Astrocyte-induced positive integrated information in neuron-astrocyte ensembles

Oleg Kanakov,[1] Susanna Gordleeva,[1] Anastasia Ermolaeva,[1] Sarika Jalan,[2] and Alexey Zaikin[1,3,4,*]
[1]*Lobachevsky State University of Nizhny Novgorod, Nizhny Novgorod, Russia*
[2]*Complex Systems Lab, Discipline of Physics, Indian Institute of Technology Indore, Simrol, Indore 453552, India*
[3]*Institute for Women's Health and Department of Mathematics, University College London, London, United Kingdom*
[4]*Department of Pediatrics, Faculty of Pediatrics, Sechenov University, Moscow, Russia*

(Received 29 October 2017; revised manuscript received 15 November 2018; published 23 January 2019)

Integrated information is a quantitative measure from information theory of how tightly all parts of a system are interconnected in terms of information exchange. In this study we show that astrocytes, playing an important role in regulation of information transmission between neurons, may contribute to a generation of positive integrated information in neuronal ensembles. Analytically and numerically we show that the presence of astrocytic regulation of neurotransmission may be essential for this information attribute in neuroastrocytic ensembles. Moreover, the proposed "spiking-bursting" mechanism of generating positive integrated information is shown to be generic and not limited to neuron-astrocyte networks and is given a complete analytic description.

DOI: 10.1103/PhysRevE.99.012418

I. INTRODUCTION

The integrated information (II) concept introduced in Ref. [1] marked a milestone in the ongoing effort to describe activities of neural ensembles and brain by means of information theory. II was proposed as a quantitative measure of how tightly all parts of a system are interconnected in terms of information exchange (for example, a combination of two noninteracting subsystems implies zero II). The ambitious aim of the II concept was to quantify consciousness [2]—in particular, for medical applications to detecting consciousness in a completely immobilized patient by electroencephalographic data. Several mathematical definitions of II [3–6] have been proposed since the original work, all in line with the initial idea. The perturbational complexity index, linked to II as its proxy, has reliably discriminated the level of consciousness in patients during wakefulness, sleep, anesthesia, and even in patients who has emerged from coma with minimal level of consciousness [7]. Although the relation of II to consciousness has been debated [8–10], II itself is by now widely adopted as a quantitative measure for complex dynamics [11–13]. Accordingly, the understanding of particular mechanisms producing positive II in neural ensembles is of topical interest.

The experiments have shown that astrocytes play an important role by regulating cellular functions and information transmission in the nervous system [14,15]. It was proposed that astrocyte wrapping a synapse implements a feedback control circuit which maximizes information transmission through the synapse by regulating neurotransmitter release [16]. The involvement of astrocytes in neuron-astrocyte network dynamics was quantified by estimating functional connectivity between neurons and astrocytes from time-lapse Ca^{2+} imaging data [17]. In contrast with neuronal cells the astrocytes do not generate electrical excitations (action potentials). However, their intracellular dynamics have shown similar excitable properties for changes of calcium concentration [18]. These signals can remarkably affect neuronal excitability and the efficiency of synaptic transmission between neurons by Ca^{2+}-dependent release of neuroactive chemicals (e.g., glutamate, ATP, D-serine and GABA) [19]. Networks of astrocytes accompanying neuronal cells generate collective activity patterns that can regulate neuronal signaling by facilitating or by suppressing synaptic transmission [14,15,20].

In this study we show that astrocytes may conduce to positive II in neuronal ensembles. We calculate II in a small neuro-astrocytic network with random topology by numerical simulation and find that positive II is conditioned by coupling of neurons to astrocytes and increases with spontaneous neuronal spiking activity. We explain this behavior using simplified spiking-bursting dynamics, which we implement both in the neuroastrocytic network model with all-to-all connectivity between neurons, showing astrocyte-induced coordinated bursting, and as well in a specially defined stochastic process allowing analytical calculation of II. The analytical and simulation results for the all-to-all network are in good agreement. That said, nontrivial dynamics of the random version of the network, although not being directly compatible with our analytical treatment, turns out to be even more favorable for positive *II* than the spiking-bursting dynamics of the all-to-all network. We speculate that the presence of astrocytic regulation of neurotransmission may be essential for generating positive II in larger neuroastrocytic ensembles.

II. METHODS AND MODEL

Neural network under study consists of 6 synaptically coupled Hodgkin-Huxley neurons [21]. We use 2 variants of neural network architecture: (i) network of 1 inhibitory and 5 excitatory neurons with coupling topology obtained by randomly picking 1/3 of the total number of connections out of the full directed graph, excluding self-connections [the particular instance of random topology for which the

*alexey.zaikin@ucl.ac.uk

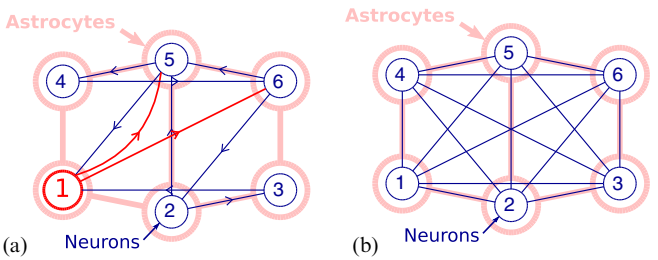

FIG. 1. Schemes of neuroastrocytic networks under study: (a) instance of random neuronal network topology; (b) all-to-all neuronal network. Inhibitory neuron is shown with an enlarged symbol and highlighted in red. Connections without arrows are bidirectional. Each astrocyte is coupled to one corresponding neuron and acts by modulating outgoing connections of the neuron.

presented data have been obtained is shown in Fig. 1(a)]; (ii) all-to-all network of 6 excitatory neurons [Fig. 1(b)].

The membrane potential of a single neuron evolves according to the following ionic current balance equation:

$$C\frac{dV^{(i)}}{dt} = I^{(i)}_{\text{channel}} + I^{(i)}_{\text{app}} + \sum_j I^{(ij)}_{\text{syn}} + I^{(i)}_P, \qquad (1)$$

where the superscript $(i = 1, \ldots, 6)$ corresponds to a neuronal index and (j) corresponds to an index of input connection. Ionic currents (i.e., sodium, potassium, and leak currents) are expressed as follows:

$$I_{\text{channel}} = -g_{Na}m^3h(V - E_{Na}) - g_K n^4(V - E_K) - g_{\text{leak}}(V - E_{\text{leak}}),$$
$$\frac{dx}{dt} = \alpha_x(1 - x) - \beta_x x, \quad x = m, n, h. \qquad (2)$$

Nonlinear functions α_x and β_x for gating variables are taken as in original Hodgkin-Huxley model with membrane potential V shifted by 65 mV. Throughout this paper we use the following parameter values: $E_{Na} = 55$ mV, $E_K = -77$ mV, $E_{\text{leak}} = -54.4$ mV, $g_{Na} = 120$ mS/cm^2, $g_K = 36$ mS/cm^2, $g_{\text{leak}} = 0.3$ mS/cm^2, $C = 1$ μF/cm^2. The applied currents $I^{(i)}_{\text{app}}$ are fixed at constant value controlling the depolarization level and dynamical regime that can be either excitable, oscillatory or bistable [22]. We use $I^{(i)}_{\text{app}} = -5.0$ μA/cm^2 which corresponds to the excitable regime. The synaptic current I_{syn} simulating interactions between the neurons obeys the equation

$$I^{(ij)}_{\text{syn}} = \frac{g_{\text{syneff}}(V^{(j)} - E_{\text{syn}})}{1 + \exp\left(\frac{-(V^{(i)} - \theta_{\text{syn}})}{k_{\text{syn}}}\right)}, \qquad (3)$$

where $E_{\text{syn}} = -90$ mV for the inhibitory synapse and $E_{\text{syn}} = 0$ mV for the excitatory. Neural network composition of one inhibitory and five excitatory neurons is in line with the experimental data showing that the fraction of inhibitory neurons is about 20% [23]. Variable g_{syneff} describes the synaptic weight in mS/cm^2 modulated by an astrocyte [as defined by Eq. (8) below], parameters $\theta_{\text{syn}} = 0$ mV and $k_{\text{syn}} = 0.2$ mV describe the midpoint of the synaptic activation and the slope of its threshold, respectively.

Each neuron is stimulated by a Poisson pulse train mimicking external spiking inputs $I^{(i)}_P$ with a certain average rate λ. Each Poisson pulse has constant duration 10 ms and constant amplitude, which is sampled independently for each pulse from uniform random distribution on interval $[-1.8, 1.8]$. Sequences of Poisson pulses applied to different neurons are independent.

Note that the time unit in the neuronal model Eqs. (1) and (2) is 1 ms. Due to a slower timescale, in the astrocytic model (see below) empirical constants are indicated using seconds as time units. When integrating the joint system of differential equations, the astrocytic model time is rescaled so that the units in both models match up.

We consider astrocytic network in the form of a two-dimensional square lattice with only nearest-neighbor connections [24]. Such topology for the Ca^{2+}- and IP_3-diffusion model is justified by experimental findings stating that astrocytes occupy "nonoverlapping" territories [25]. The neuroastrocyte network of real brain has a 3D structure with one astrocyte interacting with several neurons and vise versa. However, in our modeling we use a simplified approach. The latter reflects the fact that, throughout area CA1 of the hippocampus, pyramidal (excitatory) cells are arranged in a regular layer and surrounded by a relatively uniform scatter of astrocytes [26]. According to the experimental data [26,27], modeled astrocytes are distributed evenly across the neural network, with a total cell number equaled to the number of neurons (due to small size of networks, astrocyte network is modeled as a 2D lattice in our study). Astrocytes and neurons communicate via a special mechanism modulated by neurotransmitters from both sides. The model is designed so that when the calcium level inside an astrocyte exceeds a threshold, the astrocyte releases neuromodulator (e.g., glutamate) that may affect the release probability (and thus a synaptic strength) at neighboring connections in a tissue volume [28]. Single astrocyte can regulate the synaptic strength of several neighboring synapses which belong to one neuron or several different neurons, but since we do not take into account the complex morphological structure of the astrocyte, we assume for simplicity that one astrocyte interacts with one neuron.

In a number of previous studies a biophysical mechanism underlying calcium dynamics of astrocytes has been extensively investigated [29,30]. Calcium is released from internal stores, mostly from the endoplasmic reticulum (ER). This process is regulated by inositol 1,4,5-trisphosphate (IP_3) that activates IP_3 channels in the ER membrane resulting in a Ca^{2+} influx from ER. IP_3 acting as a second messenger is produced when neurotransmitter (e.g., glutamate) molecules are bound by metabotropic receptors of the astrocyte. In turn IP_3 can be regenerated depending on the level of calcium by the phospholipase C-δ (PLC-δ). State variables of each cell include IP_3 concentration IP_3, Ca^{2+} concentration Ca, and the fraction of activated IP_3 receptors h. They evolve according to the following equations [29,30]:

$$\frac{d\text{Ca}^{(m,n)}}{dt} = J^{(m,n)}_{\text{ER}} - J^{(m,n)}_{\text{pump}} + J^{(m,n)}_{\text{leak}} + J^{(m,n)}_{\text{in}} - J^{(m,n)}_{\text{out}} + J^{(m,n)}_{\text{Cadiff}},$$
$$\frac{d\text{IP}^{(m,n)}_3}{dt} = \frac{\text{IP}^*_3 - \text{IP}^{(m,n)}_3}{\tau_{\text{IP3}}} + J^{(m,n)}_{\text{PLC}} + J^{(m,n)}_{\text{IP3diff}}, \qquad (4)$$
$$\frac{dh^{(m,n)}}{dt} = a_2\left[d_2\frac{\text{IP}^{(m,n)}_3 + d_1}{\text{IP}^{(m,n)}_3 + d_3}(1 - h^{m,n}) - \text{Ca}^{m,n}h^{m,n}\right],$$

with $m = 1, \ldots, 3$, $n = 1, 2$. Currents J_{ER} is Ca^{2+} current from the ER to the cytoplasm, J_{pump} is the ATP pumping current, J_{leak} is the leak current, J_{in} and J_{out} describe calcium exchanges with extracellular space, J_{PLC} is the calcium-dependent PLC-δ current and are expressed as follows:

$$
\begin{aligned}
J_{\mathrm{ER}} &= c_1 v_1 \mathrm{Ca}^3 h^3 \mathrm{IP}_3^3 \frac{(c_0/c_1 - (1 + 1/c_1)\mathrm{Ca})}{[(\mathrm{IP}_3 + d_1)(\mathrm{IP}_3 + d_5)]^3}, \\
J_{\mathrm{pump}} &= \frac{v_3 \mathrm{Ca}^2}{k_3^2 + \mathrm{Ca}^2}, \\
J_{\mathrm{leak}} &= c_1 v_2 [c_0/c_1 - (1 + 1/c_1)\mathrm{Ca}], \\
J_{\mathrm{in}} &= v_5 + \frac{v_6 \mathrm{IP}_3^2}{k_2^2 + \mathrm{IP}_3^2}, \\
J_{\mathrm{out}} &= k_1 \mathrm{Ca}, \\
J_{\mathrm{PLC}} &= \frac{v_4 [\mathrm{Ca} + (1 - \alpha) k_4]}{\mathrm{Ca} + k_4}.
\end{aligned} \tag{5}
$$

Biophysical meaning of all parameters in Eqs. (4) and (5) and their values determined experimentally can be found in Refs. [29,30]. For our purpose we fix $c_0 = 2.0 \mu\mathrm{M}$, $c_1 = 0.185$, $v_1 = 6\ \mathrm{s}^{-1}$, $v_2 = 0.11\ \mathrm{s}^{-1}$, $v_3 = 2.2 \mu\mathrm{Ms}^{-1}$, $v_5 = 0.025\ \mu\mathrm{Ms}^{-1}$, $v_6 = 0.2\ \mu\mathrm{Ms}^{-1}$, $k_1 = 0.5\ \mathrm{s}^{-1}$, $k_2 = 1.0\ \mu\mathrm{M}$, $k_3 = 0.1\ \mu\mathrm{M}$, $a_2 = 0.14 \mu\mathrm{M}^{-1}\mathrm{s}^{-1}$, $d_1 = 0.13\ \mu\mathrm{M}$, $d_2 = 1.049\ \mu\mathrm{M}$, $d_3 = 0.9434\ \mu\mathrm{M}$, $d_5 = 0.082\ \mu\mathrm{M}$, $\alpha = 0.8$, $\tau_{\mathrm{IP3}} = 7.143$ s, $\mathrm{IP}_3^* = 0.16\ \mu\mathrm{M}$, $k_4 = 1.1\ \mu\mathrm{M}$ [31]. Parameter v_4 describes the rate of IP_3 regeneration and controls the dynamical regime of the model Eqs. (4) and (5) that can be excitable at $v_4 = 0.3\ \mu\mathrm{Ms}^{-1}$, or oscillatory at $v_4 = 0.5\ \mu\mathrm{Ms}^{-1}$ [30]. Here we limit ourselves to the oscillatory case.

Currents J_{Cadiff} and J_{IP3diff} describe the diffusion of Ca^{2+} ions and IP_3 molecules via gap junctions between astrocytes in the network and can be expressed as follows [24]:

$$
\begin{aligned}
J_{\mathrm{Cadiff}}^{(m,n)} &= d_{\mathrm{Ca}} (\Delta \mathrm{Ca})^{(m,n)}, \\
J_{\mathrm{IP3diff}}^{(m,n)} &= d_{\mathrm{IP3}} (\Delta \mathrm{IP}_3)^{(m,n)},
\end{aligned} \tag{6}
$$

where parameters $d_{\mathrm{Ca}} = 0.001\ \mathrm{s}^{-1}$ and $d_{\mathrm{IP3}} = 0.12 \mathrm{s}^{-1}$ describe the Ca^{2+} and IP_3 diffusion rates, respectively. $(\Delta \mathrm{Ca})^{(m,n)}$ and $(\Delta \mathrm{IP}_3)^{(m,n)}$ are discrete Laplace operators:

$$
\begin{aligned}
(\Delta \mathrm{Ca})^{(m,n)} = (&\mathrm{Ca}^{(m+1,n)} + \mathrm{Ca}^{(m-1,n)} \\
&+ \mathrm{Ca}^{(m,n+1)} + \mathrm{Ca}^{(m,n-1)} - 4\mathrm{Ca}^{(m,n)}).
\end{aligned} \tag{7}
$$

Astrocytes can modify release probability of nearby synapses in tissue volume [14], likely by releasing signaling molecules ("gliotransmitters") in a Ca^{2+}-dependent manner [15]. We proposed that each astrocyte from the network interacts to the one neuron from the neural network by modulation of the synaptic weight. For the sake of simplicity, the effect of astrocyte calcium concentration Ca upon synaptic weight of the affected synapses g_{syneff} [which appears in Eq. (3)] has been described with the simple formalism based on earlier suggestions [32–34]:

$$
g_{\mathrm{syneff}} = \begin{cases} g_{\mathrm{syn}}(1 + g_{\mathrm{astro}} \mathrm{Ca}^{(m,n)}), & \text{if } \mathrm{Ca}^{(m,n)} > 0.2, \\ g_{\mathrm{syn}}, & \text{otherwise}, \end{cases} \tag{8}
$$

where $g_{\mathrm{syn}} = 0.04\ \mathrm{mS/cm}^2$ is baseline synaptic weight, parameter $g_{\mathrm{astro}} > 0$ controls the strength of synaptic weight modulation, and $\mathrm{Ca}^{(m,n)}$ is the intracellular calcium concentration in the astrocyte Eq. (4).

In general, phenomena of astrocytic neuromodulation are highly versatile and depend upon the actual gliotransmitter and its target [15], which in particular may lead to the inhibition of synaptic transmission instead of its potentiation. In this sense, the model Eq. (8) is not universal, but we anticipate its at least qualitative applicability to cases where synaptic strength potentiation by astrocytes has been confirmed experimentally. These include the impact of astrocytic glutamate upon presynaptic terminals leading to potentiating excitatory transmission in the hippocampal dentate gyrus [35], and both excitatory [28,36,37] and inhibitory [38,39] synaptic transmission in the CA1 area of hippocampus. In addition, glutamate action on postsynamtic terminals was also shown to improve neuronal synchrony [40].

The time series of neuron membrane potentials $V^{(i)}(t)$ are converted into binary-valued discrete-time processes according to Ref. [41] as follows. Time is split into windows of duration T which become units of the discrete time. If inequality $V^{(i)}(t) > V_{\mathrm{thr}} = -40.0$ mV is satisfied for at least some t within a particular time window (essentially, if there was a spike in this time window), than the corresponding binary value (bit) is assigned 1, and 0 otherwise. The size of time window is chosen so that spontaneous spiking activity produces time-uncorrelated spatial patterns, but a burst shows as a train of successive 1's in the corresponding bit.

We use the definition of II according to Ref. [3] as follows. Consider a stationary stochastic process $\xi(t)$ (binary vector process), whose instantaneous state is described by $N = 6$ bits. The full set of N bits ("system") can be split into two nonoverlapping nonempty subsets of bits ("subsystems") A and B, such splitting further referred to as bipartition AB. Denote by $x = \xi(t)$ and $y = \xi(t + \tau)$ two states of the process separated by a specified time interval $\tau \neq 0$. States of the subsystems are denoted as x_A, x_B, y_A, y_B.

Mutual information between x and y is defined as

$$
I_{xy} = H_x + H_y - H_{xy}, \tag{9}
$$

where $H_x = -\sum_x p_x \log_2 p_x$ is entropy (base 2 logarithm gives result in bits), $H_y = H_x$ due to stationarity which is assumed. Next, a bipartition AB is considered, and "effective information" as a function of the particular bipartition is defined as

$$
I_{\mathrm{eff}}(AB) = I_{xy} - I_{x_A, y_A} - I_{x_B, y_B}. \tag{10}
$$

II is then defined as effective information calculated for a specific bipartition AB^{MIB} ("minimum information bipartition") which minimizes specifically normalized effective information:

$$
II = I_{\mathrm{eff}}(AB^{\mathrm{MIB}}), \tag{11a}
$$

$$
AB^{\mathrm{MIB}} = \mathrm{argmin}_{AB} \left\{ \frac{I_{\mathrm{eff}}(AB)}{\min[H(x_A), H(x_B)]} \right\}. \tag{11b}
$$

Note that this definition prohibits positive II, when I_{eff} turns out to be zero or negative for at least one bipartition AB.

In essence, entropy H_x generalizes the idea of measuring an effective number of independent bits in x. For example, if

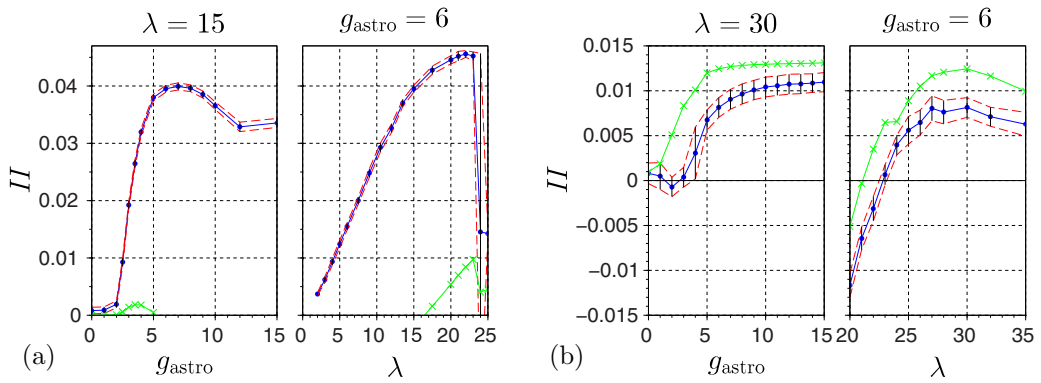

FIG. 2. Dependence of II upon neuroastrocytic interaction g_{astro} and upon average stimulation rate λ: (a) in random network [instance shown in Fig. 1(a)]; (b) in all-to-all network [Fig. 1(b)]. Blue solid lines with dot marks—direct calculation by definition from simulation data; red dashed lines—error estimation; green lines with cross marks—analytical calculation for spiking-bursting process with parameters estimated from simulation data.

all N bits in x are independent and are "fair coins" (have equal probability 1/2 of getting 0 or 1), then $H_x = N$. If x consists of m independent groups of bits which are fully synchronized within each group (or all bits in x are uniquely expressed in terms of $m < N$ independent fair coins), then $H_x = m$.

In the same conceptual sense, mutual information I_{xy} Eq. (9) measures the degree of dependence (effective number of dependent bits) between two random events x and y. In case of causality, when dependence is unidirectional, one can speak of degree of predictability instead. For example, if y exactly repeats x (full predictability) with all bits in x being independent fair coins, then in Eq. (9) $H_x = H_y = H_{xy} = N$, which gives $I_{xy} = N$. If instead y and x are totally independent (absence of predictability), then $H_x = H_y = N$, $H_{xy} = 2N$, $I_{xy} = 0$. In an intermediate situation, when only m bits in y exactly repeat the corresponding bits in x (like perfect transmission lines), with other $N - m$ bits in y (acting as randomly failing transmission lines) and all bits in x being independent fair coins, then again $H_x = H_y = N$, but now $H_{xy} = 2N - m$, because m out of total $2N$ bits in the combination xy are expressed in terms of others, which leaves $2N - m$ independent bits. According to Eq. (9), this yields $I_{xy} = m$. The definition of mutual information Eq. (9) generalizes this idea, retaining its applicability in case of arbitrary dependence between two random events, even when this dependence can not be attributed to specific bits.

In turn, effective information Eq. (10) measures how much the system is more predictable as a whole than when trying to predict the subsystems separately. Trivial cases when I_{eff} is zero are (i) independent subsystems (then system as a whole is equally predictable as a combination of the parts) and (ii) complete absence of predictability (when all mutual informations are zero). When the system is fully synchronized (all bits are equal in any instance of time), for any bipartition we get $I_{xy} = I_{x_A,y_A} = I_{x_B,y_B}$, which implies $I_{\text{eff}} < 0$ according to Eq. (10). From Eqs. (11a) and (11b) we conclude that II is zero or negative in the mentioned cases. The idea behind the choice of "minimal information bipartition" AB^{MIB} in Eq. (11), according to Ref. [3], is to identify the worst-case partition in terms of information interconnection, but with preference to nontrivial partitions with roughly similarly sized subsystems, which is achieved by normalization in Eq. (11b). For more detail on the rationale behind the used definition of II we refer the reader to the original paper [3], and for the general concept of II—to the papers cited in the Introduction.

III. RESULTS

We calculated II directly, according to the definition above, using empirical probabilities from binarized time series of simulated neuroastrocytic networks of both mentioned architectures [Figs. 1(a) and 1(b)]. For each architecture we performed two series of simulation runs: (i) with constant Poissonian stimulation rate λ (equal 15.0 Hz for the random network and 30.0 Hz for the all-to-all network) and neuroastrocytic interaction g_{astro} varied, (ii) with constant $g_{\text{astro}} = 6.0$ and λ varied, other model parameters as indicated above. Time window T used in binarization and time delay τ used in computation of II are $\tau = T = 0.1$ s for the random network, and $\tau = T = 0.2$ s for the all-to-all network. The length of time series to calculate each point is 5×10^5 s, taken after 2×10^3 s transient time. The estimate of II shows convergence as the length of time series is increased. Error due to finite data (shown as half-height of errorbar in the graphs) is estimated as maximal absolute difference between the result for the whole observation time and for each its half taken separately. Obtained dependencies of II upon g_{astro} and λ are shown in Fig. 2.

For the random topology [Fig. 2(a)] we observe that (i) positive II is greatly facilitated by nonzero g_{astro} (i.e., by the effect of astrocytes), although small positive quantities, still exceeding the error estimate, are observed even at $g_{\text{astro}} = 0$; (ii) II generally increases with the average stimulation frequency λ which determines the spontaneous activity in the network [42].

The visible impact of astrocytes on the network dynamics consists in the stimulation of space-time patterns of neuronal activity due to astrocyte-controlled increase in the neuronal synaptic connectivity on astrocyte timescale. An instance of such pattern of activation for the random network is shown as a raster plot in Fig. 3(a). The pattern is rather complex, and we only assume that II must be determined by properties of this pattern, which in turn is controlled by astrocytic interaction (as well as by network topology and by external inputs to the

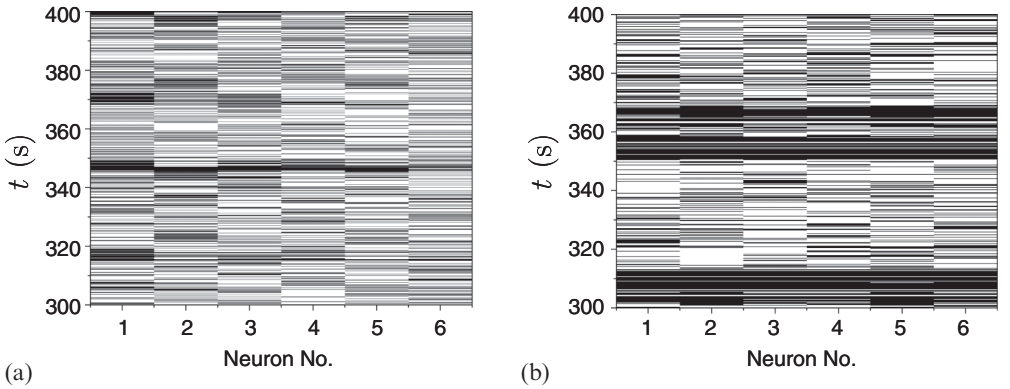

FIG. 3. Raster plots of neuronal dynamics at $g_{\text{astro}} = 6$, $\lambda = 20$: (a) in random network [instance shown in Fig. 1(a)]; (b) in all-to-all network [Fig. 1(b)]. White and black correspond to 0 and 1 in binarized time series.

neurons represented by Poissonian processes in the model). We currently do not identify specific properties of activation patterns linked to the behavior of II in the random network; however, we do it (see below) for the all-to-all network of identical (all excitatory) neurons, due to its simpler "spiking-bursting" type of spatiotemporal dynamics consisting of coordinated system-wide bursts overlaid upon background spiking activity, see raster plot in Fig. 3(b). As seen in Fig. 2(b), this network retains the generally increasing dependence of II upon g_{astro} and λ, with the most notable difference being that II is negative until λ exceeds a certain threshold.

To confirm the capacity of II as a quantitative indicator for properties of complex dynamics in application to the system under study, we additionally consider graphs of mutual information I_{xy} in the same settings, see Fig. 4 [note a greater range over λ in Fig. 4(b) as compared to Fig. 2(b)]. Comparing Figs. 2 to 4 we observe a qualitative difference in dependencies upon λ in case of all-to-all network [Figs. 2(b) and 4(b)]: while mutual information decreases with the increase of λ, II is found to grow, and transits from negative to positive values before reaching its maximum. It means that even while the overall predictability of the system is waning, the system becomes more integrated in the sense that the advantage in this predictability when the system is taken as a whole over considering it by parts is found to grow. This confirms the capability of II to capture features of complex dynamics that are not seen when using only mutual information.

Our analytical consideration is based upon mimicking the spiking-bursting dynamics by a model stochastic process which admits analytical calculation of effective information. We define this process $\xi(t)$ as a superposition of a time-correlated dichotomous component which turns system-wide bursting on and off, and a time-uncorrelated component describing spontaneous activity which occurs in the absence of a burst, in the following way.

At each instance of time the state of the dichotomous component can be either "bursting" with probability p_b, or "spontaneous" (or "spiking") with probability $p_s = 1 - p_b$. While in the bursting mode, the instantaneous state of the resulting process $x = \xi(t)$ is given by all ones: $x = 11..1$ (further abbreviated as $x = 1$). In case of spiking, the state x is a random variate described by a discrete probability distribution s_x, so that the resulting one-time state probabilities read

$$p(x \neq 1) = p_s s_x, \tag{12a}$$

$$p(x = 1) = p_1, \quad p_1 = p_s s_1 + p_b, \tag{12b}$$

where s_1 is the probability of spontaneous occurrence of $x = 1$ in the absence of a burst (all neurons spontaneously spiking within the same time discretization window).

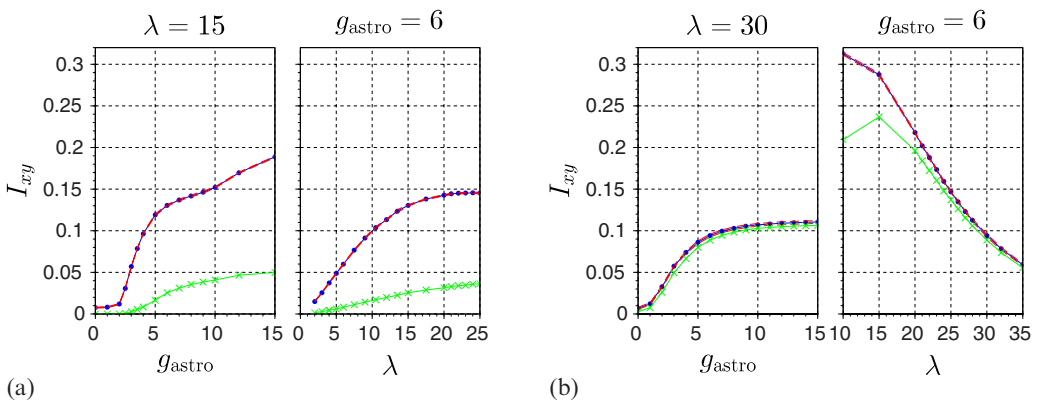

FIG. 4. Dependence of mutual information I_{xy} upon neuroastrocytic interaction g_{astro} and upon average stimulation rate λ: (a) in random network (instance shown in Fig. 1(a)); (b) in all-to-all network [Fig. 1(b)]. Legend as in Fig. 2.

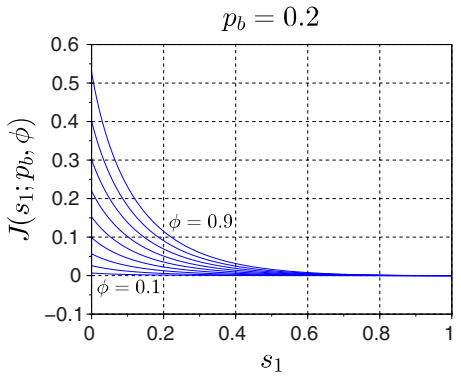

FIG. 5. Family of plots of $J(s_1; p_b, \phi)$ at $p_b = 0.2$ and ϕ varied from 0.1 to 0.9 with step 0.1.

To describe two-time joint probabilities for $x = \xi(t)$ and $y = \xi(t + \tau)$, we consider a joint state xy which is a concatenation of bits in x and y. The spontaneous activity is assumed to be uncorrelated in time: $s_{xy} = s_x s_y$. The time correlations of the dichotomous component are described by a 2×2 matrix of probabilities p_{ss}, p_{sb}, p_{bs}, p_{bb}, which denote joint probabilities to observe the respective spiking and/or bursting states in x and y. The probabilities obey $p_{sb} = p_{bs}$ (due to stationarity), $p_b = p_{bb} + p_{sb}$, $p_s = p_{ss} + p_{sb}$, thereby allowing to express all one- and two-time probabilities describing the dichotomous component in terms of two quantities, for which we chose p_b and correlation coefficient ϕ defined by

$$p_{sb} = p_s p_b (1 - \phi). \tag{13}$$

The two-time joint probabilities for the resulting process are then expressed as

$$p(x \neq 1, y \neq 1) = p_{ss} s_x s_y, \tag{14a}$$

$$p(x \neq 1, y = 1) = \pi s_x, \quad p(x = 1, y \neq 1) = \pi s_y, \tag{14b}$$

$$p(x = 1, y = 1) = p_{11}, \tag{14c}$$

$$\pi = p_{ss} s_1 + p_{sb}, \quad p_{11} = p_{ss} s_1^2 + 2 p_{sb} s_1 + p_{bb}. \tag{14d}$$

Note that the above notations can be applied to any subsystem instead of the whole system (with the same dichotomous component, as it is system-wide anyway).

For this spiking-bursting process, the expression for mutual information of x and y Eq. (9) after substitution of probabilities Eqs. (12) and (14) and algebraic simplifications reduces to

$$I_{xy} = 2(1 - s_1)\{p_s\} + 2\{p_1\} - (1 - s_1)^2\{p_{ss}\} - 2(1 - s_1)\{\pi\} - \{p_{11}\} = J(s_1; p_b, \phi), \tag{15}$$

where we denote $\{q\} = -q \log_2 q$ for compactness. With expressions for p_1, p_{11}, π from Eqs. (12b) and (14d) taken into account, I_{xy} can be viewed as a function of s_1, denoted in Eq. (15) as $J(\cdot)$, with two parameters p_b and ϕ characterizing the dichotomous (bursting) component.

A typical family of plots of $J(s_1; p_b, \phi)$ versus s_1 at $p_b = 0.2$ and ϕ varied from 0.1 to 0.9 is shown in Fig. 5. Important particular cases are

$$J(s_1 = 0) = 2\{p_s\} + 2\{p_b\} - \{p_{ss}\} - 2\{p_{sb}\} - \{p_{bb}\} > 0,$$

which is the information of the dichotomous component alone; $J(s_1 = 1) = 0$ (degenarate case—"always on" deterministic state); $J(s_1) \equiv 0$ for any s_1 when $p_b = 0$ or $\phi = 0$ (absent or time-uncorrelated bursting). Otherwise, $J(s_1)$ is a positive decreasing function on $s_1 \in [0, 1)$.

Derivation of Eq. (15) does not impose any assumptions on the specific type of the spiking probability distribution s_x. In particular, spikes can be correlated across the system (but not in time). Note that Eq. (15) is applicable as well to any subsystem A (B), with s_1 replaced by s_A (s_B) which denotes the probability of a subsystem-wide simultaneous (within the same time discretization window) spike $x_A = 1$ ($x_B = 1$) in the absence of a burst, and with same parameters of the dichotomous component (here p_b, ϕ). Effective information Eq. (10) is then written as

$$I_{\text{eff}}(AB) = J(s_1) - J(s_A) - J(s_B). \tag{16}$$

Since as mentioned above $p_b = 0$ or $\phi = 0$ implies $J(s_1) = 0$ for any s_1, this leads to $I_{\text{eff}} = 0$ for any bipartition, and, accordingly, to zero II, which agrees with our simulation results [left panels in Figs. 2(a) and 2(b)], where this case corresponds to the absence of coordinated activity induced by astrocytes ($g_{\text{astro}} = 0$).

Consider the case of independent spiking with

$$s_1 = \prod_{i=1}^{N} P_i, \tag{17}$$

where P_i is the spontaneous spiking probability for an individual bit (neuron). Then $s_A = \prod_{i \in A} P_i$, $s_B = \prod_{i \in B} P_i$, $s_1 = s_A s_B$. Denoting $s_A = s_1^{\nu}$, $s_B = s_1^{1-\nu}$, we rewrite Eq. (16) as

$$I_{\text{eff}}(s_1; \nu) = J(s_1) - J\left(s_1^{\nu}\right) - J\left(s_1^{1-\nu}\right), \tag{18}$$

where ν is determined by the particular bipartition AB.

Figure 6 shows typical families of plots of $I_{\text{eff}}(s_1; \nu = 0.5)$ at $p_b = 0.2$ and ϕ varied from 0.1 to 0.9 in Fig. 6(a) [with increase of ϕ, maximum of $I_{\text{eff}}(s_1)$ grows], and at $\phi = 0.2$ with p_b varied from 0.02 to 0.2 in Fig. 6(b) [with increase of p_b, root and maximum of $I_{\text{eff}}(s_1)$ shift to the right].

Hereinafter assuming $\phi \neq 0$ and $p_b \neq 0$, we notice the following: first, $I_{\text{eff}}(s_1 = 0) = -J(0) < 0$, which implies $II < 0$; second, $I_{\text{eff}}(s_1 = 1) = 0$; third, at $\phi > 0$ function $I_{\text{eff}}(s_1)$ has a root and a positive maximum in interval $s_1 \in (0, 1)$. It implies that absent or insufficient spontaneous spiking activity leads to negative II, while the increase in spiking turns II positive. This is exactly observed in the all-to-all network simulation results, where spiking is determined by λ, see Fig. 2(b) (right panel). It can be additionally noticed in Fig. 6 that the root of $I_{\text{eff}}(s_1)$ (which is essentially the threshold in s_1 for positive II) shows a stronger dependence upon the burst probability p_b than upon correlation coefficient of bursting activity ϕ.

Furthermore, expanding the last term of Eq. (18) in powers of ν yields

$$I_{\text{eff}} = -J\left(s_1^{\nu}\right) + \nu \cdot s_1 \log s_1 J'(s_1) + \cdots \tag{19}$$

Consider the limit of large system $N \to \infty$ and a special bipartition with subsystem A consisting of only one bit (neuron). Assuming that individual spontaneous spike probabilities of neurons P_i in (17) retain their order of magnitude (in

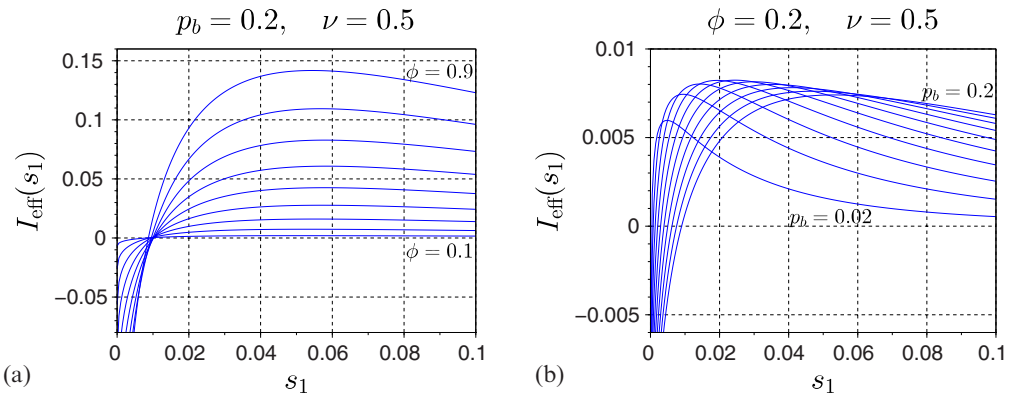

FIG. 6. Families of plots of $I_{\text{eff}}(s_1; \nu = 0.5)$: (a) at $p_b = 0.2$ and ϕ varied from 0.1 to 0.9 with step 0.1; (b) at $\phi = 0.2$ and p_b varied from 0.02 to 0.2 with step 0.02.

particular, do not tend to 0 or 1), we get

$$s_1 \to +0, \quad s_1^{\nu} = s_A = O(1), \quad \nu \to +0, \tag{20}$$

and finally $I_{\text{eff}} < 0$ from Eq. (19), which essentially prohibits positive II in the spiking-bursting model for large systems.

The mentioned properties of I_{eff} dependence upon parameters can also be deduced from purely qualitative considerations in the sense of the reasoning in the end of Sec. II. Absence of time-correlated bursting ($p_b = 0$ or $\phi = 0$), with only spiking present (which is time-uncorrelated), implies absence of predictability and thus zero II. Absence of spontaneous spiking [$s_1 = 0$ in Eq. (18)] implies complete synchronization (in terms of the binary process), and consequently highest overall predictability (mutual information), but negative II. The presence of spontaneous activity decreases the predictability of the system as a whole, as well as that of any subsystem. According to Eq. (10), favorable for positive I_{eff} (and thus for positive II) is the case when the predictability of subsystems is hindered more than that of the whole system. Hence the increasing dependence upon s_1: since in a system with independent spiking we have $s_1 = s_A s_B < \min\{s_A, s_B\}$, spontaneous activity has indeed a greater impact upon predictability for subsystems than for the whole system, thus leading to an increasing dependence of I_{eff} upon s_1. This may eventually turn I_{eff} positive for all bipartitions, which implies positive II.

To apply our analytical results to the networks under study, we fitted the parameters of the spiking-bursting process under the assumption of independent spiking Eq. (17) to the empirical probabilities from each simulation time series. The calculated values of s_1, p_b, ϕ in case of the all-to-all neuronal network are plotted in Fig. 7 versus g_{astro} and λ (results for the random network not shown due to an inferior adequacy of the spiking-bursting analytical model in this case, see below). As expected, spontaneous activity (here measured by s_1) increases with the rate of Poissonian stimulation λ [Fig. 7(a), right panel], and time-correlated component becomes more pronounced (which is quantified by a saturated increase in p_b and ϕ) with the increase of astrocytic impact g_{astro} [Fig. 7(b), left panel].

In Figs. 2 and 4 we plot the (semi-analytical) result of Eqs. (15) and (16) with the estimates from Fig. 7 substituted for s_1, p_b, ϕ, and with bipartition AB set to the actual minimum information bipartition found in the simulation. For the all-to-all network [Figs. 2(b) and 4(b)] this result is in good agreement with the direct calculation of I_{xy} and II [failing only in the region $\lambda < 20$, see Fig. 4(b)], unlike in case of random network [Figs. 2(a) and 4(a)], where the spiking-bursting model significantly underestimates both I_{xy} and II, in particular, giving negative values of II where they are actually positive.

IV. DISCUSSION

We have demonstrated the generation of positive II in neuroastrocytic ensembles as a result of interplay between

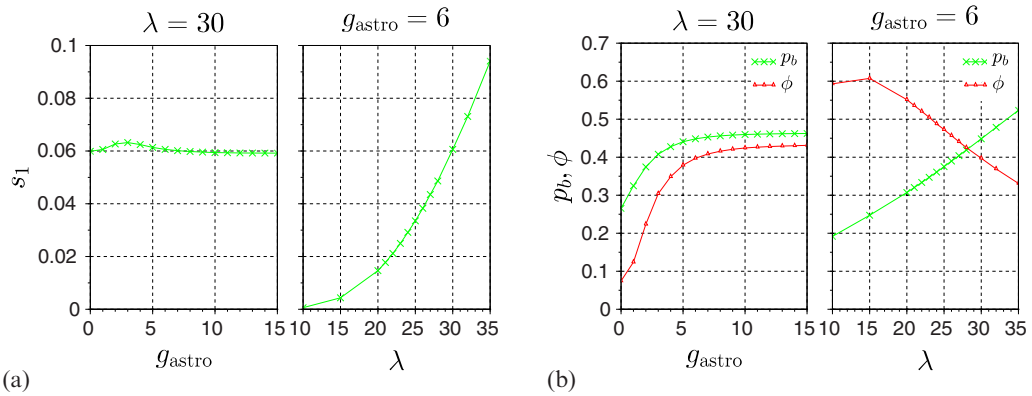

FIG. 7. Parameters of the spiking-bursting model s_1, p_b, ϕ fitted to simulated time series in case of the all-to-all neuronal network.

spontaneous (time-uncorrelated) spiking activity and astrocyte-induced coordinated dynamics of neurons. The analytic result for spiking-bursting stochastic model qualitatively and quantitatively reproduces the behavior of II in the all-to-all network with all excitatory neurons [Fig. 2(b)]. In particular, the existence of analytically predicted threshold in spontaneous activity for positive II is observed.

Moreover, the spiking-bursting process introduced in this paper may be viewed as a simplistic but generic mechanism of generating positive II in arbitrary ensembles. Complete analytic characterization of this mechanism is provided. In particular, it is shown that time correlated system-wide bursting and time uncorrelated spiking are both necessary ingredients for this mechanism to produce positive II. Due to the simple and formal construction of the process, thus obtained positive II must have no connection to consciousness in the underlying system, which may be seen as a counterexample to the original intent of II. That said, it was also shown that II of the spiking-bursting process is expected to turn negative when system size is increased. Aside from consciousness considerations, it means at least that positive II in a large system requires a less trivial type of spatiotemporal patterns than one provided by the spiking-bursting model.

The increasing dependence of II upon neuroastrocytic interaction g_{astro} and upon the intensity of spiking activity determined by λ in a range of parameters is also observed in a more realistic random network model containing both excitatory and inhibitory synapses [Fig. 2(a)], for which our analysis is not directly applicable though. Remarkably, the decrease of λ in the random network, in contrast to the all-to-all network, does not lead to negative II. In this sense the less trivial dynamics of the random network appears to be even more favorable for positive II than the spiking-bursting dynamics of the all-to-all network. This may be attributed to more complex astrocyte-induced space-time behavior, as compared to coordinated bursting alone, although we have not established specific connections of II with properties of activation patterns in the random network. Nonetheless, based on this observation we also speculate that the limitation on network size which was predicted above for spiking-bursting dynamics may be lifted, thus allowing astrocyte-induced positive II in large neuroastrocytic networks. This is in line with the hypothesis that the presence of astrocytic regulation of neurotransmission may be crucial in producing complex collective dynamics in brain. Note, however, that our conclusions are based upon (and, thus, limited to) the assumption of the positive impact of astrocytic calcium upon synaptic interactions Eq. (8), which is not universal, but was found to hold in certain areas of brain and for specific gliotransmitters [15].

The extension of our study to large systems is currently constrained by computational complexity of direct calculation of II which grows exponentially with system size. Methods of entropy estimation by insufficient data [10,41] may prove useful in this challenge, but will require specific validation for this task.

ACKNOWLEDGMENT

This work was supported by the Russian Science Foundation Grant No. 16-12-00077.

[1] G. Tononi, BMC Neurosci. **5**, 42 (2004).
[2] G. Tononi, Biol. Bull. **215**, 216 (2008).
[3] A. B. Barrett and A. K. Seth, PLoS Comput. Biol. **7**, e1001052 (2011).
[4] G. Tononi, Archives Italiennes de Biologie **150**, 56 (2012).
[5] M. Oizumi, L. Albantakis, and G. Tononi, PLoS Comput. Biol. **10**, e1003588 (2014).
[6] M. Tegmark, PLoS Comput. Biol. **12**, e1005123 (2016).
[7] A. G. Casali, O. Gosseries, M. Rosanova, M. Boly, S. Sarasso, K. R. Casali, S. Casarotto, M.-A. Bruno, S. Laureys, G. Tononi, and M. Massimini, Sci. Trans. Med. **5**, 198ra105 (2013).
[8] A. Peressini, J. Conscious. Stud. **20**, 180 (2013).
[9] N. Tsuchiya, S. Taguchi, and H. Saigo, Neorosci. Res. **107**, 1 (2016).
[10] D. Toker and F. T. Sommer, arXiv:1708.02967 (2017).
[11] G. Tononi, M. Boly, M. Massimini, and C. Koch, Nat. Rev. Neurosci. **17**, 450 (2016).
[12] D. Engel and T. W. Malone, PLoS ONE 13, e0205335 (2018).
[13] R. Norman and A. Tamulis, J. Comput. Theor. Nanosci. **14**, 2255 (2017).
[14] G. Perea and A. Araque, Brain Res. Rev. **63**, 93 (2010).
[15] A. Araque, G. Carmignoto, P. G. Haydon, S. H. Oliet, R. Robitaille, and A. Volterra, Neuron **81**, 728 (2014).
[16] S. Nadkarni, P. Jung, and H. Levine, PLoS Comput. Biol. **4**, e1000088 (2008).
[17] K. Nakae, Y. Ikegaya, T. Ishikawa, S. Oba, H. Urakubo, M. Koyama, and S. Ishii, PLoS Comput. Biol. **10**, e1003949 (2014).
[18] S. Nadkarni and P. Jung, Phys. Rev. Lett. **91**, 268101 (2003).
[19] V. Parpura and R. Zorec, Brain Res. Rev. **63**, 83 (2010).
[20] M. D. Pitta, N. Brunel, and A. Volterra, Neuroscience **323**, 43 (2016).
[21] A. L. Hodgkin and A. F. Huxley, J. Physiol. **117**, 500 (1952).
[22] V. B. Kazantsev and S. Y. Asatryan, Phys. Rev. E **84**, 031913 (2011).
[23] V. Braitenberg and A. Schüz, *Anatomy of the Cortex* (Springer, Berlin, 1991).
[24] V. B. Kazantsev, Phys. Rev. E **79**, 010901 (2009).
[25] M. M. Halassa, T. Fellin, H. Takano, J.-H. Dong, and P. G. Haydon, J. Neurosci. **27**, 6473 (2007).
[26] M. Ferrante and G. A. Ascoli, Front. Cell. Neurosci. **9**, 439 (2015).
[27] L. P. Savtchenko and D. A. Rusakov, Philos. Trans. R. Soc. London B **369**, 20130614 (2014).
[28] M. Navarrete and A. Araque, Neuron **68**, 113 (2010).
[29] G. W. De Young and J. Keizer, Proc. Natl. Acad. Sci. USA **89**, 9895 (1992).
[30] G. Ullah, P. Jung, and A. Cornell-Bell, Cell Calcium **39**, 197 (2006).
[31] For aligning the time units of the neuronal and astrocytic parts of the model it is sufficient to re-express the numerical values of all dimensional constants using time unit of 1 ms.
[32] V. Volman, E. Ben-Jacob, and H. Levine, Neural Comput. **19**, 303 (2007).
[33] M. De Pitta, V. Volman, H. Berry, and E. Ben-Jacob, PLoS Comput. Biol. **7**, 1 (2011).

[34] S. Y. Gordleeva, S. V. Stasenko, A. V. Semyanov, A. E. Dityatev, and V. B. Kazantsev, Front. Comput. Neurosci. **6**, 92 (2012).

[35] P. Jourdain, L. H. Bergersen, K. Bhaukaurally, P. Bezzi, M. Santello, M. Domercq, C. Matute, F. Tonello, V. Gundersen, and A. Volterra, Nat. Neurosci. **10**, 331 (2007).

[36] M. Navarrete, G. Perea, D. F. de Sevilla, M. Gómez-Gonzalo, A. Núñez, E. D. Martín, and A. Araque, PLoS Biol. **10**, e1001259 (2012).

[37] G. Perea and A. Araque, Science **317**, 1083 (2007).

[38] J. Kang, L. Jiang, S. A. Goldman, and M. Nedergaard, Nat. Neurosci. **1**, 683 (1998).

[39] Q.-S. Liu, Q. Xu, G. Arcuino, J. Kang, and M. Nedergaard, Proc. Natl. Acad. Sci. USA **101**, 3172 (2004).

[40] T. Fellin, O. Pascual, S. Gobbo, T. Pozzan, P. G. Haydon, and G. Carmignoto, Neuron **43**, 729 (2004).

[41] E. W. Archer, I. M. Park, and J. W. Pillow, in *Advances in Neural Information Processing Systems* (MIT Press, Cambridge, MA, 2013), pp. 1700–1708.

[42] The abrupt drop of II at high λ is associated with a change of minimum information bipartition and currently has no analytical explanation.