

Alexei A. Koulakov · Dmitry A. Rinberg
Dmitry N. Tsigankov

How to find decision makers in neural networks

Received: 16 July 2004 / Accepted: 9 September 2005 / Published online: 5 November 2005
© Springer-Verlag 2005

Abstract Nervous systems often face the problem of classifying stimuli and making decisions based on these classifications. The neurons involved in these tasks can be characterized as sensory or motor, according to their correlation with sensory stimulus or motor response. In this study we define a third class of neurons responsible for making perceptual decisions. Our mathematical formalism enables the weighting of neuronal units according to their contribution to decision making, thus narrowing the field for more detailed studies of underlying mechanisms. We develop two definitions of a contribution to decision making. The first definition states that decision making activity can be found at the points of emergence for behavioral correlations in the system. The second definition involves the study of propagation of noise in the network. The latter definition is shown to be equivalent to the first one in the cases when they can be compared. Our results suggest a new approach to analyzing decision making networks.

1 Introduction

The nervous system is confronted by megabytes of information representing light, sound, smell, and other sensory inputs. This information is compiled by the brain into a set of decisions, representing the behaviors of living organisms. The mechanisms involved in this reduction have been under investigation for many years (Parker and Newsome 1998; Glimcher 2003; Romo and Salinas 2003). In this study we address a question complementary to the issue of decision making (DM) mechanisms; we develop a method for defining the neuronal units involved in making decisions. Such

analysis has practical significance, since once units involved in making a particular decision are located, further efforts could focus on uncovering their underlying mechanisms. For the purpose of this analysis we determine DM activity in surrogate networks, defined mathematically, in which complete control is present over stimuli, mechanisms, and responses. The main goal of our study is to formulate a rigorous definition of DM for simple surrogate systems, which opens the way to tackle more realistic problems in the future.

In this study, a DM task is defined as evaluation of a function in the multidimensional stimulus space (Fig. 1a). This function has a discrete set of values, representing the repertoire of responses available to the organism. The decisions may, of course, be stochastic, to reflect the uncertainty intrinsic to behavior. This definition is suitable for experiments where subjects perform poly-alternative forced-choice tasks, such as saccadic response to the direction of stimulus motion (Shadlen and Newsome 2001).

Let us consider the direction-discrimination task in more detail. Figure 1b lists some visual areas involved in this task. The areas are arranged along a rough sensory–motor axis so that the areas on the left are more “sensory”, while those on the right are more “motor”. This implies that the responses in these areas are more correlated with stimulus or response, respectively. Where on this sensory–motor axis should one position DM elements? One could argue that the elements most correlated with the decision itself are the decision makers, following the analogy with the definition of sensory and motor elements. It is, however, difficult, if not impossible, to distinguish such a definition from the definition of purely motor units (Shadlen and Newsome 2001). The latter relay the results of the decision making process without involvement in the formation of the decision. Thus, an approach is needed to distinguish pure motor from DM activity. This is one of the purposes of our investigation.

In this study we develop an approach in which DM components are positioned on the interface between sensory and motor areas. More precisely, the *first* element in the sensory–motor chain, which bears a significant correlation with the response, is identified as the decision maker. We develop

A. A. Koulakov (✉) · D. N. Tsigankov
Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
E-mail: akula@cshl.edu
Tel.: +516-3678470
Fax: +516-3678389

D. A. Rinberg
Monell Chemical Senses Center, Philadelphia, PA 19104, USA

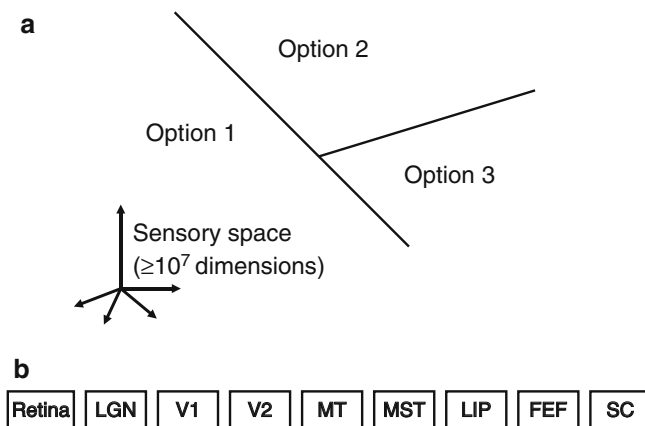


Fig. 1 **a** Definition of decision making task. The nervous system evaluates a function, whose values represent discrete decisions, in the multidimensional sensory space. **b** Some of the visual areas involved in a motion-discrimination task. The areas on the *left* are more sensory (response is correlated with the sensory input), while those on the *right* are more motor (correlated with the response)

this idea into a rigorous mathematical formulation and find a special correlation function that determines the contributions of units to DM. Such formalism allows us to approach two problems pertaining to the identities of DM units. First, we consider the case when not one but *several* elements are simultaneously involved in one decision. Our approach allows us to evaluate the relative importance of various units in such distributed DM. Second, we consider systems with loops in connectivity. For such systems the concept of ‘the first element’ becomes ill-posed. We succeed in defining command elements in such networks by considering the impact of variability in an element’s response on the behavioral outcome of the system.

2 Stochastic versus deterministic decision making

Studies of DM have identified two large classes of decisions: deterministic and stochastic (Glimcher 2003). The former category includes responses which are reliably triggered by specific stimuli. In Fig. 1a these responses correspond to regions far from the interface between two decisions. Stochastic decisions, on the other hand, are produced by stochastic sources within the nervous system. For such decisions, no deterministic connection can be established between stimuli and responses. If the stimuli and experimental conditions can be reproduced exactly, decisions of this type vary significantly from trial to trial. Such behaviors are fundamentally unpredictable and are often considered voluntary (Glimcher 2003). In terms of the decision function illustrated in Fig. 1a, stochastic decisions correspond to the regions close to the interface between domains of the same decision.

The distinction between deterministic and stochastic cases is important for the problem of identifying the decision makers. In the deterministic case, the responses are causally linked to sources outside the system. Thus, Parker and Newsome (1998) suggest that “neural responses and psychophysical

decisions will covary trivially if the stimulus strength changes from trial to trial, possibly due to fluctuations in extraneous factors.” More informative is the stochastic case, in which decisions can be linked to the internal causes. Thus, the same report suggests that “the more revealing case occurs when internal neural noise accounts for observed psychophysical variability upon successive presentations of externally identical stimuli. This sort of covariation cannot be driven by the environmental stimulus, but must arise from causal relationship within the nervous system” (Parker and Newsome 1998).

In the present study, following the ideas mentioned above, we study stochastic DM only. Our analysis will therefore concentrate on the gray zone surrounding the interface between single-decision domains in Fig. 1a. We give a precise definition of our system below. Henceforth, the term decision denotes stochastic decisions only, unless otherwise noted.

Stochastic decisions naturally arise when the strength of the stimulus is not sufficient to render deterministic judgment. In the case of the direction of a motion discrimination task (reviewed in Parker and Newsome 1998) this condition corresponds to the zero-coherence case. Generically speaking, this regime is defined by a zero or small signal-to-noise ratio (SNR) case, where ‘signal’ represents the sensory information and ‘noise’ is the variability due to intrinsic with respect to the nervous system sources. Below we use this regime for our simple system to study the identities of decision makers.

Stochastic decisions also arise in the context of game theory, when optimal behaviors are given by so-called mixed strategies (Glimcher 2003). Let us illustrate the concept of a mixed strategy with the example of C-start escape behavior in a teleost fish (Eaton and Emberley 1991). This behavior serves as the short-latency evasion of predators and, in open water conditions, is directed away from danger. Thus, in open water, a pure strategy dominates, whereby fish choose to escape at a fixed angle from the predator. A different picture emerges when the escape route is obstructed by a wall. In this case, the fish has the option of selecting one of the escape angles as a pure strategy. Instead, the fish uses an escape direction which is not predictable from the stimulus angle (Eaton and Emberley 1991). Thus, the fish chooses to mix responses from trial to trial, presumably to confuse the predator, implementing a mixed strategy. Another more familiar example is the rock–paper–scissors game (Smith 1982). The pure strategies in this case are given by three individual responses. However, the optimal strategy is to mix all three responses with equal probability. This mix should be truly random, because a slightly biased decision can be detected by an opponent and used to make predictions, ultimately leading to that opponent’s advantage. Other examples of mixed strategies are reviewed by Glimcher (2003), including more sophisticated escape behaviors.

We suggest here that the zero/small-coherence case is similar to the mixed strategy case in many respects. It could be argued that the signal in the mixed strategy case is the utility of a single pure strategy. Noise is the internal variability used

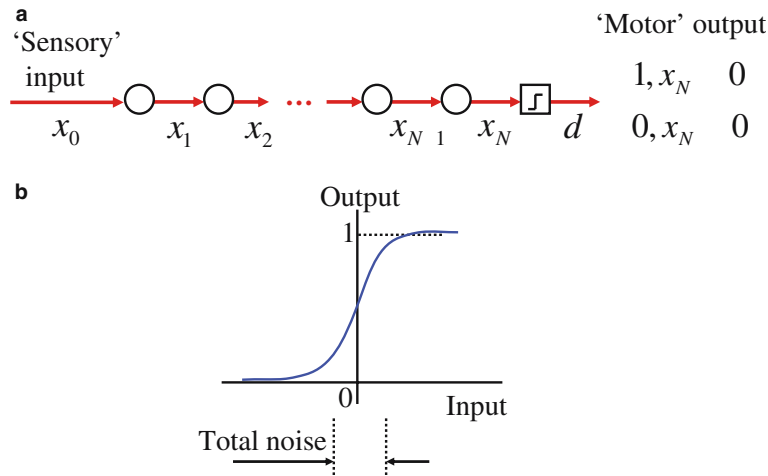


Fig. 2 **a** A simple ‘nematode’ network consists of a chain of units. All units in the chain are linear, except the last. The last unit, shown by the *square*, is non-linear and returns zero or one depending on the sign of the response of the preceding unit. **b** The average input–output relationship (psychometric function) for the ‘nematode’ chain, is given by the sigmoid function. The spread of the sigmoid is determined by the net noise in the chain

to generate random behaviors. Rigorously speaking, different mixed strategies can be implemented by our system, defined below, by tuning the SNR to generate the required probabilities of responses. Thus, our study may be useful for defining identities of decision makers in both the zero/small-coherence case and the mixed strategy case.

The small-coherence case has been extensively used before to explore neuronal correlates of decision (Celebrini and Newsome 1994; Britten et al. 1996; Shadlen and Newsome 1996; Thiele and Hoffmann 1996). It was ultimately argued that decision making units implement the synthesis of sensory and motor correlations, and can therefore be found on the interface between sensory and motor units (Shadlen and Newsome 1996; Newsome 1997; Salinas and Romo 1998; Shadlen and Newsome 2001). Here we examine a different paradigm for identifying decision makers. We argue that DM may be found at the points where motor activity emerges. DM activity is, therefore, related to the rate of emergence of the motor correlations.

3 Linear chains and trees

3.1 The ‘nematode’ network

In this subsection we consider a network, which we call ‘nematode’. We first define the model, then we show that it can make simple decisions, and finally we define the positions of decision makers in the network.

Consider a linear chain of units, whose response is characterized by a set of real numbers x_i where $i = 1, \dots, N$ is the position of the element in the chain (Fig. 2). The response of each element does not depend on time. This assumption is introduced here to simplify the analysis and can be relaxed as described below (Sect. 3.2). Each unit performs a simple linear transformation between the unit’s input and the output.

Thus, for element number i

$$x_i = x_{i-1} + \eta_i \quad (1)$$

Here η_i is noise associated with the element. In this work we assume that noise has zero mean, is individual to each unit, and, therefore, is uncorrelated between units, i.e.

$$\overline{\eta_i} = 0, \quad \overline{\eta_i \eta_j} = \overline{\eta_i^2} \delta_{ij} \quad (2)$$

Here δ_{ij} is the Kronecker delta. We further assume that noise has a Gaussian distribution. The chain of linear elements is thus completely specified by a set of noise variances $\overline{\eta_i^2}$. The model described by (1) and (2) yields the following solution for the response of the last element in the chain

$$x_N = x_0 + \eta_1 + \eta_2 + \dots + \eta_{N-1} + \eta_N. \quad (3)$$

Thus, the response of the N th element is just a sum of the input into network x_0 and noise contributions from all units, independently on the order of unit in the chain. The last element in the chain has non-linear response properties. Its response is defined by

$$d = H(x_N), \quad (4)$$

where $H(x)$ is the Heaviside stepfunction, which is equal to one/zero if the argument is positive/negative.

It follows then that our ‘nematode’ network is capable of making decisions based on the values of input variable x_0 . Thus, we interpret variable d , which is equal to either 0 or 1, as the result of DM process, as defined in Fig. 1a. The decisions are made stochastically and depend upon the instantiations of random variables η_i , which vary from trial to trial.

Our model is completely defined by the set of noise variances, pertinent to each unit $\overline{\eta_i^2}$. Our next goal is to develop a sensible definition of contributions to DM based on the

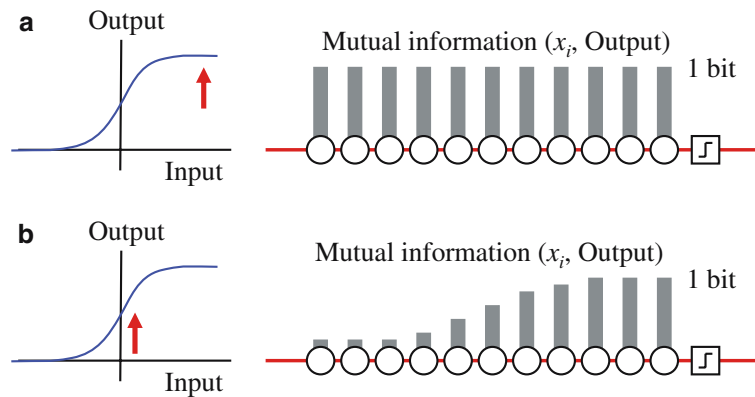


Fig. 3 **a** If the signal-to-noise ratio (SNR) is high, the responses of all units are well correlated with the output, as shown on the *right* by the mutual information between the response of a given unit and the output. **b** In the case of a low SNR, the output is better correlated with the motor units (*right*) than with the sensory ones (*left*)

vector of variances $\overline{\eta_i^2}$. Before doing so, we describe general input–output properties of the chain.

Since the decision made by the network varies from trial to trial, one can define the response of the system averaged over trials as $\overline{d(x_0)}$. As shown in Fig. 2b, it has a sigmoid shape, smeared by the total amount of noise in the system. One can, therefore, consider two cases, depending on whether the SNR ratio for the chain is large or small. These two regimes correspond to the small and large coherence cases in the motion-discrimination task (Parker and Newsome 1998).

To analyze responses of units in these two cases, we define their correlation with the decision. This correlation is defined for each element in the chain (Fig. 3, right). We initially choose mutual information (MI) as a measure of correlation between response of the i th unit, x_i , and the decision, d . MI has the advantage of being unitless (it is measured in bits) and having clear intuitive properties, as described below.

Mutual information describes the information transmitted from the i th unit to the output of the system d . Since the output variable can only have values 0 or 1, MI cannot exceed one bit. We now consider two cases, depending on the network’s SNR. If network input $|x_0|$ is large compared to the noise level, as in Fig. 3a, the response of the system is well correlated with the input. Hence, activities of all units are well correlated with both input and output, and $MI(x_i, d) \approx 1$ for all of the units. At the opposite limit of the low SNR, $|x_0|$ is smaller than the noise level, and the system’s response is weakly correlated with the input (Fig. 3b). In this case, MI as a function of a unit’s position displays a structure, shown in Fig. 3b (right). This structure, as shown below, has a key to the definition of DM components.

Let us consider the low SNR case in more detail. The units, which are close to the exit from the network, show a strong correlation with the decision. These units are analogous to the motor units. Their MI is therefore close to 1 bit. On the other hand, more ‘sensory’ units in the beginning of the chain are strongly correlated with the input. Since the input–output correlation is weak in the low SNR case, the

‘sensory’ units display virtually *no* relation to the output and $MI(x_i, d) \approx 0$ for input units (Fig. 3b, right). Thus, MI, as a function of i , displays a transition from 0 to 1 in the low SNR case.

One could suggest that the elements perfectly correlated with the output of the system, such as exit elements from the chain, are the ones that make the decision. However, such elements may be argued to be the relay or ‘motor’ units, in which case their contribution to the formation of a decision is small. Thus, despite their high correlation with the output, exit elements cannot be called decision makers. Input elements, having no correlation with the decision, are responsible for DM to an even lesser degree. We thus need to analyze the dependence of MI on position in more detail and suggest another scheme for defining DM units.

Our elimination of motor units as possible decision makers can be further extended to the entire high SNR case (Fig. 3a). In high SNR case all units have perfect correlation with the response, therefore behaving similar to pure motor units. Thus, following Parker and Newsome (1998) we suggest that the deterministic regime is not descriptive from the point of view of DM analysis. Since all units in this case have similar correlations, it is hard, if not impossible, to distinguish them and assign differential contributions. It could be argued that in the deterministic case, the external environment acts as the decision maker by providing the input variable x_0 . We conclude, following Parker and Newsome (1998), that to find decision making activity one has to concentrate on the low SNR case. We note however that there is no binary quantitative distinction between high and low SNR cases, i.e., deterministic and stochastic DM. The low SNR case is used here for convenience of the analysis.

We show below that the identities of decision making units can be deduced from the shape of the transition in Fig. 3b (right). To this end, we analyze a set of examples of networks with various distributions of noise $\overline{\eta_i^2}$. We start from the simplest example of a single noisy unit.

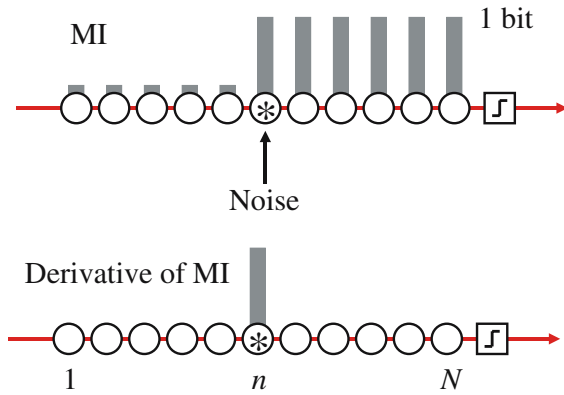


Fig. 4 The example of a ‘noisy’ neuron (marked by asterisk). *Top panel*, mutual information between given unit and the decision. *Bottom panel*, the derivative of mutual information. The derivative represents the decision making unit in this case

3.2 Example 1. ‘Noisy’ neuron

Consider a chain in which noise is absent from all units but one, whose order number in the chain is n (Fig. 4). Since, according to our previous discussion in the previous section, we need to consider the low SNR case, we will assume that

$$x_0 = 0, \quad (5)$$

i.e., network receives no input. Making the decision in this case is still possible, based on the values of noise inside the network. Since noise is only present in one neuron, from (3) we conclude that

$$x_N = \eta_n. \quad (6)$$

The decision made by the network is

$$d = H(\eta_n). \quad (7)$$

Thus, the output of the circuit is causally linked to the processes occurring in unit number n . We conclude that unit n is the decision maker.

Paradoxically, the noisiest unit in this simple formulation makes the largest impact. All noiseless elements, even non-linear, are deterministic, and work as simple relays, which transmit information along the chain. We address this apparent paradox in the Discussion section. Here we mention that our definition follows from the earlier suggestions (Parker and Newsome 1998) that “a given neuron may or may not yield a spike (or may yield a variable number of spikes) in response to repeated presentations of a particular stimulus. . . . in the extreme where a single neuron is used for the psychophysical judgment, there should be a 100% correlation between fluctuations in the firing of the neuron and the behavioral choices of the observer.”

One could suggest that the non-linear element is actually the decision maker in this case. We argue that the non-linear element does not have a causal effect on output from the circuit; therefore its role is just to relay response from neuron n to the output. In this respect the non-linear element is not different from other noiseless elements.

To link this example to our previous discussion (Fig. 3b) we plot MI as a function of position in the chain in Fig. 4 (top). As we discussed, MI is high for exit (‘motor’) units and low for input (‘sensory’) elements. Figure 4 also shows the derivative of MI with respect to a position in the chain. This derivative represents the decision making element. Thus, not a correlation with the decision but rather the rate of change of the latter along the network is the indicator of DM. Therefore, contribution to decision making may be characterized by the rate of emergence of the motor response along the chain.

3.3 Example 2: uniformly distributed noise

Our next example shows that the conclusion about the derivative of MI is qualitatively correct, but has to be amended to be numerically precise. Consider the chain in which all elements are noisy and the variance of noise is the same for each element. In this case (3) and (4) yield

$$d = H(\eta_1 + \eta_2 + \dots + \eta_N). \quad (8)$$

Equation 8 does not distinguish the order in which contributions from the units are added, and all contributions are of equal strength on average. Thus, we conclude that all units contribute to the decision *equally*.

Figure 5a shows MI as a function of position in the chain for this case (see Appendix A for derivation). It increases smoothly from 0 to 1 resulting in a non-zero derivative for all units. The derivative of MI turns out to be non-uniform, as seen in Fig. 5a. However, Eq. 8 suggests that all units participate in decision equally. The non-uniform increase of MI can be corrected if a non-linear function of MI, denoted $F(\text{MI})$, is considered rather than MI itself. This non-linear function is calculated in Appendix A and is shown in Fig. 5b. The new correlator $F(\text{MI})$ has the same basic properties as the MI. It rises from 0 to 1 monotonically when passing through the array (Fig. 5c). Its derivative is uniform as shown in Fig. 5c (bottom). This is consistent with equal participation in DM among all units in the uniformly distributed noise case. Thus, the motor response of the i th element is more adequately characterized by function $F(\text{MI}_i)$, while contributions to DM, which we denote as D_i , are given by the rate of emergence of motor response $F(\text{MI}_i)$ when moving through the array

$$D_i = F(\text{MI}_i) - F(\text{MI}_{i-1}). \quad (9)$$

Equation 9 is the main result of this paper. *Decision index* D_i represents contribution of the i th element to the decision for networks of simple connectivity, such as chains.

Equation 9 implies that, from point of view of DM contributions, not mutual information, but another correlator, given by $F(\text{MI})$, is more accurate. Equation 9 also reproduces the result obtained in the previous example of ‘noisy’ neuron. Indeed, the mutual information rises from 0 to 1 on the ‘noisy’ neuron in Fig. 4. $F(\text{MI})$ coincides with MI at these values, as follows from its plot in Fig. 5b. Thus, the derivative of $F(\text{MI})$ is also given by a single spike at the position of ‘noisy’ neuron, as in Fig. 4 (bottom). Therefore, both examples considered so

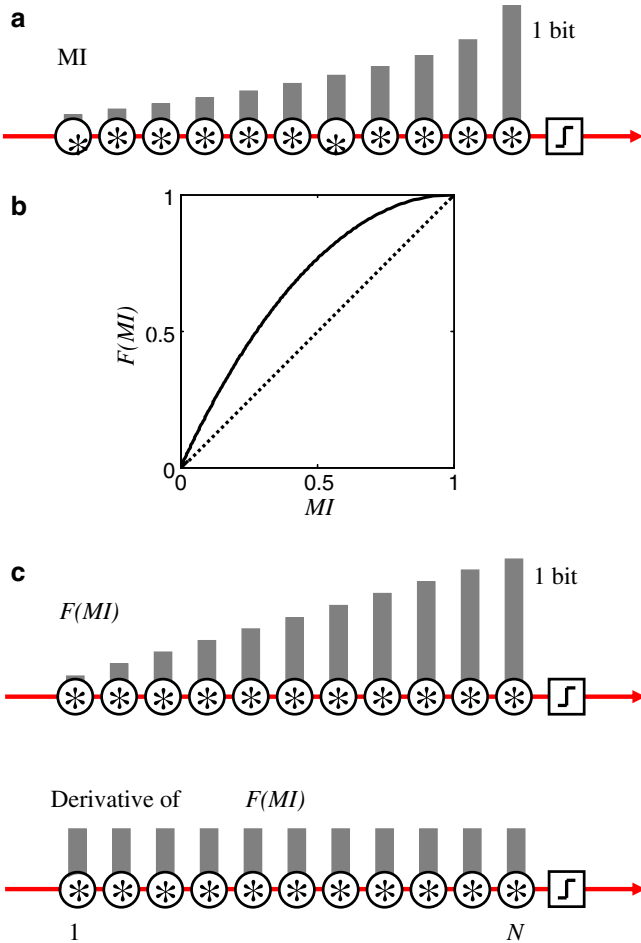


Fig. 5 The example with uniformly distributed noise. **a** Mutual information between response of a given unit and the decision. The dependence has a non-uniform increase, suggesting that mutual information is not a good measure of decision making. **b** If one applies a non-linear function (*solid curve*) to the mutual information in **a** one obtains a uniformly increasing correlator in **c**. This non-linear function, called $F(MI)$, is close to linear, shown by the *dotted line*. **c** the new correlator $F(MI)$ (it top panel) has a uniform derivative (*bottom panel*). Thus, the derivative of $F(MI)$ is a sensible measure of decision making in the case of uniform noise

far are adequately described by Eq. 9. In the next subsection we argue that this equation can handle more complex cases with uneven strength of contributions.

3.4 Example 3: the ‘loud’ neuron

In this example, we modify the initial chain model (1) by adding a strong amplification link between units 5 and 6 ($K \gg 1$), keeping the noise variances for all neurons the same as in the previous case:

$$x_6 = Kx_5 + \eta_6 \tag{10}$$

The network output in this case is given by

$$d = H(x_N) = H(K[\eta_1 + \eta_2 + \dots + \eta_5] + \eta_6 + \dots + \eta_{11}) \tag{11}$$

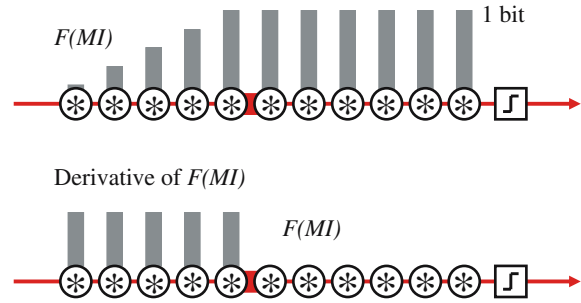


Fig. 6 The ‘loud’ neuron example. The link between units 5 and 6 is strengthened. Compare to Fig. 5a

The units are divided into two groups: before the critical link and after. Units in each group contribute equally to the decision. The impact of the neurons from the first group (units 1 through 5) is multiplied by a large number ($K \gg 1$). These neurons contribute much stronger to the decision than the units from the second group.

Figure 6 shows both $F(MI)$ and its derivative along the chain (calculated in Appendix A). Similarly to the previous cases, the derivative of $F(MI)$ is the correct indicator of units contribution to decision: the first five units have the same large DM index, while the last five units have a small contribution when $K \gg 1$.

Changing one link in the chain produces large effect on the distribution of DM. The units downstream from the link contribute little to decisions, while the units upstream contribute more. The ratio of contributions of upstream and downstream links should not depend on the sign of the link. This is because the distributions of noise variables η_i are even functions, and, therefore, changing the sign of K to the opposite in (11) should not change the decision index, which implies that even for inhibitory link ($K < 0$) the upstream units contribute more when $|K| \gg 1$. Thus, to describe the ratio of contributions from the upstream and downstream units an even function of K needed, which is shown in Appendix A to be K^2 .

3.5 Trees

Our studies indicate that information-theoretical analysis (Eq. 9) can be further extended to tree-like topologies. The latter are defined as connectivities with no loops in them, each network element having only one outgoing link (Fig. 7).

For each element of the network we can calculate the difference between incoming and outgoing flux of information about the decision [$F(MI)$], which characterizes the impact of the given unit to the decision, similarly to the case of chains. The absence of loops guarantees the independence of incoming signals and, thus, allows the summation of incoming information across multiple links.

To generalize Eq. 9 we define a column-vector \vec{f} such that $f_i = F(MI_i)$. The decision index in this case is derived in Appendix B and given by

$$\vec{D} = (\hat{I} - \hat{S})\vec{f}. \tag{12}$$

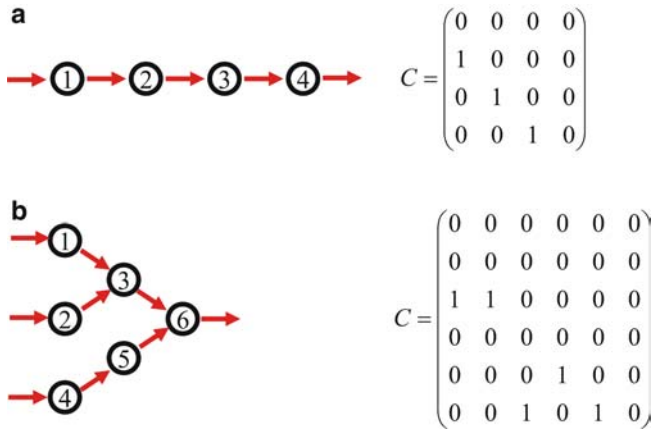


Fig. 7 The mutual information approach can be extended to connectivities other than linear chains (a). Thus, decision makers in the case of trees (b) can also be found. Arbitrary networks can be specified by connection matrices, which are provided for illustration purposes. The non-zero entries in a connection matrix indicate a connection between two elements numbered on the left. An entry value describes the strength of the connection and does not have to be unitary or positive

Here \hat{I} is the unity matrix and \hat{S} is the structure matrix defined as follows. An element S_{ij} of the structure matrix is equal to 1 if there is a connection from unit number j to i . Matrix $\hat{I} - \hat{S}$ thus implements the evaluation of differences between outgoing and incoming information, similar to (9). The structure matrix is related to the connection matrix, containing network's weights through $S_{ij} = |\text{sign}(C_{ij})|$. The connection matrices for some networks are shown in Fig. 7. Note that the connection matrices employed in this study are equal to transposed weighted adjacency matrix used in graph theory.

Information-theoretical approach can be further extended to the cases wherein signals propagate on the tree in time resulting thus in delays between units' responses and the decision. In this case the modified mutual information $F(\text{MI})$ is different from zero only at certain moments preceding the decision. For example, for unit 2 in Fig. 7b the $F(\text{MI})$ is non-zero only three time steps before the decision, because at other times the response of this unit and the decision are not correlated. To compensate for the presence of such delays one should understand \vec{f} as a sum of correlations over all times preceding the decision (Appendix B). Correlator \vec{f} then can be used in Eq. 12. This approach can only be applied in tree topologies.

Equation 12 [or (9)] cannot be used for topologies other than trees because the signal coming to the same unit through different links cannot be considered independent. For such connectivities a dynamic approach introduced below should be used to identify decision makers.

3.6 Alternative definition of DM

So far we have used definition (9), which involves the calculation of a function $F(\text{MI})$. It is, however, possible to reproduce the results shown above in another way. This alternative

approach also allows generalization of our definition to the case of arbitrary network topologies.

To establish alternative definition we prove in Appendix A the following proposition: The decision index defined in Eq. 9 for a network element is proportional to the product of the variance of noise on this element and its contribution to the variability of the output variable x_N^2 , i.e.,

$$D_i \equiv F(\text{MI}_i) - F(\text{MI}_{i-1}) = \frac{\overline{\eta_i^2}}{Z} \frac{\partial x_N^2}{\partial \eta_i^2}. \quad (13)$$

Here Z is the normalization constant introduced to ensure that $\sum_i D_i = 1$.

Let us illustrate the new definition using the examples considered above. In the 'noisy' neuron case, from Eq. 6 we obtain

$$\overline{x_N^2} = \overline{\eta_n^2}, \quad (14)$$

where n is the position of 'noisy' neuron. Hence, according to Eq. 13,

$$D_i = \begin{cases} 1, & i = n, \\ 0, & i \neq n. \end{cases} \quad (15)$$

In the case of uniform noise, Eq. 8 yields

$$\overline{x_N^2} = \overline{\eta_1^2} + \overline{\eta_2^2} + \dots + \overline{\eta_n^2}. \quad (16)$$

The contributions to DM as given by Eq. 13 are the same for all units, i.e., $D_i = 1/N$. This is in accord with definition (9).

Let us illustrate how Eq. 13 works in the 'loud' neuron case. From (11) we obtain

$$\overline{x_N^2} = K^2(\overline{\eta_1^2} + \dots + \overline{\eta_5^2}) + \overline{\eta_6^2} + \dots + \overline{\eta_{11}^2}. \quad (17)$$

By applying Eq. 13 we obtain

$$D_{1..5} = \frac{\overline{\eta_{1..5}^2}}{Z} K^2, \quad D_{6..11} = \frac{\overline{\eta_{6..11}^2}}{Z}. \quad (18)$$

Thus, each element's contribution to DM is proportional to the variance of noise on a given element, multiplied by the square of the gain from this element to the output. In the case that $K \gg 1$, contributions from elements, 1, ..., 5 are much larger than from elements 6, ..., 11, as shown in Fig. 6.

Finally, we give another definition of DM contributions, which could be useful when noise in the system is the same for all units. Such quantity depends only on network connectivity and on the position of the unit in the network. We therefore call this quantity *topological decision index (tD)*.

$$tD_i = \frac{1}{Z} \frac{\partial \sigma^2(x_N)}{\partial \eta_i^2} \quad (19)$$

Here by $\sigma^2(x_N)$ we denote the variance of x_N . If $\sigma^2(x_N)$ is a linear function of the set of noise variances η_i^2 , as in the case of linear networks considered here, topological DM does not depend on the strength of noise. This quantity is thus useful in describing the network's connectivity and shows how strongly each element of the circuit affects the output.

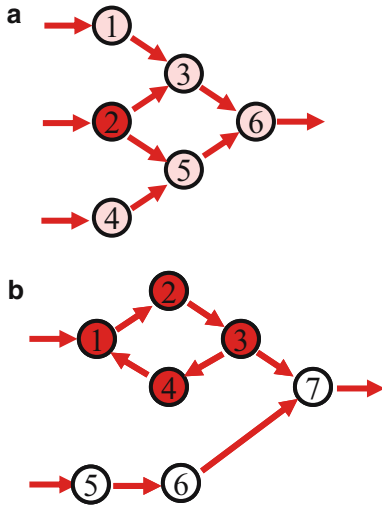


Fig. 8 Two cases in which the identities of decision makers can be found using discrete-time approach. The variance of noise on all elements is the same; all network links have unitary strength. The degree of decision making is shown by the *intensity of red*. (a) The fan-out effect. (b) The temporal integrator

4 Dynamic models

To apply our analysis to systems with loops, and, in general, to networks with arbitrary connectivity (Fig. 8), here we consider time-dependent models. This allows us to observe the propagation of noise around loops explicitly and to make accurate conclusions about contributions to DM.

We limit ourselves to linear dynamical systems, where the single nonlinear element is the last one, transforming an analog system output to a binary response. We first consider temporal dynamics in a discrete-time approximation, which contains all essential features of our approach. We then extend discrete model to the continuous-time case.

4.1 Discrete-time model

We consider a system of N elements, whose activity at each instant is described by an N dimensional column-vector $\vec{x}(t)$. Time has discrete values separated by an interval τ . The values of the activity at two moments are related by the connection matrix \hat{C}

$$\vec{x}(t + \tau) = \hat{C}\vec{x}(t) + \vec{\eta}(t) + \vec{s}(t) \quad (20)$$

Here $\vec{\eta}(t)$ is the vector describing internal noise. The variable $\vec{s}(t)$ represents sensory input. The rules of the temporal evolution of activities described by this equation are sufficient to capture many interesting phenomena and to mimic modeling of real systems on digital computers.

We assume that noise $\vec{\eta}(t)$ has a zero mean and is uncorrelated in time:

$$\overline{\eta_i(t_1)\eta_j(t_2)} = N_{ij}\delta_{t_1,t_2} \quad (21)$$

This assumption can be easily relaxed and is used here to simplify the analysis. It becomes rigorously valid when the

time-interval τ is longer than the correlation time of noise. If noise is specific to each neuron, the same-time correlation matrix \hat{N} is diagonal

$$N_{ij} = \overline{\eta_i^2}\delta_{ij} \quad (22)$$

This takes place, i.e., when stochasticity is induced by probabilistic nature of synaptic vesicle release, in which case every two neurons receive uncorrelated fluctuating inputs.

At some moment t the system is forced to make a decision through the following process. First, a scalar quantity

$$y = \vec{v}^T \cdot \vec{x}(t) \quad (23)$$

is evaluated. The output vector \vec{v} describes the way in which the system's activity affects motor response. In the simplest case, which was considered in the previous section, a single element number n evokes responses, $v_i = \delta_{in}$. In a more complex situation, when multiple areas/neurons have direct influence on decision, vector \vec{v} has more than one non-zero element. In the second step, a decision is made based on the sign of y

$$d = H(y) \quad (24)$$

Thus, this model describes a two-alternative forced-choice task.

Our system is completely defined by the following set of parameters: \hat{C} , $\vec{s}(t)$, \hat{N} , and \vec{v} . As we have shown in the previous section, the presence of the stimulus is not required to define DM elements (Eq. 5). We therefore set $\vec{s}(t)$ to zero and are left with three parameters \hat{C} , \hat{N} , and \vec{v} . We now determine DM elements in this model.

Following definition (13) the decision index becomes:

$$D_i = \frac{N_{ii}}{Z} \frac{\partial \sigma^2(y)}{\partial N_{ii}} \quad (25)$$

Therefore, we need to evaluate the variability on the output from the system $\sigma^2(y)$. To this end we notice that $y = \vec{v}^T \cdot \vec{x}(t) = \vec{x}^T(t) \cdot \vec{v}$ and

$$\sigma^2(y) = \vec{v}^T \cdot \overline{\vec{x}(t)\vec{x}^T(t)} \cdot \vec{v} = \vec{v}^T \hat{X}(t, t) \vec{v} \quad (26)$$

Here we introduced the cross-correlation matrix defined as follows:

$$\hat{X}(n, k) \equiv \overline{\vec{x}(n)\vec{x}^T(k)} = \begin{pmatrix} x_1(n) \\ \vdots \\ x_N(n) \end{pmatrix} (x_1(k) \cdots x_N(k)) \quad (27)$$

We replace here the time variable with the integers, specifying the time-slice number. The averaging in (26) and (27) is assumed over different instantiations of noise (trials).

As follows from (26), we are particularly interested in the same-time correlator $\hat{X}_0(n) \equiv \hat{X}(n, n)$, which determines fluctuations in y . Using (20) we obtain

$$\begin{aligned} \hat{X}_0(n+1) &= \overline{\vec{x}(n+1)\vec{x}^T(n+1)} \\ &= \overline{[\hat{C}\vec{x}(n) + \vec{\eta}(n)] [\vec{x}^T(n)\hat{C}^T + \vec{\eta}^T(n)]} \quad (28) \end{aligned}$$

We then notice that the correlator $\overline{\vec{x}(n)\vec{\eta}^T(n)}$ is identically zero, since $\vec{x}(n)$ is a linear combination of values of noise at times $k < n$ (Eq. 20). Thus

$$\hat{X}_0(n+1) = \hat{C}\hat{X}_0(n)\hat{C}^T - \hat{N} \quad (29)$$

This is a linear equation on $\hat{X}_0(n)$, which has solution in the form of a series

$$\hat{X}_0(n) = \sum_{k=0}^n \hat{C}^k \hat{N} (\hat{C}^T)^k. \quad (30)$$

This solution allows us to determine the same-time correlator \hat{X}_0 from connectivity and noise cross-correlogram, defined in (21).

Finding decision makers can be simplified, if correlator $\hat{X}_0(n)$ does not depend on time. The condition for this can be found from Eq. 30. We obtain

$$\hat{R}(n) \equiv \hat{X}_0(n+1) - \hat{X}_0(n) = \hat{C}^n \hat{N} (\hat{C}^T)^n \quad (31)$$

The difference matrix $\hat{R}(n)$ goes to zero if all of the eigenvalues of the connectivity matrix \hat{C} are less than 1 in absolute value, in which case $\hat{C}^n \rightarrow 0$ when $n \rightarrow \infty$. This can be seen by reducing \hat{C} to diagonal or Jordan normal form. If the latter condition is true, the same-time correlator $\hat{X}_0(n)$ can be replaced by a constant matrix \hat{X}_0 in (29), which leads to

$$\hat{X}_0 - \hat{C}\hat{X}_0\hat{C}^T = \hat{N} \quad (32)$$

All of the examples analyzed below will be based on this equation.

Let us examine the properties of this equation. In the most generic case, Eq. 32 allows us to determine \hat{X}_0 from \hat{C} and \hat{N} uniquely. Indeed, Eq. 32 is a system of N^2 linear equations for N^2 unknown entries in \hat{X}_0 . Hence, this system, in most cases, can be solved uniquely. However, with few exceptions, \hat{X}_0 cannot be expressed explicitly in terms of matrices \hat{C} and \hat{N} .

The contribution to DM from a given element can be determined from Eq. 26

$$D_i = \frac{N_{ii}}{Z} \frac{\partial \sigma^2(y)}{\partial N_{ii}} = \frac{1}{Z} \vec{v}^T \frac{\partial \hat{X}_0}{\partial \ln N_{ii}} \vec{v} \quad (33)$$

The topological decision index is

$$tD_i = \frac{1}{Z} \vec{v}^T \frac{\partial \hat{X}_0}{\partial N_{ii}} \vec{v} \quad (34)$$

Using Eqs. 32 and 33 one can analyze a variety of network connectivities. Some new effects emerging for non-tree systems are described next.

4.2 Example 4: fan-out hub

Let us consider network shown in Fig. 8a, in which all elements have the same variance of noise and all connections have unitary strength. Figure 8a shows two pathways from

unit 2 to the exit unit 6. The resulting network gain from unit 2 to unit 6 is thus equal to two. All other units' gains at the exit are equal to one. The contribution to DM from unit 2 is thus four times larger than from other units. This is because noise at this unit is multiplied by a factor of two and the variance of noise by a factor of four (Eq. 18). We conclude that there may be some special elements in the network which occupy hub-like positions gaining large influence due to the abundance of their outputs. These elements are somewhat similar to the 'loud' neuron in the example above.

4.3 Example 5: temporal integrator

Figure 8b shows the example of a network with a loop in connectivity. All links in this network have unitary strength and noise is the same for all units. The presence of a loop affects DM drastically: Our discrete-time model marks units belonging to the loop as decision makers.¹ This is easy to understand, since noise generated by each unit on each time-step, cannot leave the loop and therefore builds up there without limits. This case is somewhat analogous to the 'noisy' neuron example. Of course, the growth of variance of noise is stopped by non-linearities pertinent to a real system.

What is the possible role of loops in biological networks? Loops, similar to the one shown in Fig. 8b, have many useful properties. For instance, they can act as parametric memory systems. Imagine that responses of all units in the loop have the same values, equal to x . Assume that no other inputs are received from the outside of the system. It follows that in the absence of noise on all elements and unit gain on all connections in the loop, the value of the response will reverberate around the loop forever. Such loops can thus memorize a graded parameter, such as x , functioning as parametric memory elements.

Suppose, in addition, that a non-zero constant input s is applied to element number 1. Since this element acts as a summator, its response on the next step is $x_1(1) = x + s$. The signal s propagates around the loop and in four steps it reaches the first element again, at which time its response is $x_1(5) = x + 2s$. In four more steps $x_1(9) = x + 3s$. Thus, not only noise, but also signal can be accumulated by the system. Therefore, a loop can operate as a temporal integrator. The integration is not perfect if one of the links has a non-unitary strength, in which case the integrator becomes leaky (Robinson 1989). Temporal integrators play a special role in DM since they act as accumulators of sensory evidence (Shadlen and Newsome 2001; Gold and Shadlen 2002; Roitman and Shadlen 2002; Mazurek et al. 2003).

4.4 Continuous-time model

We finally consider a model in which time runs continuously. This model has potential relevance to real biological net-

¹ The set of Eqs. 32 and 33 does not have a valid solution for the loop with unit connections. One has to set one of the connections as a parameter, $\alpha < 1$, solve the equations, and consider the limit $\alpha \rightarrow 1$

works. The responses of units satisfy the following equation

$$\frac{d\vec{x}(t)}{dt} = -\hat{A}\vec{x}(t) + \vec{\eta}(t) + \vec{s}(t). \quad (35)$$

The network weight matrix \hat{A} is determined by network connectivity. It can be related to connection matrix from the discrete-time model in (20) through $\hat{C} = e^{-\hat{A}\tau}$ (see Appendix D). We assume that all of the eigenvalues of the weight matrix \hat{A} have positive real parts, i.e., the system (35) has no instabilities. Noise is defined by its cross-correlation

$$\overline{\eta_i(t_1)\eta_j(t_2)} = N_{ij}\delta(t_1 - t_2), \quad (36)$$

where

$$N_{ij} = \overline{\eta_i^2}\delta_{ij} \quad (37)$$

is a diagonal cross-correlogram of noise. Similarly, we define the output scalar and the decision variable

$$\begin{aligned} y &= \vec{v}^T \cdot \vec{x}(t) \\ d &= H(y). \end{aligned} \quad (38)$$

Here t is time when the system makes the decision.

Our model is thus completely defined by Eqs. 35–38. We will now use Eq. 33 to find decision makers. As in the discrete-time case we need to know the variance of the output variable, $\sigma^2(y)$, which is related to the same-time correlation matrix \hat{X}_0 by Eq. 33. The equation for \hat{X}_0 is derived in Appendix C and has similar properties to Eq. 29 describing the discrete case

$$\hat{A}\hat{X}_0 + \hat{X}_0\hat{A}^T = \hat{N} \quad (39)$$

This equation is the central tool for the continuous-time theory. It is known in optimal control theory as algebraic Riccati equation (Brogan 1991). The contribution to DM from each unit (decision index) is found by differentiating $\sigma^2(y) = \vec{v}^T \hat{X}_0 \vec{v}$ with respect to noise, similarly to Eq. 33

$$D_i = \frac{1}{Z} \frac{\partial \sigma^2(y)}{\partial \ln N_{ii}} = \frac{1}{Z} \vec{v}^T \frac{\partial \hat{X}_0}{\partial \ln N_{ii}} \vec{v} \quad (40)$$

The cross-correlation for arbitrary time is found in Appendix C to be

$$\hat{X}(t_1, t_2) = \begin{cases} e^{\hat{A}(t_2-t_1)} \hat{X}_0, & t_1 \geq t_2 \\ \hat{X}_0 e^{\hat{A}^T(t_1-t_2)}, & t_1 < t_2 \end{cases} \quad (41)$$

This equation suggests a useful strategy for determining noise matrix \hat{N} . Indeed, (39) and (41) imply that

$$\hat{N} = \left. \frac{\partial \hat{X}(t_1, t_2)}{\partial t_1} \right|_{t_1=t_2-\varepsilon} - \left. \frac{\partial \hat{X}(t_1, t_2)}{\partial t_1} \right|_{t_1=t_2+\varepsilon} \quad (42)$$

Here ε is an infinitesimally small positive number. In other words, the noise matrix is equal to discontinuity in the time-derivative of cross-correlation at $t_1 = t_2$. Since in our case the noise correlation matrix is diagonal, the non-zero elements are

$$\overline{\eta_i^2} = \left. \frac{\partial X_{ii}(t_1, t_2)}{\partial t_1} \right|_{t_1=t_2-\varepsilon} - \left. \frac{\partial X_{ii}(t_1, t_2)}{\partial t_1} \right|_{t_1=t_2+\varepsilon} \quad (43)$$

Equations 42 and 43 may be useful if noise variances cannot be measured directly. These equations provide a way to evaluate them from the time-dependent cross-correlation.

Equations 39, 40, and 43 represent a set of tools which allows to find DM components for various connectivities. We present here two possible cases in which decision makers can be found. They differ in what is known about the system.

Scenario 1. Assume we know the network connectivity \hat{A} , the output metrics vector \vec{v} , and the autocorrelation for each unit $X_{ii}(t_1, t_2)$. The steps below allow us to find the decision makers.

1. Since the noise matrix is diagonal, as per (37), it can be found via autocorrelation using (43).
2. Solving (39) allows the determination of $\partial \hat{X}_0 / \partial N_{ii}$, the derivative of equal-time cross-correlation with respect to noise in each element.
3. Decision makers are found from (40).

Scenario 1 does not require simultaneous measurements from all units. However, it does require knowledge of the network connectivity. The next scenario is complementary in this respect.

Scenario 2. Suppose we have measured the full cross-correlation matrix $\hat{X}(t_1, t_2)$ by simultaneous recordings from all units. Suppose also that we know how the output of the system is evaluated (vector \vec{v}). These are the steps to determine DM units.

1. Use (43) [or (42)] to find noise matrix \hat{N} .
2. Use (39) to find the weight matrix \hat{A} .
3. Solve (39) to calculate $\partial \hat{X}_0 / \partial N_{ii}$ for each element.
4. Use (40) to find decision makers.

5 Analysis using electric stimulations

Our previous analyses relied on the knowledge of complete network connectivity (Scenario 1) or cross-correlation between network elements (Scenario 2). The latter approach involves simultaneous recording from all pairs of elements in the network. In this section we study an approach, which neither involves the knowledge of network connectivity nor requires the simultaneous recordings from several units. The idea is to stimulate network elements electrically with randomly changing current to directly implement definition (25). This operation can be performed sequentially for each element of the network in question, and, therefore, involves only single-unit recordings.

In this section we assume that the network receives external input. This assumption is necessary to implement stimulation method, as becomes clear from the discussion below. Thus, the output variable y in (38) is a linear function of the inputs and the noise components, variable from trial to

trial. Since noise in each network unit is Gaussian, the output variable is described by the Gaussian distribution

$$\rho(y) = \exp\left[-(y - \bar{y}(s))^2/2\sigma^2(y)\right] / \sqrt{2\pi\sigma^2(y)} \quad (44)$$

Here $\bar{y}(s)$ is the mean response to stimulus s and $\sigma^2(y)$ is the standard deviation. The response of the system is equal to one if y is positive, and zero otherwise (38). The probability to obtain response equal to one is given by the error function (Abramowitz and Stegun 1972)

$$p_1(s) = \int_0^\infty \rho(y)dy = \left[1 + \operatorname{erf}\left(\bar{y}(s)/\sigma^2(y)\sqrt{2}\right)\right] / 2 \quad (45)$$

The stimulation method described here is based on the direct calculation of the partial derivative $\partial\sigma^2(y)/\partial\eta_i^2$ by injecting a random Gaussian current with variance \bar{s}_i^2 into each cell. The partial derivative is determined by an increase in the variance of the output variable due to added stimulation

$$\partial\sigma^2(y)/\partial\eta_i^2 = \Delta\sigma^2(y)/\bar{s}_i^2. \quad (46)$$

In practice one has no access to the variable y , so one cannot measure directly the change in variance $\Delta\sigma^2(y)$. Instead, one could measure the change in the probability of correct responses under the influence of distracting current. Indeed, assume that for a given stimulus p_1 describes the probability of correct response. The change in this probability due to the stimulation of neuron number i is

$$\Delta p_1(i) = \frac{\partial p_1}{\partial\sigma^2(y)} \Delta\sigma^2(y) \quad (47)$$

Combining the last two equations, for the important derivative we obtain

$$\frac{\partial\sigma^2(y)}{\partial\eta_i^2} = \frac{\Delta p_1(i)}{\bar{s}_i^2} \left(\frac{\partial p_1}{\partial\sigma^2(y)}\right)^{-1} \quad (48)$$

Since the probability of correct responses always decreases under the influence of distracters, the derivative $\partial p_1/\partial\sigma^2(y)$ is a negative constant, independent of the unit, which is stimulated. We arrive therefore to the expression for contributions to DM, which follows from (25)

$$D_i \sim -\bar{\eta}_i^2 \Delta p_1(i) / \bar{s}_i^2 \quad (49)$$

Here $\Delta p_1(i)$ is the decrease in probability of correct responses produced by electric stimulation with variance of the random current equal to \bar{s}_i^2 . The variance of noise on each unit $\bar{\eta}_i^2$ can be found from autocorrelation using (43). This procedure is valid for any topology in our simplified model.

The procedure, which we just described, permits further optimization. Indeed, assume that the probability of correct responses is exactly 1/2, which occurs when there is no external stimulus. Adding distracting stimulation current will not change this probability, i.e., $\Delta p_1(i) = 0$ no matter what unit is stimulated. In the opposite limiting case when $p_1 \simeq 1$, the effect of distracter on performance is exponentially small. Hence, behavioral response to stimulation has an optimum

between $p_1 = 1/2$ and 1. To find the optimum we observe from (47) that $\Delta p_1(i)$ has a maximum for the same $\Delta\sigma^2(y)$ when $\partial p_1/\partial\sigma^2(y)$ is maximum. The maximum of this probability can be calculated from (45) and corresponds to the probability of correct responses

$$p_{\text{optimal}} \simeq 0.841 \quad (50)$$

If the network performs at this level, one should expect the maximum yield from the stimulation experiment. This result is discussed in greater detail in Koulakov et al. (2004) along with other stimulation strategies including those affecting the average response to the stimulus.

To summarize this section we present a possible scenario of obtaining the decision index without the knowledge of network connectivity using electric stimulations.

Scenario 3. Assume that we know autocorrelation function for each unit $X_{ii}(t_1, t_2)$. The steps below allow finding the decision-makers.

1. Prepare stimulus so that the probability of correct responses is close to the value given by (50).
2. Stimulate one unit with random current, whose variance is \bar{s}_i^2 , and measure the decrease in probability of correct responses $\Delta p_1(i)$.
3. Record autocorrelation and evaluate noise variance $\bar{\eta}_i^2$ for this unit using (43).
4. Find decision index for this unit using equation (49).
5. Repeat steps 1 through 4 for all units in the system.
6. Normalize decision index so that $\sum_i D_i = 1$.

6 Discussion

In this work we defined decision makers in networks, which behave in a well-controlled fashion. We argued that DM can be measured by the emergence of behavioral correlations in neural responses. As with any definition, there is more than one way to quantify the properties we were attempting to describe. Therefore we had to make decisions about the features of the model system we choose to describe. We demonstrated these features in a set of examples. Future studies will show whether these features can be used as the basis for a more complete model-independent theory.

6.1 Noise as a source of decisions

In this study we concentrated on stochastic rather than deterministic DM. Therefore, we adopted the postulate that variability and noise, causally linked to decisions, are descriptors of DM (Parker and Newsome 1998). We analyzed DM in the absence of sensory input. Such an analysis was used before to describe DM activity, including the studies of error trials (Celebrini and Newsome 1994; Britten et al. 1996; Shadlen and Newsome 1996; Thiele and Hoffmann 1996). It was previously argued that to describe decision

making activity one has to find traces of both sensory and motor correlations in the same elements (Shadlen and Newsome 1996; Shadlen and Newsome 2001; Newsome 1997; Salinas and Romo 1998), because DM involves transformation of sensory information to behavioral responses. Here we implement a different strategy. We suggest that decisions can be measured by observing the emergence of behavioral correlations alone. Our strategy may include both cases when stochastic decisions are triggered by a weak stimulus and when those decisions do not require external inputs (Glimcher 2003).

Although the association of noise with decisions may seem paradoxical, we suggest two arguments in its favor. First, when external stimulus does not provide sufficient information to render judgment the decisions may have to be made on the basis of internal sources of information. Thus Schall (2001) suggests that “a decision process precedes choices that involve some perplexity, that is, when the alternatives are difficult to distinguish, have uncertain pay-offs or require prior knowledge to resolve them. Thus, in contrast to simple choices, decisions are more effortful, take more time, require attention and deliberation, and are more error prone.” Such an internal variability may also reflect additional information needed to make a decision in case of uncertainty. This information may take the form of inputs from other modalities, memories, or some other relevant inputs, for example, supplying continuously changing utility values (Glimcher 2003). Second, many behaviors, such as C-start escape responses in fish (Eaton and Emberley 1991) and other organisms (Glimcher 2003), are stochastic in character, despite the presence of significant sensory input. This makes the task of the pursuer more difficult. The decisions made under such conditions are produced by intrinsic sources on noise.

6.2 Linear versus nonlinear

Not all of the network elements are linear in our model, since DM is a non-linear task. However, our model is essentially based on linear elements. The motivation for this model is that it is easy to analyze. The study of simple models is a necessary step before analysis can proceed to more complex cases. One of the important questions resolved here is that a completely linear element can be a decision maker, despite the presence of non-linear units in the network.

6.3 Multi-unit correlations

The decision index introduced here describes the contributions of individual cells, thus assuming point-like localized causes for behaviors. This model contrasts with the notion of DM activity’s being synergetic between cells. Alternatively, correlations between two or even more units could be the sources of decision. Our approach can naturally be generalized to include such multi-unit correlations. Indeed, we assumed that the noise crosscorrelation matrix is diagonal, as in Eq. 22. This assumption is introduced to model noise

localized on individual units. If variability is produced by common input to pairs of neurons, the noise matrix is non-diagonal, and the decision index (36) becomes a matrix

$$D_{ij} = \frac{1}{Z} \frac{\partial \sigma^2(y)}{\partial \ln N_{ij}} \quad (51)$$

Alternatively, if the noise structure is more complex and determined by a multi-unit matrix $N_{ij..k}$, the decision index which reflects these correlations is

$$D_{ij..k} = \frac{1}{Z} \frac{\partial \sigma^2(y)}{\partial \ln N_{ij..k}} \quad (52)$$

Similar generalization could be made for the case of non-stationary noise with delayed correlations.

6.4 Possible applications of this analysis

The approach described here could be applied to the following neuronal systems. First, our mutual information-based definition (12) could be used to find decision makers in compartmental models of dendritic trees with no backpropagation (Poirazi and Mel 2001). Second, Scenario 1 may be applied when a wiring diagram of the system is known, such as the nervous system of *Caenorhabditis elegans*, for which the identities of all 302 neurons and most synapses between them have been discovered by the serial-section electron microscopy techniques (White et al. 1986; Durbin 1987; Hall and Russell 1991). Third, some imaging methods based on voltage-sensitive dyes allow monitoring responses of many neurons simultaneously, such as responses of neurons in medicinal leech ganglia (Taylor et al. 2003). Here the Scenario 2 could be useful. Finally, Eq. (13) could be applied operationally by electrically stimulating neurons with stochastic currents and observing changes in the probabilities of responses. In this case no information about connectivity or simultaneous measurement of neuronal activity is needed (Sect. 4).

A decision task, as formulated in Fig. 1a, is similar to a general object discrimination task. The representation of motor response in our model is not distinguishable mathematically from the representation of an abstract object/decision category (Horwitz and Newsome 1998; Shadlen and Newsome 2001). The latter does not necessarily lead to a motor command. Thus, our analysis may uncover the identities of units responsible for categorization of sensory inputs. In terms of this analysis we emphasize the distinction between units representing the object category and the units in which this representation is actually formed. The former are analogous to motor units in the decision task, while the latter are similar to decision makers. As follows from this study, the analysis is dependent upon the topology of the network involved. For simple linear sensory chains our conclusion is that the *first* unit, spatially or temporally, in which the representation of the object is correlated with final outcome of the discrimination process is responsible for casting the stimulus in one of the abstract classes. In case of recurrent networks a

more detailed quantitative analysis is needed to draw conclusions about identities of categorizing units. Thus, our analysis may find broader use in identifying units representing abstract object's percepts.

Our study may be relevant to other biological systems. Possible applications include the analysis of molecular networks, such as protein binding or regulatory genetic networks; studies of structural networks of connectivities between different brain areas; and analysis of social networks.

7 Conclusion

In this study we define network elements responsible for making decisions. We obtain two equivalent definitions. According to one, decisions are made by elements in which correlations with the decision are first formed. According to the second definition, decision making activity is measured by the impact of variability in a given unit on the response. We give examples of network motifs, especially potent from a decision making prospective, such as fan-out hubs and recurrent loops. The latter can function as the temporal integrators of sensory inputs.

Appendix A: The linear chain model

Here we solve a more general version of linear chain model than considered in the text. The responses of neighboring neurons are related linearly

$$x_i = C_{i-1}x_{i-1} + \eta_i \quad (53)$$

This is a generalization of (1). The response of the n th unit is

$$x_n = \sum_{i=1}^n \alpha_{ni} \eta_i + \alpha_{n0} x_0 \quad (54)$$

where coefficients $\alpha_{ni} = C_{n-1}C_{n-2}, \dots, C_i$, $\alpha_{nn} = 1$. The external signal x_0 is assumed to be zero in this appendix, due to (5). For the last element in the chain we have

$$x_N = \sum_{i=1}^N \alpha_{Ni} \eta_i. \quad (55)$$

Comparing (54) and (55) we conclude that

$$x_N = \alpha_{Nn} x_n + \xi, \quad (56)$$

where ξ is a variable, which describes noise in the networks downstream from unit n . It is thus not correlated with x_n . This is the point where the chain topology enters our solution, since in case of loops, x_n and ξ are correlated. Our goal now is to calculate MI between the decision variable $d = H(x_N)$ and x_n . We will use the definition for MI

$$MI(d, x_n) = \sum_{d=0,1} \int_{-\infty}^{\infty} dx_n \rho(d, x_n) \log_2 \left[\frac{\rho(d, x_n)}{\rho(d) \rho(x_n)} \right] \quad (57)$$

Here $\rho(d) = 1/2$, since there is no signal;

$$\rho(x_n) = \exp\left(-x_n^2/2\overline{x_n^2}\right) / \left(2\pi\overline{x_n^2}\right)^{1/2} \quad (58)$$

and

$$\rho(d; x_n) = \frac{\rho(x_n)}{2} \left[1 \pm \operatorname{erf}\left(\frac{\alpha_{Nn} x_n}{\sigma(\xi)\sqrt{2}}\right) \right]. \quad (59)$$

The upper/lower sign is assumed for $d = 0$ or 1 in (59); $\sigma(\xi)$ is the standard deviation of Gaussian variable ξ defined in (56). The expression for MI (57) results in a simpler expression

$$MI_n = M(s_n) \\ M(s_n) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dz e^{-z^2} [1 + \operatorname{erf}(zs_n)] \log_2 [1 + \operatorname{erf}(zs_n)] \quad (60) \\ s_n = \sigma(\alpha_{Nn} x_n) / \sigma(\xi).$$

Here $\operatorname{erf}(x)$ is the error function (Abramowitz and Stegun 1972). MI is therefore a function of the SNR s_n . Inversely,

$$s_n^2 = \frac{\alpha_{Nn}^2 \overline{x_n^2}}{\xi^2} = [M^{-1}(MI_n)]^2 \quad (61)$$

On the other hand, (56) leads to

$$\overline{x_N^2} = \alpha_{Nn}^2 \overline{x_n^2} + \xi^2, \quad (62)$$

since x_n and ξ are statistically independent. Solving (61) and (62) with respect to $\alpha_{Nn}^2 \overline{x_n^2}$ we have

$$\frac{\alpha_{Nn}^2 \overline{x_n^2}}{\overline{x_N^2}} = \frac{[M^{-1}(MI_n)]^2}{1 + [M^{-1}(MI_n)]^2} \equiv F(MI_n) \quad (63)$$

Function M^{-1} here is inverse to M defined in (60). Function $F(MI)$ numerically calculated from (60) and (63) is shown in Fig. 5. It can be effectively fitted with polynomials

$$F(x) = \sum_{n=1}^5 a_n x^n + \varepsilon(x). \quad (64)$$

The use of the set of coefficients $a_{1..5} = 2.17654, -1.4974, 0.51567, -0.24682, \text{ and } 0.05202$ ensures that $|\varepsilon(x)| < 5 \cdot 10^{-5}$.

Lastly, we relate the variances $\alpha_{Nn}^2 \overline{x_n^2}$ to the strength of noise $\overline{\eta_i^2}$ through (54). We have

$$\alpha_{Nn}^2 \overline{x_n^2} = \sum_{i=1}^n \alpha_{Ni}^2 \overline{\eta_i^2} \quad (65)$$

Equations (63) and (65) are used below to prove a variety of statements about function $F(MI)$ used in the main text.

A.1 In the uniform noise example $F(MI)$ is a linear function of a position in the chain

In this case $C_1 = \dots = C_{N-1} = 1$, and, consequently, $\alpha_{N1} = \dots = \alpha_{NN} = 1$. Noise variance is the same on every node, i.e., $\overline{\eta_i^2} \equiv \eta^2$. As follows from (65) that $\overline{x_n^2} = \eta^2 n$, which results in

$$F(MI_n) = n/N \quad (66)$$

It follows that contributions to DM defined by (9) are the same for all units.

A.2 In the ‘loud’ neuron example the contributions of units upstream from the strong link are larger by a factor of K^2 than contribution from the downstream units

In this case $\alpha_{1,\dots,k} = K$, while $\alpha_{k+1,\dots,N} = 1$, assuming that the link from unit k to $k+1$ is strengthened. In the example in the text, $k = 5$ [cf. (10)]. Equation 65 leads us to the values for variances of responses

$$\alpha_{Nn}^2 \overline{x_n^2} = \begin{cases} \eta^2 K^2 n, & n \leq k \\ \eta^2 K^2 k + \eta^2 (n - k), & n > k \end{cases} \quad (67)$$

Applying (63) we obtain the expression for $F(\text{MI})$

$$F(\text{MI}_n) = \begin{cases} \frac{K^2 n}{N - k + K^2 k}, & n \leq k, \\ \frac{n - k + K^2 k}{N - k + K^2 k}, & n > k, \end{cases} \quad (68)$$

which is a piece-wise linear function of n . Equation 9 determines the decision index as

$$D_n = \begin{cases} \frac{K^2}{N - k + K^2 k}, & n \leq k \\ \frac{1}{N - k + K^2 k}, & n > k \end{cases} \quad (69)$$

This confirms that the upstream units ($n \leq k$) are K^2 times more potent than the downstream ones ($n > k$).

A.3 Proof of Equation 13

Let us start by evaluating the right-hand side of Eq. 13. According to (55)

$$\overline{x_N^2} = \sum_{i=1}^N \alpha_{Ni}^2 \overline{\eta_i^2}. \quad (70)$$

Right-hand side of Eq. 13 implies that

$$D_i = \frac{\overline{\eta_i^2}}{Z} \frac{\partial \overline{x_N^2}}{\partial \overline{\eta_i^2}} = \frac{\alpha_{Ni}^2 \overline{\eta_i^2}}{Z}. \quad (71)$$

After normalization we obtain using (65)

$$D_i = \frac{\alpha_{Ni}^2 \overline{\eta_i^2}}{x_N^2}. \quad (72)$$

Let us derive the same result from Eq. 9 (left-hand side of Eq. 13). As follows from (63)

$$\begin{aligned} D_i &= F(\text{MI}_n) - F(\text{MI}_{n-1}) = \frac{1}{x_N^2} \left(\alpha_{Nn}^2 \overline{x_n^2} - \alpha_{Nn-1}^2 \overline{x_{n-1}^2} \right) \\ &= \frac{\alpha_{Nn}^2}{x_N^2} \left(\overline{x_n^2} - C_{n-1}^2 \overline{x_{n-1}^2} \right) = \frac{\alpha_{Nn}^2 \overline{\eta_n^2}}{x_N^2} \end{aligned} \quad (73)$$

Thus (71) and (73) are identical, which proves Eq. 13 and the theorem in Sect. 3.6.

Appendix B: Decision makers on trees

Let us now consider a tree connectivity specified by a connection matrix C_{ij} . An element of the connection matrix specifies the strength of connection from node j to i . The constraint on connectivity is that each node has only one outgoing link. There may be a few incoming links. Consider a node number n . There are k incoming links for this node with indexes j_1 through j_k . The response of the n th node is

$$x_n = C_{nj_1} x_{j_1} + \dots + C_{nj_k} x_{j_k} + \eta_n. \quad (74)$$

This is similar to Eq. 53. Because of the tree topology, responses of different branches are uncorrelated, i.e. $\overline{x_{j_k} x_{j_m}} = 0$. We have therefore

$$\alpha_{Nn}^2 \overline{x_n^2} = \alpha_{Nj_1}^2 \overline{x_{j_1}^2} + \dots + \alpha_{Nj_k}^2 \overline{x_{j_k}^2} + \alpha_{Nn}^2 \overline{\eta_n^2}. \quad (75)$$

Here we multiplied both sides of the equation by α_{Nn}^2 , which is given by a product of links’ strength squared on the path leading from node n to the output of the chain N .

The next step is to notice that Eq. 56 also applies to the tree case with $x_n \xi = 0$ similarly to the chain topology. Equation 56 is not valid for arbitrary topology, because responses from the upstream units to node n can also contribute to the residue ξ , thus leading to a non-zero correlation between x_n and ξ . If (56) is valid for each node in the graph, so does Eq. 63, from which we obtain

$$\begin{aligned} F(\text{MI}_n) &= \frac{\alpha_{Nn}^2 \overline{x_n^2}}{x_N^2}, \quad F(\text{MI}_{j_1}) = \frac{\alpha_{Nj_1}^2 \overline{x_{j_1}^2}}{x_N^2}, \dots \\ F(\text{MI}_{j_k}) &= \frac{\alpha_{Nj_k}^2 \overline{x_{j_k}^2}}{x_N^2}. \end{aligned} \quad (76)$$

Combined with (75) these equations lead to

$$D_n = \frac{\alpha_{Nn}^2 \overline{\eta_n^2}}{x_N^2} = F(\text{MI}_n) - F(\text{MI}_{j_1}) - \dots - F(\text{MI}_{j_k}). \quad (77)$$

The identification with the decision index here is made on the basis of Eq. 13. Another way to write this is

$$D_n = \sum_{j=1}^N (I_{nj} - S_{nj}) F(\text{MI}_j), \quad (78)$$

where I_{nj} is the unity matrix and S_{nj} is the network structure matrix, equal to 1 if there is a connection from j to n and 0 otherwise independently of the strength/sign of the connection. After denoting $F(\text{MI}_j) \equiv f_j$ Eq. 12 immediately follows.

We would like to address time-dependent problem on a tree now. All of the equations derived above can be used in the discrete time-dependent case (20). In this case however the propagation of signal on the tree introduces delays. In particular for the modified mutual information between response of the n th node and decision we have

$$F[\text{MI}_n(t)] = F(\text{MI}_n) \delta_{t+d_n,0} \quad (79)$$

Here d_n is the discrete time delay for the propagation of signal from node n to the output; decision is assumed to be made

at $t = 0$. To obtain correlators $F(\text{MI}_n)$ needed for Eq. 12 we could e.g. sum (79) over discrete time to eliminate the Kronecker symbol

$$f_n = \sum_{t=-\infty}^{\infty} F[\text{MI}_n(t)] \quad (72)$$

This quantity can be used in (12) to calculate the decision index.

Appendix C: Solution of the continuous-time model

Important for us is the time-dependent correlator

$$\hat{X}(t_1, t_2) = \overline{\vec{x}(t_1)\vec{x}(t_2)}, \quad (80)$$

which we now evaluate. The solution for (35) is obtained using matrix exponentials:

$$\vec{x}(t) = \int_{-\infty}^t dt' e^{\hat{A}(t-t')} [\vec{\eta}(t') + \vec{s}(t')] \quad (81)$$

If the external stimulus is zero or constant in time, the correlator at $t_1 > t_2$ is

$$\hat{X}(t_1, t_2) = \int_{-\infty}^{t_2} dt' e^{\hat{A}(t_1-t')} \hat{N} e^{\hat{A}^T(t'-t_2)} \quad (82)$$

We then seek $\hat{X}(t_1, t_2)$ in the form

$$\hat{X} = e^{\hat{A}(t_2-t_1)} \hat{X}_0, \quad (83)$$

where \hat{X}_0 is equal-time cross-correlation. To find an equation for \hat{X}_0 we differentiate (82) and (83) as follows

$$\frac{\partial \hat{X}}{\partial t_2} = \hat{A} e^{\hat{A}(t_2-t_1)} \hat{X}_0 = e^{\hat{A}(t_2-t_1)} \hat{N} - \hat{X} \hat{A}^T \quad (84)$$

From (83) and (84), Eq. 39 immediately follows.

8 Appendix D: The connection between discrete- and continuous-time models

Here we show that the discrete-time model can be derived from the continuous-time model. Starting from Eq. (74) for the unit responses in the continuous case we obtain the relation for solutions at two different time points separated by the time-interval τ , analogous to (20) in the discrete-time description.

From (81) we obtain

$$\vec{x}(t + \tau) = e^{-\hat{A}\tau} \vec{x}(t) + \int_t^{t+\tau} e^{-\hat{A}(t+\tau-t')} \vec{\eta}(t') dt'. \quad (85)$$

This equation can be rewritten as $\vec{x}(t + \tau) = \hat{C} \vec{x}(t) + \vec{\eta}'(t)$, where

$$\hat{C} = e^{-\hat{A}\tau} \approx \hat{I} - \hat{A}\tau. \quad (86)$$

Thus it has the same form as (20). Using (86) we obtain the new noise cross-correlation matrix

$$\hat{N}' = \int_0^{\tau} e^{-\hat{A}t'} \hat{N} e^{-\hat{A}^T t'} dt'. \quad (87)$$

The solution of the continuous-time problem satisfies the equations of the discrete-time model for an arbitrarily large time interval τ , but the new noise cross-correlation matrix \hat{N}' is non-diagonal in this case. In the limiting case $\tau \rightarrow 0$ it becomes diagonal. Indeed (80) implies that in this limit

$$\hat{N}' = \hat{N}\tau + O(\tau^2). \quad (88)$$

The matrix $\hat{N}\tau$ is diagonal by the definition of the continuous-time model. Thus, in the limit $\tau \rightarrow 0$ the matrix \hat{N}' is diagonal as needed in our formulation of discrete-time model. If the discrete model applies at small τ , Eq. 29 should also apply. This leads to $\tau(\hat{A}\hat{X} + \hat{X}\hat{A}^T - \hat{N}) = O(\tau^2)$, which is another way of deriving Eq. 39.

References

- Abramowitz M, Stegun IA (1972) Handbook of mathematical functions with formulas, graphs, and mathematical tables, 10th edn., US Govt Print Off, Washington
- Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13:87–100
- Brogan WL (1991) Modern control theory. 3rd edn. Prentice Hall, Englewood Cliffs, NJ, USA
- Celebrini S, Newsome WT (1994) Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J Neurosci* 14:4109–4124
- Durbin RM (1987) Studies on the development and organisation of the nervous system of *Caenorhabditis elegans*. In: PhD Thesis, Cambridge University
- Eaton RC, Emberley DS (1991) How stimulus direction determines the trajectory of the Mauthner-initiated escape response in a teleost fish. *J Exp Biol* 161:469–487
- Glimcher PW (2003) The neurobiology of visual-saccadic decision making. *Annu Rev Neurosci* 26:133–179
- Gold JI, Shadlen MN (2002) Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36:299–308
- Hall DH, Russell RL (1991) The posterior nervous system of the nematode *Caenorhabditis elegans*: serial reconstruction of identified neurons and complete pattern of synaptic interactions. *J Neurosci* 11:1–22
- Horwitz GD, Newsome WT (1998) Neurophysiology: sensing and categorizing. *Curr Biol* 8:R376–378
- Koulakov AA, Rinberg DA, Tsigankov DN (2004) How to find decision-makers in neural circuits? Preprint at <http://xxx.lanl.gov/abs/q-bio/0401005>
- Mazurek ME, Roitman JD, Ditterich J, Shadlen MN (2003) A role for neural integrators in perceptual decision making. *Cereb Cortex* 13:1257–1269
- Newsome WT (1997) The King Solomon lectures in neuroethology. Deciding about motion: linking perception to action. *J Comp Physiol [A]* 181:5–12
- Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 21:227–277
- Poirazi P, Mel BW (2001) Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron* 29:779–796

- Robinson DA (1989) Integrating with neurons. *Annu Rev Neurosci* 12:33–45
- Roitman JD, Shadlen MN (2002) Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J Neurosci* 22:9475–9489
- Romo R, Salinas E (2003) Flutter discrimination: neural codes, perception, memory and decision making. *Nat Rev Neurosci* 4:203–218
- Salinas E, Romo R (1998) Conversion of sensory signals into motor commands in primary motor cortex. *J Neurosci* 18:499–511
- Salzman CD, Murasugi CM, Britten KH, Newsome WT (1992) Microstimulation in visual area MT: effects on direction discrimination performance. *J Neurosci* 12:2331–2355
- Shadlen MN, Newsome WT (1996) Motion perception: seeing and deciding. *Proc Natl Acad Sci USA* 93:628–633
- Shadlen MN, Newsome WT (2001) Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J Neurophysiol* 86:1916–1936
- Smith JM (1982) *Evolution and the theory of games*. Cambridge University Press, Cambridge, New York
- Thiele A, Hoffmann KP (1996) Neuronal activity in MST and STPp, but not MT changes systematically with stimulus-independent decisions. *Neuroreport* 7:971–976
- White JG, Southgate E, Thomson JN, Brenner S (1986) The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci* 314:1–340