Spontaneous Scale-free Structure of Spike Flow Graphs in Recurrent Neural Networks

Filip Piękniewski [a,*], Tomasz Schreiber [a]

[a] *Nicolaus Copernicus University, Faculty of Mathematics and Computer Science, ul. Chopina 12/18, 87-100 Toruń, Poland*

Abstract

In this paper we introduce a simple and mathematically tractable model of an asynchronous spiking neural network which to some extent generalizes the concept of a Boltzmann machine. In our model we let the units contain a certain (possibly unbounded) charge, which can be exchanged with other neurons under stochastic dynamics. The model admits a natural energy functional determined by weights assigned to neuronal connections such that positive weights between two units favor agreement of their states whereas negative weights favor disagreement. We analyze energy minima (ground states) of the presented model and the graph of charge transfers between the units in the course of the dynamics where each edge is labelled with the count of unit charges (spikes) it transmitted. We argue that for independent Gaussian weights in low enough temperature the large-scale behavior of the system admits an accurate description in terms of a winner-take-all type dynamics which can be used for showing that the resulting graph of charge transfers, referred to as the *spike flow graph* in the sequel, has scale-free properties with power law exponent $\gamma = 2$. Whereas the considered neural network model may be perceived to some extent simplistic, its asymptotic description in terms of a winner-take-all type dynamics and hence also the scale-free nature of the spike flow graph seem to be rather universal as suggested both by a theoretical argument and by numerical evidence for various neuronal models. As establishing the presence of scale-free self-organization for neural models, our results can also be regarded as one more justification for considering neural networks based on scale-free graph architectures.

Key words: Scale-free graph, stochastic neural network, winner-take-all dynamics

1. Introduction

The concept of a scale-free network has gathered a lot of attention in recent years providing a unified description of a wide variety of complex network topologies displaying the evidence of strong structuring principles co-existent with a considerable degree of randomness, see (Albert and Barabási, 2002) for a comprehensive survey. A distinctive feature of a *scale-free* network is that the degree distribution of its nodes follows a power law, thus lacking a characteristic scale in the language of statistical mechanics, whence the name. The presence of such power laws has been observed for a broad class of networks, prominent examples including the World Wide Web (Albert et al., 1999), science collaboration networks (Barabási et al., 2002), citation networks (Redner, 1998), ecological networks (Montoya and V., 2002), linguistic networks (i Cancho and Solé, 2001) as well as cellular metabolic networks (Jeong et al., 2000) and many other ones, see (Albert and Barabási, 2002). Many instances of structuring principles resulting in scale-free networks have been proposed, with a prominent collection of examples stemming from Barabási-Albert model (Barabási and Albert, 1999; Albert and Barabási, 2002) and its variants modeling a variety of scale-free networks with different power law exponents by exploiting the fundamental properties of growth and preferential attachment.

Recently, considerable interest has been attracted by neural networks built on scale-free graphs and it turned out that a hierarchical scale-free network architecture is in many cases beneficial for efficiency of neuronal information processing. A scale-free graph is relatively sparse, and so the memory needed to store a neural network built on such a graph, as well as computational effort required to provide certain tasks are significantly reduced, see (Perotti et al., 2006; Stauffer et al., 2003). In this context it is natural to ask whether these advantages are reflected in some mech-

* Corresponding author.
Email addresses: philip@mat.umk.pl (Filip Piękniewski), tomeks@mat.umk.pl (Tomasz Schreiber).

anisms inherent to the usual recurrent neural network dynamics and resulting in emergence of power laws. In general this does not seem to be necessarily the case for neural networks with simple processing units, for instance the neural network of C.elegans worm exhibits rather exponential decay (Amaral et al., 2000; Koch and Laurent, 1999). Numerical simulation based on the simple model by E. Izhikevich (Izhikevich, 2003) also did not reveal any scale-free structure, as long as the processing units were single neurons (Piekniewski, 2007). However, things are very different when more complicated individual unit architecture is assumed, in which case a single *formal neuron* can be interpreted as modeling a computational unit exhibiting some non-trivial internal structure and memory, for instance a group of biological or artificial neurons, see (Piekniewski, 2007) for related numerical study.

In this paper we propose a simple and tractable mathematical model for such a situation. The units are allowed to exchange stored charge under stochastic dynamics, which is modeled as neuronal spikes being transmitted along the edges of a fully connected network. Next, each edge is labelled with the count of spikes it transmitted, which results in a graph with weighed edges, called the *spike flow graph* in the sequel. Our theoretical results below, further confirmed by numeric evidence, state roughly speaking that if we remove those neural connections which are only relatively seldom used for spike transfers and we keep only those often used and relevant to the dynamics, the resulting graph is with overwhelming probability scale free with power law exponent $\gamma = 2$. The proof goes by showing that in low enough temperatures the large-scale behavior of the system admits an accurate description in terms of a particular winner-take-all type dynamics. Whereas the considered neural network model may be regarded to some extent simplistic, its asymptotic description in terms of a winner-take-all type dynamics and hence also the scale-free nature of the spike flow graph seem to be rather universal, as suggested by numeric evidence (Piekniewski, 2007).

We find the results of this paper important and interesting as showing how scale-free structures spontaneously emerge in neural information processing, arguably for rather general models and with no special assumptions aimed at stimulating this kind of self-organization. Apart from their fundamental theoretical interest the results established in this paper provide a further justification for considering neural architectures based on small-world and scale-free graphs, as has become popular in the literature in recent years, see (Perotti et al., 2006; Stauffer et al., 2003) and the references therein.

The remaining part of the paper is organized as follows. In Section 2 below we introduce our basic theoretical model sharing certain features with the standard Boltzmann machines (Aarts and Korst, 1989), yet admitting a richer state space and assuming a rather different dynamics for individual neurons which are simple spiking units here. Next, in Section 3 we describe the behavior of this model in large system size and long evolution time limit and argue it can be represented via a kind of a winner-take-all dynamics whose particular features enable us to establish explicit results on the scale-free properties of the spike-flow graph in the following Section 4. In the further Section 5 we present numeric evidence supporting our theoretical claims. Finally in Section 6 we conclude the paper and conjecture that in spite of the rather simple nature of our spiking network model as designed to allow for exact calculations, the asymptotic behavior of its large scale dynamics seems to be quite universal for a wide range of realistic spiking networks in that a power law is present in their respective spike flow graphs, even though the precise values of the corresponding exponents may vary between different models. This universality claim is also supported by some numeric evidence, as already previously reported in (Piekniewski and Schreiber, 2007) and (Piekniewski, 2007).

2. Basic model

In our research we sought for a model whose dynamics would in its essence resemble that encountered in usual recurrent neural networks and, while being simple in terms of its statistical mechanics, would exhibit a scale-free structure as a natural consequence of its construction. These considerations resulted in the following *spike flow model* originally introduced in (Piekniewski and Schreiber, 2007). We consider a simple stochastic recurrent neural network consisting of N neurons assuming states labeled by natural numbers $\sigma_i \in \{0, 1, \ldots, M_i\}$, $i = 1, \ldots, N$, interpreted as *neuronal charges* below, and with natural or possibly infinite numbers M_i standing for maximum admissible values for the respective charges σ_i, $i = 1, \ldots, N$. The network is built on a complete graph in that there is a connection between each pair of neurons σ_i, σ_j, $i \neq j$, carrying a real-valued weight $w_{ij} \in \mathbb{R}$ satisfying the usual symmetry condition $w_{ij} = w_{ji}$, moreover $w_{ii} := 0$. The values of w_{ij} are drawn independently from the standard Gaussian distribution $\mathcal{N}(0, 1)$ and are assumed to remain fixed in the course of the network dynamics. A configuration $\bar{\sigma} = (\sigma_i)_{i \leq N}$ of the network is assigned its Hamiltonian given by

$$\mathcal{H}(\bar{\sigma}) := \frac{1}{2} \sum_{i \neq j} w_{ij} |\sigma_i - \sigma_j| \qquad (1)$$

if $0 \leq \sigma_i \leq M_i$, $i = 1, \ldots, N$, and $\mathcal{H}(\bar{\sigma}) = +\infty$ otherwise. The dynamics of the network is defined as follows: at each step we randomly choose a pair of neurons (σ_i, σ_j), $i \neq j$, and denote by $\bar{\sigma}^*$ the network configuration resulting from the original configuration $\bar{\sigma}$ by decreasing σ_i by one and increasing σ_j by one, that is to say by *letting a unit charge transfer from σ_i to σ_j*, whenever $\sigma_i > 0$ and $\sigma_j < M_j$. Next, if $\mathcal{H}(\bar{\sigma}^*) \leq \mathcal{H}(\bar{\sigma})$ we accept $\bar{\sigma}^*$ as the new configuration of the network whereas if $\mathcal{H}(\bar{\sigma}*) > \mathcal{H}(\bar{\sigma})$ we accept the new configuration $\bar{\sigma}^*$ with probability $\exp(-\beta[\mathcal{H}(\bar{\sigma}^*) - \mathcal{H}(\bar{\sigma})])$, $\beta > 0$, and reject it keeping the original configuration $\bar{\sigma}$ otherwise, with $\beta > 0$ standing for an extra parameter of the dynamics, in the sequel referred to as the inverse temperature conforming to the usual language of statistical mechanics and assumed fixed and large (low temperature) throughout. Observe that the sum $\sum_i \sigma_i$ of neuronal charges is preserved by the dynamics and that, in the course of dynamics with some initial configuration $\bar{\sigma}^0$, any other $\bar{\sigma}$ with $\sum_i \sigma_i^0 = \sum_i \sigma_i$ is eventually reached with positive (although possibly very small) probability. Consequently, upon standard verification of the usual detailed balance conditions, we readily see that the collection of stationary states of the above dynamics are precisely the distributions

$$\mathbb{P}_n(\bar{\sigma}) = \begin{cases} \dfrac{\exp(-\beta\mathcal{H}(\bar{\sigma}))}{\left(\sum_{\bar{\sigma}',\, \sum_i \sigma_i'=n} \exp(-\beta\mathcal{H}(\bar{\sigma}'))\right)}, & \text{if } \sum_i \sigma_i = n, \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

and their convex combinations. In particular, our model bears some resemblance to the usual stochastic Boltzmann machines (Aarts and Korst, 1989), with the weights w_{ij} indicating the extent to which the system favors the agreement (for positive w_{ij}) or disagreement (for negative w_{ij}) of the neuronal states σ_i and σ_j. There are evident differences though, one of them being the possibly unbounded state space whenever $M_i = \infty$, the other one that precisely two neurons are affected in each update with clearly determines the source and destination of the charge flow. Whereas the latter difference does not lead far away from the concept of a Boltzmann machine, as yielding a rather similar form of the stationary distribution, the former one is crucial – indeed, if M_i is a large number, the behavior of the corresponding i-th unit becomes quite complex and arguably it can be regarded as exhibiting some kind of *memory* of charge transfers undergone in the course of the dynamics. In this paper we shall concentrate on the cases where M_i's will be all infinite or of order only slightly smaller than the overall charge stored in the system, thus conforming to our leading assumption of a complex neuronal structure. We will prove below that in this set-up the network exhibits natural scale-free features. On the other extreme one can impose all M_i's very small, which makes our model resemble classical Bolztmann machines. Taking all $M_i \equiv 1$ and i.i.d. Gaussian weights yields a network which can be regarded as a somewhat modified version of the well-known Sherrington-Kirkpatrick spin glass model, see Chapter 2 in (Talagrand, 2003). In general, our model interpolates between both extremes and can exhibit a wide range of behaviors depending on the choice of the M_i's. For the network dynamics running during a period $[0, T]$ we are now in a position to define the *spike flow graph* to be a directed graph with vertices corresponding to the neurons $\sigma_i,\ i = 1, \ldots, N$ and whose edges carry numbers (edge multiplicities) $F_{i\to j}$ indicating how many times in the course of the dynamics the charge flow occurred from σ_i to σ_j. If β is large, which is always going to be assumed in this paper, after a long enough simulation run the system freezes in some ground state whereupon any further charge flow becomes very rare and consequently the numbers $F_{i\to j}$ also freeze undergoing virtually no further changes. The in-degree of a neuron σ_i is now defined as $d_{\text{in}}(j) := \sum_i F_{i\to j}$. The main question considered in this paper is whether the so-defined spike flow graph is scale-free in that its in-degree distribution follows a power law, that is to say $\mathbb{P}(d_{\text{in}}(i) \approx x) \sim c_{\text{in}} x^{-\gamma_{\text{in}}}$ for a randomly picked node i. Choosing M_i's large enough we shall establish a positive answer to this question. It should be noted at this point, as will become clear from our discussion below, that the asymptotic behavior of the corresponding out-degree distribution is the same as that of the in-degrees.

3. Winner-take-all dynamics and ground states

For the scope of this research we will limit ourselves to very low temperature regime, which amounts to assuming that the overwhelming majority of network updates are just jumps towards lower energy configurations, as in the corresponding zero-temperature (infinite β) approximation.

We assume first that all M_i's are infinite and thus no upper bounds are imposed on individual charges. In this extreme set-up we argue that with overwhelming probability with respect to the choice of the weights w_{ij}, the unique ground state (lowest energy state) of the network, and hence also the unique attractor of its dynamics, is a configuration in which all charge present in the system is stored in a single *best* unit with all the remaining units devoid of charge. To see this, for each unit σ_i consider the *support* S_i it gets from the remaining units, given by

$$S_i := -\sum_{j\neq i} w_{ij}.$$

Clearly, all the S_i's so defined are Gaussian random variables $\mathcal{N}(0, N-1)$ and are virtually independent – indeed, S_i and S_j for $i \neq j$ share just one summand w_{ij} whereas the remaining ones are independent. With $S_{:k}$ standing for the k-th largest value among S_i's, it is known by extreme value theory, see e.g. Section 1.2 in (Talagrand, 2003), that the order statistics $S_{:k}$ are well approximated by

$$S_{:k} \approx \sqrt{2N \log_2 N}\left(\sqrt{\log 2} + \frac{\xi_k}{\sqrt{\log_2 N}}\right) \tag{3}$$

where the sequence $\xi_1 > \xi_2 > \ldots$ is chosen according to a Poisson point process with intensity $\frac{1}{\pi}\exp(-2t\sqrt{\log 2}),\ t \in \mathbb{R}$, in particular the p-th ξ_i above 0 is of order $\log p$ and $S_{:k}$'s are of order $\sqrt{N \log N}$ which is much higher than the order of the typical S_i being $\sqrt{N}$. To proceed, assume we run our spike-flow dynamics for some long enough amount of time to get close to equilibrium, whereupon we consider a small number $o(N)$ of neurons which store the highest charge, considerably higher than the remaining units, and we call these *elite neurons* while granting the term *bulk neurons* to the remaining units in the system. Since the number of elite neurons is a negligible fraction of N, the formula (1) becomes then

$$\mathcal{H}(\bar{\sigma}) \approx -\sum_{i\in\text{elite}} \sigma_i S_i + \frac{1}{2}\sum_{j,l\in\text{bulk}} w_{jl}|\sigma_j - \sigma_l|. \qquad (4)$$

Thus, whenever in the course of the network dynamics a charge transfer is proposed from a bulk neuron σ_j to an elite neuron σ_i, the resulting energy change is seen to be well approximated by $-S_i$ plus a term due to the interaction between σ_j and other bulk neurons. In general, we have no control of this term, yet if σ_i is one of the neurons with the highest support as in (3), this offending term of order at most $\sqrt{N}$ is very likely to be negligible compared to $-S_i$ which is of order $\sqrt{N\log N}$, thus making the energy change strongly negative and the proposed transfer extremely likely to be accepted. Clearly, the inverse transfer becomes then almost impossible. Consequently, whenever a neuron with a very high support enters the elite, it virtually never leaves it; moreover it continuously drains charge from the bulk losing it only to other elite members if at all (see Figure 1 for numerical support of these claims). Furthermore, should a neuron with a small support value happen to enter the elite at the early stages of the dynamics, it will soon leave it having its charge drained by other higher supported neurons. Thus, after running our dynamics long enough we end up with a picture where the elite consists of neurons with the highest support. Although the elite neurons do struggle for charge between themselves, they cooperate in draining it from the bulk. Therefore eventually almost no charge will be present in the bulk and hence the Hamiltonian will admit a particularly simple approximation

$$\mathcal{H}(\bar{\sigma}) \approx -\sum_{i\in\text{elite}} \sigma_i S_i \qquad (5)$$

and all further updates in the system will only happen due to charge transfers within the elite. Note now that the interactions between elite neurons as determined by their connectivities (weights) are of order $o(\sqrt{N})$ since the cardinality of the elite is $o(N)$ whereas the differences between the highest consecutive support values are of higher order $\Theta(\sqrt{N})$ in view of (3) which makes the former negligible compared to the latter. Thus, the dynamics between the elite neurons takes eventually a particularly simple form: a pair of elite neurons is chosen by random and if the one with smaller support attempts to transfer a unit charge to the one with higher support, the attempt is accepted, otherwise it is rejected. The only ground state of the system is then obtained by putting all charge into the unit of the highest support. It should be noted that at intermediate stages of the dynamics it may happen that elite members show up with charges whose order is inverse to that determined by the supports rather than consistent with it. This is an artifact due to the fact that if we admitted negative charges here, a twofold sign-flip symmetry would be present in the system in full analogy to usual networks with no external field and such inverse ordering would compete with the standard one on equal rights. This is not the case here though because negative charges are not allowed and therefore such inversely ordered structures are unstable and do not persist in the course of the dynamics.

In view of the above discussion, the highest in-degrees of the spike-flow graph are observed in elite units enjoying the highest support from the system, and the corresponding charge flows $F_{i\to j}$ are mainly due to the internal charge transfers within the elite. Thus, we have shown that the asymptotic behavior of our network model is accurately described by the following *winner-take-all* model:

- the system consists of K neurons $u_i,\ i = 1,\dots,K$, representing the elite units and ordered according to decreasing supports,
- n units of charge are sequentially introduced into the system, each time according to the following dynamics
  - · first, a unit charge is transferred to a randomly chosen neuron $u_{k_0},\ k_0 \le K$,
  - · thereupon it starts jumping to further neurons u_{k_l}, where $k_{l+1} < k_l$ is randomly chosen in $\{1,\dots,k_l - 1\}$,
  - · eventually the unit charge reaches u_1 and gets frozen there.
- the in-degrees of the elite neurons in the original network are approximated by the numbers D_i indicating how many charge units have visited u_i on their way to u_1.

In other words, in this model the charge transfers always occur from a neuron with smaller support to a randomly chosen better supported one, whence the term *winner-take-all* dynamics. Curiously enough, the winner-take-model is easily seen to exhibit a *consistency property* – if we take some $K' < K$ and observe the behavior of the model restricted to K' neurons of highest support only, this exactly coincides with what we get if we run our original dynamics on the restricted set $\{u_1,\dots,u_{K'}\}$ of units. Consequently, from the viewpoint of our asymptotic analysis of the in-degree sequence $D_1, D_2,\dots$ the precise value of K is irrelevant as long as $K \ll N$ but $K \to \infty$ as $N \to \infty$.

Now, repeating the argument presented in this section for M_i's large but finite, we end up with the following modification of the above winner-take-all dynamics. Assume first that the elite neurons of the highest support are not yet saturated, that is to say their capacity has not yet been reached. In such a case the dynamics follows exactly as previously. Once a certain elite neuron gets saturated, it becomes *inactive*, since it cannot accept any more incoming charge. If this happens to be a neuron of high support, then it is very unlikely to get unblocked prior than possibly at the very final stages of the dynamics, since with overwhelming probability only units of higher support drain any charge from the considered unit, and their number is negligible compared to that of lower supported neurons pumping their charges upwards the support hierarchy. Therefore, should a neuron of a very high support get saturated, it will most likely stay inactive for the most of the simulation thereafter, and it could be removed from any additional consideration as playing no relevant role anymore. Further evolution of the so reduced system follows the same pattern: at any stage the winner-take-all dynamics is present among the set of best unsaturated neurons. Consequently, in case M_i's are of the same order as total charge present in the system, the deviation from the *unbounded* version of the dynamics is negligible, which can be easily noted in simulations. If M_i's are much smaller though, the saturation factor becomes significant and the *winner-take-all* dynamics breaks down. Some models exhibiting this property will also be discussed in the next Section 4.

4. Power law for spike-flow in-degrees

We are now in a position to exactly characterize the asymptotic behavior of the in-degree sequence $D_i,\ i \geq 1$. Again, we begin with the extreme set-up $M_i \equiv \infty$ first, passing to more general choice of charge constraints thereafter. It is worth noting that asymptotically the out-degree sequence behaves in exactly the same way as the in-degrees since for most units save the highest support neuron and very low support neurons their in- and out-degrees are almost equal. Some insignificant disagreements may occur in finite numerical simulations, where all units start with some fixed amount of charge and proceed according the dynamics. In such case the out-degree sequence is disturbed by the single (therefore insignificant) unit that eventually receives and keeps the whole charge present in the system, whereas the in-degree sequence is disturbed within some range of low degrees (units which received far less charge than they gave away to others). We can avoid these fluctuations by only looking at the tail of the distribution (in practice, say, degrees higher than 5-10 times the initial charge per neuron) or by simulating larger systems.

To proceed, consider a single charge unit introduced into the system and write k_l for the number of neuron u_{k_l} it visits after its l-th jump, $l = 0, 1, \ldots$. Recall from Section 3 that k_0 is drawn uniformly from $\{1, \ldots, K\}$. Further, consider also a sequence $X_0, X_1, X_2, \ldots$ of continuous $(0,1)$-valued random variables such that X_0 is uniform in $(0,1)$ and X_{l+1} is chosen uniformly from $(0, X_l)$ for all $l \geq 0$. Then it is easily seen that for K large enough we can safely approximate in law

$$k_l = \lceil K X_l \rceil$$

with $\lceil \cdot \rceil$ standing for the upper integer value of its argument. In particular, defining $\pi_i,\ i = 1, \ldots$ to be the probability that the charge unit visits u_i, we have $\pi_i = \mathbb{P}(\exists_l k_l = i)$ and hence for K large enough we get the approximation

$$\pi_i \approx \mathbb{E}|\{l,\ X_l \in [(i-1)/K, i/K]\}|,\ \ i > 1 \qquad (6)$$

and, clearly, $\pi_1 = 1$. The values of in-degrees D_i are then binomially distributed $b(\pi_i, n)$ with parameters π_i and n, the latter standing for the number of charge units present in the system.

To proceed with our asymptotic analysis we observe that X_l's form a so-called *record sequence* in the sense of classical extreme-value theory, see Chapter 4 in (Resnick, 1987). Consequently, by Section 4.1 ibidem, the sequence $T_l := -\log X_l$ is simply a unit intensity homogeneous Poisson point process in $\mathbb{R}_+$. Thus, using (6) we get

$$\pi_i \approx \mathbb{E}|\{l,\ T_l \in [-\log(i/K), -\log((i-1)/K)]\}| \approx 1/i.$$

Hence, for large values of n we have by the law of large numbers

$$D_i \approx n/i.$$

It means that, for large k,

$$|\{i,\ D_i > k\}| \approx n/k$$

or,

$$|\{i,\ D_i \approx k\}| \approx n/k^2.$$

We have thus proven the main result of this paper.

Theorem 1 *For the basic spike-flow model with $M_i \equiv \infty$ the resulting spike-flow graph is scale-free with exponent $\gamma = 2$.*

In analogy to this argument assume now that M_i's instead of being infinite are finite, indepedent from the weights w_{ij}, independent among themselves, and drawn from a power-law distribution

$$\mathbb{P}(M_i > k) \approx ck^{-\alpha} \qquad (7)$$

for some $\alpha > 0$. In such a set-up, if c in (7) is not very large, a non-negligible fraction of units will get saturated in the course of the dynamics and therefore would stop accepting any more incoming charge at some stage of the network evolution. This may considerably alter the behavior characterized by Theorem 1. In fact, it is natural to expect that three groups of units will emerge:

- Units of highest support, *elite of the elite*, which can be sure to reach their capacities. By the independence of M_i's from the weights w_{ij}'s and hence also from the supports S_i's, when choosing at random among such units the probability of exceeding in-degree k is given by the product of the probability of the chosen unit exceeding the in-degree k in the unconstrained dynamics ($M_i \equiv +\infty$) times the probability of its capacity being higher than k. Consequently, in view of Theorem 1 the in-degrees of the highest support units should follow a power law with exponent $\alpha + \gamma = \alpha + 2$.
- Units of intermediate supports, *lower elite*, still falling into the elite and reaching rather high in-degrees, but not exceeding or even reaching their capacities. Such units do not feel the constraints M_i's and constitute a portion of the network where in-degrees should follow a power law with exponent $\gamma = 2$ as in the unconstrained dynamics.
- Units of rather high but not highest supports, *medium elite*, for which the capacity and in-degrees they would reach under the unconstrained dynamics are of a comparable order. Their behavior should interpolate between the above two extremes.

These observations would suggest that two principal regimes should be observable for the in-degree distribution of such networks: highest in-degrees should follow a power law with exponent $\alpha + 2$ whereas the lower elite in-degrees should behave as in Theorem 1 stating power law with exponent 2. The region separating these regimes should interpolate between these two behaviors, possibly exhibiting very complicated properties due to the presence

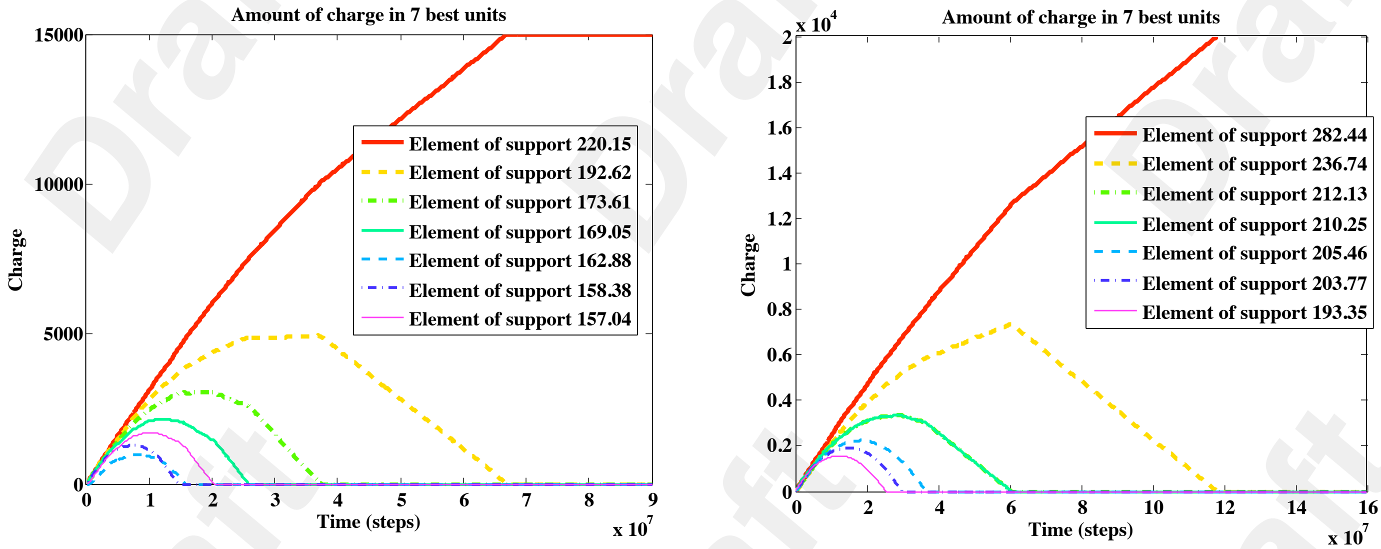

Fig. 1. Typical evolution of the charge stored in seven units of the highest support. The above figure is a result of a simulation run of 3000 units (left) and 4000 units (right). Note that while in the beginning of the simulation all seven elite members compete for charge, by the end the single best unit gets everything.

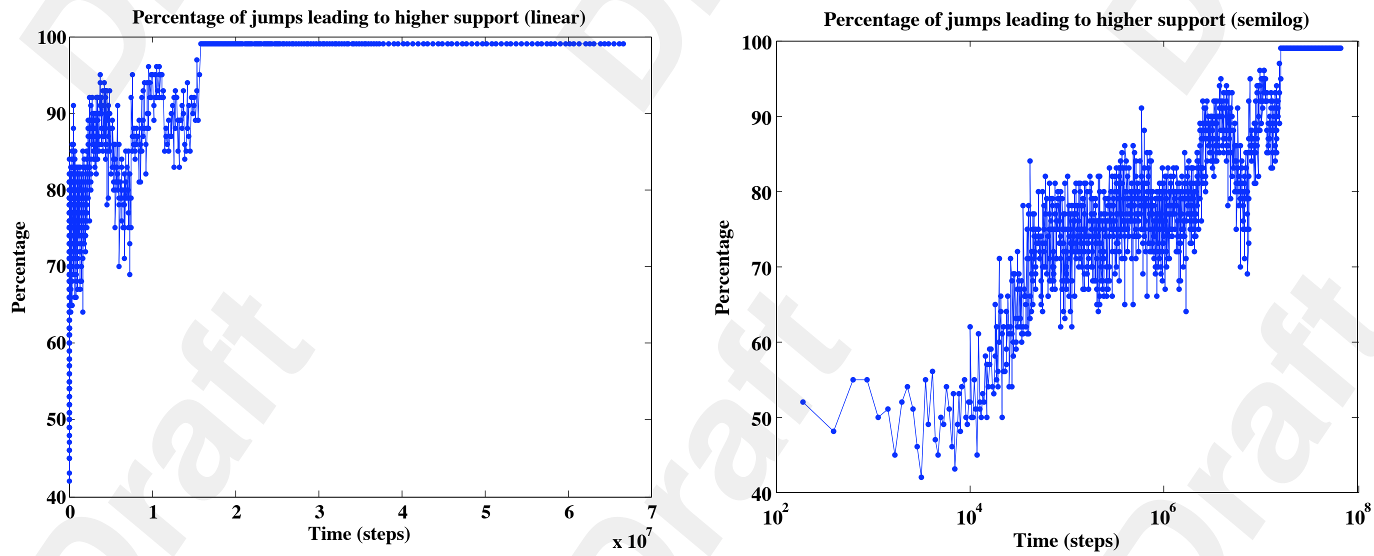

Fig. 2. Percentage of charge jumps leading to a unit of higher support (sampled every 100 jumps) in a simulation run of 3000 units. The plot on the left is scaled linearly, the one on the right is semi-log. Note that after initial unstable phase (about 10^4 steps), the plot increases steadily until about $2 * 10^6$ steps where again some fluctuations occur. These fluctuations are caused by increased significance of stochastic term (tiny energy modifications leaving room for thermal fluctuations). It is worth noting that by that time jumps are already infrequent while the state is near the energy optimum. By the step $2 * 10^7$ the system freezes completely in the ground state.

to *traffic jams* at medium elite units, which are no more negligible unlike in case of lower elite units, but which only temporarily disable the blocked units and may be eventually discharged in contrast to the case of high elite units. We cannot claim to have confirmed these conjectures by numerical results though because the realistic system sizes we were able to reach in our simulation were too small to ensure statistically significant collection of units in each of the afore-mentioned regimes.

5. Numeric results

The above considerations were accompanied by a numerical simulation implemented in Matlab, letting us continuously verify our assumptions, and giving valuable hints for further investigation. The simulations were usually carried out for systems of about 3000-6000 neurons with the basic dynamics (as described in Section 2) – the only speed-up was that the neuron to pass a unit charge to some other one was chosen randomly only among those containing any charge at all. The total energy computation required a quadratic time in the number of neurons, but during the

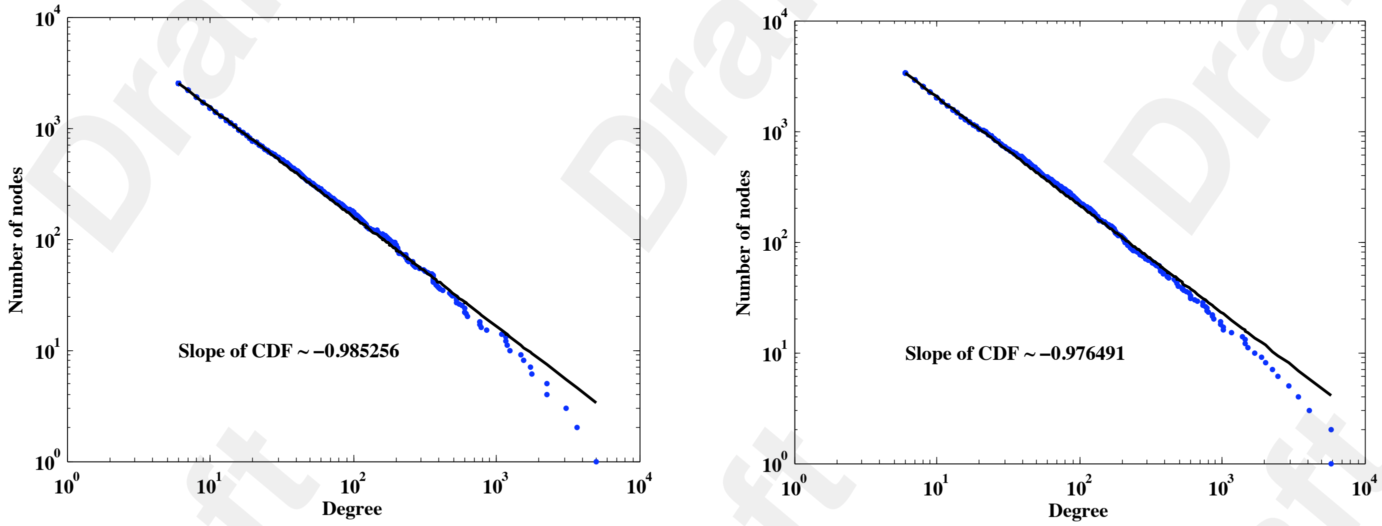

Fig. 3. Cumulative distribution function (CDF) of the out-degree in the spike flow graph (in-degree yields a similar plot) of 3000 units (left) and 4000 units (right). Presented CDF slopes correspond to the power law exponent $\gamma = 2 +/- 0.03$. The slopes were approximated by the least squares method.

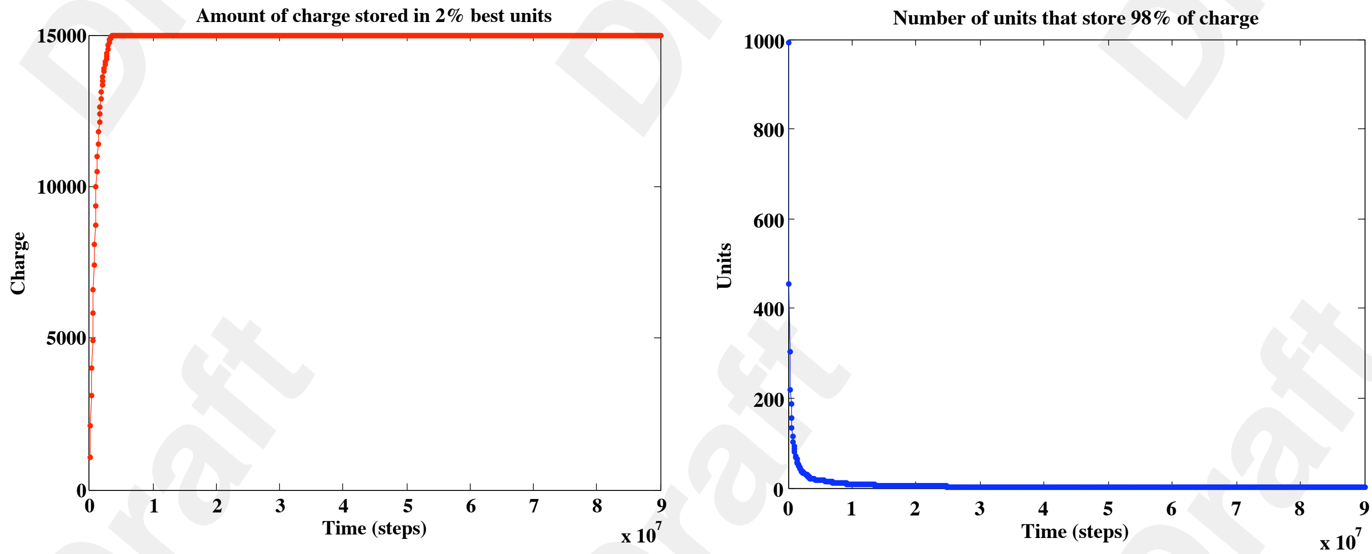

Fig. 4. The charge stored in 2% units of highest support (left) and number of units storing 98% of total charge. These figures give strong support to the idea of dividing the units into elite and bulk, and treating these groups separately.

simulation we only needed to compute local energy updates, which took only linear time. Despite of these straightforward enhancements, larger systems ($\approx$ 10000 units) became problematic due to memory consumption and did not give any qualitatively better results. In the future we plan to simulate much bigger systems based on the simplified version of dynamics (the winner-take-all asymptotic version) that would allow us to avoid the need for explicit connectivity matrix, in order to confirm the intuitions described in final paragraphs of Section 4. In the course of the present simulation β (the inverse temperature) was fixed at $\beta = 10$ which, since the average energy updates in the simulation were of order ≈ 1 per step, places us in the low temperature regime. The temperature only became more significant by the end of the simulations when the energy modifications were of much smaller order leaving place for thermal fluctuations, but by that time the system usually had already converged to the expected "winner-take-all" configuration (Figure 2 gives some insight into temperature based fluctuations). The results of the simulations confirmed our theoretical predictions about the "winner-take-all" dynamics (see Figures 1,2,4), as well as the scale-free properties of the *spike flow graph* (Figure 3). The number of steps was 10 times the number of neurons squared, which was about the number of steps required for full convergence to the ground state. Rarely the system converged to a state in which two units of highest support shared the whole charge. This is possible, whenever the weight between the two competing units is comparable to the difference of their supports, thus forming a local energy minimum (pumping charge to the better unit requires temporary energy increase). Since the experimental system is finite such unusual configurations may appear with some small probability. Evidently as the system size increases, such energy minima become less probable (asymptotically negligible).

6. Conclusions

The model introduced in Section 2 above is sufficiently simple to allow for theoretical approach, yet its dynamics is rich enough to clearly exhibit scale-free properties. Moreover, the argument in the crucial Section 3 providing a description of the large-size behavior of our model in terms of a winner-take-all dynamics, seems to be rather robust and independent of many specific features of the model, which makes us believe it is universal for a broad class of similar models with not necessarily Gaussian and not necessarily i.i.d. weights, and even admitting possible modifications to the dynamics etc. These and related issues are a subject of our present research in progress.

A further important issue, already signalled in the introduction above, is that our results are by no means contradictory to the lack of scale-free properties reported in certain real-world neural networks, e.g. negative results for C.elegans worm nervous system, whose connectivity reveals rather exponential decay – see (Amaral et al., 2000; Koch and Laurent, 1999). The point is that we only expect scale-free properties to arise in presence of sufficiently complicated information processing units, corresponding for instance to neuronal groups rather than single neurons, and exhibiting a kind of (collective) state memory. This is often not the case for individual biological neurons or their models, where only the presence of a short refractory period carries some information about the history of previous excitations which cannot be stored for later use. Things can be different however if a recurrent group of neurons is taken into account as a single computational unit. In this case the rich dynamics inside the group can develop a kind of a collective state memory we have in mind. In recent paper (Piekniewski, 2007) we show a numerical experiment based on the Eugene Izhikewich simple spiking neuron model (Izhikevich, 2003) that supports the above claims. Shortly, the experiment shows that by substituting small neuronal groups in place of single neurons, the resulting activity graph (analogous to the spike-flow graph discussed above) becomes a scale-free network, even for fairly small neuronal groups (less than 20 neurons per group). The questions whether such structures can emerge among more biologically plausible, spontaneously created groups of synchronized neurons is currently a subject of our ongoing research. Other more complex variants of the *spike flow model* are also being investigated.

7. Acknowledgement

We gratefully acknowledge the support from the Polish Minister of Scientific Research and Higher Education grant N N201 385234 (2008-2010).

References

Aarts, E., Korst, J., 1989. Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing. John Wiley & Sons, Inc., New York, NY, USA.

Albert, R., Barabási, A.-L., January 2002. Statistical mechanics of complex networks. Reviews of modern physics (74), 47–97.

Albert, R., Jeong, H., Barabási, A.-L., Septmeber 1999. Diameter of the world-wide web. Science 401, 130–131.

Amaral, L. A., Scala, A., Barthelemy, M., Stanley, H. E., October 2000. Classes of small-world networks. Proc Natl Acad Sci U S A 97 (21), 11149–11152.
URL http://dx.doi.org/10.1073/pnas.200327197

Barabási, A.-L., Albert, R., October 1999. Emergence of scaling in random networks. Science (286), 509–512.

Barabási, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek, T., 2002. Evolution of the social network of scientific collaborations. Physica A 311 (4), 590–614.

i Cancho, R. F., Solé, R. V., November 7 2001. The small-world of human language. Proceedings of the Royal Society of London B 268 (1482), 2261–2265.

Izhikevich, E. M., 2003. Simple model of spiking neurons. IEEE Transactions on Neural Networks (14), 1569–1572.
URL http://www.nsi.edu/users/izhikevich/publications/spikes.pdf

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., Barabási, A.-L., October 5 2000. The large-scale organization of metabolic networks. Nature 407 (6804), 651–653.

Koch, C., Laurent, G., 1999. Complexity and the Nervous System. Science 284 (5411), 96–98.
URL http://www.sciencemag.org/cgi/content/abstract/284/5411/96

Montoya, J. M., V., R. V. S., February 7 2002. Small world patterns in food webs. Journal of Theoretical Biology 214 (3), 405–412.

Perotti, J. I., Tamarit, F. A., Cannas, S. A., Nov. 2006. A scale-free neural network for modelling neurogenesis. Physica A Statistical Mechanics and its Applications 371, 71–75.

Piekniewski, F., August 2007. Emergence of scale-free graphs in dynamical spiking neural networks. In: Proc. IEEE International Joint Conference on Neural Networks. IEEE Press, Orlando, Florida, USA, pp. 755–759.

Piekniewski, F., Schreiber, T., April 2007. Emergence of scale-free spike flow graphs in recurrent neural networks. In: Proc. IEEE Symposium Series in Computational Intelligence - Foundations of Computational Intelligence. Honolulu, Hawai USA, pp. 357–362.

Redner, S., 1998. How popular is your paper? an empirical study of the citation distribution. European Physical Journal B 4 (2), 131–134.

Resnick, S. I., 1987. Extreme Values, Regular Variation, and Point Processes. Springer Series in Operations Research and Financial Engineering. Springer-Verlag New York, Inc., Secaucus, NJ, USA.

Stauffer, D., Aharony, A., da Fontoura Costa, L., Adler, J., 2003. Efficient hopfield pattern recognition on a scale-free neural network. The european physical journal B (32), 395–399.

Talagrand, M., 2003. Spin Glasses: A Challenge for Mathematicians, Cavity and Mean Field Models. Vol. 46 of Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics. Springer-Verlag New York, Inc., Secaucus, NJ, USA.