

## Analytical estimates of efficiency of attractor neural networks with inborn connections

Ksenia Solovyeva<sup>1,2,3</sup>, Iakov Karandashev<sup>1,2</sup> and Witali Dunin-Barkowski<sup>1,2</sup>

<sup>1</sup>Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, 141700, Russia

<sup>2</sup>Scientific Research Institute for System Analysis, RAS, Moscow, Nakhimovskiy Prospect 36, b.1, 117218, Russia

<sup>3</sup>Moscow Technological Institute, Moscow, Leninskii Prospect 38a, 119334, Russia

**Abstract.** The analysis is restricted to the features of neural networks endowed to the latter by the inborn (not learned) connections. We study attractor neural networks in which for almost all operation time the activity resides in close vicinity of a relatively small number of attractor states. The number of the latter,  $M$ , is proportional to the number of neurons in the neural network,  $N$ , while the total number of the states in it is  $2^N$ . The unified procedure of growth/fabrication of neural networks with sets of all attractor states with dimensionality  $d=0$  and  $d=1$ , based on model molecular markers, is studied in detail. The specificity of the networks ( $d=0$  or  $d=1$ ) depends on topology (i.e., the set of distances between elements) which can be provided to the set of molecular markers by their physical nature. The neural networks parameters estimates and trade-offs for them in attractor neural networks are calculated analytically. The proposed mechanisms reveal simple and efficient ways of implementation in artificial as well as in natural neural networks of multiplexity, i.e. of using activity of single neurons in representation of multiple values of the variables, which are operated by the neural systems. It is discussed how the neuronal multiplexity provides efficient and reliable ways of performing functional operations in the neural systems.

### 1 Introduction

The analysis and multiple applications of neural networks are currently on a steep rise. Among the recent achievements in this field are: human-like recognition of images, human-level spoken words identification, machine translation between multiple images [1]. There emerges general understanding that constructing the neuromorphic devices might present the easiest way to obtain smart informational technologies (deep learning and related methods). These impressive results increased visibility of and attracted tens of thousands of scientists to the neural network research, both modern and classical. In this novel atmosphere of research of particular importance become the topics of computational neuroscience, which have not yet been included into modern neural network toolkit. Potentially, they can improve particular aspects of neural network functioning.

In present work, we complement computer simulations of [2] with a thorough mathematical analysis, which enables us to make analytical estimates of the network operational parameters.

The structure of the present note is following.

First, we consider a notion of neural attractors and their types. In this work, we restrict the analysis with stationary attractors, i.e. stable in time states of activities of neural networks. Our attention is focused on relations

between these states: the distances between the states in configurational space of the neural networks and properties of the set of attractor states of a neural network. Our contribution to the field is in an analysis of the preformed, inborn connections between the neurons in the network. Most of previous works, following John Hopfield, considered networks, in which the connections were established, using the Hebb rules [3]. That means that the connections were formed, depending on coincidence of activities of the neurons, which are to be connected. We deal with the connections, established by the mechanism, which does not depend on the neural activity. Unlike previous works, which simply postulated the existence of inborn connections with needed properties [4], we introduce in Section 2 biologically realistic details of connection setting mechanism, the method of model molecular markers. The gains, obtained with this method are analyzed in detail throughout our work.

Section 3, consisting of four subsections, presents details of the analysis. The first two subsections are devoted to the networks with  $d = 0$ , the last two – to the networks with  $d = 1$ .

The subsection 3.1 deals with evaluation of the number of attractor states for  $d = 0$ . The number to find is determined by the condition that every ‘theoretical’ or predetermined by the connection formation process

<sup>a</sup> Corresponding author: kasya726@gmail.com

attractor state is stable. The analysis yields a relation between the number of the stable theoretical attractor states and the number of neurons in the network. The analytical estimates are compared to the results of computational experiments [2].

Section 3.2 continues the analysis to get the trade-off relations for the numbers of attractor states, the number of neurons in the network and the number of active neurons in each attractor state. The results are compared with the computational experiments data [2], which are presented in Figs. 1 and 2.

In section 3.3 the important property of  $d=1$  attractors is evaluated. The overall structure of these attractors is described by the ‘snake-in-a-box’ metaphor [3]. In this presentation, the points of attractor are supposed to make the axis of a snake, which is put into the box<sup>1</sup>. In this case, the points at the snake axis, which are at a large distance from each other along the snake axis, in the configurational space of the neural network, cannot be closer to each other than some positive value, which is dubbed the ‘snake thickness’. The dependence of the distance between the points at the snake axis in configurational space on the distance between these points along the snake axis is effectively calculated in this section. The calculations are performed for the algorithm of obtaining  $d=1$  attractors, described in section 2. Again, the results are compared to computational experiments showing a good agreement.

Section 3.4 treats the problem of the upper limit of the length (the size) of a continuous  $1-d$  attractor in the network. The estimate is in accordance with the computational experiments.

As it is argued in Conclusion, the most important results of this work are: (1) a formulation of an algorithm for forming neural networks, which specificity is determined by topology model molecular markers and (2) the finding that the number of attractor points in neural networks of biologically realistic size can exceed the number of neurons in 100-1000 times.

## 2 Inborn attractors

We consider neural networks, which contain  $N$  neurons. The behavior of the  $i$ -th ( $i=1, \dots, N$ ) neuron is described by its phase function,  $\varphi_i(t)$ . There are two constants: the duration of excitation,  $w$ , and the duration of refractoriness,  $r$ . If  $\varphi_i(t) \geq r$  and  $(\sum_{j=1}^N (w_{ji} \cdot x_j(t)) - \theta_i) \geq 0$ , then  $\varphi_i(t+1) = 0$ , otherwise  $\varphi_i(t+1) = \varphi_i(t) + 1$ . The output of the neuron  $x_i(t) = 1$ , if  $\varphi_i(t) < w$ , otherwise  $x_i(t) = 0$ . Usually,  $w = 1$  and  $r = 0$ .

In computational experiments [2] we have explored the methods for making inter-neuronal connections, which enable the neural networks to have attractor states (states which are transformed into themselves under the specified above rules of neural dynamics). The states of the neural network are  $N$ -dimensional vectors of ones

(neuron is excited) and zeros (neuron is quiet). Our approach is based on using model molecular markers, which we first distribute randomly between neurons. Then connections between the neurons are established, depending on properties of markers, which the neurons receive at the previous stage. The set of all attractor states (SAAS) of the network can be characterized by pairwise distances between the states and properties of the whole set of the attractor states. Two cases are of particular importance: (1) when all attractor states are at large distances from each other. We say that these SAAS have dimensionality  $d$  equal to zero ( $d=0$ ); (2) each attractor state has two other attractor states close to it and all states compose a closed chain, in which each state has immediate neighbor, then neighbor of neighbor, etc. The distance to neighbors grows linear with the degree of neighboring relations until some limit, beyond which the distances do not systematically increase and fluctuate around some value, which is about the same as average distances between states in SAAS with  $d = 0$ . If the set of attractor states has these properties, we say that the SAAS of the neural network has dimensionality  $d = 1$ .

Let  $M$  be the total number of molecular markers. To obtain a neural network having SAAS with  $d = 0$  we divide the markers in  $n = M/L^2$  groups. Inside a group, the distance between markers is zero; between the groups, it is “very large”, say that it is close to  $N$ .

To obtain a neural network having SAAS with  $d = 1$  we endow the markers with the other metrics. We consider that all  $M$  markers are placed equidistantly on a ring. The distance between any two markers is defined as the minimum number of other markers, which are located between the selected ones.

Once the distances between markers are established, the further procedures, leading to formation of neural networks with the desirable properties of the SAAS do not depend on the type of SAAS. The markers are distributed between  $N$  neurons randomly subject to only one restriction: the distance between markers inside one neuron should exceed some pre-determined value,  $\Delta$ . As soon as the markers are distributed between the neurons, the latter are connected with excitatory synapses. The connections are symmetric and are set between two neurons, if they have markers, the distance between which is less, than another pre-determined value,  $\delta$ . Computational experiments have demonstrated that the two types of metrics marks, described above, yield the neural networks with the SAAS with  $d = 0$  and the SAAS with  $d = 1$ , when the same for both cases conditions for marker distribution and connection setting have been used [2].

## 3 Parameters estimation

In this section we obtain parameters estimation for both types of attractors, which have been described in Section 2.

<sup>1</sup> To avoid complications, connected to boundary conditions, “the snake” is considered to be a closed elastic ring, like the bike wheel inner tube.

<sup>2</sup> When applicable in this paper, the ratio expressions are considered to yield integer values. Generalization to the non-integer values is not difficult, but is not treated here.

### 3.1 Number of inborn attractor states for $d=0$

Consider  $T_{ij}$ , the neurons interconnections matrix  $N \times N$ , with values 1, at  $i$  and  $j$ , where excitatory connections are formed as described, and values 0 otherwise. The mean value of matrix elements is:

$$\gamma = \overline{T_{ij}} = \Pr\{T_{ij} = 1\} = 1 - \Pr\{T_{ij} = 0\} = 1 - \left(1 - \frac{L^2}{N^2}\right)^n \quad (1)$$

We suppose that  $n$  is proportional to  $N^2/L^2$ , i.e.  $n = q \cdot (N^2/L^2)$ . As  $n \rightarrow \infty$ , we have:

$$\gamma = 1 - (1 - q/n)^n = 1 - e^{-q} \quad (2)$$

Now, let the network have at the input one of its theoretical attractor patterns. Then, the probability, that a ‘foreign’ neuron have  $L$  excitatory inputs is:

$$P\{h_i = L\} = \gamma^L.$$

So, the probability that at least one of  $(N-L)$  foreign neurons will get  $L$  units of excitation is:

$$P_{err} = P\{\exists i_{inactive} : h_i = L\} = 1 - (1 - \gamma^L)^{N-L} \approx 1 - (1 - \gamma^L)^N$$

Taking  $P_{err} = 1/N$ , we have:

$$1 - (1 - \gamma^L)^N = 1/N,$$

and, after non-complicated transformations, leaving only first order terms (by  $1/N$ ), we have:

$$\gamma = N^{-2/L}. \quad (3)$$

Comparison of (2) and (3) finally yields:

$$q = -\ln(1 - N^{-2/L}). \quad (4)$$

Table 1 gives numerical values of  $q$  for a set of  $N$  values at  $L=20$ . Although computational estimates of  $q$  give the constant value  $q=1$ , one can see that the analytical reasoning does not diverge too far from the computational experiments estimate.

**Table 1.** Computational estimates of  $q$

$N$	100	500	1000	2000	3000	5000
$q$	1.0	0.77	0.69	0.63	0.6	0.56

### 3.2 Trade-off relations for attractors with $d=0$

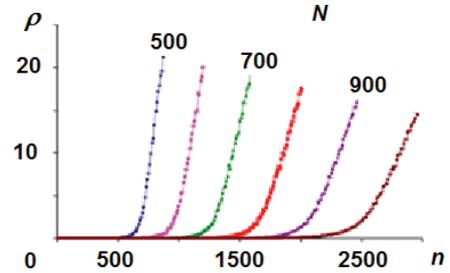
For neural network applications, it is important to know how the network behaves depending on the values of its parameters. In particular, it is important to know if the attractor states are stable, i.e. after small perturbations the states of the system returns to the original attractor state. Fig. 1 shows the ‘error’ in states, as a function of  $N$ ,  $n$  at  $L=20$ , obtained in Monte Carlo computational experiments. The synchronous  $L$  winners dynamics of the network was used [5] and the ‘error’ was considered to be the Hamming distance between the experimentally obtained stable state and the ‘theoretical’ attractor point which served as the initial condition. It can be seen that the error grows with  $n$ . For  $n$  less than a critical value,  $n_{cr}$ , the ‘theoretical’ attractor points are stable.

Fig. 2 shows dependence of  $\alpha_{cr} = n_{cr}/N$  on  $N$ . The linear empirical approximation shows that with increasing

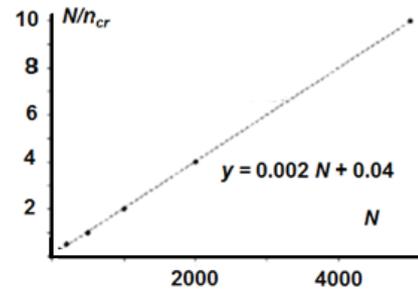
$N$ , the value of  $n_{cr}$  increases  $\approx N^2$  (when  $L$  is constant):

$n_{cr} \approx (N/L)^2$ . The mean value of matrix elements is

$$\overline{T_{ij}} \cong ML^2/N^2. \quad (5)$$



**Figure 1.** The ‘error’ in neural network stable states.  $N$  and  $n$  are the numbers of neurons and markers, correspondingly. The number of excited elements,  $L=20$ . In each case, plotted are values, averaged over all attractor points.



**Fig. 2.** The critical values of ratio  $n_{cr}/N$  as a function of  $N$ . Computational experiments (dots); broken line is the least square regression;  $L=20$ .

When the network resides at attractor state,  $S^m$ , the  $L$  neurons which are active get the following inputs:

$$h_{act} = \sum_{j=1, j \neq i}^N \mathbf{T}_{ij} S_j^m = L - 1. \quad (6)$$

Input to the rest  $(N-L)$  of the neurons is approximately:

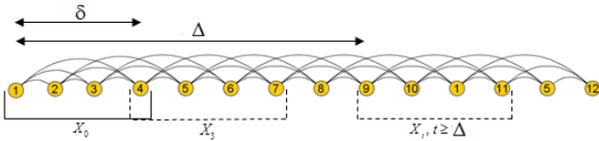
$$h_{inact} = L \cdot \overline{\mathbf{T}_{ij}}. \quad (7)$$

Here  $\overline{\mathbf{T}_{ij}}$  is the average value of matrix element of  $\mathbf{T}$ . The distinction between the right parts of (6) and (7) enables the neural network discriminate between attractor and non-attractor states. Thus, we obtain the trade-off relation  $n_{cr} \approx (N/L)^2$  for the number of attractor states,  $n$  ( $n < n_{cr}$ ), the number of neurons in the network,  $N$ , and the number of active neurons in each attractor state  $L$ . Besides, we get an analytical reasoning, which qualitatively explains the data of computational experiments, displayed in Figs. 1 and 2.

### 3.3 Distances between attractor states in 1 d continuous attractor obtained with help of model molecular markers

Fig. 3 presents a fragment of the layout of all neurons in order of the order numbers of the markers, which they

contain (as each neuron has  $k$  markers, each neuron is presented  $k$  times in this layout). The neurons become excited in the same order, while the activity propagates over the attractor<sup>3</sup>. We now consider the distance between an initial state,  $X_0$  and states which are following it,  $X_1, X_2, X_3, \dots, X_t, (t=1, 2, \dots, M)$ .



**Fig. 3.** Layout of neurons in accordance to their markers order number

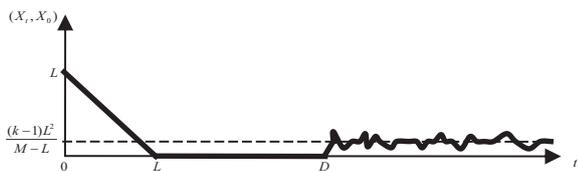
Fig. 4 shows the inner product  $(X_0, X_t)$  of two states plotted against  $t$ . It is obvious that in the beginning, the intersection between  $X_0$  and  $X_t$  decreases linearly from  $L$  to 0. Then, it stays equal to 0 up to  $t = \Delta$ . For  $t > \Delta$ , the intersection is random. Its mean value can be obtained as follows. Each neuron of the first state takes later part in  $(k-1)$  states. That means that for any neuron which is excited in  $X_0$ , the probability to be excited in  $X_t$  (for large  $t$ ) is

$$p = \frac{(k-1)L}{M-L}$$

That means that the average intersection between  $X_0$  and  $X_t$  will be  $pL \approx (k-1)L^2 / M$ . The Hamming distance between the states is connected an intersection of them according to the relation  $H(X_t, X_0) = 2(L - (X_t, X_0))$ , where from we get for  $t \geq \Delta$ :

$$\underline{D} \approx 2L \left( 1 - \frac{(k-1)L}{M} \right)$$

With  $L=15, k=3, M=900$ , we have  $\underline{D} \approx 29$ , which coincides with the results of the computational experiments [2].



**Fig. 4.** Schematic plot of the scalar product  $(X_0, X_t)$

### 3.4 Evaluation of $k_c$ in one-dimensional continuous attractors

As each neuron takes part in  $k$  attractor states, each line of the matrix  $T$  contains about  $2\delta k$  positive (equal to 1) matrix elements. The probability of positive matrix

<sup>3</sup> To study the structure of attractors with  $d=1$ , the artificial neural network dynamics has been used in [2]. In this case, the neuron threshold increases, when the neuron is active for a long time. Due to this modification of the neuron model, the activity of the network with 1d 'snake-in-a-box' attractor, can run for indefinite time over the ring of the attractor states.

element is  $2\delta k / N$ . That means that the probability of firing of one excessive neuron, which is not active in  $X_0$ , is  $(2\delta k / N)^L$ . On the contrary, the probability that this will not happen for any of the remaining  $N-L$  neurons is  $[1 - (2\delta k / N)^L]^{N-L}$ . In computational experiments, the critical value  $k_c$  was defined as the value of  $k$ , such that five random networks with a given  $k$  show perfect cycles. Thus, for  $k_c$  we obtain an equation:

$$1 - [1 - (2\delta k_c / N)^L]^{N-L} < 1/5.$$

And finally, for the critical value of  $k$ , we have:

$$k_c < N / (2\delta \sqrt[4]{5N}). \tag{8}$$

Factor  $\sqrt[4]{5N}$  is of the order of 1 and changes very slowly with  $N$ . So, (8) yields practically linear dependence of  $k_c$  on  $N$ .

In particular, we have:

at  $L = 15, N = 300, \sqrt[4]{5 \cdot 300} \approx 1.62, k = 7.7$  (in experiment,  $k = 5$ );

at  $L = 15, N = 1200, \sqrt[4]{5 \cdot 1200} \approx 1.79, k = 28$  (in experiment,  $k = 21$ ).

Thus, we have obtained analytical justification of results of computational experiments [2].

## 4. Conclusion

This work belongs to the series of analytical and computational studies of data formats in neuronal systems [2, 5, 6]. In this series we explore neural networks properties, which can be provided by the mechanisms of inborn connection formation. Here we for the first time explicitly state that using a single mechanism for connection formation we obtain two different important classes of neural networks SAAS, i.e. the attractor neural networks with  $d=0$  and the attractor neural networks with  $d=1$ . The particular class of neural network depends on the properties of the set of molecular markers. If this set is endowed with continuously changing distance function, we obtain  $d=1$ . In the discrete case, we get  $d=0$ . The model molecular markers, postulated in our approach, can be actually implemented (naturally or artificially) with bio-molecular techniques (as discussed in [2]).

Further, we have demonstrated the workability of one of the important principles of the work of natural multi-component systems – the multiple using of single elements. Thus, one can see from Fig. 2 that in attractor neural network with  $d = 0$ , each neuron can take part in up to 10 isolated points of the neural attractor. When extrapolated to the higher possible sizes of elementary neural ensembles ( $10^3$  elements) this figure may be increased up to 200. This fact has a practical value as it is well-known that in neural systems meaningful objects are represented with activities of many neurons (appropriate citations are given in [2]). If so, then the number of represented objects would be severely restricted to a number, small compared to the number of neurons in the neural system. Our simulations and analytical

calculations demonstrate that the number of attractor points in neural system can substantially exceed the number of neurons, by a factor of up to 100.

A slightly different and also important in practice is the fact that the number of attractor points in continuous neural attractors can exceed the number of neurons in the neural network. In this case the “multiplexing” factor attains the values of 21 for the number of neurons  $N = 1200$ . When, again, extrapolated to the largest possible natural neural networks we come to the factor of 900. On one side, this is natural, that representation of single values of continuous variables demands less resources than representation of discrete variables (cf. figures 900, here and 100, above). On the other side, the large number of  $d=1$  attractor states enable precise representation of the external world in neural systems.

Another important principle, demonstrated analytically and computationally in our work is the exceptional role of the random structural codes in neural systems. This principle was first revealed (based on heuristic approximate reasoning) by Sir Giles Brindley, F.R.S. [7]. Recently, results of [7] were thoroughly checked in thorough computational experiments [8].

This principle appeals to well-known high efficiency of random codes in systems of noise-resistant communications [9]. In few words, the communicational randomized systems become “large” and demonstrate “asymptotic” noise-tolerant deterministic features when the number of constituent elements is measured by hundreds and thousands (like in neural systems) and not by  $10^{20}$ , as in statistical physics.

We hope that our work might be helpful in understanding natural neural systems – brains and also can be used for synthesis of artificial neural constructions.

The work was supported by the Russian Foundation of Basic Research grant # 16-07-01059.

## References

1. Y. LeCunn, Y. Bengio and G.Hinton, *Nature*, **521**, 436 (2015)
2. K. Solovyeva, I. Karandashev and W. Dunin-Barkowski, *Front. Syst. Neurosci.*, **9**, 178 (2016)
3. W. Dunin-Barkowski and N. Osovets, *Neural Processing Letters*, **2**, 28 (1995)
4. S. Romani and M. Tsodyks, *PLoS Comput Biol*, **6**, e1000869 (2010)
5. I. Karandashev, *Proc. XV Conf. “Neuroinformatics 2013”*, part 2, p. 125 (Moscow, 2013)
6. K. Solovyeva, *Proc. XVI Conf. “Neuroinformatics 2014”*, part 1, p. 155 (Moscow, 2014)
7. G. Brindley, *Proc. Roy. Soc. Lond. B*, **174**, 173 (1969)
8. I. Karandashev and W. Dunin-Barkowski, *Optical Memory and Neural Networks*, **24**, 8 (2015)
9. G. Battail, *Lecture Notes in Computer Science*, **1133**, 76 (1996)