Do neural models scale up to a human brain?

Roman V Belavkin

Abstract—**Models of cognition generally operate either at a micro or a macro level. It is not clear, however, if the micro models can predict the macroscopic properties of biological neural systems, such as the human brain. Here, I evaluate some hypotheses about the main functions of neural processing by scaling them to higher levels. Using neurobiological literature, I estimate the numbers of inputs and outputs of the entire nervous system. Then, I apply optimal control and information theories to predict the numbers of neurons required to implement these functions. The addition of constraints on connectivity leads to numerical estimates comparable to the numbers of neurons and synapses in human brain.**

I. INTRODUCTION

Human brain operates with approximately 100 billions of neurons, which are believed to be the fundamental information processing units. Many different approaches have been applied to model the mind. Some are based on models of neurons (e.g. [1], [2], [3]); others use higher order representations to model higher level cognition (e.g. [4], [5]). Although some convergence between these two approaches can be observed, there are doubts whether the models of neurons will stand up at the macroscopic level.

Perhaps, the key to understanding the organisation and implementation of any device or a system is its function. For example, there are many possible implementations of an engine, all of which, however, fulfil the same function. The differences in implementations are due to some additional constraints. Likewise, different animals have different organisations of nervous systems. These may be due to different habitats, senses and other constraints. However, the main function is most likely the same.

The unified theories of cognition approach due to [4] was proposed as the way to test different theories in one computational architecture. Some implementations have been very successful in modelling various psychological phenomena [6]. Yet, these models largely rely on symbolic representations of higher level cognition, and it is not clear how can they be represented by neurons. The research in artificial neural networks, on the other hand, has often been driven by engineering and signal processing problems resulting in systems with little connection to biological systems.

Recently, there has been progress in using biologically inspired neural architectures to simulate symbolic processing and higher level cognition [7]. The CABOT project, currently in progress, aims at incorporating elements of natural language, symbol grounding and planning in a system, based entirely on cell–assemblies of fatiguing leaky integrate–and–fire neurons [8]. Although this project may prove the sufficiency of using current neural models for implementing some higher order cognitive tasks, it is not clear how to test which models are necessary.

In this paper, we propose the idea and report the results of testing the models of neural function against macroscopic physiological data of human brain. In the next section, we overview some basic facts about the organisation of human nervous system, and use them to estimate the total number of inputs and outputs of the central nervous system (CNS). In the following two sections, we investigate two functions of neural processing: Optimal control and optimal coding of information. It will be shown that although standard mathematical neural models correspond well to these functions, a fully connected system would require significantly fewer units than the number of neurons in human CNS. Several mathematical relations are derived in order to account for partial connectivity in neural networks. When these relations are used with the constraints on connectivity based on the average number of synapses of human CNS, then our estimates for both the optimal control and optimal coding functions are comparable with the corresponding parameters of human CNS. The influence of additional constraints on particular implementation of these functions is discussed. The paper concludes by discussing the limitations, future directions and possible applications of this research.

II. COUNTING THE AFFERENT AND EFFERENT FIBRES

In this section, we shall overview some main facts about the nervous system, its organisation and estimate the total numbers of *afferent* (input or ascending) and *efferent* (output or descending) fibres of the CNS. The main objective is to estimate the dimensionality of the input and the output spaces, in which CNS operates. These estimates will help us validate our hypotheses about the function of the CNS.

A. Overview of Human Nervous System

The nervous system of human, as well as of other vertebrates, is usually organised into the central nervous system (CNS), which consists of the brain and the spinal cord, and the peripheral nervous system (PNS), which is the complement of CNS, and it consists of the somatic and autonomic nervous systems (Table I). PNS is responsible for collecting all the sensory information and sending all the control signals to the body, which include voluntary actions, sympathetic and parasympathetic processes. Thus, most of this (if not all) information is processed by the CNS.[1]

Roman V. Belavkin is with School of Computing Science, Middlesex University, London NW4 4BT, UK (phone: +442084116263; fax: +442084116943; email: R.Belavkin@mdx.ac.uk).

[1]Enteric system contains approximately one billion neurons, which operate using the same neurotransmitters as in the brain. It is only connected with CNS via the vagus nerve and can operate quite independently of the rest of the nervous system. Its role is not yet well understood.

Central (CNS)	Peripheral (PNS)
Brain (10^{11} neurons) • Forebrain • Midbrain • Hindbrain Spinal cord (10^9)	Somatic voluntary control Autonomic (ANS) • Sympathetic (fight or flight) • Parasympathetic (rest and digest) • Enteric (10^9)

Nerve:	Afferent (IN)	Efferent (OUT)	Fibres
olfactory	smell		$1.2 \cdot 10^7$
optic	vision		$1.2 \cdot 10^7$
vestibulocochlear	hearing, balance		$3.1 \cdot 10^4$
oculomotor		eye, pupil size	$3 \cdot 10^4$
trochlear		eye	$3 \cdot 10^3$
abducens		eye	$3.7 \cdot 10^3$
hypoglossal		tongue	$7 \cdot 10^3$
spinal-accessory		throat, neck	?
trigeminal	face	chewing	$8.1 \cdot 10^3$
facial	2/3 taste	face	10^4
glossopharyngeal	1/3 taste, blood pressure	throat, soliva secreation	?
vagus	pain	heart, lungs, abdominal, throat	?

CNS is separated from the rest of the body by three layers of tissue, called meninges, which make CNS largely insulated from the rest of the body [9]. The main 'interface' between CNS and its environment (the body) is provided by nerves, which carry all the afferent and efferent fibres between CNS and PNS. The information flow is represented on the diagram below

$$\text{PNS} \longrightarrow (\text{inputs})^m \, \text{CNS} \, (\text{outputs})^n \longrightarrow \text{PNS}$$

where m is the total number of afferent, and n is the total number of efferent fibres. Let us also denote by S the total number of neurons in the CNS.

The brain is by far the largest collection of neurons in the body with some estimates on the order of 10^{11} neurons, while the spinal cord contains approximately 10^9 neurons, many of which aggregate and relay the information into and from the brain [10]. The brain is likely to fulfil the majority of information processing and control functions in the body.

In this paper, we shall consider mathematical models of a neuron to hypothesise about the major function of the CNS. By using the estimates of m, n and S, we can test whether these functions can be scaled up to the entire CNS.

B. Overview of the Connections

Although many sources in the neurological literature suggest that the number of afferent fibres to the CNS is greater than the number of efferent, unfortunately, we were unable to find the estimates of these figures. Thus, in order to estimate m and n, we had to survey the literature for the data about the number of afferent and efferent fibres in all the nerves connecting CNS with PNS, and then sum these estimates together.

There are 12 pairs of cranial nerves that connect directly to the brain, and 31 pair of spinal nerves that connect to the spinal cord. The majority of the nerves carry both afferent and efferent fibres. The data about the number of fibres varies across different sources and is quite noisy. In some cases, we could not find the exact figures and had to rely on estimates, an improvement of which is desirable. However, our current estimates, although quite noisy, turned out to be very useful as a starting point of this research.

1) Cranial Nerves (12 pairs): Table II provides the summary of afferent and efferent fibres in cranial nerves together with the approximate number of fibres. One can see that there

are three nerves with only afferent, five with only efferent and four nerves with mixed fibres. The optic and olfactory nerves account for the majority of afferent fibres among cranial nerves, and they provide information for the most complex perceptual systems in the body.

Unfortunately, we could not find information regarding some of the nerves, and had to use estimations. For example, we are still looking for a good count of fibres in the spinal-accessory nerve. Based on the counts of fibres in the other four efferent–only cranial nerves, our estimate for the spinal-accessory nerve is $\frac{1}{4}(30 + 3 + 3.7 + 7) \cdot 10^3 = 1.09 \cdot 10^4$. Similarly, the estimates for glossopharyngeal and vagus are based on the other two mixed cranial nerves, and they are $\frac{1}{2}(8.1 + 10) \cdot 10^3 = 9.05 \cdot 10^3$.

The other information we did not find was the ratio between afferent and efferent fibres in the mixed nerves. One way to estimate this ratio is to take into account the ratio between the numbers of afferent and efferent fibres in all other cranial nerves. However, the olfactory and the optic nerves have significantly higher numbers of fibres than any other nerves [10], [9], and this method would probably result in a very biased estimate. For this reason, we decided to estimate that half of fibres in a mixed nerve are afferent and the other half is efferent. Below are the estimates of the total numbers of afferent and efferent fibres in 24 cranial nerves:

$$m_c \approx 4.81 \cdot 10^7, \qquad n_c \approx 1.45 \cdot 10^5$$

2) Spinal Nerves (31 pairs): Spinal nerves are both sensory and motor. Thus, each spinal nerve carries both afferent and efferent fibres. Table III shows the summary of all 31 pairs.

Unfortunately, we could not find good estimates of the number of fibres in the spinal nerves, and therefore had to use estimates. There are 10^9 neurons in the spinal cord. The afferent sensory fibres from the nerves connect to one second and then to two third order spinal neurons, which then project afferent fibres into the brain. Thus, there at least three neurons in the spinal cord for each afferent fibre. The

TABLE III
ORGANISATION OF SPINAL NERVES.

Nerves:	Number
cervical	8
thoracic	12
lumbar	5
sacral	5
coccyx	1

efferent fibres in the spinal nerves are axons of the α and γ motor neurons, which receive efferent fibres directly from the brain. Given the numbers of second, third and motor neurons in the spinal cord, one could estimate the numbers of afferent and efferent fibres in the spinal nerves. Unfortunately, we had insufficient data to be satisfied with this method at time of writing this text. Here, we provide the estimates based on the mixed cranial nerves:

$$m_s = n_s \approx 2 \cdot 31 \cdot 4.5 \cdot 10^3 = 2.8 \cdot 10^5$$

C. The Estimates

Adding up the results for cranial and spinal nerves, we can now estimate the total numbers of afferent and efferent fibres of the CNS:

$$m = m_c + m_s \approx 4.84 \cdot 10^7$$
$$n = n_c + n_s \approx 4.26 \cdot 10^5$$

These estimates are based on rather imprecise data, and they are clearly very noisy and not final. Nevertheless, they provide some important information, which will be used in this paper. Firstly, the number of afferent fibres, m, does indeed exceed greatly the number of efferent, n. Secondly, m, which is the dimensionality of the input space of the CNS, is much greater than the number of synapses, k, of the average neuron (estimated $k \approx 10^3 - 10^4$). Thus, CNS is not a fully connected network, and each neuron operates in the space of a much smaller dimensionality as compared to the CNS. Thirdly, the number of efferent fibres, n, is much smaller than the total number of neurons in the CNS.

Some additional data about internal structures of CNS can be evaluated using the above estimates of m and n. For example, the number of fibres in corpus callosum, which connects the left and right cerebral hemispheres, is approximately $2.5 \cdot 10^8$. Interestingly, half of this number is quite comparable with our estimate of m. One might speculate that each hemisphere receives m inputs and then relays some information to the other hemisphere via corpus callosum. Another interesting fact is the number of fibres in pyramidal decussation, which are the motor fibres that pass from the brain to medulla. This number is approximately $1.1 \cdot 10^6$, and it is also comparable with the our estimate of number of outputs from the spinal nerves, n_s.

III. OPTIMAL ESTIMATION AND CONTROL

Let us consider the assumption that CNS implements optimal estimation and control of the body. Let us assume

that X is a set of all states of the environment, and there is a preference relation ordering these states (total preorder). This order can be induced by some non–constant numerical cost function $c : X \rightarrow \mathbf{R}$. If states $x \in X$ are observed indirectly through some measurements $y \in Y^m$, then the optimal function $u : Y^m \rightarrow U^n$ is found by minimising the conditional expected cost

$$
\begin{aligned}
u^*(y) &= \arg\min_{u(y)} E\{c(x, u(y)) \mid y\} \\
&= \arg\min_{u(y)} \int c(x, u(y)) P(\mathrm{d}x \mid y)
\end{aligned}
$$

where minimisation is done over all functions $u(y)$. The choice of the optimal function depends on the cost functional, $c(x, u)$. For example, the optimal estimation function minimising quadratic error $c(x, u) = |x - u(y)|^2$ is conditional expectation $u^*(y) = E\{x \mid y\}$. Its linear approximation is the linear mean–square regression

$$E\{x \mid y\} \approx E\{x\} + B^T (y - E\{y\}), \qquad (1)$$

where B is a matrix of the regression coefficients, which depend on the correlations between different components of vectors x and y.

Linear algorithms are known to be optimal for Gaussian distributions [11], [12], and in this case, linear regression defines the optimal hyperplane minimising the quadratic distance between x and its estimate $u(y)$. Similar optimal linear algorithms can be formulated for control problems. Their implementation requires only the knowledge of the first and second order statistics. Optimal non–linear algorithms, in general, rely on higher order statistics. However, algorithms optimal in the sense of maximum probability can be significanly simplified [13]. Interestingly, some popular models of a neuron and neural networks have great similarities with the optimal linear algorithms.

A. Neural Linear Transformations

The very first mathematical model of a neuron due to [1] is a function of the inner product between the weights the input vectors:

$$u(y) = \varphi \left(a + w^T y \right) = \varphi \left(a + \sum_{i=1}^{k} w_i y_i \right)$$

where a is the threshold (bias), y is the input vector, w is the weight vector, and k is the number of inputs (synapses) of a neuron. Often, the activation function, φ, is the Heaviside threshold function. Notice the similarities between the argument of φ in the above model and linear regression (1). The bias, a, can be seen as the 'intercept' coefficient of the linear transformation, $E\{x\} - B^T E\{y\}$; the weights, w, can be seen as the regression coefficients, b. Furthermore, the latter association is supported by many theories of synaptic learning.

The idea, proposed by Hebb [2], that synaptic weights increase when neurons 'co–fire' was extended into the covariance learning rule [14], [3], according to which the

weights are learnt as correlations between the pre and post synaptic activities. Similar ideas can be found in the works on principal or independent components analysis using neural nets [15], [16] as well as the self–organising maps [17], which linearly transform signals into lower dimensional spaces minimising some cost functions (e.g. quadratic or Kullback distance). Furthermore, it is well known that multilayer networks with a large number of linear units can be used to approximate non–linear functions. Let us investigate further these similarities between neural models and the optimal estimation and control algorithms. For this, however, we shall need to develop some additional mathematical apparatus.

B. Constraints on Connectivity

Suppose that a system of artificial neurons implements a control algorithm transforming m–dimensional input signal $y \in Y^m$ into an n–dimensional control signal $u \in U^n$. A linear algorithm can be implemented using a single layer percerptron consisting of n neurons with m synapses each. Indeed, such a network would implement a linear transformation, and if the weights matrix of this network equals to the regression coefficients matrix, B, then the system will be optimal for Gaussian input and quadratic cost. Thus, according to the previous section, such a system with m and n equivalent to human CNS would only require $n \approx 4.26 \cdot 10^5$ motor neurons connected directly to $m \approx 4.84 \cdot 10^7$ input nodes. This, obviously, contradicts the fact that there are billions of neurons in the CNS between the inputs and the outputs (approximately 10^{11} in the human brain).

In the 'equivalent' single layer system, however, the output neurons would have $m \approx 4.84 \cdot 10^7$ synapses, which exceeds by far the estimated number of synapses of an average neuron in the human brain (between 10^3 and 10^4, or $k \approx 4 \cdot 10^3$, [10], [9]). Furthermore, each of the sensory neurons would have to be connected to $n \approx 4.26 \cdot 10^5$ neurons in the output layer. Thus, a single layer system cannot account for the numerical data about human CNS.

Let us consider a system with m input and n output nodes, but with a constraint that the average number of synapses cannot be larger than k. Thus, n output nodes can be connected to maximum $n \times k$ nodes. If $n \times k < m$, then there must be nodes between the input and the output layers, which we shall refer to as hidden nodes.

Let us denote by $i = 1, 2, \ldots$ the distance of a node from the output nodes, measured in the number of forward connections. Thus, $i = 0$ corresponds to the output nodes, and $i = 1$ corresponds to all nodes that connect directly to the output nodes. In analogy with feed–forwards networks, i is the number of a hidden layer starting from the output layer.

Let r_i be the number of nodes in the ith layer. It is clear that $\max r_{i+1} = r_i \times k$ can only be achieved if each of the nodes in layer $i + 1$ is connected to a different single node in layer i. In other words, each of the nodes in ith layer is connected to a different set of k nodes in layer $i + 1$. In this case, layer i consists of r_i independent nodes, and the input space of layer i is an r_{i+1}–dimensional space consisting of r_i disjoint k–dimensional subsets.

Suppose that the k–dimensional subsets of the r_{i+1}–dimensional input space are not disjoint. This means that the nodes in layer r_{i+1} are connected on average to $h > 1$ nodes in layer i. By noticing that the outgoing connections from layer $i + 1$ to i and connections received by layer i from $i + 1$ are the same, one can write $r_{i+1}h = r_ik$, and the following recurrent relation holds

$$r_{i+1} = r_i \frac{k}{h}$$

For a system with l hidden layers, m inputs and n outputs, we have

$$r_0 = n\,, \quad r_i = n \left(\frac{k}{h} \right)^i\,, \quad r_{l+1} = m \qquad (2)$$

Thus, $m \left(\frac{h}{k} \right)^{l+1} = n$, and the number of hidden layers can be computed as follows

$$l = \frac{\ln m - \ln n}{\ln k - \ln h} - 1 \qquad (3)$$

We shall call l the *order of connectivity*. Note that because $l \geq 0$ and $m \geq n$, it follows from (3) that $k \geq h$. Moreover, $l \leq 0$ when $\frac{h}{k} \leq \frac{n}{m}$. Note that $l = 0$ corresponds to a single layer percerptron, discussed above. The order of connectivity, l, increases infinitely as $\frac{h}{k} \to 1$:

$$\lim_{\frac{h}{k} \to 1} \frac{\ln m - \ln n}{\ln k - \ln h} - 1 = \infty$$

Let us denote by S the total number of hidden nodes. Using equations (2) and (3), it follows that

$$S = \sum_{i=1}^{l} r_i = n \sum_{i=1}^{l} \left(\frac{k}{h} \right)^i = m \sum_{i=1}^{l} \left(\frac{h}{k} \right)^i \qquad (4)$$

Note that l and S do not depend on the values of h and k explicitly, but rather on the $\frac{h}{k}$ ratio. However, since both h and k are natural numbers, maximisation of $\frac{h}{k}$ and minimisation of k suggests that $h = k - 1$.

Using estimates of m and n from the previous section, let us compute the $\frac{h}{k}$ ratio such that the total number of neurons in hidden layers, S, is approximately the same as in human CNS (10^{11}). For $\frac{h}{k} = .9995$, equations (3) and (2) give us

$$l = 9461\,, \qquad S = 0.96 \cdot 10^{11}$$

Thus, the $\frac{h}{k}$ ratio in this model has to be very close to one in order to achieve the number of hidden neurons comparable to that of a human brain. Also, we observe a very high order of connectivity. Interestingly, because h and k are natural numbers, the minimal value of k that can achieve ratio $\frac{h}{k} = .9995$ ($h = k - 1$) is $k = 2 \cdot 10^3$, which matches well the average number of synapses in the human brain (between 10^3 and 10^4).

These results, however, should be carefully interpreted. First, the model used to derive equations (3) and (2) assumed the average numbers of synapses, k, and postsynaptic

connections, h. In biological systems, these quantities have quite a high variance. The model also took into account only forward connections, and the feedback connections are not considered. This is, of course, an idealisation. Perhaps, the results, obtained above, should be treated as the estimates of the average values of the parameters, which may have a high variance.

In the conclusion of this section, let us discuss the validity of linear models. Indeed, for non–Gaussian distributions, linear algorithms are not optimal, and it is unlikely that the input of the CNS is Gaussian. However, as discussed earlier, a multilayer system with high number of linear units can successfully approximate a non–linear function. Furthermore, any distribution can be approximated as a mixture of Gaussians with small variance (indeed, the Dirac δ–function can be seen as a Gaussian with zero variance). Also, because for Gaussian distributions the expected value corresponds to the maximum of probability, such neural networks can be nearly optimal in the sense of maximum likelihood. Thus, partially connected multilayer networks, such as human CNS, can be applied to much wider range of problems than a simple linear optimisation.

IV. Optimal Coding

In this section, we consider neural systems from the information theoretic point of view. Let us consider the input vector, $y \in Y^m$, and the output vector, $u \in U^n$, as random variables. Thus, a neural system, such as CNS, can be viewed as a function of random variable $u(y)$. If cardinalities $|Y|$ and $|U|$ are the numbers values each input and output neuron can transmit (e.g. if one assumes a binary model of neural communication, then $|Y| = |U| = 2$) , then $|Y|^m$ and $|U|^n$ are the cardinalities of the input and the output spaces.

In general, only an isomorphic mapping can transform a random variable without information loss. However, $u(y)$ is not an invertible function if $|U|^n < |Y|^m$, and some information is lost. However, we know from information theory that under some conditions on y, it is possible to find an optimal code with an alphabet of size $|U|^n < |Y|^m$ and no loss of information.

This condition is entropical stability, which is defined as follows. Let $H(y) = -\ln P(y)$ denote the random (or Hartley) entropy, and $H_y = -E\{P(y)\}$ denote the expected (or Boltzmann) entropy. A family of random variables $\{y_i\}$ is called *entropically stable* if the ratio $H(y_i)/H_{y_i}$ converges in probability to one as $i \to \infty$. It can be shown that many random processes are entropically stable [13] (e.g. any Markov process with constant transition probabilities).

A very deep result of information theory is the theorem about asymptotic properties of entropically stable processes. This theorem [13] suggests that if $\{y_i\}$ is entropically stable, then the set of all possible realisations of y_i can be divided into two disjoint subsets A_i and B_i such that as $i \to \infty$

1) $P(A_i) \to 0$;
2) $P(y_i)/P(y_i') \to 1$ for y_i, $y_i' \in B_i$;
3) The number of realisations in B_i approaches $e^{H_{y_i}}$.

Thus, for large m, some realisations of entropically stable $y \in Y^m$ are extremely rare, while some realisations are more common and have asymptotically similar probabilities. Furthermore, the number or realisations of these, more probable y can be approximated by e^{H_y}. Let us apply these general ideas to a neural system with m inputs and n outputs.

Let H_y and H_u denote the expected entropies of the input and the output vectors. If y has uniform distribution, then $H_y = \ln |Y|^m$, which is the maximum entropy; otherwise, $H_y < \ln |Y|^m$. If input is an entropically stable process, then we only require to encode $e^{H_y} < |Y|^m$ realisations of y. Moreover, these realisations have asymptotically equal probabilities. Thus, an optimal code can be implemented using $|U|^n = e^{H_y}$ uniformly distributed realisations of u. If $|U| = |Y|$, then $n < m$.

Interestingly, many algorithms, based on artificial neural networks, implement transformations, such as discussed above. Indeed, the independent component analysis (ICA) is a transformations of y into u, such that the components of u are statistically independent. This means minimisation of Kullback divergence between joint probability $P(u)$ and the product of marginal probabilities $P(u_1) \times \cdots \times P(u_n)$. This condition can be written in terms of entropies as follows

$$u^*(y) = \arg\min_{u(y)} \left(\sum_{i=1}^{n} H_{u_i} - H_u \right)$$

where minimisation is done over all functions $u(y)$. Some ICA algorithms are based on minimisation of the sum of marginal entropies [18]. However, one can see from above that minimisation can also be achieved by maximising H_u, which results in almost uniform distribution of u. Thus, ICA algorithms can be used for optimal coding.

Neural implementations of linear ICA require only one layer fully connected networks. This means n neurons with m synapses each. As in previous section, we observe a discrepancy with biological systems, where many more neurons are used with fewer synapses. Let us discuss the implications of the upper bound constraint on the number of synapses.

Each neuron with k synapses transforms k–dimensional random vector into a scalar (one axon). This constitutes to an average information loss of dimension $k - 1$. More information can be communicated using additional neurons, connected to the same input. Thus, a system of $h = k$ neurons, connected to the same input, can communicate with no loss of information. In previous section, we estimated that $\frac{h}{k} = .9995$, which suggests almost perfect communication.

A neural system of S units has the capacity to communicate $|U|^S$ realisations. However, there are h units on average with the same input. Also, if each unit is connected to h other units, then these h units receive the same information. Thus, the total number of realisations should be reduced to $|U|^{S/h}$. Perfect communication of information to the output through a channel of S hidden units suggests that the channel should have the capacity for $|U|^m$ realisations. Thus,

$$m = \frac{S}{h}$$

One can see from above that S is an increasing function of h. If the number of synapses, k, has to be minimised, while the ratio $\frac{h}{k}$ maximised, because h and k are natural numbers, it follows that $h = k - 1$. Using our estimates from previous section for $m \approx 4.84 \cdot 10^7$ and $h \approx 2 \cdot 10^3$, we can estimate

$$S \approx 0.97 \cdot 10^{11}$$

The above result is quite close to our previous calculations, even though a different approach has been used.

V. CONCLUSIONS

In this study, we attempted to hypothesise about the main function of nervous system and the brain, and evaluate these hypotheses by using numerical data about human CNS and the estimates of the input and output spaces it operates in. It is unfortunate that the data and the estimates are quite noisy and imprecise, and we hope that better approximations will soon be available. However, current figures allowed us to develop and test some preliminary ideas.

Two main mathematical models have been evaluated: Optimal control and optimal information coding. Both functions can be implemented in artificial neural systems very different from their biological counterparts. For example, both the optimal linear algorithms and optimal communication can be implemented using a simple percerptron. Although some advantages of introducing hidden nodes have been known for decades (e.g. non–linear separation), we found that a constraint on the number of synapses also leads to a solution with hidden nodes. In this case, however, the network is not fully connected.

Several relations have been derived for networks with bounded connectivity. In these relations, the ratio between the numbers of post and pre synaptic connections, $\frac{h}{k}$, plays an important role. Systems with ratios close to one can implement functions similar to single layer transformations with no constraints on connectivity. Using two different approaches, we obtained very similar results, which predict the estimated number of neurons in human CNS as well as the average number of synapses of a neuron.

Our estimates and computations support the idea that CNS implements an optimal control and encoding of information. The redundancy in the input allows for a dimensionality reduction of the control signal. Thus, the system makes an optimal transformation from Y^m to U^n, but with additional constraints on connectivity.

Minimisation of the number of synapses is probably justified, as it may reduce the number and the length of connections, and possibly the size and energy requirements of each cell in the nervous system. In addition, each neuron operates in a subspace of a lower dimensionality, which simplifies the computations. However, the increase of the $\frac{h}{k}$ ratio, which reduces information loss, sets the lower bound on the number of synapses $k = h+1$. In addition, we observe an increases of the number of hidden neurons, S. Such systems can be applied for non–linear problems. The total number of hidden neurons, however, is bounded above due to the physical constraints. Thus, the topology and organisation of nervous systems, observed in nature, can be due to a complex multicriteria optimisation.

The mathematical models, presented here, should not be translated literally to biological systems. They, however, introduce the ideas about how models of a single neuron can be scaled up, even to levels as high as human CNS. The macroscopic data about biological systems can be used to evaluate these microscopic models and reason about the higher level functions using mathematical models. Understanding of how these functions are implemented in biological systems as a result of evolution, can help us in the design and applications of artificial systems.

REFERENCES

[1] W. McCulloch and W. Pitts, "A logical calculus of ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–133, 1943.

[2] D. O. Hebb, "Drives and the C.N.S. (conceptual nervous system)," *Psychological Review*, vol. 62, pp. 243–254, 1955.

[3] T. J. Sejnowski and P. K. Stanton, "Covariance storage in the hippocampus," in *An introduction to neural computing*. ch. 19, pp. 365–377, Academic Press, 1995.

[4] A. Newell, *Unified theories of cognition*. Cambridge, Massachusetts: Harvard University Press, 1990.

[5] J. R. Anderson and C. Lebiere, *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum, 1998.

[6] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An integrated theory of the mind," *Psychological Review*, vol. 111, no. 4, pp. 1036–1060, 2004.

[7] C. Huyck and R. V. Belavkin, "Counting with neurons, rule application with nets of fatiguing leaky integrate and fire neurons," in *Proceedings of the Seventh International Conference on Cognitive Modeling*, D. Fum, F. D. Missier, and A. Stocco, Eds. Trieste, Italy: Edizioni Goliardiche, April 2006.

[8] C. Huyck, "Cell assemblies and neural network theory: From correlators to cell assemblies," Middlesex University, Tech. Rep. ISSN 1462-0871 CS-02-02, 2002.

[9] R. Poritsky, *Neuroanatomy: A Functional Atlas of Parts and Pathways*. Hanley & Belfus Inc., 1992.

[10] M. F. Bear, B. W. Connors, and M. Paradiso, *Neuroscience: Exploring the Brain*, 3rd ed. PA: Lippincott Williams & Wilkins, 2007.

[11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transaction of the ASME Journal of Basic Engineering*, vol. 82, pp. 35–45, March 1960.

[12] R. L. Stratonovich, "Conditional Markov processes," *Theory of Probability and its Applications (USSR)*, vol. 5, no. 2, pp. 156–178, 1960.

[13] ——, *Information Theory (in Russian)*. Moscow, USSR: Sovetskoe Radio, 1975.

[14] T. J. Sejnowski, "Storing covariance with nonlinearly interacting neurons," *Journal of Mathematical Biology*, vol. 4, pp. 303–321, 1977.

[15] E. Oja, "A simplified neuron model as a principal component analyzer," *Journal of Mathematical Biology*, vol. 15, pp. 267–273, 1982.

[16] A. Hyvärinen and E. Oja, "Independent component analysis by general nonlinear Hebbian–like learning rules," *Signal Processing*, vol. 64, pp. 301–313, 1998.

[17] T. Kohonen, "Self–organized formation of topologically correct feature maps," *Biological Cybernetics*, vol. 43, pp. 59–69, 1982.

[18] E. G. Learned-Miller and J. W. F. III, "ICA using spacings estimates of entropy," *Journal of Machine Learning Research*, vol. 4, pp. 1271–1295, 2003.