Biological conservation law as an emerging functionality in dynamical neuronal networks

Boris Podobnik[a,b,c,d,e,f,1], Marko Jusup[g], Zoran Tiganj[h], Wen-Xu Wang[i], Javier M. Buldú[j,k], and H. Eugene Stanley[a,b,1]

[a]Center for Polymer Studies, Boston University, Boston, MA 02215; [b]Department of Physics, Boston University, Boston, MA 02215; [c]Faculty of Civil Engineering, University of Rijeka, 51000 Rijeka, Croatia; [d]Zagreb School of Economics and Management, 10000 Zagreb, Croatia; [e]Luxembourg School of Business, 2453 Luxembourg, Luxembourg; [f]Faculty of Information Studies, SI-8000 Novo Mesto, Slovenia; [g]Center of Mathematics for Social Creativity, Research Institute for Electronic Science, Hokkaido University, Sapporo 060-0808, Japan; [h]Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215; [i]School of Systems Science, Beijing Normal University, 100875 Beijing, China; [j]Laboratory of Biological Networks, Center for Biomedical Technology, Universidad Politécnica de Madrid, 28223 Pozuelo de Alarcon, Madrid, Spain; and [k]Complex Systems Group, Universidad Rey Juan Carlos, 28933 Móstoles, Madrid, Spain

Contributed by H. Eugene Stanley, August 28, 2017 (sent for review April 6, 2017; reviewed by Mario Chavez and Drazen Prelec)

Scientists strive to understand how functionalities, such as conservation laws, emerge in complex systems. Living complex systems in particular create high-ordered functionalities by pairing up low-ordered complementary processes, e.g., one process to build and the other to correct. We propose a network mechanism that demonstrates how collective statistical laws can emerge at a macro (i.e., whole-network) level even when they do not exist at a unit (i.e., network-node) level. Drawing inspiration from neuroscience, we model a highly stylized dynamical neuronal network in which neurons fire either randomly or in response to the firing of neighboring neurons. A synapse connecting two neighboring neurons strengthens when both of these neurons are excited and weakens otherwise. We demonstrate that during this interplay between the synaptic and neuronal dynamics, when the network is near a critical point, both recurrent spontaneous and stimulated phase transitions enable the phase-dependent processes to replace each other and spontaneously generate a statistical conservation law—the conservation of synaptic strength. This conservation law is an emerging functionality selected by evolution and is thus a form of biological self-organized criticality in which the key dynamical modes are collective.

complex networks | emerging properties | homeostatic plasticity | neuroscience | neuronal noise

What accounts for the functionality of complex systems? Is it component availability or the layout of the available components? Recent discoveries emphasize the role of the layout (i.e., the topological structure) by showing that some complex systems—both living and artificial—preserve functionality after the random removal of components. In this context, biologists debate whether nature creates new functionalities by adding components through mutations or by reorganizing and rewiring already existing components (1–3). For example, a particularly well-known biological case study that illustrates the importance of topological structure is of brain networks in which the small-world organization (4) is pervasive in both an anatomical (5) and a functional (6) sense. Irrespective of scale—from neurons (4, 7) to cortical regions (8)—small-world topology is a fundamental organizational principle in the brain that extends beyond the specific role a particular brain region may have. Deviations from this specific network organization are, in fact, related to brain dysfunction or the emergence of neurodegenerative diseases (9). Understanding the role of the small-world topological structure and how it emerges is extremely difficult (10) because it requires analyzing the interplay between at least four fundamental brain properties: structure, dynamics, function, and evolution.

Our goal is to understand how functionalities spontaneously emerge in a complex system. Working in the framework of network science, we focus less on topological structure and more on the collective dynamical state of a complex system as it responds to the dynamics of its fundamental components, the interaction between them, and the evolution of their network of interactions (11). This notion is exemplified in ref. 12 in which weak correlations between pairs of neurons may contribute to strongly collective behavior in a neural population. Accordingly, in building our model, we draw inspiration from neuroscience and designate network nodes and links as neurons and synapses, respectively, and their dynamics as a recurrent flipping between two different phases. This phase flipping is ubiquitous in the complex system modeling of neuroscience. Examples include flipping between collective depolarization and polarization phases with predominant excitatory cortical neurons and inhibitory interneurons, respectively (13), enhancing and blocking sensory-motor processing, and the generation of pathological rhythms associated with psychiatric or neurological disorders (14).

Here we use the synaptic homeostasis hypothesis (15), which states that to generate the long-lasting synaptic homeostasis needed for the optimal functioning of the brain, the brain flips between a phase with dominant memory-formation processes and a phase with dominant memory-consolidation processes. This formation of new memories is associated with synaptic potentiation during wakefulness, while the consolidation is thought to result in synaptic downscaling during sleep/rest. In fact, a restorative role of deep sleep for the brain's capacity to undergo neuroplastic changes during wakefulness was proved in a recent experiment (16). The dynamics of our network model illustrate how such a homeostasis may spontaneously arise as a consequence of simple interactions between nodes and links. Specifically, by combining processes in two mutually exclusive phases between which the network flips, while downplaying biochemical processes at the neuronal level (17), a statistical conservation law of synaptic strength emerges without being a priori ingrained in either of the phases. This and similar spontaneously emerging new functionalities in complex biosystems are still inadequately understood within network science (18).

Significance

Living complex systems create functionalities by pairing up complementary processes, one to build and another to correct and clean the space for more building. When these processes are embedded into mutually exclusive phases, establishing a balance between them requires a phase flipping mechanism. For a simple neuronal network composed of neurons and synapses, we demonstrate that flipping between excited and resting phases gives rise to a statistical conservation law of synaptic strength. Furthermore, this law is selected by evolution regardless of the network's initial state, thus reconciling biology with the physical concept of self-organized criticality.

Author contributions: B.P., M.J., Z.T., W.-X.W., J.M.B., and H.E.S. designed research, performed research, and wrote the paper.

Reviewers: M.C., CNRS; and D.P., Massachusetts Institute of Technology.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. Email: bp@phy.hr or hes@bu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1705704114/-/DCSupplemental.

Accordingly, our contribution demonstrates how the evolution of such functionalities follows a self-organized criticality pattern (17) whereby, irrespective of the system's initial state, the section of the phase space in which the conservation law of synaptic strength holds gets selected.

Model

Theoretical research (19) has depicted the propagation of action potentials through neurons using the Hodgkin–Huxley model based on detailed biophysical and biochemical processes. More recently, percolation theory (20) was used to emulate information transfer between neurons such that a presynaptic neuron can excite a postsynaptic neuron by firing an action potential. Further developments have led to models of signal processing, including synfire propagation with synchronous neuronal activity and firing-rate propagation (21).

We use a dynamical neuronal network approach in which neurons are either excitatory or inhibitory. To minimize the number of parameters, the dynamics of the inhibitory neurons are conditional on the excitations of the excitatory neurons. This network model is a highly stylized version of a realistic neuronal network in that we neglect the existence of a brain substructure or the fact that memory formation assumes interaction between different brain regions (22). In addition, excitable cortical neurons in realistic networks process and transmit information in the form of electrical and chemical signals in response to outside stimuli (i.e., hearing, sight, touch, smell, and taste). When this neuronal network receives a time-dependent input, signals spread along pathways and cause different neurons to become excited and fire action potentials. Because each neuron thus generates its own time series of excitations, we can estimate the probability that a neuron has become excited. These probabilities play the role of model parameters and reflect the average behavior of neurons when they are triggered by outside stimuli.

In our dynamical neuronal network, links between nodes are synapses that change their strength, depending on neural activity. Ever since the work of Cajal, it has been known that learning changes the strength of synaptic connections between neurons. In the face of changing synaptic strengths, how does the brain achieve homeostatic plasticity (23, 24)? How does it maintain conditions necessary for normal and robust functioning with respect to fluctuations in the environment? To answer these questions we start by locating the part of the phase space where the neuronal network flips between an excited phase and a resting phase. Then we look at what each neuron does locally to produce the desired collective behavior (17, 25) and generate a statistical conservation law that preserves the average synaptic strength. This approach leads to the model's three fundamental assumptions:

Assumption *i*. Although neurons activate in response to external stimuli by firing action potentials, their activations are also subject to stochastic processes governed by, e.g., the reliability of synaptic transmission and the stochastic fluctuations of the membrane potential (26–28). Averaging external stimuli over all neurons, we thus assume that as a result of past external stimulations any neuron can excite spontaneously with probability p and remain active for τ time steps. Each neuron has the probability pdt of this excitation during time interval dt, similar to the internal failure described in refs. 29 and 30. This is also similar to the concept in percolation theory in which a connected cluster in a random graph is constructed by assuming that the edge or "bond" between two neighbors can be open (allowing the signal through) with probability p or closed with probability $1-p$. The difference is that we are modeling neuron excitability rather than link excitability.

Assumption *ii*. As stated above, neurons usually activate in response to external stimuli. We thus assume that each (postsynaptic) neuron can excite (fire action potentials) in response to activity in its immediate neighborhood (presynaptic neurons). If a single neighbor of neuron i is excited at time $t-1$, then neuron i will fire at time t with a probability $p_1\,dt$. When there is more than one excited neighbor at time $t-1$, the probability that i will fire at t increases to $p_2\,dt$ ($p_1 < p_2$). We assume that a neuron can be excited externally only by its nearest neighbors. Once a neuron fires in response to another neuron, it stays excited for τ' time steps (without any loss of generality we set $\tau' = \tau$).

Assumption *iii*. To mimic synaptic and interneuronal dynamics, two neighboring neurons in the excited state, i and j, generate an excitatory postsynaptic potential with probability p_s and an inhibitory postsynaptic potential with remaining probability $1-p_s$. If, by contrast, either of neighbors i and j is in the resting state, only an inhibitory postsynaptic potential is generated. The excitatory potential strengthens the synapse between i and j by an amount $\Delta\epsilon_U = \epsilon$, where ϵ is taken from an exponential distribution with SD σ. The inhibitory potential weakens the same synapse by an amount $\Delta\epsilon_D = z\epsilon$, where z is a small number such that $\Delta\epsilon_D << \Delta\epsilon_U$. Practice is required for learning (31), but once we learn it is not easy to forget.

To make the exposition easier to follow, we initially focus on the relatively rich dynamics arising from assumption *i* and assumption *ii*, while disabling the model's features specified in assumption *iii*. This last assumption plays a central role once our attention is turned toward synaptic plasticity and the conservation of synaptic strength.

Results

Spontaneous Phase Flipping. Note that over a large region of the phase space given by (p, p_1, p_2) the neuronal network stays in either a primarily excited or a resting phase (29). Assuming a fixed value of p_2, over a small part of the (p, p_1) phase plane (Fig. 1) the neuronal network exhibits spontaneous phase flipping (Fig. 2) (29). This spontaneous phase flipping is not in response to a specific stimulus but is intrinsically generated by the brain (32–34).

Rarely is phase flipping in living complex systems purely stochastic. Regulatory stimuli are also present, and our formalism accommodates both. For example, let us set a network at point ($p_1 = P_1, p_2 = P_2$) in the phase space. Here flipping is spontaneous, but regulation decreases the p_1 parameter to $p_1 = (1-\Delta p)P_1$ during the excited phase and increases it to $p_1 = (1+\Delta p)P_1$ during the resting phase where $0 \le \Delta p < 1$ is a real number. Fig. 2 shows that the larger the value of Δp is, the smaller the spontaneous contribution to phase flipping. As the stimulus-driven phase flipping becomes dominant, the timing of phase transitions gets more predictable. Positioning itself in the phase space to secure spontaneous phase flipping, the system ensures an evolutionary advantage in that even small stimuli are sufficient to achieve considerable regulation. This advantage may furthermore be important in the context of aging, assuming that the ability to generate reliable stimuli gradually degrades.

The described behavior is, in the mean-field limit, independent of the detailed topology. A crucial parameter is only the average degree of nodes. Numerical simulations, however, require working with a particular network structure. Here we simply place the nodes in a 2D lattice and create links from any focal node to its first and second physical neighbors. The lattice is given toroidal boundaries, resulting in a network as visualized in Fig. 2. The phase-flipping mechanism in our framework gives rise to a new higher-order emerging functionality, the preservation of the average synapse strength. How the network dynamically sets itself in this narrow part of phase space is explained later by means of evolution.

The equilibrium fraction of excited neurons can also be approximated analytically. If internal (X) and external (Y) excitations are considered independent (only approximately possible), according to probability theory $P(X \cup Y) = P(X) + P(Y) - P(X)P(Y)$. In our model, a randomly chosen neuron i will become externally excited with probability p_1 if it is surrounded by a single excited neuron, while the same will happen with

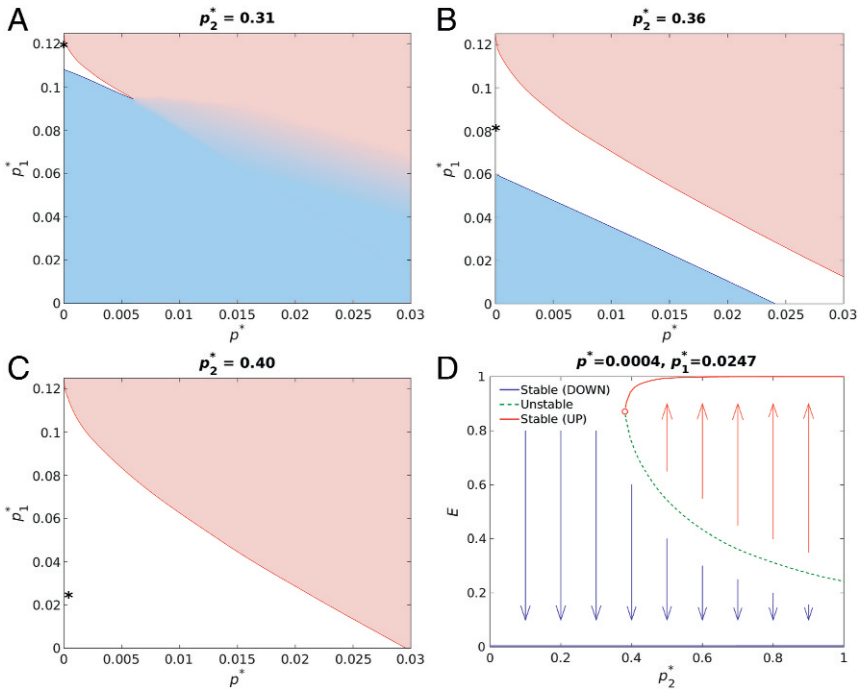

Fig. 1. Model of bistability and the phase-flipping mechanism. Generally, our model has an equilibrium of low (light blue) and an equilibrium of high (light red) activity, depending on the parameter values. (*A*) As the value of parameter p_2 (and hence $p_2^* = 1 - e^{-p_2\tau}$) increases, a region of bistability opens in the (p^*, p_1^*) phase space (white). The existence of this region is a necessary condition for phase flipping to occur. The bistable region and the corresponding spinodals are predicted theoretically, but the validity of these predictions is confirmed by the fact that the numerical implementation of the model exhibits phase flipping when positioned inside the bistable region (asterisk-marked point). (*B* and *C*) Increasing p_2^* even further causes the bistable region to widen and split, whereby the low-activity region eventually disappears. (*D*) For fixed p^* and p_1^*, the equilibrium fraction of nodes with excited neighbors, *E*, as a function of p_2^* bifurcates. Phase flipping is possible only near the bifurcation point, where the stochasticity in the model formulation easily pushes *E* across the border (dashed green line) that separates the states of low and high activity.

probability p_2 if it is surrounded by two or more excited neurons. Thus, the probability $a \equiv P(X \cup Y)$ that a randomly chosen neuron i is excited either by neuronal noise or by i's neighbors equals

$$a = p^* + E_1 p_1^* + E_2 p_2^* - p^*(E_1 p_1^* + E_2 p_2^*), \qquad [1]$$

where $E_1 = E(k, 1, a)$ and $E_2 = \Sigma_{m\geq 2} E(k, m, a)$ with $E(k, m, a) \equiv a^m (1-a)^{k-m} \binom{k}{k-m}$ being the probability that node i with k links has m excited neighbors. We use $p^* = 1 - e^{-p\tau}$. Analogous definitions hold for p_1^* and p_2^* (below). Quantity a can alternatively be interpreted as the equilibrium fraction of excited neurons.

Several properties of the model become intuitively clear with the help of Eq. **1**. For example, if we considered the unidirected interactions by designating some links as incoming and others as outgoing, Eq. **1** would still hold as long as the node degree, k, is replaced with the in-degree $k_{in} \leq k$. Accordingly, the model's quantitative properties would change, but not its qualitative ones. A similar conclusion holds even if we changed the network topology quite considerably. This claim is substantiated by the form of Eq. **1**, which is free of any assumptions on the underlying network topology. The main reason why quantitative details may differ between models is that for more heterogeneous network topologies (e.g., scale-free), the average node degree, k, is much less representative of the true degree of a randomly chosen node than for less heterogeneous topologies (e.g., regular random). A good illustration of how a theoretically predicted bistable region may deviate from a numerically observed one is shown in ref. 29, wherein the authors consider a conceptually similar model. Furthermore, because we constructed the model in such a way that a single firing neighbor of the focal neuron is much less likely to affect this focal neuron than if there were two or more firing neighbors, the terms in Eq. **1** that include p_2^* are generally an order of magnitude higher than the terms that include p_1^*, thus causing the model to be much more responsive to p_2 than to p_1. Note that parameters p_1 and p_2 measure how responsive the network is to an external forcing which, in turn, is determined by the value of parameter p. A consequence is that as p_2 increases, the network should more easily transition into the excited state, which in the (p^*, p_1^*) plane is reflected as a downward shift of both boundaries of the bistable region in Fig. 1 *C* and *D*. However, the shift of the excited state's boundary is much less pronounced at first. As p_2 keeps increasing, the boundary of the excited state eventually touches the line $p_1^* = 0$, meaning that the resting state degenerates in the sense of being accessible only when no outside stimuli occur (i.e., $p = 0$). Finally, relatively strong external forcing compared with the network's responsiveness (i.e., high enough p relative to p_1 and p_2) transitions the network immediately into the excited state. For flipping to occur, the low state must also be accessible, which indeed happens only when external forcing is reasonably weak.

Several analytical results follow from the model assumptions. Let $0 < e(t) < 1$ be the fraction of nodes that are in the externally excited state, where t denotes the current moment in time. We decompose quantity $e(t)$ into the fraction of externally excited nodes with only one excited neighbor, $e_1(t)$, and with two or more excited neighbors, $e_2(t)$. Consequently, $e(t) \equiv e_1(t) + e_2(t)$. At any moment t, each neuron experiences new external excitation with probability $p_1 << 1$, meaning that the clock that counts the time since the last excitation is reset to 0. If we denote the time on this clock with ℓ, then at $\ell = \tau$ a node that was previously externally excited returns to the resting state. It is now possible to infer the time evolution of the fraction of nodes, $c_\ell(t)$, that experienced their last external excitation at moment $t - \ell$. Specifically, if we decompose $c_\ell(t)$ in a manner analogous to $e(t)$, then $c_{1,\ell}(t)$ evolves as

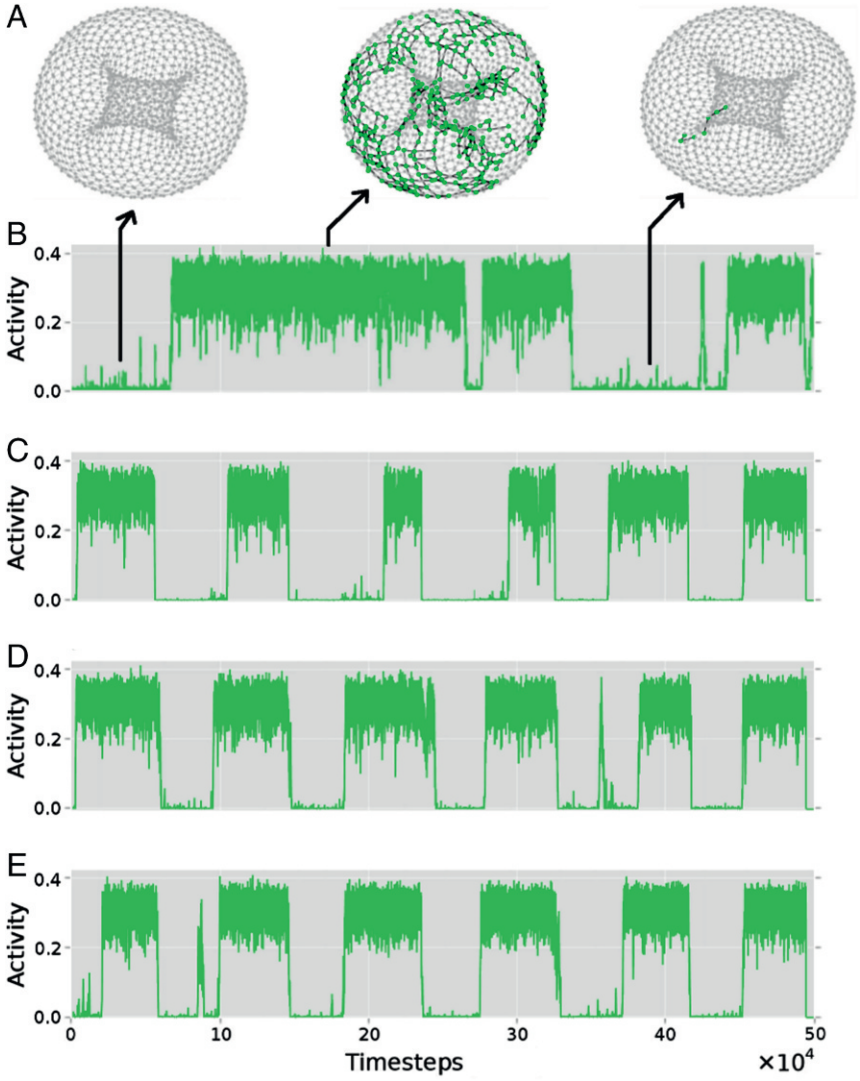

Fig. 2. Activity of a phase-flipping neuronal network. (*A* and *B*) A graphical representation of the network defined in the main text with phase flipping between a mainly excited and a mainly resting phase, which emerges if the network is set appropriately inside a bistable region of the phase space (Fig. 1). In this example, the parameter values are such that the network is at the asterisk-marked point in Fig. 1*C*, while the corresponding equilibrium fraction of nodes with one or more excited neighbors is shown in Fig. 1*D*—$p^* = 0.0004$, $p_1^* = 0.0247$, and $p_2^* = 0.40$. (*C*–*E*) In addition to completely spontaneous phase flipping, regulated stimuli may exert substantial control over the timing of transitions between the two phases. Here, the time series of activity is progressively more periodic as the value of parameter Δp increases. Generally, if Δp equals zero, the network exhibits only spontaneous phase flipping. As Δp approaches unity, stimuli start to dominate over the spontaneous flipping mechanism.

$$c_{1,\ell}(t+1) - c_{1,\ell-1}(t) = -p_1 c_{1,\ell-1}(t), \quad [2]$$

indicating that the decrease in the fraction of nodes that underwent excitation at $t-\ell$ due to their single excited neighbor is possible only because of new excitation at t. Extending the same argument to $c_{2,\ell}(t)$, we obtain

$$c_{2,\ell}(t+1) - c_{2,\ell-1}(t) = -p_2 c_{2,\ell-1}(t). \quad [3]$$

The time evolution of $e_1(t)$ and $e_2(t)$ attains a similar mathematical form that can now be written

$$e_1(t+1) - e_1(t) = p_1[E_1(t) - e_1(t)] - c_{1,\tau}(t), \quad [4]$$

and

$$e_2(t+1) - e_2(t) = p_2[E_2(t) - e_2(t)] - c_{2,\tau}(t) \quad [5]$$

with $E(t)$, $E_1(t)$, and $E_2(t)$ defined in Eq. **1**. The interpretation of the above equations is that the fraction of externally excited nodes (*i*) increases only if some unexcited nodes have excited neighbors and (*ii*) decreases due to the finite time of being in the externally excited state.

Time-evolution equations reveal steady states. In a steady state, the equation for $c_{1,\ell}$ becomes $c_{1,\ell} = c_{1,0}(1-p_1)^\ell \approx E_1 p_1 e^{-p_1 \ell}$ because the initial condition at $\ell = 0$ is $c_{1,0}(t) = E_1 p_1$. Similarly, $c_{2,\ell} = c_{2,0}(1-p_2)^\ell \approx E_2 p_2 e^{-p_2 \ell}$. The steady-state solution for e_1 satisfies condition $p_1(E_1 - e_1) = c_{1,\tau}$, resulting in $e_1 = E_1(1 - e^{-p_1 \tau})$. Analogously, $e_2 = E_2(1 - e^{-p_2 \tau})$. Combining the last two equations finally gives $e \equiv e_1 + e_2 = E_1(1 - e^{-p_1 \tau}) + E_2(1 - e^{-p_2 \tau}) \equiv E_1 p_1^* + E_2 p_2^*$, where we define $p_j^* \equiv 1 - e^{-p_j \tau}$, $j = 1, 2$.

Synaptic Plasticity and the Conservation of Synaptic Strength. To incorporate the strengthening and weakening of synapses into our dynamical network approach, synaptic strengths between neurons are initially assigned random values from an exponential distribution with mean and SD set to unity. The network is furthermore set in a phase-flipping state. To allow homeostatic plasticity to emerge dynamically and spontaneously as the system flips between a predominantly strengthening phase and a predominantly weakening phase, we use assumption *iii*.

Because there is some chance for synapse weakening to occur during the excited phase, our definition of the interaction between two neurons yields that this phase is associated with a net increase in the synaptic strength—synapse strengthening dominates over synapse weakening (15). The resting phase restores the balance because synapse weakening dominates over synapse strengthening. Quantity ϵ_D, however, is sufficiently low that a large number of steps are required until the average synapse strength falls below the baseline level. It is furthermore very unlikely that a randomly chosen synapse of both neurons will be excited during rest. The strength of each synapse thus decreases by a proportional amount almost universally,

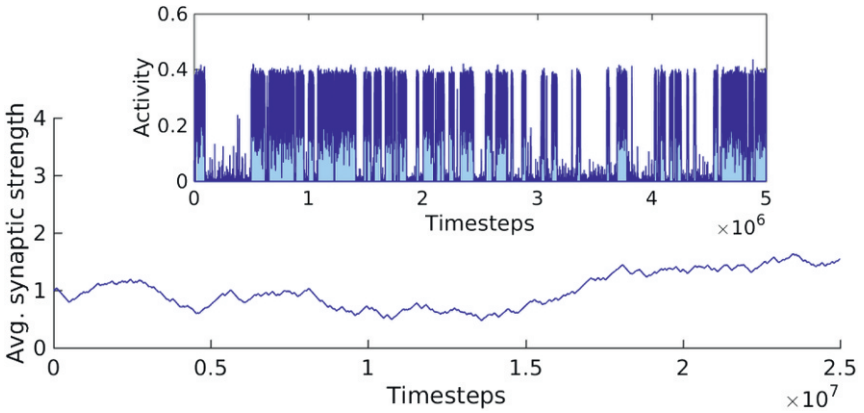

Fig. 3. A realization of the system-wide average synaptic strength in response to network activity (*Inset*). The neuronal network is seen switching between excited and resting phases. In the excited (resting) phase, synaptic strength is predominantly increasing (decreasing). Strengthening and weakening of synapses are the basic processes that exist in two mutually exclusive phases. Recurrent phase transitions of the network maintain the system-wide average synaptic strength near the expected value, thus generating a statistical conservation law that neither of these basic processes can on their own. This conservation law is a new functionality and a manifestation of homeostatic plasticity. Parameter values are $p_s = 1$, $z = 1/16$, and $\sigma = 10^{-4}$.

meaning that the decrease in synaptic strength is a near-collective phenomenon.

A small value of ϵ_D implies that resting must be sufficiently long to enable recovery and new learning, yet how long is sufficient? Because $a \equiv P(X \cup Y)$ is the probability that a randomly chosen neuron is excited, a randomly chosen synapse will have neurons at both its ends excited with probability a^2. In accordance with assumption *iii*, this synapse will strengthen by amount $E(\Delta\epsilon_U) = \sigma$ with probability $a^2 p_S$, whereas with remaining probability $1 - a^2 p_S$, the same synapse will weaken by $E(\Delta\epsilon_D) = z\sigma$. Note that $a_U \equiv a > 0$ in the excited phase differs from $a_D \equiv a \approx 0$ in the resting phase. Accordingly, the average synaptic strength is preserved if

$$\sigma[T_U(a_U^2 p_s - (1 - a_U^2 p_s)z) + T_D(a_D^2 p_s - (1 - a_D^2 p_s)z)] = 0.$$

The ratio of the average fractions of time spent in excited and resting phases—T_U and T_D, respectively—that balances the increases with the decreases in synaptic strength is

$$\frac{T_U}{T_D} = \frac{(1 - a_D^2 p_s)z - a_D^2 p_s}{a_U^2 p_s - (1 - a_U^2 p_s)z}. \qquad [6]$$

If this equality is satisfied, the phase-flipping mechanism and the local mechanism of pairwise neuron interactions lead to a biological conservation law (Fig. 3), i.e., the conservation of the average synaptic strength (15).

We explained multiple qualitative properties of the model in the context of Eq. **1**, but an unanswered question that arises naturally is how network size affects the conservation of synaptic strength. This question is best understood in relation to Eq. **6**, which illustrates the necessity to precisely balance the ratio of time spent in excited and resting phases, T_U/T_D. Namely, spontaneous phase flipping arises when the model is placed in the bistable region of the phase space because of inherent stochasticity whereby the values of parameters p, p_1, and p_2 are realized only approximately in finite networks. The law of large numbers dictates that this approximation is more accurate as the number of nodes increases. A consequence is that the range of realized parameter values becomes stifled by network size, thus making phase transitions rarer. In our model, rarer transitions manifest themselves as an increase of ratio T_U/T_D with network size (Fig. 4). Does this mean that very large networks are necessarily stuck in the excited state? Not at all, because a higher precision guaranteed by the law of large numbers is readily countered with extra stochasticity, which may be introduced into the model through, for example, parameter Δp mentioned previously in the context of the system's regulation (Fig. 2 *C–E*). An interesting implication here is that stochastic regulation is sufficient if a particular value of ratio T_U/T_D is to be maintained, whereas deterministic regulation is needed to keep T_U and T_D individually at a particular value.

Evolution of the Conservation Law of Synaptic Strength. Finally, we address the question of how nature "selects" the parameter values that lead precisely to a conservation law. For the normal activity of many complex systems, parameters must be kept within a narrow range around an optimal point. The normal activity of neurons, for example, critically depends on the constancy of pH, temperature, and electrolyte concentrations (23). In this context, a "weapon" that Nature has at its disposal is evolution or, more specifically, the process of selection. This process works against the diminished fitness of individuals whose parameters violate Eq. **6**, i.e., prevent the conservation of synaptic strength. To see why fitness of such individuals may diminish, let us consider the recent evidence that perturbed deep sleep cannot restore the normal capacity to perform neuroplastic changes associated with wakefulness (16). In turn, learning efficiency becomes impaired (16), causing the fitness to diminish. The described situation can be interpreted as an

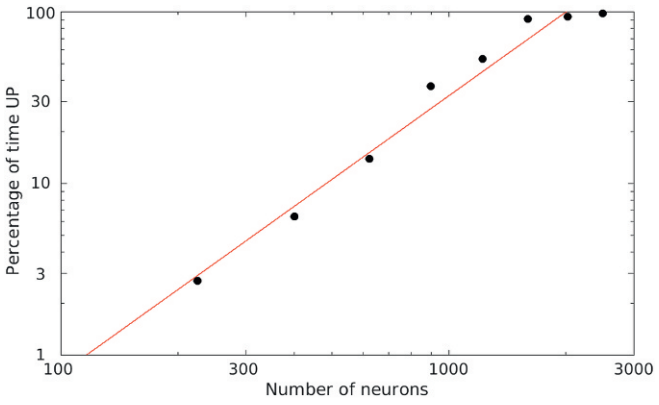

Fig. 4. Percentage of time spent in the excited state as a function of the neuronal network size. Time spent in the excited state monotonically increases with the network size. The red line in the log–log scale helps to show the power law dependence on the network size. Some deviation from this power law is observed as the percentage of time in the excited state approaches 100%. The parameter values used here are the same as in Figs. 1 *C* and *D* and 2—$p^* = 0.0004$, $p_1^* = 0.0247$, and $p_2^* = 0.40$.

imbalanced T_U/T_D ratio in the model because perturbed sleep acts as if decreasing the value of parameter z that regulates the restoration of synaptic strengths in the resting phase. By working against the diminished fitness, selection pressure pushes the system toward an optimal value of z and thus balanced T_U/T_D. An extended version of this argument with the accompanying simulations (Fig. S1) is presented in *SI Results*.

Discussion

We demonstrated that dynamically generated phase flipping between the two different modes of network activity leads to regulated neuron excitability such that a statistical conservation law emerges as a spontaneously generated functionality. In neuroscience, homeostatic plasticity is attributed to synaptic scaling, which is thought to act both globally and locally. Global adjustment implies generating a cell-wide signal that operates on all synapses proportionally, whereas local adjustment allows each synapse to regulate itself in a homeostatic manner. By setting the neuronal network in a phase-flipping state in which one phase predominantly strengthens while the other weakens synapses, homeostatic plasticity emerges dynamically and spontaneously as a result of neuronal activity. Phase flipping between two different phases, therefore, constitutes the feedback mechanism regulating the constancy in neuronal activity.

During the model construction, we paid considerable attention to motivating and/or interpreting each assumption with an underlying physiological mechanism. Thus, each of the two phases corresponds to a level of mental engagement or consciousness (35–37) whereas flipping between the two phases naturally fits the need to recover from fatigue (16, 38). The model in fact has the potential to accommodate a number of other observed or hypothesized processes in the brain. It would be of interest, for example, to couple our approach with that of ref. 39 in which excitatory and inhibitory neurons with strong synaptic couplings lead to rich internal dynamics in response to incoming stimuli, thus providing a substrate for complex information processing and learning. A coupled model could lead to further insight into the connection between homeostatic plasticity and the ability to learn, which has been confirmed in recent experiments (16).

As another example of possible model extensions, we envision two interconnected networks. One network, for instance, may represent the frontal lobe and the other the cerebellum as functional areas in charge of explicit and implicit thought, respectively (40). When attention is directed to a given problem, the excitation of neurons in one network may also excite neurons in the other network due to the interconnectedness of these networks. If explicit thought is more costly, fatigue in the corresponding network may ensue relatively early in the process of problem solving. However, the coupled system may continue working on the problem due to the network for implicit thought. In this case, the reason why the latter network, which houses an internal model of the problem (40), may be able to continue is that implicit thought costs much less energy. A cost reduction may be the result of a simplified structure of the internal model compared with the full mental model of the problem constructed in the network for explicit thought.

ACKNOWLEDGMENTS. We are grateful to M. E. Hasselmo and S. Kuroda for helpful suggestions. B.P. and H.E.S. received support from the Defense Threat Reduction Agency, the Office of Naval Research, and the National Science Foundation Grant CMMI 1125290. B.P. also received support from the Slovenian Research Agency (ARRS, project J5-8236) and the University of Rijeka. M.J. was supported by the Japan Science and Technology Agency Program to Disseminate Tenure Tracking System and the Research Grant Program of the Inamori Foundation. J.M.B. acknowledges financial support from the Ministry of Economy, Industry, and Competitiveness under Project FIS2013-41057-P.

1. Winzeler EA, et al. (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285:901–906.
2. Raman K, Wagner A (2011) The evolvability of programmable hardware. *J R Soc Interface* 8:269–281.
3. Ball P (2015) Strange inevitability of evolution. *Nautilus* 20:1–20.
4. Watts DJ, Strogatz SH (1998) Collective dynamics of small-world networks. *Nature* 393:440–442.
5. He Y, Chen ZJ, Evans AC (2007) Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. *Cereb Cortex* 17:2407–2419.
6. Bassett DS, Meyer-Lindenberg A, Achard S, Duke T, Bullmore E (2006) Adaptive reconfiguration of fractal small-world human brain functional networks. *Proc Natl Acad Sci USA* 103:19518–19523.
7. Kaiser M, Hilgetag CC (2006) Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Comput Biol* 2:e95.
8. Zalesky A, et al. (2010) Whole-brain anatomical networks: Does the choice of nodes matter? *Neuroimage* 50:970–983.
9. Stam CJ (2014) Modern network science of neurological disorders. *Nat Rev Neurosci* 15:683–695.
10. Papo D, Zanin M, Martínez JH, Buldú JM (2016) Beware of the small-world neuroscientist! *Front Hum Neurosci* 10:96.
11. Morone F, Roth K, Min B, Stanley HE, Makse HA (2017) A model of brain activation predicts the neural collective influence map of the human brain. *Proc Natl Acad Sci USA* 114:3849–3854.
12. Schneidman E, Berry MJ, II, Segev R, Bialek W (2006) Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440:1007–1012.
13. Gentet LJ, et al. (2010) Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. *Neuron* 65:422–435.
14. McCormick DA, McGinley MJ, Salkoff DB (2015) Brain state dependent activity in the cortex and thalamus. *Curr Opin Neurobiol* 31:133–140.
15. Tononi G, Cirelli C (2003) Sleep and synaptic homeostasis: A hypothesis. *Brain Res Bull* 62:143–150.
16. Fattinger S, et al. (2017) Deep sleep maintains learning efficiency of the human brain. *Nat Commun* 8:15405.
17. Millman D, Mihalas S, Kirkwood A, Niebur E (2010) Self-organized criticality occurs in non-conservative neuronal networks during up states. *Nat Phys* 6:801–805.
18. Paperin G, Green DG, Sadedin S (2011) Dual phase evolution in complex adaptive systems. *J R Soc Interface* 8:609–629.
19. Parnas I, Segev I (1979) A mathematical model for conduction of action potentials along bifurcating axons. *J Physiol* 295:323–343.
20. Bunde A, Havlin S, eds (1996) *Fractals and Disordered Systems* (Springer, New York).
21. Vogels TP, Abbott LF (2005) Signal propagation and logic gating in networks of integrate-and-fire neurons. *J Neurosci* 25:10786.
22. Hasselmo ME (1999) Neuromodulation: Acetylcholine and memory consolidation. *Trends Cogn Sci* 3:351–359.
23. Turrigiano GG, Nelson SB (2004) Homeostatic plasticity in the developing nervous system. *Nat Rev Neurosci* 5:97–107.
24. Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB (1998) Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature* 391:89–136.
25. Abbott LF, DePasquale B, Raoul-Martin M (2016) Building functional networks of spiking model neurons. *Nat Neurosci* 19:350–355.
26. Mazzoni EO, Desplan C, Blau J (2005) Circadian pacemaker neurons transmit and modulate visual information to control a rapid behavioral response. *Neuron* 45:293–300.
27. Stein RB, Gossen ER, Jones KE (2005) Neuronal variability: Noise or part of the signal? *Nat Rev Neurosci* 6:389–397.
28. Jacobson GA, et al. (2005) Subthreshold voltage noise of rat neocortical pyramidal neurones. *J Physiol* 564:145–160.
29. Majdandzic A, et al. (2014) Spontaneous recovery in dynamical networks. *Nat Phys* 10:34–38.
30. Podobnik B, et al. (2015) The cost of attack in competing networks. *J R Soc Interface* 12:20150770.
31. Kandel E (2006) *The Emergence of a New Science of Mind* (W. W. Norton, New York).
32. Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8:700–711.
33. McCormick DA (2005) Neuronal networks: Flip-flops in the brain. *Curr Biol* 15:R294–R296.
34. Holcman D, Tsodyks M (2006) The emergence of up and down states in cortical networks. *PLoS Comput Biol* 2:e23.
35. Tononi G (2008) Consciousness as integrated information: A provisional manifesto. *Biol Bull* 215:216–242.
36. Alkire MT, Hudetz AG, Tononi G (2008) Consciousness and anesthesia. *Science* 322:876–880.
37. Zhou DW, Mowrey DD, Tang P, Xu Y (2015) Control of mental activities by internal models in the cerebellum. *Phys Rev Lett* 115:108103.
38. Rönnbäck L, Hansson E (2004) On the potential role of glutamate transport in mental fatigue. *J Neuroinflammation* 1:1–22.
39. Ostojic S (2014) Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons *Nat Neurol* 17:594–600.
40. Ito M (2008) Control of mental activities by internal models in the cerebellum. *Nat Rev Neurosci* 9:304–313.
41. Halley JD, Winkler A (2008) Critical-like self-organization and natural selection: Two facets of a single evolutionary process? *Biosystems* 92:148–158.