta', s')}{\sigma^3} \right) \right]$$

$$\simeq -(N-k) + \frac{1}{\ln 2} \left[\sum_{s'} \int d\vartheta' p(\vartheta', s') \left(\sum_{i=1}^{k} \frac{\eta_i \tilde{\eta}_i(\vartheta', s')}{\sigma^2} - \sum_{i=1}^{k} \frac{\tilde{\eta}_i^2(\vartheta', s')}{2\sigma^2} - \sum_{i=k+1}^{N} \sqrt{\frac{2}{\pi}} \frac{\tilde{\eta}_i(\vartheta', s')}{\sigma} \right. \right.$$

$$\left. \left. + \frac{1}{2} \sum_{i,j=1}^{k} \frac{\eta_i \eta_j \tilde{\eta}_i(\vartheta', s') \tilde{\eta}_j(\vartheta', s')}{\sigma^4} + \frac{1}{\pi} \sum_{i \neq j} \frac{\tilde{\eta}_i(\vartheta', s') \tilde{\eta}_j(\vartheta', s')}{\sigma^2} - \sum_{i=1}^{k} \sum_{j=k+1}^{N} \sqrt{\frac{2}{\pi}} \frac{\eta_i \tilde{\eta}_i(\vartheta', s') \tilde{\eta}_j(\vartheta', s')}{\sigma^3} \right) \right.$$

$$
-\frac{1}{2}\sum_{s',s''}\int d\vartheta'd\vartheta''p(\vartheta',s')p(\vartheta'',s'')\left(\sum_{i,j=1}^{k}\frac{\eta_i\eta_j\tilde{\eta}_i(\vartheta',s')\tilde{\eta}_j(\vartheta'',s'')}{\sigma^4}+\frac{2}{\pi}\sum_{i\neq j}\frac{\tilde{\eta}_i(\vartheta',s')\tilde{\eta}_j(\vartheta'',s'')}{\sigma^2}\right.
$$

$$
\left.-2\sum_{i=1}^{k}\sum_{j=k+1}^{N}\sqrt{\frac{2}{\pi}}\frac{\eta_i\tilde{\eta}_i(\vartheta',s')\tilde{\eta}_j(\vartheta'',s'')}{\sigma^3}\right)\Bigg],
\tag{B1}
$$

where we have kept only the terms which will give a contribution of the order of $1/\sigma^2$ to the information.

This expression has to be inserted in Eq. (40) and integrated on $\{\eta_i\}$. One obtains

$$
-\left\langle\left(\sum_{s}\int d\vartheta p(\vartheta,s)\sum_{k=0}^{N}\binom{N}{k}\int_0^\infty\prod_{i=1}^{k}d\eta_i\prod_{i=1}^{k}\frac{1}{\sqrt{2\pi\sigma^2}}\exp-[\{\eta_i-\tilde{\eta}_i(\vartheta,s)\}^2/2\sigma^2]\right.\right.
$$

$$
\times\prod_{i=k+1}^{N}\{1-\mathrm{erf}[\tilde{\eta}_i(\vartheta,s)/\sigma]\}\log_2\left[\sum_{s'}p(s')\int d\vartheta'p(\vartheta')\left(\prod_{i=1}^{k}\exp[2\eta_i\tilde{\eta}_i(\vartheta',s')-\tilde{\eta}_i^2(\vartheta',s')/2\sigma^2]\right.\right.
$$

$$
\left.\left.\left.\times\prod_{i=k+1}^{N}\{1-\mathrm{erf}[\tilde{\eta}_i(\vartheta',s')/\sigma]\}\right)\right]\right)\right\rangle_{\varepsilon,\vartheta^0}
$$

$$
\simeq\sum_{s,s',s''}\int d\vartheta d\vartheta'd\vartheta''p(s,\vartheta)p(s',\vartheta')p(s'',\vartheta'')\left[\sum_{k=0}^{N-1}\binom{N}{k}(N-k)\langle\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{k}\right.
$$

$$
\times\langle1-\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k}-\sum_{k=1}^{N}\binom{N}{k}\frac{k}{\ln 2}\langle\tilde{\eta}(\vartheta,s)\tilde{\eta}(\vartheta',s')/\sigma^2\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}\langle\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{k-1}\langle1
$$

$$
-\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k}-\sum_{k=1}^{N}\binom{N}{k}\frac{k}{\ln 2}\langle\tilde{\eta}(\vartheta',s')/\sqrt{2\pi}\sigma e^{-\tilde{\eta}^2(\vartheta,s)/2\sigma^2}\rangle_{\varepsilon,\vartheta^0}\langle\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{k-1}\langle1-\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k}
$$

$$
+\sum_{k=0}^{N-1}\binom{N}{k}\frac{N-k}{\ln 2}\sqrt{\frac{2}{\pi}}\left\langle\frac{\tilde{\eta}(\vartheta',s')}{\sigma}[1-\mathrm{erf}\{\tilde{\eta}(\vartheta,s)/\sigma\}]\right\rangle_{\varepsilon,\vartheta^0}\langle\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{k}\langle1-\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k-1}
$$

$$
+\sum_{k=1}^{N}\binom{N}{k}\frac{k}{2\ln 2}\left\langle\frac{\tilde{\eta}(\vartheta',s')\tilde{\eta}(\vartheta'',s'')}{\sigma^2}\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\right\rangle_{\varepsilon,\vartheta^0}\langle\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{k-1}\langle1-\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k}
$$

$$
+\sum_{k=0}^{N-1}\binom{N}{k}\frac{N-k}{\pi\ln 2}\langle\tilde{\eta}(\vartheta',s')\tilde{\eta}(\vartheta'',s'')/\sigma^2[1-\mathrm{erf}\{\tilde{\eta}(\vartheta,s)/\sigma\}]\rangle_{\varepsilon,\vartheta^0}\langle\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{k}\langle1-\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\rangle_{\varepsilon,\vartheta^0}^{N-k-1}\Bigg],
\tag{B2}
$$

where we have assumed that quenched disorder is uncorrelated across neurons.

It must be noticed that the terms of the order of N^2 appearing in the expansion of the logarithm, Eq. (B1) do not appear in Eq. (B2) any more: it can be easily shown that after averaging across the stimuli they cancel out.

The sums on k in Eq. (B2) can be performed and the result is then inserted in the expression for the mutual information, Eq. (40), which finally can be rewritten as follows:

$$
I(\{\eta_i\},\vartheta\otimes s)=\sum_{s}\int d\vartheta p(\vartheta,s)\left[\frac{N}{\ln 2}\langle[1-\mathrm{erf}\{\tilde{\eta}(\vartheta,s)/\sigma\}]\ln[1-\mathrm{erf}\{\tilde{\eta}(\vartheta,s)/\sigma\}]\rangle_{\varepsilon,\vartheta^0}\right.
$$

$$
+\frac{N}{2\ln 2}\left\langle\frac{\tilde{\eta}^2(\vartheta,s)}{\sigma^2}\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\right\rangle_{\varepsilon,\vartheta^0}+N\langle[1-\mathrm{erf}\{\tilde{\eta}(\vartheta,s)/\sigma\}]\rangle_{\varepsilon,\vartheta^0}\Bigg]
$$

$$
-\sum_{s,s',s''}\int d\vartheta d\vartheta'd\vartheta''p(s,\vartheta)p(s',\vartheta')p(s'',\vartheta'')\left[\frac{N}{\ln 2}\left\langle\frac{\tilde{\eta}(\vartheta,s)\tilde{\eta}(\vartheta',s')}{\sigma^2}\mathrm{erf}[\tilde{\eta}(\vartheta,s)/\sigma]\right\rangle_{\varepsilon,\vartheta^0}\right.
$$

$$-\frac{N}{2\ln 2}\left\langle \frac{\widetilde{\eta}(\vartheta',s')\,\widetilde{\eta}(\vartheta'',s'')}{\sigma^2}\,\mathrm{erf}[\,\widetilde{\eta}(\vartheta,s)/\sigma]\right\rangle_{\varepsilon,\vartheta^0} -\frac{N}{\ln 2}\sqrt{\frac{2}{\pi}}\left\langle \frac{\widetilde{\eta}(\vartheta',s')}{\sigma}[\,1-\mathrm{erf}\{\widetilde{\eta}(\vartheta,s)/\sigma\}]\right\rangle_{\varepsilon,\vartheta^0}$$

$$-\frac{N}{\pi\ln 2}\left\langle \frac{\widetilde{\eta}(\vartheta',s')\,\widetilde{\eta}(\vartheta'',s'')}{\sigma^2}[\,1-\mathrm{erf}\{\widetilde{\eta}(\vartheta,s)/\sigma\}]\right\rangle_{\varepsilon,\vartheta^0}\Bigg].\tag{B3}$$

Using expansion (41) for the error function and keeping only the terms up to order N/σ^2, one arrives at Eq. (43).

---

[1] C.E. Shannon, AT&T Tech. J. **27**, 379 (1948).

[2] T.M. Cover and J.A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).

[3] F. Rieke, D. Warland, R.R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code* (MIT Press, Cambridge, MA, 1996).

[4] E.T. Rolls and A. Treves, *Neural Networks and Brain Function* (Oxford University Press, Oxford, 1998).

[5] Alessandro Treves, in *Neuro-Informatics and Neural Modelling*, edited by F. Moss and S. Gielen, Handbook of Biological Physics, (Elsevier, Amsterdam, 2000), pp. 803–829.

[6] Inés Samengo and Alessandro Treves, Phys. Rev. E **63**, 011910 (2001).

[7] V. Del Prete and A. Treves, Phys. Rev. E **64**, 021912 (2001).

[8] V. Del Prete and A. Treves, Phys. Rev. E **65**, 041918 (2002).

[9] K. Kang and H. Sompolinsky, Phys. Rev. Lett. **86**, 4958 (2001).

[10] H. Sompolinsky, H. Yoon, K. Kang, and M. Shamir, Phys. Rev. E **64**, 051904 (2001).

[11] A. Treves and E.T. Rolls, Hippocampus **2**, 189 (1992).

[12] A. Treves, J. Comput. Neurosci. **2**, 259 (1995).

[13] S. Schultz and A. Treves, Phys. Rev. E **57**, 3302 (1998).

[14] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).

[15] O. Donchin, A. Gribova, O. Steinberg, H. Bergman, and E. Vaadia, Nature (London) **395**, 274 (1998).

[16] V. Del Prete, O. Steinberg, A. Treves, and Eilon Vaadia, Neurocomputing **38–49**, 1181 (2001).

[17] Valeria Del Prete (unpublished).