# The nature of the memory trace and its neurocomputational implications

P. H. de Vries · K. R. van Slochteren

**Abstract** The brain processes underlying cognitive tasks must be very robust. Disruptions such as the destruction of large numbers of neurons, or the impact of alcohol and lack of sleep do not have negative effects except when they occur in an extreme form. This robustness implies that the parameters determining the functioning of networks of individual neurons must have large ranges or there must exist stabilizing mechanisms that keep the functioning of a network within narrow bounds. The simulation of a minimal neuronal architecture necessary to study cognitive tasks is described, which consists of a loop of three cell-assemblies. A crucial factor in this architecture is the critical threshold of a cell-assembly. When activated at a level above the critical threshold, the activation in a cell-assembly is subject to autonomous growth, which leads to an oscillation in the loop. When activated below the critical threshold, excitation gradually extinguishes. In order to circumvent the large parameter space of spiking neurons, a rate-dependent model of neuronal firing was chosen. The resulting parameter space of 12 parameters was explored by means of a genetic algorithm. The ranges of the parameters for which the architecture produced the required oscillations and extinctions, turned out to be relatively narrow. These ranges remained narrow when a stabilizing mechanism, controlling the total amount of activation, was introduced. The architecture thus shows chaotic behaviour. Given the overall stability of the operation of the brain, it can be concluded that there must exist other mechanisms that make the network robust. Three candidate mechanisms are discussed: synaptic scaling, synaptic homeostasis, and the synchronization of neural spikes.

**Action Editor: Alessandro Treves**

P. H. de Vries (✉) · K. R. van Slochteren
Department of Psychology, University of Groningen,
Gr. Kruisstraat 2/1, 9712 TS, Groningen, The Netherlands
e-mail: p.h.de.vries@rug.nl

## 1 Introduction

Already in 1893 the Italian psychiatrist Tanzi (1893) postulated that the formation of memories was carried by the growth or strengthening of interactions in the brain (D'Anguilli and Dalenoort 1996), an idea that some fifty years later was stated more explicitly by Hebb (1949). He formulated the rule that the efficiency of a synapse is increased when a pair of neurons involved are simultaneously active. Although the well-known learning rule that originated from the ideas of Tanzi and Hebb, has been extensively studied over the last fifty years, there have been relatively few studies directed at the consequence of the rule: that cell-assemblies are the carriers of our memory traces. Hebb saw this as one of his major contributions to our understanding of cognitive brain functioning. The question of the robustness of cell-assemblies was not raised before Milner (1957), who introduced inhibitory interactions in the simulations of cell-assemblies in order to make them more stable. Since then studies on cell-assemblies have remained relatively scarce.

The issue of robustness was studied in Hopfield networks (1982) by means of analytical methods. For such relatively simple networks, with simple models of neurons, it is possible to draw conclusions from analytical studies (Gerstner and Kistler 2002). Provided the model of the network is not too complex, the equations describing the dynamics of such networks can be approximately analysed, on the basis of arguments that are mainly heuristic. These equations can only be analysed for networks that are of infinite size, that have some properties of symmetries, or that consist of simplified neurons, for example such that they all have the same threshold, and the same numbers of

interconnections. Moreover, the system must be random to allow statistical arguments. The equations can then be handled in a statistical fashion, in manners known from statistical physics. Unfortunately, or perhaps fortunately, these statistical analyses cannot be used for the study of the properties of networks that are to serve as substratum for cognitive tasks that have some relevance for the study of cognition, such as doing simple arithmetic, or producing and understanding language. These networks can in principle not be analysed fully in a statistical manner. Even a basic requirement, the implementation of binding—a topic to be discussed later on—seems to be impossible in an analytical model. For these inhomogeneous and non-uniform networks, only the process itself can be simulated. This is in contrast to the numerical analysis of a simple network for which the equations can be so far approximated that the equations can be evaluated by numerical techniques for different cases of parameter values, and for different types of networks. For the architecture and dynamics of neural networks that can serve as the substratum of specific cognitive tasks, only simulations are possible of the network itself. (Dalenoort, personal communication).

The last two decades have shown an increase in the interest of cell-assemblies (Dalenoort 1985; Pulvermüller 1996, 1999; Huyck 2004) and also in analytical studies, where they are represented in terms of attractors (Amit 1995; Amit and Mongillo 2003). As we argued above an analytical representation in terms of attractors is not suitable to answer questions about the specific network structure of cell-assemblies necessary from a cognitive point of view (Dalenoort and de Vries 1995).

The approach of this paper is that cognitive requirements expressed in terms of the functional notion of memory traces are used to design simulations at the neural level. On the basis of these simulation studies new phenomena in the model can be distinguished and compared with what is known neuro-physiologically. It is possible that this will lead to the discovery of actual new phenomena. In earlier work (de Vries 1995), we referred to this approach as 'downward emergence'. Quintessential to the approach is that a strict bottom-up study of brain functioning will not be sufficient. Cognition—the top-down approach—has to be taken into account as well (Dalenoort 1990; Dalenoort and de Vries 1998a).

## 2 Hebb revisited: a generalization of the cell-assembly concept

In this paper we will generalize the concept of cell-assembly. This notion has its origin in the phenomenon that the discussed Tanzi–Hebb learning rule leads to clusters in which neurons are more strongly connected to each other than to neurons in other clusters. According to Hebb

each cluster corresponds to a memory trace. Already in the 1940s it had to be assumed that the neurons of a cell-assembly must be widely distributed since lesion experiments do not have specific effects on memory traces (Hebb 1949; Lashley 1951). Still, we consider the original notion of a cell-assembly as being too static. The neurons carrying a memory trace do not belong exclusively to a single cluster. Rather, a memory trace should be considered as an excitation pattern in a large network of neurons. In different contexts a memory trace is carried by different sets of active neurons. Accordingly a memory trace can contribute to various context-specific meanings (Dalenoort 1982, p. 176).

Another extension of the original notion of cell-assembly that has been applied in the current study, concerns the existence of a critical threshold. For almost all networks of threshold elements—such as neurons—there exists a level of activation above which the activity in the network will rise autonomously to its maximum level, whereas it will extinguish below that level. Such a critical threshold is important to relate cognitive and neural models of information-processing (Dalenoort 1985): suprathreshold activity (autonomous growth of excitation) might be hypothesized to correspond to cognitive processes that are reported to be consciously experienced, whereas subthreshold activity (followed by gradual extinction of excitation) corresponds to processes of implicit memory such as priming. According to Hebb's original notion, activity of cell-assemblies represents short term memory, whereas the network of assemblies as a whole stands for long term memory. By means of the concept of the critical threshold of a cell-assembly, in addition, explicit and implicit memory can be represented in the same network. The notion of 'critical threshold' is also present in the concept of 'ignition of a cell-assembly' (Braitenberg 1978; Pulvermüller 1999). These studies emphasize the instantaneous character of the activation of cell-assemblies. The present study focuses on the conditions necessary to create sufficient room for subthreshold excitation. When this excitation exceeds the critical threshold, a smoothly developing autonomous growth of excitation should produce an oscillation in the network.

Although the neurons participating in the excitation pattern underlying a memory trace do not form a fixed set, there must exist a nucleus of neurons that is active in each occurrence of the pattern. The memory trace obtains its identity from the specific connections with the sensory and motoric parts of the nervous system (Dalenoort 1996).

In this approach the identity or 'signature' of a memory trace is not represented by the temporal structure of the spikes in a excitation pattern, as assumed by Shastri and Ajjanagadde (1993), Sougné (2001), and Harris (2005). It is difficult to see how such purely temporal code could lead to specific behavioural responses without the intervention of some kind of 'decoding' device. The introduction of such a

device in a model of cognitive brain functioning, however, would be inconsistent with the requirement of self-organisation for such a model (Dalenoort and de Vries 1998b).

Within the generalized model of cell-assemblies the temporal structure of spikes is not without significance, however. They may play a crucial role in the solution of the binding problem. In the context of the paper a possible solution to this problem is best introduced by the following example. The fact that we can make associations between *arbitrary* words implies that between the cell-assemblies of the memory traces of these words, a resonance is established. How is it possible that such a resonance comes into being between cell-assemblies that are not specifically connected (like those for 'black' and 'white')? This resonance must come into being in a fashion that agrees with the self-organising character of the network, no top-down information can be used that is not "available" to the network. Within this approach it is assumed that for such a resonance to occur, the spike trains produced by both assemblies have to be in phase. This solution of the binding problem is compatible with the ideas of binding expressed in theories about synfire chains (Abeles 1991; Abeles et al. 2004; Hayon et al. 2005). In the discussion of the synfire chains, however, no explicit statements are made with regard to the nature of the memory trace. A wave of excitation in a synfire network is a purely temporal phenomenon. Thus the approach based on synfire chains begs the question of how memory traces obtain their identity (Dalenoort 1996), and how they are related to explicit and implicit memory.

As was already pointed out in the introduction, networks of cell-assemblies must have a very specific structure in order to carry the cognitive processes typical of human cognition. A good example of a minimal model structure needed, is exhibited by the cell-assemblies that carry the memory traces for words. As shown in Fig. 1, such a cell-assembly, the so-called 'word node', has incoming excitatory connections from the cell-assemblies corresponding
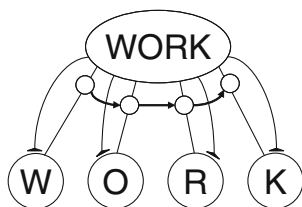


**Fig. 1** An example of the logistics of a small network of cell-assemblies. The *ellipse* is a word node denoting a cell-assembly for the memory trace of the word 'WORK', the *large circles* are letter nodes denoting cell-assemblies of the memory traces of its constituting letters. The *small circles* represent sub-assemblies of the word node, necessary for the representation of the order of letters. *Unmarked lines* indicate excitation loops between (sub)assemblies, *arrows* stand for the excitatory connections from one sub-assembly to another, and *lines ending with a bar* stand for inhibitory connections

to the memory traces of the letters constituting the word, the so-called letter nodes. There must also exist excitatory connections from the word node back to the letter nodes, otherwise the top-down-effects typical for the word superiority effect (Reicher 1969; Wheeler 1970; McClelland and Rumelhart 1981) could not occur. As a consequence the relationship between a word node and one of its letter nodes is an excitatory loop. In addition the loops between a word node and its letter nodes must reflect the order of the letters in the word. This can easily be concluded from the fact that we can quickly answer, e.g., the question 'what is the third letter of the word grass?'. At the neural level this means that within a word node there must exist subpopulations of neurons for each loop to a letter node. In turn these subpopulations, or subnodes, must be organised in a chain that reflects the serial order of the letters of the word. Propagation of excitation through this chain is necessary to find which letter is at a certain position within a word. For this purpose it is also necessary that there exists a network of cell-assemblies that represents order in a generic sense. Because of the binding occurring between the letter nodes of a word and the memory traces for 'first', 'second', etc. in this special network, the excitation process in the chain of subnodes will reach, and stop at, the subnode corresponding to a letter position asked for.

Besides the excitatory connections discussed so far, there must exist inhibitory connections. A specific mechanism of inhibition employed in the presented cell-assembly model, is referred to as *backward inhibition*. Because of this mechanism the activity of a cell-assembly that brings the excitation level in another assembly to a level above the critical threshold, is extinguished by the latter. In the word-and-letter example of Fig. 1, connections for backward inhibition thus run in the opposite direction as the excitation loops through which a letter node activates a word node. Accordingly the letter nodes are extinguished by their word node, once its excitation level has increased above the critical threshold. At the cognitive level this corresponds to the phenomenon that in normal reading we perceive a word as a gestalt without the constituting letters. This short cognitive analysis of the structure of a memory trace already requires a relatively complex network.

At present it is cumbersome to construct a computer simulation at the level of individual neurons for a cognitive task of some complexity. The top-down, or programmed, tuning of the numerous parameters is impossible. The tuning has to be done by a self-organising process of which we do not yet know the conditions (Dalenoort and de Vries 1998b). At present we still need additional insights to further specify a learning rule, so that the appropriate forms of self-organisation arise. The complexities at the neuronal level can be avoided by switching to a higher level of description. In de Vries (2004) the functioning of cell-

assemblies, including binding processes required by the task, is described and simulated at the level of conceptual networks. In such a network each node represents a cell-assembly, and accordingly the number of parameters involved is considerably reduced.

An essential property of cell-assemblies in the context of this paper concerns their robustness. The neuronal circuitry underlying cognitive processes must be very robust since our cognitive system continues to function in a broad range of biological conditions (extensive cell-death in adults, loss of sleep, consumption of alcohol, head injuries a.o.). This robustness implies that the parameters determining the functioning of networks of individual neurons must have large ranges, or there must exist stabilizing mechanisms that keep the parameters for proper functioning of a network within narrow bounds. The role of the parameters of the neurons making up this circuitry has only been understood to a limited extent. In this paper we will explore the parameter space of model neurons to see whether there exist relatively large subspaces that underlie the robustness of cell-assemblies. In order to carry out this exploration we will use a network (cf. Fig. 2(a)) of neurons that is minimally required for the specific structures of memory traces described above, and necessary to understand the heterogeneity of human cognition.

The chosen network is constructed and does not come about through self-organisation. At present there is still insufficient knowledge to specify a learning rule for networks of the required cognitive complexity. However, the results of the exploration of the parameter space may provide useful insights for the formulation of such a rule.

# 3 The choice of a minimal cognitive architecture

Many sciences have benefited from the choice of an idealized model from which general conclusions could be drawn. Since the exploration of the neuronal parameter space is not feasible for networks of neurons representing cognitive tasks, we need an idealized architecture in order to find the parameters of individual neurons that underlie the robust functioning of cell-assemblies. The architecture has to obey the following conditions:

1. an 'autonomous-growth' condition (Fig. 2(b)): if a cell-assembly is externally activated at a level above the critical threshold, then its excitation level should grow autonomously to its maximum level. As a consequence it will activate other cell-assemblies. If one of these also becomes active at a level above the critical threshold, it will extinguish the activating cell-assembly. Accordingly only one cell-assembly at a time will be active at a supra-threshold level.
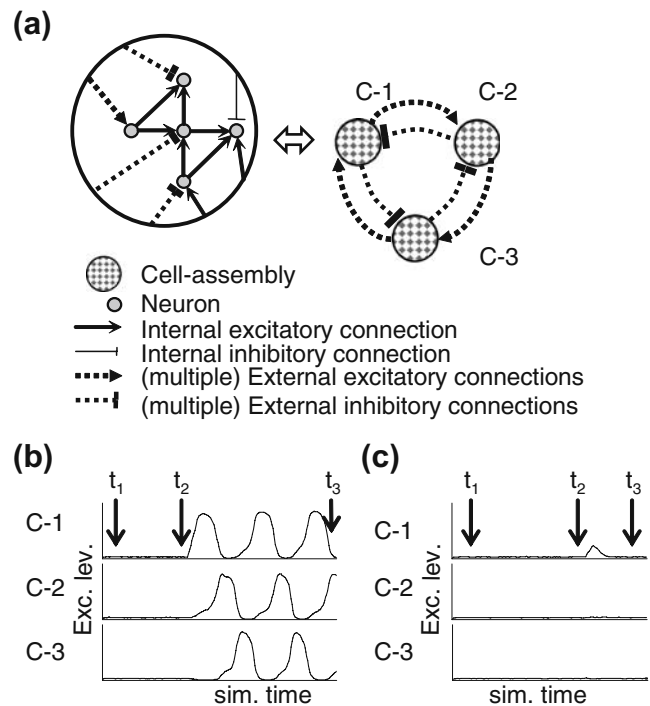


Fig. 2 The chosen minimal architecture and its computer simulation. (a) A loop of three cell-assemblies, the *large gray circles* denote cell-assemblies with dashed lines indicating the excitatory and inhibitory connections between them; the *double-sided white arrow* is used to show part of a cell-assembly at the level of neurons (*small gray circles*) with *solid lines* indicating the excitatory and inhibitory connections internal to the assembly. (b) Computer simulation of autonomous growth with subsequent oscillation around the critical threshold when sufficient input is given, i.e. activation of a random selection of 25% of the neurons in one of the cell-assemblies with a value of .25 per neuron on a scale from zero to one. (c) Computer simulation of extinction of excitation when insufficient input was given, similar as in (b) but now with an activation value of .20 per neuron. The symbols $t_{1-3}$ are time points for conducting a simulation experiment: $t_1$ marks the beginning of a start-up condition in which the network is only activated with random activation (defined by the function $R$ in formula (1)) and should settle in an equilibrium, $t_2$ marks the end of the start-up condition and the beginning of the condition in which autonomous growth (b) or extinction (c) should occur on the basis of the external inputs described above, $t_3$ marks the end of the simulation experiment

2. an 'extinction' condition (Fig. 2(c)): if a cell-assembly is externally activated at a level below the critical threshold, then its excitation level should extinguish. Although the cell-assembly will contribute to the activation of other assemblies, it alone will not bring them above the critical threshold.

These conditions are derived from cognitive arguments for task performance. The observation that humans can only be conscious of one thing at a time, corresponds to the first condition that only one assembly can be active at a level above the critical threshold. The observation that memories can remain subconscious corresponds to the second condition, in which excitation of a cell-assembly

does not exceed the critical threshold. The approach in which observations at a certain level of description are used to introduce hypotheses about new phenomena at a lower level, was referred to as 'downward emergence' earlier in this paper.

The *minimal* architecture obeying these conditions is a loop of three cell-assemblies (cf. Fig. 2(a)). In the architecture excitatory connections from one cell-assembly to the next are required, as well as inhibitory connections in the reverse direction (cf. the mechanism of backward inhibition discussed in the previous section). In this way the rise of excitation above the critical threshold in one cell-assembly will extinguish the excitation in the preceding assembly. Consequently one assembly will remain active in an oscillating loop.

Moreover, if this propagation of excitation is to be plausible from a cognitive point of view then the corresponding excitation curves should show a smooth increase and decline during oscillation. Within a cell-assembly inhibitory connections may therefore be necessary to prevent too sharp an increase in excitation.

A loop of two (instead of three) cell-assemblies is too small to allow the excitation level of an extinguished cell-assembly to decrease at a level below the critical threshold. This would not lead to the required oscillation, in which only one cell-assembly is active at a level above the critical threshold.

A fundamental assumption underlying the proposed minimal architecture is that one must first acquire insight into its parameter space before the problem of its self-organisation (on the basis of an extension of the Tanzi–Hebb rule) can be solved.

When reduced to its minimal form, the architecture may also be studied according to analytical methods, such as those proposed in Van Vreeswijk and Sompolinsky (1996), Amit and Brunel (1997), and Brunel (2000). This would provide a useful input to studies about the self-organisation of larger, heterogeneous networks for specific cognitive tasks, which are not amenable to analytical methods, and have to be simulated.

In the proposed minimal network we do not include the role of binding and therefore no spike trains are modeled. Modeling a neuron at the level of spikes would require too many parameters for a first approach to solve the problem of robustness for a minimal network necessary for cognitive tasks.

We focus on the robustness of the oscillation of an excitation wave in a loop of three cell-assemblies. Such an oscillation represents a steady state, that at the cognitive level is supposed to correspond to a condition that something becomes known in the network. As an example of a cognitive task corresponding to the described network, one can think of the rehearsal of three items, e.g. three letters of the alphabet. It is crucial that the critical threshold of the cell-assemblies in the loop has an appropriate value from a cognitive point of view. This means that this threshold should not be too low, for then a cell-assembly would start its autonomous growth too quickly, and there would be little room for subthreshold excitation corresponding to priming and other processes of implicit memory. Neither should a critical threshold be too high, for then autonomous growth will hardly develop.

## 4 Choice of parameters

Each of the three cell-assemblies in the loop is composed of model neurons, compatible with the 'state-of-the-art' knowledge of neural functioning. The number of these neurons and their properties constitute the 12 main parameters of the computer simulation, cf. Table 1. Most of the parameters are stochastic in nature, which means that their values are specified by two components: a mean, and a standard deviation. In any concrete realization of the network actual values have to be drawn from the normal distributions specified by these components. The generation of a series of actual values from a distribution is determined by a random seed. This implies that a network is determined by its parameter values and its set of random seeds (one seed for each distribution parameter). Accordingly, different versions of a network can be created when the same parameter values are combined with different sets of random seeds.

The model of the neuron and its parameters is based on the processes in a synapse. The effectors of presynaptic neuron $A$ contain neurotransmitters that in interaction with the properties of receptors on postsynaptic neuron $B$ can either produce an excitation or an inhibition of $B$. Since the impact of released neurotransmitters is dependent on the types of receptors in the postsynaptic neuron, each model neuron can have excitatory as well as inhibitory connections to other neurons. The processes of excitation and inhibition for each neuron have different absolute thresholds in the model.

A model neuron can have excitatory as well as inhibitory connections, each with a separate threshold. At the neural level the inhibitory connections can be represented by the combination of an excitatory neuron and an inhibitory interneuron, each with its own firing threshold. For the computational purposes of this paper we refrained from such a representation with an interneuron because it would increase the number of parameters to an extent that the exploration of the parameter space would no longer be computationally feasible. Also in analytical studies (Brunel 2000) parameters are taken together in order to make a solution possible.

**Table 1** The 12 parameters of the chosen architecture

| Description | Name mean | Range | Name SD | Range |
|---|---|---|---|---|
| Parameters specified as single values | | | | |
| 1. Number of neurons per cell assembly | NrCA | (135, 240) 15 | | |
| 2. Decay of excitation level of a neuron per time step | Decay | (.2,.55).05 | | |
| Parameters specified as normal distributions | | | | |
| 3. Excitatory threshold of a neuron | mThrE | (.1,.8).1 | sThrE | (.150,.325).025 |
| 4. Inhibitory threshold of a neuron | mThrI | (.3, 1.0).1 | sThrI | (.35,.525).025 |
| 5. Number of excitatory connections of a neuron in a cell assembly | mNrEI | (5.0, 7.8).4 | sNrEI | (.0, 2.8).4 |
| 6. Strength of excitatory connections of a neuron in a cell assembly | mStrEI | (.11,.18).01 | sStrEI | (.0,.07).01 |
| 7. Number of inhibitory connections of a neuron in a cell assembly | mNrII | (.0, 3.5).5 | sNrII | (.0, 1.75).25 |
| 8. Strength of inhibitory connections of a neuron in a cell assembly | mStrII | (.0,.35).05 | sStrII | (.0,.175).025 |
| 9. Number of excitatory forward connections of a neuron external to a cell assembly | mNrEfE | (.65, 1.0).05 | sNrEfE | (.05,.04).05 |
| 10. Strength of excitatory forward connections of a neuron external to a cell assembly | mStrEfE | (.05,.12).01 | sStrEfE | (.0,.07).01 |
| 11. Number of inhibitory backward connections of a neuron external to a cell assembly | mNrIbE | (1.0, 3.8).4 | sNrIbE | (.4, 3.2).4 |
| 12. Strength of inhibitory backward connections of a neuron external to a cell assembly | mStrIbE | (.5,.85).05 | sStrIbE | (.0,.35).05 |

The parameters can be distinguished in those having a single value for the entire architecture (rows 1 and 2), and parameters characterizing each individual neuron, the value of which is drawn from a normal distribution (rows 3–12). For the latter the mean, and the standard deviation of the distribution are given. For each parameter the range and step size used in the exploration of the parameter space is specified. Since the exploration turned out to be particularly sensitive to the value of the decay, this parameter was set to a fixed value during the initial stage of the exploration and varied later on, cf. Fig. 3.

The behaviour of each neuron is characterized by its excitation level, that is based on a commonly used integrate-and-decay-mechanism, cf. formulas (1), (2) and (3).

$$E_i(t) = (1 - D) \times E_i(t-1) + \sum_{j=1}^{n} \text{ExcIn}_{ij}(t)$$
$$+ \sum_{j=1}^{n} \text{InhIn}_{ij}(t) + R_i(t) \quad 0 \leq E_i(t) \leq 1$$
(1)

$$\text{ExcIn}_{ij}(t) = \begin{cases} 0 \text{ if } E_j(t-1) < \text{ThrE}_j \\ E_j(t-1) \times W_{ij} \text{ if } W_{ij} \geq 0 \wedge E_j(t-1) \geq \text{ThrE}_j \end{cases}.$$
(2)

$$\text{InhIn}_{ij}(t) = \begin{cases} 0 \text{ if } E_j(t-1) < \text{ThrI}_j \\ E_j(t-1) \times W_{ij} \text{ if } W_{ij} < 0 \wedge E_j(t-1) \geq \text{ThrI}_j \end{cases}$$
(3)

In formulas (2) and (3) the values of the variable $\text{ThrE}_j$, resp. $\text{ThrI}_j$, are drawn for the distribution determined by mThrE and sThrE, resp. mThrI and sThrI, (see Table 1) and give the actual threshold for excitation, resp. inhibition, for neuron $j$.

In the simulation, excitation and inhibition correspond to the increase, resp. decrease, of the number of spikes per time frame in the spontaneous firing rate of a biological neuron. The parameter $D$ expresses the decay of the excitation level $E$ of a neuron. In a biological neuron decay corresponds to the phenomenon that the number of spike trains per time frame produced by a neuron will gradually return to the rate of spontaneous firing if no excitatory input is received.

The function $R_i(t)$ in formula (1) denotes random fluctuations occurring in a neuron's excitation level. It was introduced to control for the sensitivity of the simulation to fluctuations. Throughout all simulation experiments its value was an amount drawn from a normal distribution with mean 0.0 and standard deviation .05 for each neuron at every time step. The actual value of a neuron's excitation level was increased or decreased by this amount.

# 5 Simulation experiments based on 'downward emergence'

The exploration of the parameter space of the minimal architecture is accomplished on the basis of simulation experiments carried out by a Neural Network Simulator (NNS). Each experiment consists of the described conditions, 'autonomous-growth' and 'extinction', and is given the values of the parameters of a selected point in the parameter space, together with $N$ sets of random seeds. The exploration aims to find the values of parameters that produce the effects expected in both conditions. As such the search for the appropriate parameter values is an example of downward emergence since it is guided by knowledge from a higher level of description.

In each simulation experiment $N$ networks are constructed according to the given parameter values and the $N$ sets of random seeds. The latter are necessary to minimize the

probability that the occurrence of the desired effects is due to an idiosyncratic sample of parameter values drawn from their specified distributions. Moreover, certain specific connection structures in the network may have unwanted effects. In order to prevent the occurrence of such artefacts, NNS can generate different versions of a network by combining the same parameter values with different sets of random seeds. With the given parameter values NNS 'conducts' a simulation *run* for each set of random seeds. Each run consists of one condition for autonomous growth and one for extinction. A *simulation experiment* therefore contains as many runs, and as many conditions for autonomous growth and for extinction, as the number of sets of random seeds that were specified.

The exploration of the parameter space took place in two stages. In the first stage the exploration of the parameter space was done on the basis of a set of fixed parameter values for which one could estimate statistically that the desired effects would occur. If these estimations were confirmed by simulation experiments, the ranges around these values were explored. For each experiment the number of sets of random seeds was set to five. Although it was possible to find ranges of parameter values that produced the desired effects, it turned out that these were relatively small and fragmented.

In order to explore larger parts of the parameter space a computerized search was used in the second stage. A program (GA) based on a genetic algorithm was used (Goldberg 1989; Coley 1999). GA selects parameter values from prespecified ranges and determines value sets of optimal fitness, i.e. sets of parameter values for which the loop of cell-assemblies manifests to a sufficient degree autonomous growth and extinction, given the appropriate input. The ranges from which the values were chosen, were obtained from prior estimates based on the size of the simulated cell-assemblies and the number and strength of their internal and external connections—obtained in the first stage of parameter search. To determine the fitness values GA calls NNS, which constructs the minimal architecture described above, according to the parameter values selected by GA from the specified ranges and the sets of random seeds. According to the principles underlying genetic algorithms this selection of parameter values takes place according to a quasi-random process, again determined by a random seed. In order to find regions in the parameter space that give a sufficiently high fitness, it is necessary to use different random seeds. In this way the parameter space will be explored from different starting points, and one avoids hitting only local maxima.

With the introduction of the genetic algorithm we must now distinguish two kinds of random seeds: those necessary for the selection of parameter values by GA and those necessary for the construction of different versions of a network within the simulation experiments conducted by NNS. These different kinds of random seeds will be referred to as GA-random-seeds

and NNS-random-seeds, respectively. In the second stage of exploration of the parameter space we increased the number of NNS-random-seeds per simulation experiment from five to ten. Accordingly the effects of idiosyncratic network structures in this stage of exploration should be decreased.

For each set of parameter values it selects, GA computes a fitness measure. This measure reflects the extent to which each of the ten generated versions of the architecture—one for each of the ten sets of NNS-random seeds in a simulation experiment—leads to:

a. a stable oscillation in the autonomous-growth condition, and
b. an excitation curve that approaches zero in the extinction condition.

For a formalization of these conditions such that they can be used by the genetic algorithm, the reader is referred to the appendix.

By means of the approach of genetic algorithms one can carry out a simultaneous optimization of all the parameters of the chosen minimal architecture. As a result one obtains sets of 12 parameter values—one for each GA-random-seed—each of which indicates a point of maximal fitness in the parameter space.

GA gives us the points in the parameter space that produce the desired effects, for all ten sets of random seeds provided. In view of the paper's general question on the robustness of cell-assemblies, however, we are looking—per parameter—for a region of values that satisfies the requirements for both the condition of autonomous-growth and the condition of extinction. In order to find these regions, we used a third computer program, GF, that generates a list of symbolic codes for a prespecified range of parameter values. Each symbolic code indicates to which extent the conditions of autonomous-growth and extinction produce the desired effects for each of the ten runs in a simulation experiment per parameter value. For this purpose the code is composed of two triples. Each triple consists of three counters $<i_1, i_2, i_3>$ expressing the number of runs in which did occur respectively:

1. autonomous growth (see Fig. 2(b)).
2. an extinction. (see Fig. 2(c))
3. neither 1 nor 2 (in this case e.g. the network can be in a chaotic state, or its excitation levels can remain at a fixed level above zero.)

The first triple in the symbolic code is used to expresses the extent to which the autonomous-growth condition is satisfied, whereas the second reflects the extent to which the extinction condition is met. For example, the triple <10, 0, 0> represents that 10 runs in a simulation experiment led to autonomous growth, whereas none of the runs produced an extinction or other state of the network. For the
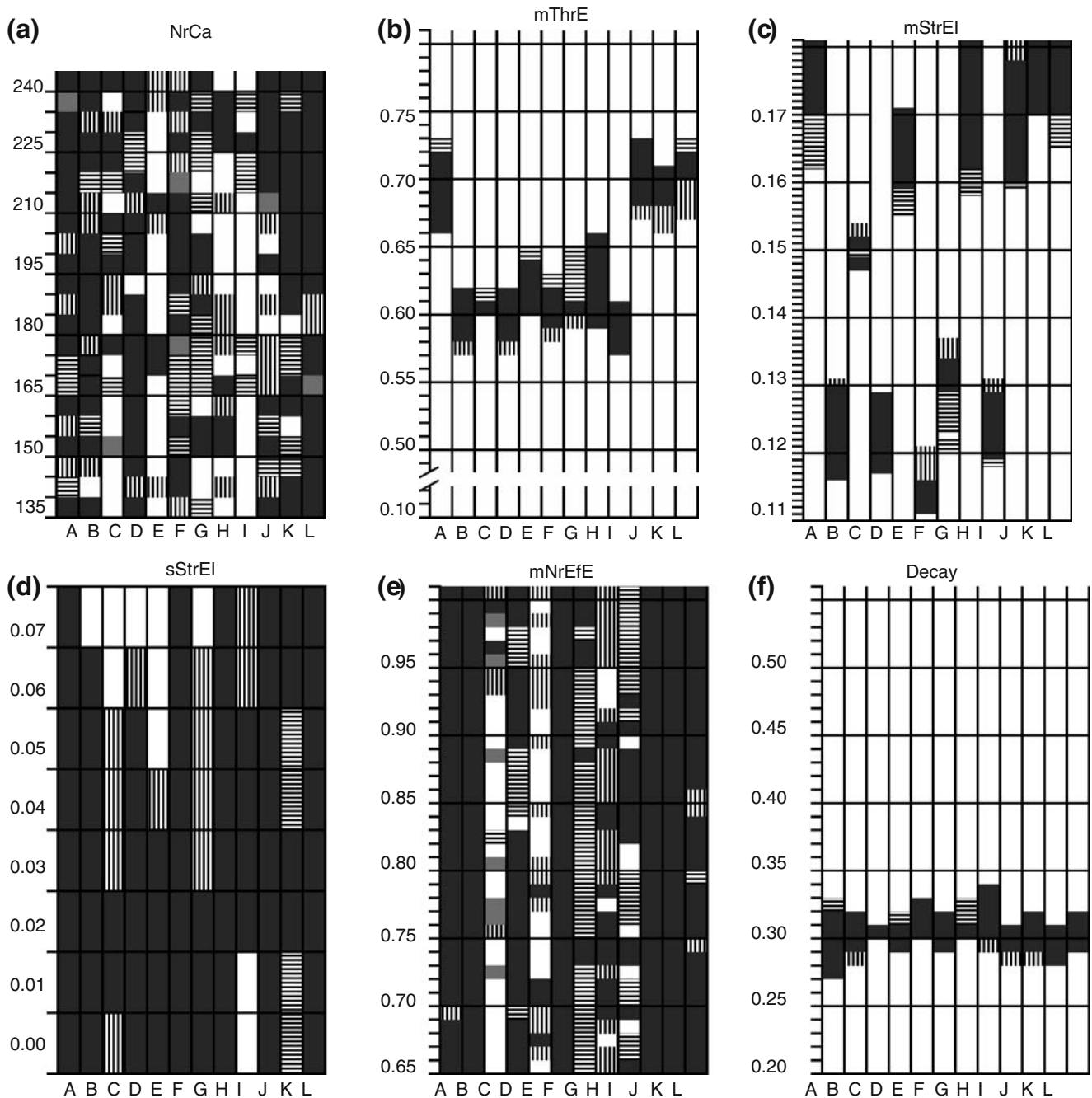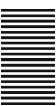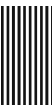
**Fig. 3** Parameter graphs for 6 of the 12 parameters in Table 1: the graphs (**a**–**f**) display the extent to which the architecture exhibits the required autonomous growth and extinction behaviour for 6 sets of choices of the parameter values (computations were done for all 12 parameters, leading to 12 optimal sets); each set consists of (accidentally also) 12 parameter values, corresponding to the 12 GA-random-seeds (A–L) used to initialize the different searches of the parameter space; each parameter graph contains 12 vertical bars, one for each GA-random-seed; the differences in shading in each bar $\beta$ ($\beta$ = A–L) in a parameter graph represent the changes in behaviour of the minimal architecture of Fig. 2(a) when the values of that parameter are varied along the vertical axis (only 6 cases of 12 are shown); the other parameters of the architecture keep the optimal value found in the parameter search corresponding to the random-seed of $\beta$; in each bar five different types of outcome are distinguished, each indicated by a different character of shading in the graphs (see Table 2); the parameters shown are (**a**) *NrCA*, number of neurons per cell-assembly, (**b**) *mThrE*, mean excitatory threshold of a neuron, (**c**) *mStrEI*, per neuron the mean strength of excitatory connections internal to a cell-assembly, (**d**) *sStrEI*, per neuron the standard deviation of strength of excitatory connections internal to a cell-assembly, (**e**) *mNrEfE*, per neuron the mean number of excitatory forward connections with neurons external to its cell-assembly, (**f**) decay of excitation level (*D*) of a neuron per time step (cf. Table 1, parameters of the chosen architecture)

**Table 2** Different shadings used in Fig. 3 to characterize the behaviour of the minimal architecture shown in Fig. 2

| | |
|---|---|
| ■ | dark grey: autonomous-growth and extinction condition both achieve their maximum score of 10: the 10 runs in both conditions all gave the expected outcome. |
| ☰ | horizontal shading: 9 out of the 10 runs in the autonomous-growth condition gave the expected result, all 10 runs in the extinction condition were successful. |
| ‖ | vertical shading: 9 out of 10 runs in the extinction condition gave the expected result, all 10 runs in the autonomous-growth condition were successful. |
| ▦ | light grey: for both the autonomous-growth and the extinction condition 9 of the 10 runs gave the expected outcome. |
| □ | white: for the autonomous-growth or the extinction condition (or both) less than 9 of the 10 runs gave the expected outcome. |

oscillation condition in the experiment this example triple is the desired outcome, whereas for the extinction condition it is just the opposite. Our discussion of the results of the simulation experiments will be based on these symbolic codes. Fitness values generated by GA, do not reveal which of the two conditions in a simulation experiment are met.

GF accomplishes the generation of symbolic codes by calling NNS with the appropriate parameter values and ten sets of random seeds. NNS then carries out the simulation experiments. GF can thus be used to characterize the behaviour of the architecture in simulation experiments based on parameter values which vary around the values of maximal fitness found by GA. Within each set of 12 parameter values found on the basis of a GA-random-seed, either individual or pairs of parameters can be selected for processing by GF. In both cases the not-selected parameters keep their optimal value for that GA-random-seed.

## 6 Results of the simulation experiments

Table 1 lists the 12 parameters of the architecture that were optimized and the ranges from which GA selected the values in the exploration of the parameter space. In the present study 12 different GA-random-seeds were used to start separate searches of the parameter space. In these

explorations the criterion for sufficient fitness was set to 90% of the maximum fitness. For each of the GA-random-seeds the genetic algorithm found a set of 12 parameter values that produce a fitness above the criterion level. For all the parameter values in prespecified ranges GF computed the symbolic codes reflecting the outcome of the simulation experiments. Given the relatively large number of parameters and of GA-random seeds involved, our discussion will be limited to the ranges found for 6 of the 12 parameters in each of the 12 explorations initialized by a GA-random-seed. These six parameters in particular are relevant to the conclusions on the robustness of cell-assemblies. In Fig. 3 each graph (a–f) corresponds to one of these six parameters. In each of these parameter graphs the successive vertical bars correspond to the 12 different GA-random-seeds (A–L) used in the optimization process. The differently shaded areas within a bar indicate the extent to which the outcomes of the simulation experiments approach the desired behaviour, based on the triples discussed in the previous section. Each bar in a parameter graph in Fig. 3 corresponds to a GA-random-seed (indicated by the capitals A–L) used in a parameter search. A bar in a parameter graph reflects the changes in behaviour of the architecture when the values of that parameter are varied while the other parameters keep their optimal value, found in the parameter search with the GA-random seed corresponding to that bar.

Let us now consider what the obtained results of the simulation experiments mean for the robustness of cell-assemblies. We will first focus on the five parameters specifying a mean: the number of neurons per cell-assembly (NrCA), the mean excitatory threshold (mThrE), the mean strength of excitatory connections internal to a cell-assembly (mStrEI), the mean number of excitatory forward connections external to the cell-assembly (mNrEfE), and the decay of the excitation level of a neuron ($D$), displayed in Fig. 3(a), (b), (c), (e), and (f) respectively. Of primary interest are parameter ranges in which the desired behaviour did occur, i.e. in which all 10 oscillation conditions in a simulation experiment indeed produced an oscillation, and all 10 extinction conditions led to an extinction. Such regions, the dark grey areas in the graphs of Fig. 3, will be qualified as indicating ideal behaviour.

### 6.1 Parameters specifying a mean

*Proliferation of subspaces* The graphs of these parameters display rather large differences between the ranges produced by the different GA-random-seeds. When we look in the graph for the parameter NrCA (Fig. 3(a)), we see that the ranges of the GA-random-seeds J, K, and L overlap to a large extent, especially if we allow one of the simulation runs to fail on the autonomous-growth or extinction condition (the horizontally, resp. vertically, shaded areas). The same holds for these

three random seeds if the parameters mThrE, mStrEI, and mNrEfE are considered (Fig. 3(b),(c), and (e); the value of the decay parameter was set to a fixed value for all parameter searches). The other GA-random-seeds, however, do not occupy the same regions in the parameter space. The ranges of GA-random-seeds A and B, for example, do overlap for parameters NrCA and mNrEfE (Fig. 3(b)) but are widely dispersed for parameters mThrE and mStrEI (Fig. 3(b) and (c), respectively). In addition, the ranges of GA-random-seeds B–I overlap for the mean excitatory threshold (mThrE in Fig. 3(b)) but for the other parameters in Fig. 3 representing a mean, these GA-random-seeds have quite different ranges in the parameter space. These findings are not compatible with the hypothesis that the robustness of cognitive brain functioning can be explained on the basis of clearly distinguishable and relatively large subspaces of the parameter space, in which the architecture produces the desired behaviour. There seem to exist several, relatively small subspaces that are appropriate.

*Narrow ranges* Another issue concerns the relative narrowness of most of the ranges of ideal behaviour displayed in Fig. 3. For the used GA-random-seeds the width of the ranges of mThrE, mStrEI, and $D$ is at most 10% of the average of their lower and upper bound (Fig. 3(b),(c), and (f)). Besides their narrowness, these ranges also seem to be rather steep. This can be deduced from the horizontally and vertically shaded areas indicating failures in the autonomous growth and extinction conditions. For mThrE, the excitatory threshold (Fig. 3(b)), and $D$, the decay (Fig. 3(f)), the horizontally shaded areas above the ranges of ideal behaviour are all referring to a failure in the autonomous-growth condition. This is due to the relatively high value of mThrE, resp. $D$, in these areas, which makes it hard for autonomous growth to develop. Similarly, the vertically shaded areas below the ranges of ideal behaviour in the graphs of mThrE and $D$ all indicate a failure in the extinction condition. This stems from a lower excitatory threshold, resp. decay, in these areas which makes it difficult for the excitation in the network to extinguish.

Just the opposite pattern can be observed for the mean strength of excitatory connections external to the cell-assembly (mStrEI). For this parameter the vertically shaded areas above the ranges of ideal behaviour all concern a failure in the extinction condition. The relatively high connection strengths in these ranges are favourable for the extinction of excitation in the network. Similarly, the horizontally shaded areas below the ranges of ideal behaviour in mStrEI all represent failures in the autonomous-growth condition. Relatively weak connections do not promote autonomous growth of excitation in the network.

The horizontally or vertically shaded areas all represent a failure for one of the ten conditions of autonomous-growth or extinction in each simulation experiment. However, their width is again very small, in most cases only a single step size in the total range selected for the parameters. The architecture does therefore not display a graceful degradation. The narrow and steep nature of the ranges of ideal behaviour can not serve as a basis for robust cognitive brain functioning.

*Fragmentation per parameter* A third issue to be observed in Fig. 3 is the fragmentation of the value space of parameters. This phenomenon occurs for the number of neurons per cell-assembly (NrCA, Fig. 3(a)) and the mean number of excitatory, forward connections external to the cell-assembly (mNrEfE, Fig. 3(e)). In the corresponding parameter graphs we see ranges of ideal behaviour mixed with ranges displaying one or more failures in the autonomous growth or in the extinction condition (the shaded, light grey, or white areas). These ranges of suboptimal behaviour do not exhibit a systematic pattern as was the case for the parameters mThrE, mStrEI, and $D$. Like the proliferation of subspaces and the narrowness of the ranges of ideal behaviour, this fragmentation of the value space of a parameter does not support the hypothesis that the robustness of cell-assemblies is based on large ranges of parameter values, producing stable behaviour.

## 6.2 Parameters specifying a standard deviation

A different picture comes up when we focus on parameters for standard deviations. In contrast to the observed narrow and fragmented ranges of parameters specifying a mean is the finding that parameters for standard deviations do have large ranges. The graph of the standard deviation of the strength of excitatory connections internal to a cell-assembly (sStrEI in Fig. 3(d)) provides a good example here. Apparently a relatively large variation in the actual parameter values of individual neurons is possible, as long as the mean of their distribution lies within a relatively narrow margin.

## 6.3 Compensatory relationships between parameters

The robustness of the chosen architecture can also depend on interactions between parameters if these have compensatory relationships. For two parameters that have such a relationship, GF can determine the area in the parameter space that produces the required behaviour, for a single GA-random-seed. If the variation of the values of both parameters in such a pair would lead to robust behaviour, one would expect to see areas of ideal behaviour that are relatively large for certain combinations of parameter values. Evidently such a compensatory relationship holds

among the parameters for decay of excitation in a single neuron ($D$) and the mean excitatory threshold in a neuron (mThrE), cf. Fig. 4(a). In addition it holds between the decay parameter and the average strength per neuron of the excitatory connections internal to a cell-assembly (mStrEI), cf. Fig. 4(b). A decrease in the decay parameter $D$ leads to a quicker build-up of excitation in the network, which can be compensated for by an increase in mThrE or a decrease in mStrEI. *Mutatis mutandis* the same holds when the values change in the opposite direction.

An inspection of the effects of pairs of compensatory parameters on the behaviour of the architecture, did not reveal the existence of large ranges of parameter values necessary for robust cognition. It can be observed in Fig. 4(a) and (b) that the width of the area of ideal behaviour remains fairly narrow under the variation of both parameters. Remarkably, the shaded areas of suboptimal behaviour—in which either a single of the ten autonomous-growth, resp. extinction, conditions fails—are nearly absent. When the effects of the same pairs of parameters are analyzed for different GA-random-seeds similar patterns of ideal behaviour emerge. These findings do not support the idea that the robustness of cognitive brain functioning rests on large ranges of parameter values.

### 6.4 A stabilizing mechanism

The results plotted in Figs. 3 and 4 raise the question whether the introduction of stabilizing mechanisms is required to make the network more robust. One candidate mechanism is a so-called arousal control mechanism (Dalenoort 1985). It is also used in many other programs for the simulation of neural networks. Such a mechanism controls the level of total excitation in the network. If this level exceeds a certain criterion value $C$, the excitation level of each neuron is lowered to an extent that produces a decrease of the total excitation level of the network to reset value $S$. With the introduction of this mechanism the excitation level of neuron $i$ will be expressed as $E'_i(t)$, cf. formulas (4) and (5), with

$0 \leq E'_i(t) \leq 1$ and in which $E_i(t)$ is defined in formula (1).

$$E'_i(t) = \begin{cases} E_i(t) - (E_{Net}(t) - S) & \text{if } E_{Net}(t) > C \\ E_i(t) & \text{if } E_{Net}(t) \leq C \end{cases} \quad (4)$$

$$E_{Net}(t) = \frac{\sum_{i=1}^{n} E_i(t)}{n} \quad (5)$$

The two parameters $C$ and $S$ in formulas (4) determining the criterion and the reset excitation level, were optimized by means of the genetic algorithm GA. The ranges from which GA could select the values to be optimized, were chosen such that activation of the arousal control mechanism would be likely. Accordingly the ranges used by GA for the parameter $C$ and $S$ were 0.20–0.55 with step size 0.05, and 0.05–0.40 with step size 0.05, respectively. For the criterion parameter $C$ the values producing ideal behaviour were distributed across the entire prespecified range. For parameter $S$, giving the level for the reset excitation, these values turned out to be 0.35 or higher within the prespecified range. In Fig. 5 the ranges of ideal behaviour are shown for the simulations, where the arousal control mechanism of formulas (4) and (5) was used. From Fig. 5 it can be concluded that the suppression of extremely high levels of activation in the network (as defined by formulas (4) and (5)) does not enlarge the parameter ranges necessary for robust behaviour of the architecture. However, there do exist other possibilities for the functioning of an arousal control mechanism, which will be examined in future research.

### 7 Discussion

We have explored the parameter space of a minimal model of cognitive brain functioning. The results of the simulation experiments indicate that the desired behaviour of the model occurs in several subspaces of the parameter space. In every subspace, however, two typical phenomena can be observed. On the one hand the behaviour strongly depends on the



Fig. 4 Areas indicating for two pairs of parameters the extent to which the required behaviour is approached (*shaded areas*, see the legend in Table 2): (a) the mean excitatory threshold of a neuron (*mThrE*) and the *decay*, and (b) per neuron the mean strength of excitatory connections internal to a cell-assembly (*mStrEI*) and the *decay*
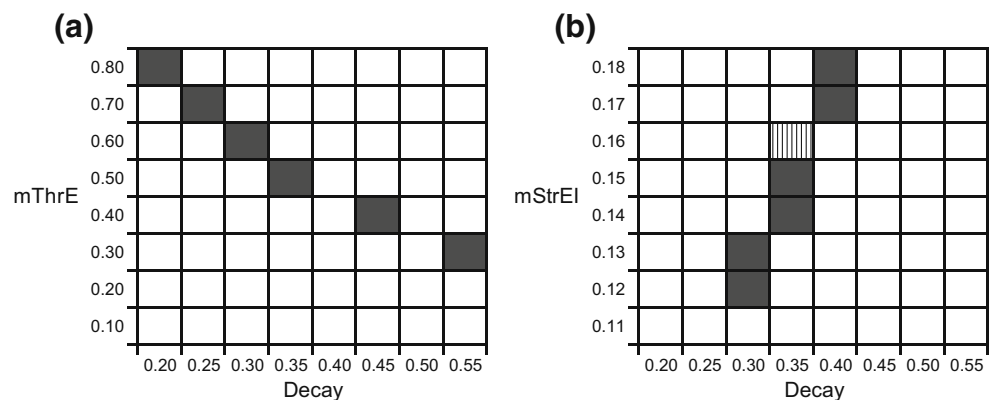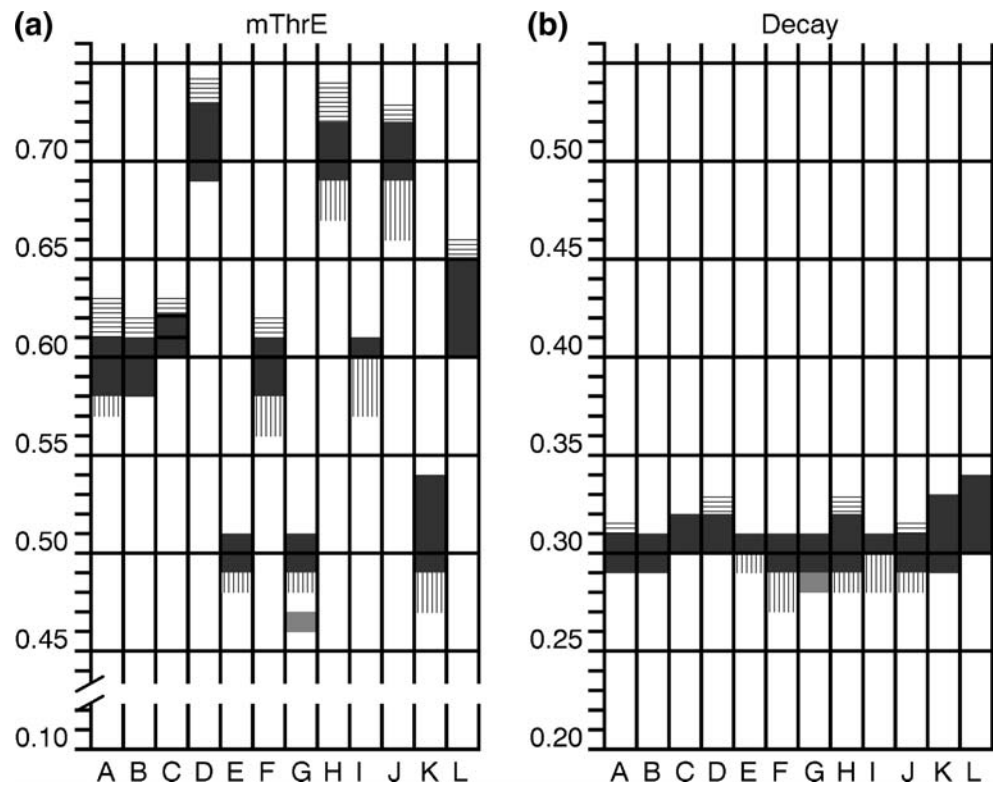
**Fig. 5** Ranges of parameter values approaching the required behaviour obtained after the introduction of an arousal control mechanism (*shaded areas*: see legend in Table 2); the graphs shown concern parameters that had relatively narrow ranges in Fig. 3: (**a**) the mean excitatory threshold of a neuron (*mThrE*), and (**b**) the decay (*D*)



selection of specific parameter values from a single narrow range, like the mean excitatory threshold, the mean strength of excitatory connections internal to a cell-assembly, and the decay. On the other hand, value ranges of parameters can be quite fragmented. For a single parameter there exist value ranges that produce adequate behaviour, arbitrarily mixed with ranges exhibiting unexpected effects. These findings indicate that the key to the solution of the robustness question is not to be found in the parameter space of the brain. In the following we will pursue some alternative answers.

### 7.1 The robustness/flexibility dilemma

Suppose that we would have found the large parameter ranges necessary for the robust behaviour of the minimal architecture. This would then raise the question of how the architecture could adapt itself to new changes in its environment or could develop new cognitive structures, e.g. corresponding to creative thought. A necessary condition for these things to happen is that the brain is capable of producing a sufficient variation of excitation patterns. The observed deviations in the autonomous-growth and extinction conditions may be part of this variation.

### 7.2 Stabilizing mechanisms

Even if we take into account that the unexpected behaviours observed in simulation experiments may have a function, the role of stabilizing mechanisms—other than the discussed arousal control system—is not excluded. We will review three candidate mechanisms, of which the first two are discussed in Turrigiano (1999) whereas the last one is a hypothesis of the authors. The three mechanisms have to be distinguished from learning mechanisms because they do not depend on any external events playing a role in learning.

*Synaptic scaling* is the regulation by cortical and hippocampal neurons of their own firing rates by scaling their synaptic inputs up or down as a function of activity. This mechanism operates relatively slowly: requiring hours or days of altered activity to modify synaptic strengths. As a solution to the robustness problem formulated this paper, it may therefore be insufficient. A mechanism that keeps the architecture within its bounds should also be able of an instantaneous reaction since sudden changes in parameter values should not disrupt cognitive functioning.

*Synaptic homeostasis* refers to the capability of a neuron to maintain relative constant firing properties although it is subject to many changes: growth, changes in shape, loss and gain of synapses, and the constant turnover of the ion-channels that determine its electrical-firing properties. The underlying mechanism probably makes use of the intracellular concentration of certain ions. A change in this concentration triggers a compensatory reaction that modifies ionic conductance such that the level of neuronal activity remains constant. Accordingly, distortions—like the failures occurring in the described simulation experiments—are immediately repaired.

However, the effects of the mechanism of synaptic homeostasis have only been found in neuro-muscular synapses.

*Synchrony of firing* implies that the spikes produced by two or more neurons are in phase. This phenomenon is relevant to the issue of robustness because phase synchrony may be a condition for the propagation of neural excitation: two neurons will only activate a third one if their spikes are in phase. Such a propagation may become robust if it takes the form of a loop, in which the spikes are interlocking. This synchronous firing of neurons could then trigger a biochemical process that compensates for changes in parameters of neural functioning. If robustness is based on synchronous firing, one would expect specific spike patterns on the presentation of a stimulus, although not every stimulus needs to have a unique pattern (see the discussion on the identity of a memory trace in the second section of this paper). In addition, repeated presentations of the same stimulus should reproduce the same spike patterns. Data compatible with this hypothesis have been reported by Fellous et al. (2004). Accordingly, synchronization could also play a role in the robustness of permanent memory structures next to its role in the temporal coupling of neuronal activity ('binding'), a hypothesis fundamental to the already cited work on synfire chains and to many neurophysiological studies such as Singer et al. (1994), Roelfsema et al. (1997), and Freiwald et al. (2001).

The list of stabilizing mechanisms presented here, is not meant to be exhaustive. Moreover, the three mechanisms on the list are not mutually exclusive and each of them requires further study. Their discussion is the offspring from the exploration of the parameter space of a minimal architecture, by means of which we have tried to lay down some important questions on cognitive brain functioning.

## Appendix

Formalization of Fitness

The genetic algorithm GA computes a fitness measure, which reflects the extent to which each of ten generated versions of the network architecture in Fig. 2 leads to:

a.  a stable oscillation in the autonomous-growth condition, and
b.  an excitation curve that approaches zero in the extinction condition.

Each of the ten versions of the network is constructed on the basis of a separate NNS-random seed in order to control for effects due to an idiosyncratic connectivity.

The degree to which the mentioned conditions are satisfied, is expressed by the fitness measure $F_i$ defined for cell-assembly $i$ in formulas (A1), (A2), (A3), (A4), (A5) and (A6) of this appendix; cell-assembly $i$ being the one that is given the external input. These formulas make use of three time points that apply to the condition of autonomous growth as well as to the extinction condition in each simulation experiment (cf. Fig. 2 b and c):

- $t_1$: beginning of the start-up condition of the simulation experiment, the neurons only produce a noise as excitation. Noise is characterized by a normal distribution with a mean of 0, and standard deviation of .05, i.e. 5% of the maximal excitation, as specified by the function $R_i(t)$ in formula (1).
- $t_2$: end of the start-up condition and the beginning of the condition of autonomous growth, resp. extinction, for each simulation experiment.
- $t_3$: end of the simulation experiment.

The fitness measure $F_i$, cf. formulas (A1), (A2), (A3), (A4), (A5) and (A6), is the sum of the fitness values for the conditions start-up, autonomous growth, and extinction, provided that the parameters under which the simulation experiment is run are such that the mean excitatory threshold is lower then, or equal to the inhibitory one. If this is not the case the sum is multiplied by 0.1. Accordingly, these implausible combinations of parameter values will produce a fitness value that rules out their selection in a set of optimal parameter values.

$$F_i = \begin{cases} \text{Start}_i + \text{Aut}_i + \text{Ext}_i \text{ if mThrE} \leq \text{mThrI} \\ 0.1 \times (\text{Start}_i + \text{Aut}_i + \text{Ext}_i) \text{ if mThrE} > \text{mThrI} \end{cases}$$

$$\tag{A1}$$

$$\text{Start}_i = 1 - F_i'(t_2) + F_i'(t_1) \tag{A2}$$

$$\text{Aut}_i = \begin{cases} \text{AvACor}_i \text{ if } F_i'(t_2) < c \times F_i'(t_1) \\ 0 \text{ if } F_i'(t_2) \geq c \times F_i'(t_1) \end{cases} \tag{A3}$$

$$\text{Ext}_i = \begin{cases} 1 - G_i'(t_3) \text{ if } F_i'(t_2) < c \times F_i'(t_1) \\ 0 \text{ if } F_i'(t_2) \geq c \times F_i'(t_1) \end{cases} \tag{A4}$$

$$G_i'(t) = \max_{k=1}^{\text{NrRdSd}}[\text{LpExc}(k, t, \text{CA}_i)] \tag{A5}$$

$$F_i'(t) = \frac{\sum_{k=1}^{\text{NrRdSd}} \text{AveExc}(k, t, \text{CA}_i)}{\text{NrRdSd}} \tag{A6}$$

Formulas (A1), (A2), (A3), (A4), (A5) and (A6) are based on aggregated variables. They have been obtained from the

excitation levels produced in the simulation of 10 different versions of the architecture (NRdSd = 10), one for each of the ten NNS-random-seeds.

The constant c is used to evaluate whether the start-up condition was terminated successfully. A successful termination means that at the end of the start-up condition the excitation level of a cell-assembly has not increased above the noise level. This restriction was formalized by means of the auxiliary fitness measure $F_i'(t)$, which determines, for the ten versions of the architecture, the excitation level in cell-assembly $i$ at time $t$ averaged over these different versions. The constant c specifies the degree to which $F_i'(t_2)$ may exceed $F_i'(t_1)$ in order for Aut$_i$, resp. Ext$_i$ to be larger than zero. In the current study $c$ was set to 1.1.

In formula (A3) the variable AvACor$_i$ expresses the autocorrelation of the excitation curve of cell-assembly $i$ averaged over the ten NNS-random-seeds in the autonomous-growth condition of the simulation experiment (where an oscillation in the excitation curve is expected).

In formula (A5) the function LpExc($k$, $t$, CA$_i$) determines the height of the last peak in the excitation level of cell-assembly $i$, constructed on the basis of random seed $k$, in the 50 time steps before time $t$. Because of the random fluctuation in the excitation level of the neurons of a cell-assembly (cf. the function $R_i(t)$ in formula (1)) such a peak will always exist. From the simulation runs with different versions of the network NRdSd measurements of this excitation value are obtained. The maximal value of them gives $G'(t)$. Accordingly, the function produces a decrease in fitness when an oscillation did occur in the extinction condition or if the excitation level did not decrease to zero.

In formula (A6) the function AveExc($k$, $t$, CA$_i$) determines the average excitation of the neurons in cell-assembly $i$, constructed on the basis of random seed $k$ at time $t$.

The programs used in the simulation experiments (the neural network simulator NNS, the genetic algorithm GA, and the generator of fitness values GF) were written in PASCAL (Delphi) and when linked to the appropriate modules for input and output, run on common operating systems (WINDOWS, Unix, Linux). Further specifications of these programs, the simulated networks, and their parameters are available upon request.

# References

Abeles, M. (1991). *Corticonics—Neural circuits of the cerebral cortex*. Cambridge: Cambridge University Press.

Abeles, M., Hayon, G., & Lehmann, D. (2004). Modeling compositionality by dynamic binding of synfire chains. *Journal of Computational Neuroscience*, 17, 179–201.

Amit, D. J. (1995). The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behavioral and Brain Sciences*, 18, 179–201.

Amit, D. J., & Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex*, 7, 237–252.

Amit, D. J., & Mongillo, G. (2003). Spike-driven synaptic dynamics generating working memory states. *Neural Computation*, 15, 565–596.

Braitenberg, V. (1978). Cell assemblies in the cerebral cortex. In R. Heim, & G. Palm (Eds.) *Theoretical approaches to complex systems* (pp. 171–188). Berlin: Springer.

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience*, 8, 183–208.

Coley, D. A. (1999). *An introduction to genetic algorithms for scientists and engineers*. Singapore: World Scientific.

D'Anguilli, & Dalenoort, G. J. (1996). Tanzi's contribution to modern neuroscience. *Brain and Cognition*, 32, 252–254.

Dalenoort, G. J. (1982). In search of the conditions for the genesis of cell-assemblies: A study in self-organization. *Journal of Social Biological Structures*, 5, 161–187.

Dalenoort, G. J. (1985). The representation of tasks in active cognitive networks. *Cognitive Systems*, 1, 253–272.

Dalenoort, G. J. (1990). Towards a general theory of representation. *Psychological Research*, 52, 229–237.

Dalenoort, G. J., & de Vries, P. H. (1995). What's in a cell-assembly? *Behavioral and Brain Sciences*, 18, 629–630.

Dalenoort, G. J. (1996). Johannes Müller and the systems approach to the problem of specificity of the memory trace. In *Proceedings of the 15th Cheiron Conference*, Leiden: Department of Psychology, Leiden University.

Dalenoort, G. J., & de Vries, P. H. (1998a). Understanding in terms of correspondences between levels of description. In C. Taddei-Ferretti, & C. Musio (Eds.) *Downward processes in the perception representation mechanisms* (pp. 497–503). Singapore: World Scientific.

Dalenoort, G. J., & de Vries, P. H. (1998b). Self-organization in the brain. In C. Taddei-Ferretti, & C. Musio (Eds.) *Downward processes in the perception representation mechanisms* (pp. 504–520). Singapore: World Scientific.

de Vries, P. H. (1995). Downwards emergence in conceptual networks. *Cognitive Systems*, 4, 233–246.

de Vries, P. H. (2004). Effects of binding in the identification of objects. *Psychological Research*, 1, 41–66.

Fellous, J. M., Tiesinga, P. H. E., Thomas, P. J., & Sejnowski, T. J. (2004). Discovering spike patterns in neuronal responses. *Journal of Neuroscience*, 24, 2989–3001.

Freiwald, W. A., Kreiter, A. K., & Singer, W. (2001). Synchronization and assembly formation in the visual cortex. *Progress in Brain Research*, 130, 111–140.

Gerstner, W., & Kistler, W. (2002). *Spiking neuron models*. Cambridge: Cambridge University Press.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*. Reading: Addison-Wesley.

Harris, K. D. (2005). Neural signatures of cell assembly organization. *Nature Reviews: Neuroscience*, 6, 399–406.

Hayon, G., Abeles, M., & Lehmann, D. (2005). A model for representing the dynamics of a system of synfire chains. *Journal of Computational Neuroscience*, 18, 41–53.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences USA*, 79, 2558.

Huyck, C. (2004). Overlapping cell assemblies from correlators. *Neurocomputing: An International Journal*, 56, 435–439.

Lashley, K. S. (1951). In search of the engram. *Symposia of the Society of Experimental Biology*, 4, 454–482.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*, 375–407.

Milner, P. M. (1957). The cell assembly: Mark II. *Psychological Review, 64*, 242–252.

Pulvermüller, F. (1996). Hebb's concept of cell assemblies and the psychophysiology of word processing. *Psychophysiology, 33*, 317–333.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22*, 253–336.

Reicher, G. M. (1969). Perceptual recognition as a function of the meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*, 274–280.

Roelfsema, P. R., Engel, A. K., Koenig, P., & Singer, W. (1997). Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature, 385*, 157–161.

Shastri, L., & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioural and Brain Sciences, 16*, 417–494.

Singer, W. (1994). Putative functions of temporal correlations in neocortical processing. In C. Koch, & J. L. Davis (Eds.) *Large-scale neuronal theories of the brain* (pp. 201–237). Cambridge, MA: MIT-Bradford.

Sougné, J. (2001). Binding and multiple instantiation in distributed networks of spiking nodes. *Connection Science, 13*, 99–126.

Tanzi, E. (1893). I fatti e le induzione nell'odierna istologia del sistema nervoso. *Rivista Sperimentale di Freniatria, 19*, 419–472.

Turrigiano, G. G. (1999). Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. *Trends in Neurosciences, 22*, 221–227.

Van Vreeswijk, C., & Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science, 274*, 1724–1726.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1*, 59–85.