Dynamic Flux Tubes Form Reservoirs of Stability in Neuronal Circuits

Michael Monteforte* and Fred Wolf†

Max Planck Institute for Dynamics and Self-Organization (MPIDS), 37077 Goettingen, Germany
Bernstein Focus Neurotechnology Goettingen, 37077 Goettingen, Germany
Bernstein Center for Computational Neuroscience Goettingen, 37077 Goettingen, Germany
Faculty of Physics, University of Goettingen, 37077 Goettingen, Germany

(Received 23 February 2011; revised manuscript received 31 May 2012; published 1 November 2012)

Neurons in cerebral cortical circuits interact by sending and receiving electrical impulses called spikes. The ongoing spiking activity of cortical circuits is fundamental to many cognitive functions including sensory processing, working memory, and decision making. London et al. [Sensitivity to Perturbations In Vivo Implies High Noise and Suggests Rate Coding in Cortex, Nature (London) 466, 123 (2010)] recently argued that even a single additional spike can cause a cascade of extra spikes that rapidly decorrelate the microstate of the network. Here, we show theoretically in a minimal model of cortical neuronal circuits that single-spike perturbations trigger only a very weak rate response. Nevertheless, single-spike perturbations are found to rapidly decorrelate the microstate of the network, although the dynamics is stable with respect to small perturbations. The coexistence of stable and unstable dynamics results from a system of exponentially separating dynamic flux tubes around stable trajectories in the network’s phase space. The radius of these flux tubes appears to decrease algebraically with neuron number $N$ and connectivity $K$, which implies that the entropy of the circuit’s repertoire of state sequences scales as $N \ln(KN)$.

DOI: 10.1103/PhysRevX.2.041007

Subject Areas: Biological Physics, Interdisciplinary Physics, Nonlinear Dynamics

I. INTRODUCTION

The dynamics of neuronal networks is fundamental to the brain’s ability to perform cognitive functions such as sensory processing, working memory, and decision making. In the cerebral cortex—the part of the mammalian brain that expanded the most during the evolution of humans [1] and that is critically and specifically involved in all of these functions [2]—dynamically generated activity patterns are, in general, irregular and spatiotemporally complex [3–8]. The prevailing theoretical explanation for this irregular activity is that its origin lies in strong fluctuations in inputs that arise from the dynamical balance of excitation and inhibition—known as the balanced state [9–11]. It is a long standing theme of dynamic brain theories that such complex activity patterns might serve as a rich encoding and processing space for neural computations [12–17]. In particular, if the precise timing of every nerve impulse is used for information encoding, the processing capacity of neuronal circuits might be amazingly rich.

At a fundamental level, brains are network dynamical systems, evolved to process information. It has long been conjectured that dynamical systems optimally operate as information processing devices near a transition between order and chaos—the so-called edge of chaos. This edge-of-chaos hypothesis was originally raised in studies of cellular automata [18,19]. The dynamics at the edge of chaos is expected to support a relatively rich set of distinct state sequences and, at the same time, preserve information about initial conditions over intermediate time scales, both of which appear as important prerequisites for complex computations. In contrast, deep in a chaotic regime, information is rapidly destroyed by the dynamical entropy production, while deep in a stable regime, state sequences are often stereotyped and impoverished. Even in cellular automata, however, the edge of chaos has been found neither necessary nor sufficient for optimal performance [20]. In the context of neuronal-circuit dynamics, the notion of computations near the edge of chaos has been emphasized in studies of so-called reservoir computing [21–24]. In such applications, time-varying “sensory” inputs are represented by the instantaneous state of the entire network that can be “read out” by subsequent processing stages. Many such applications use models of spiking neural networks [22,25–30]. The basic state-space features and information capacity of such networks, however, have been analyzed only for firing-rate models [31,32]. For spiking neural networks, it has not been possible, so far, to obtain the most elementary characteristics such as the repertoire of state sequences or the way in which the size of this repertoire depends on fundamental variables, such as the neuron number or synaptic connectivity.
With regard to biological neural networks, it is not obvious that networks of spiking neurons can operate in an edge-of-chaos regime. Recent studies, both experimental [33] and theoretical [34], suggest that chaos in typical irregular firing states can be very intense. In particular, London and co-workers have argued that, from the responses to artificially induced single spikes in the rodent barrel cortex, even single additional spikes in a cortical firing sequence can induce substantial decorrelation of the network’s microstate [33]. Studies of the state-space structure and the dynamics of spiking neural networks have uncovered a plethora of exotic behaviors that raise doubts as to whether notions from the theory of smooth dynamical systems, such as the edge of chaos, are generally appropriate to describe such behavior. For instance, it has been found that networks of spiking neurons can robustly exhibit unstable limit-cycle attractors [35], which are an oxymoron in classical dynamical-systems theory. Other exotic behaviors that arise from the basic structural features of spiking neuron networks are time-varying phase-space dimensionality [36], topological speed limits [37], or topologically induced chaotic transients [38]. These and other recent observations indicate that understanding information processing by neuronal-circuit dynamics requires a specific characterization of the collective dynamics of large neural networks.

Here, we analyze the state-space structure of spiking neural networks, namely, sparse random networks of inhibitory leaky integrate-and-fire (LIF) neurons [39,40] in the balanced state [11]. Although capable of generating complex irregular spike sequences, we show that these networks actually exhibit negative-definite Lyapunov spectra. The spectra are invariant to the network size, hence this stable dynamics is extensive and preserved in the thermodynamic limit. These results confirm and extend previous studies that described the occurrence of stable chaos [41,42] in networks of LIF neurons [43–45]. We find that various state perturbations are predicted to decay extremely fast. In particular, in the limit of high connectivity, small perturbations to the membrane potentials of the neurons decay as quickly as in isolated cells. In addition, single-spike perturbations induce only minute responses in the population firing rates that decay on a millisecond time scale. Surprisingly, however, single-spike perturbations always lead to exponential state separation, causing complete decorrelation of the networks’ microstates. This rapid decoherence is also established within only a few milliseconds. By examining the dynamics for arbitrary perturbation sizes, we explain this behavior by a picture of tangled exponentially separating flux tubes composing the networks’ phase space.

These dynamic flux tubes form reservoirs of stability around stable trajectories with unique spike sequences and determine the repertoire of state sequences. A perturbation that leads the microstate of the network from one flux tube to another induces a cascade of new spikes that replaces the original spike sequence but obeys the same statistical and dynamical properties. The flux-tube structure of the phase space demonstrates that the fading-memory property (information about differences in the microstate decay over time) and the separation property (distinguishable inputs lead to significantly different states) can coexist in the dynamics of spiking neuron networks. We numerically measure the flux-tube radius and estimate the total flux-tube length. Studying the scaling of the flux-tube radius and length with the number of neurons \( N \), synapses per neuron \( K \), and firing rate \( \bar{\nu} \) leads us to surmise that the entropy of distinct multispike sequences grows faster rather than more extensive: \( H \sim N \ln(\sqrt{KN\bar{\nu}\tau_m}) \).

Intriguingly, the flux-tube structure of the phase space calls into question the applicability of notions of classical dynamical systems even in the thermodynamic limit. The decrease of the flux-tube radius with the system size leads to an ambiguity in the characterization of the dynamics in the thermodynamic limit. It can appear either as superchaotic or as stable depending on the order in which the weak-perturbation limit and the large-system limit are taken. To optimize the phase-space structure for computational applications, network structures that increase the flux-tube radius relative to random networks are needed.

## II. MODEL AND METHOD

We studied large sparse networks of \( N \) LIF neurons arranged on directed Erdős-Rényi random graphs of mean indegree \( K \). The neurons’ membrane potentials \( V_i \in (-\infty, V_F] \) with \( i = 1, \ldots, N \) satisfy

\[
\tau_m \frac{dV_i(t)}{dt} = -V_i(t) + I_i(t)
\]

between spike events. When \( V_i \) reaches the threshold \( V_F = 1 \), neuron \( i \) emits a spike and \( V_i \) is reset to \( V_R = 0 \). The membrane time constant is denoted \( \tau_m \). The synaptic input currents are

\[
I_i(t) = \sqrt{K}I_0 - \frac{J_0}{\sqrt{K}} \sum_{j \in \text{pre}(i)} \sum_s \delta(t - t_s^{ij}) \tau_m
\]

They are composed of constant excitatory external currents \( \sqrt{K}I_0 \) and inhibitory nondelayed \( \delta \) pulses of strength \( -J_0/\sqrt{K} \), received at the spike times \( t_s^{ij} \) of the presynaptic neurons \( j \in \text{pre}(i) \). The external currents \( I_0 \) are chosen to obtain a target network-averaged firing rate \( \bar{\nu} = (1/N)\sum_i \nu_i \).

For numerical simulations, we used a phase representation of the network dynamics equivalent to the voltage representation, Eqs. (1) and (2). In the phase representation, each neuron’s state is described by a phase \( \phi_i \in (-\infty, 1] \), defined by
\[
\phi(V) = -\frac{\tau_m}{T_{\text{free}}} \ln \left( \frac{V - \sqrt{K}I_0}{V_R - \sqrt{K}I_0} \right),
\]

with
\[
T_{\text{free}} = -\tau_m \ln \left( \frac{V_T - \sqrt{K}I_0}{V_R - \sqrt{K}I_0} \right).
\]

Between successive spikes in the network \((t_{s-1}, t_s)\), all phases evolve with the constant phase velocity \(1/T_{\text{free}}\). When receiving a spike at \(t_s\), a neuron’s phase is updated with the phase-transition curve
\[
U(\phi) = -\frac{\tau_m}{T_{\text{free}}} \ln \left[ \exp \left( -\phi \frac{T_{\text{free}}}{\tau_m} \right) + \frac{J_0}{K I_0} \right].
\]

The phase-transition curve is related to the phase-response curve by \(Z(\phi) = U(\phi) - \phi\). The resulting map of all neurons’ phases reads
\[
\phi_i(t_s) = \begin{cases} 
\phi_i(t_{s-1}) + (t_s - t_{s-1})/T_{\text{free}} & \text{for } i \neq i^* \\
U[\phi_i(t_{s-1}) + (t_s - t_{s-1})/T_{\text{free}}] & \text{for } i = i^*,
\end{cases}
\]

where \(i^* \in \text{post}(j^*)\) denote the postsynaptic neurons of the spiking neuron \(j^*\) at spike time \(t_s\).

The exact map (3) was used for event-based simulations [46–48] and to analytically calculate the single-spike Jacobians describing the evolution of infinitesimal phase perturbations. For the derivation of the single-spike Jacobian \(D(t_s) = \frac{d\phi(t_s)}{d\phi(t_{s-1})}\), we refer to the Supplemental Material [49] (see, also, Refs. [45–48,50–53] therein):
\[
D_{ij}(t_s) = \begin{cases} 
U'[\phi_j(t_s)] & \text{for } i = j = i^* \\
1 - U'[\phi_i(t_s)] & \text{for } i = i^*, j \neq j^* \\
\delta_{ij} & \text{otherwise}
\end{cases}
\]

The single-spike Jacobian is determined by the derivative of the phase-transition (phase-response) curve \(U'[\phi_j(t_s)] = 1 + Z'[\phi_j(t_s)]\), evaluated at times \(t_s\) just before the spike reception. They were used for numerically exact calculations of all \(N\) Lyapunov exponents as described in [51]. Positive Lyapunov exponents represent the rate of exponential growth of perturbations, characteristic of a chaotic dynamics. Negative Lyapunov exponents represent the rate at which perturbations decay exponentially. A stable dynamics is characterized by all Lyapunov exponents being negative definite. The Lyapunov exponents \(\lambda_1 \geq \cdots \geq \lambda_N\) thus comprehensively characterize the network dynamics with respect to infinitesimal perturbations [52].

In the asynchronous, irregular balanced state, neurons are driven by strong input fluctuations that result from a dynamical balance of excitatory and inhibitory inputs. The existence of a balanced-state fixed point in exclusively inhibitory coupled networks, Eqs. (1) and (2), is implied by the equation for the average input current [11]. The neurons receive exclusively inhibitory recurrent inputs \(-J_0/\sqrt{K}\) from \(K\) presynaptic neurons that fire on average with firing rate \(\bar{v}\). These inputs counteract the constant excitatory current \(\sqrt{K}I_0\). Thus, the average input current reads \(\bar{I} = \sqrt{K}(I_0 - J_0\bar{v}\tau_m)\). It is easy to see that \((I_0 - J_0\bar{v}\tau_m)\) cannot exhibit a nonzero large \(K\) limit. On the one hand, if \(\lim_{K \to \infty}(I_0 - J_0\bar{v}\tau_m) > 0\) was true, then \(\bar{I}\) would diverge to \(-\infty\) and the neurons would fire at their maximal rate. The resulting strong inhibition would break the inequality, leading to a contradiction. If, on the other hand, \(\lim_{K \to \infty}(I_0 - J_0\bar{v}\tau_m) < 0\) was true, then \(\bar{I}\) would diverge to \(+\infty\) and the neurons would be completely silenced. The resulting lack of inhibition again breaks the inequality. Thus, neither of these two possibilities for a nonzero limit of \((I_0 - J_0\bar{v}\tau_m)\) is self-consistent. Instead, self-consistency requires the balance of excitation and inhibition: \(\lim_{K \to \infty}(I_0 - J_0\bar{v}\tau_m) = 0\). From this, the network-averaged firing rate in the balanced state can be predicted as
\[
\bar{v}_{\text{bal}} = \frac{I_0}{J_0\tau_m}.
\]

While the average input current vanishes in the balanced state, the input-current fluctuations remain strong in the large-\(K\) limit [11]. This can be calculated from the input-current autocorrelation \(C(\tau) = \int \delta I(t)\delta I(t + \tau)dt\). Assuming that inputs from different presynaptic neurons are only weakly correlated, a received compound spike train can be modeled by a Poisson process with average rate \(K\bar{v}\), yielding
\[
C(\tau) = K\bar{v} \int \left( \frac{J_0}{\sqrt{K}} \right)^2 \delta(t - t')\delta(t + \tau - t')\tau_m^2 dt
= J_0^2 \bar{v}\tau_m^2 \delta(\tau).
\]

The input current thus resembles white noise of magnitude \(\sigma = J_0\sqrt{\bar{v}\tau_m}\).

III. RESULTS

As expected from the construction of the networks, Eqs. (1) and (2), the dynamics converged to a balanced state. Figures 1(a) and 1(b) show a representative spike pattern and voltage trace illustrating irregular and asynchronous firing and strong membrane-potential fluctuations. A second characteristic feature of balanced networks is a substantial heterogeneity in the spike statistics across neurons, which is indicated by broad distributions of firing rates (\(\nu\)) and coefficients of variation (\(\text{cv}\)) in Fig. 1(c). The good agreement of the predicted network-averaged firing rate in the balanced state, Eq. (5), with the numerically obtained firing rate \(\bar{v}\) confirms the dynamical balance of excitation and inhibition in the networks [Fig. 1(d)].
A. Stable chaos

Although the network state was temporally irregular, the collective dynamics of the networks was formally stable. The entire spectrum of Lyapunov exponents (disregarding the zero exponent for perturbations tangential to the trajectory) was negative [Fig. 2(a)]. This confirms the occurrence of stable chaos in LIF networks [43–45]. Furthermore, the Lyapunov spectra were invariant to the network size $N$, which demonstrates that this type of dynamics is extensive and thus presumably preserved in the thermodynamic limit.

For high connectivity, all Lyapunov exponents approached the inverse membrane time constant $\lambda_i = -1/\tau_m$. This behavior can be understood analytically from a random matrix approximation of the mean Lyapunov exponent. The mean Lyapunov exponent is the rate of phase-space volume compression and is given by

$$\lambda_{\text{mean}} = -\frac{1}{\tau_m} + O\left(\frac{1}{K^{7/2}}\right).$$

Assuming the single-spike Jacobians to be independent and identically distributed, random matrices of the form (4) with matrix elements $d(\phi) = U'Z(\phi) = 1 + Z'\phi$, sampled according to the stationary phase distribution $P(\phi)$, yields the random matrix approximation [49]:

$$\lambda_{\text{mean}} = Kp \int [\ln d(\phi)]P(\phi)d\phi.$$  

The asymptotic expansion of $d(\phi) = 1 + -1/\tau_m + O(1/K^{7/2})$ results in

$$\lambda_{\text{mean}} = \frac{1}{\tau_m} + O\left(\frac{1}{K^{7/2}}\right).$$

This and the numerical observation that the largest Lyapunov exponent approaches the mean Lyapunov exponent for large $K$ indicate that in this limit infinitesimal perturbations decay exponentially with the single-neuron membrane time constant [Fig. 2(b)]. As will become clear in the following, however, the assessment of stability to general perturbations in such networks is quite subtle.

B. Single-spike perturbations

In practice, how would one realize a small state perturbation to the dynamics of a cortical network?
Experimentally realizable and well-controlled minimal state perturbations to the dynamics of cortical networks include the addition or suppression of individual spikes. Pioneered by studies of Brecht and co-workers, experimental neuroscientists have, in fact, developed approaches to perform single-neuron, single-spike perturbations in intact cortical networks in vivo and even in awake behaving animals [54–56]. Such minimalistic neurostimulations can elicit complex animal behaviors [54–56] and can trigger measurable firing-rate responses in the local circuit [33]. We therefore examined theoretically how such single-spike perturbations affect the collective dynamics of networks in the balanced state. In the networks defined by Eqs. (1) and (2), the simplest single-spike perturbation is the suppression of one spike. Figure 3(a) illustrates the response of the network if one spike, occurring at a reference time defined as \( t = 0 \), is skipped. The missing inhibition immediately triggered additional spikes in the \( K \) postsynaptic neurons such that the network-averaged firing rate increased abruptly by \( \delta \nu = K \bar{\nu} / N \) [Fig. 3(b)]. Since the induced extra spikes inhibited neurons in the network, the overshoot in the firing rate quickly settled back to the stationary state within a time of order \( \delta t \sim 1 / (K \bar{\nu}) \). The overall number of additional spikes in the networks was therefore \( S_{\text{extra}} = N \delta \nu \delta t = 1 \) [Fig. 3(c)]. That is, the one skipped spike was immediately compensated by one extra spike.

The compensation of the one skipped spike by one additional spike in the network can be understood from a mean-field approximation of the firing-rate response. In fluctuation-driven neurons, there are generally two distinct input channels contributing to the change in the firing rate \( \delta \nu(t) \). This rate response results from changes in the mean and in the variance of the input current, respectively [50,57]. In the wake of a spike perturbation, the mean input current deviates from the stationary value \( \mu = \sqrt{K}(I_0 - J_0 \bar{\nu} \tau_m) \) by

\[
\delta \mu(t) = \sqrt{K} \left( \pm J_0 \frac{g(t) \tau_m}{N} - J_0 \delta \nu(t) \tau_m \right).
\]

The first term reflects the average change due to one skipped (extra) spike: an effectively positive (negative) current to \( K \) postsynaptic neurons. The postsynaptic current is assumed to have the form \( g(t) \), which is, in our case, a delta pulse \( \delta(t) \). The second term reflects the change in the recurrent inhibition due to the deviating average firing rate \( \delta \nu(t) \) that develops within the network. This rate response \( \delta \nu(t) \) also implies that the magnitude of the input fluctuations \( \sigma(t) = J_0 \sqrt{\nu \tau_m} + \delta \nu(t) \tau_m \) is not stationary and deviates from the stationary value \( \sigma = J_0 \sqrt{\nu \tau_m} \) for \( \delta \nu(t) \ll \bar{\nu} \) by

\[
\delta \sigma(t) = \frac{J_0}{2 \bar{\nu} \tau_m} \delta \nu(t) + O(\delta \nu(t)^2).
\]

To be self-consistent, the rate response to one skipped spike must equal the summed convolutions of \( \delta \mu(t) \) and \( \delta \sigma(t) \) with the response functions in linear response theory [50,57,58] for the mean \( \nu^\mu(t) \) and the variance \( \nu^\sigma(t) \) responses:

\[
\delta \nu(t) = \nu^\mu(t) \ast \delta \mu(t) + \nu^\sigma(t) \ast \delta \sigma(t).
\]

This equation can be solved for the rate response in the frequency domain,

\[
\delta \tilde{\nu}(\omega) = \pm \frac{\tilde{g}(\omega)}{N} \frac{1}{1 + \frac{f(\omega)}{\sqrt{K}}},
\]

with

\[
f(\omega) = \frac{1}{J_0 \tau_m} \frac{\tilde{\nu}^\mu(\omega)}{\tilde{\nu}^\mu(\omega)} - \frac{1}{2 \bar{\nu} \tau_m} \tilde{\nu}_1(\omega).
\]

As expected, the rate response, Eq. (7), vanishes in the large-\( N \) limit. In the large-\( K \) limit, the rate response can be approximated as \( \delta \tilde{\nu}(\omega) \approx \pm \frac{\tilde{g}(\omega)}{N} \), and after transforming back into the time domain, it can be approximated as \( \delta \nu(t) \approx \pm \frac{g(t)}{N} \). The rate response hence follows the form of a postsynaptic current \( g(t) \). Thus, the average number of additional (missed) spikes in the network becomes one: \( S_{\text{extra}} = \pm N \int \delta \nu(t) dt = \pm 1 \). Note that this derivation and the result are in fact independent of the specific form of the response functions \( \nu^\mu(t) \) and \( \nu^\sigma(t) \), which depend on the neuron model used. We thus generally expect that, in balanced networks with exclusively inhibitory coupling, one skipped (extra) spike induces just one additional (missed) spike and
the network returns basically immediately to a statistically stationary state.

These results show that the state of the networks, Eqs. (1) and (2), is not only stable to infinitesimal perturbations. The firing-rate dynamics, in addition, is capable of rapidly restoring the stationary firing rate in response to finite single-spike perturbations. A radically different picture emerges, however, when one considers the detailed microstate. Figure 4(a) displays the average distance $D_\phi(t) = \frac{1}{2} \sum \phi_i(t) - \phi_j(t)$ between a perturbed trajectory, in which one spike is skipped at $t = 0$, and the reference trajectory. All perturbed trajectories separate exponentially fast at a surprisingly high rate. Because this exponential separation of nearby trajectories is reminiscent of deterministic chaos, we call its separation rate the pseudo-Lyapunov-exponent $\lambda_p$. Varying the network size $N$, connectivity $K$, and network-average firing rate $\bar{\nu}$, we find that the pseudo-Lyapunov-exponent is invariant to the network size and generally positive, which is distinctly different from the classical Lyapunov exponent. With increasing connectivity, it apparently diverges linearly as $\lambda_p \sim KN$ [Fig. 4(b) and 4(c)]. The pseudo-Lyapunov-exponent is thus expected to grow to infinity in the high-connectivity limit, which is reminiscent of binary neuron networks that exhibit an infinite Lyapunov exponent in this limit [11].

### C. Flux-tube structure of phase space

In the same network, we find stable dynamics in response to infinitesimal perturbations, rapid restoration of the network’s mean firing rate, but yet a highly unstable dynamics for the microstate in response to single-spike perturbations. To characterize the transition between these completely opposite behaviors, we analyze the network states under general finite perturbations. We apply perturbations of variable size $\epsilon$ perpendicular to the unperturbed reference trajectories. Depending on the perturbation strength $\epsilon$ and direction $\delta \phi$ (with $\sum \delta \phi_i^2 = 1$), the perturbed trajectories either converge back to the reference trajectory or diverge exponentially fast [Fig. 5(a)].

The probability that a perturbation of strength $\epsilon$ induces the exponential state separation is very well described by the function $P_s(\epsilon) = 1 - \exp(-\epsilon/\epsilon_{\text{FT}})$, defining the characteristic scale $\epsilon_{\text{FT}}$ [Fig. 5(b)]. For small perturbation sizes $\epsilon \ll \epsilon_{\text{FT}}$, the perturbed trajectory almost always converges back to the reference trajectory. Increasing the perturbation size $\epsilon$ increases the probability that the perturbed trajectory starts to diverge exponentially and becomes decorrelated from the reference trajectory. For large $\epsilon \gg \epsilon_{\text{FT}}$, this decorrelation process starts almost immediately after the perturbation is applied. The characteristic scale $\epsilon_{\text{FT}}$, separating stable from unstable dynamics, scales as $\epsilon_{\text{FT}} \sim 1/(\sqrt{KN\bar{\nu} \tau_m})$ [Fig. 5(c)].

Intriguingly, the vanishing flux-tube radius for large $N$ implies that the dynamics in the thermodynamic limit ($N \to \infty$) would be unstable even to arbitrarily small perturbations ($\epsilon \to 0$). However, the analysis of the Lyapunov spectra has shown that taking the limit $\epsilon \to 0$ first and $N \to \infty$ second yields stable dynamics. Thus, the order of these limits is crucial for characterizing the dynamical nature of such balanced networks. The noncommutativity of these limits might be the origin of the contradicting previous predictions of stable [43–45] or suprachaoic dynamics [11].

While single-spike failures always induce exponential state separation, the probability for this to happen after single-synaptic failures is relatively low. As expected from the size of such perturbations $\epsilon_{\text{syn}} \sim 1/(\sqrt{KN\bar{\nu} \tau_m})$ compared to the flux-tube radius $\epsilon_{\text{FT}} \sim 1/(\sqrt{KN\bar{\nu} \tau_m})$, the probability of exponential state separation after synaptic failures is independent of $K$ and $N$ and increases linearly with $\bar{\nu}$ [49].

---

**FIG. 4.** High sensitivity to single-spike failures: (a) Distance $D_\phi(t) = \frac{1}{2} \sum |\phi_i(t) - \phi_j(t)|$ between trajectories after skipped spike and reference to log-log plots for different connectivities $K$ and average firing rates $\bar{\nu}$. The small bars indicate the distances $D_\phi^m = \int |\phi - \varphi| P(\phi) d\phi$ of uncorrelated states obtained from self-consistent mean-field solutions $P(\phi)$ [49,50]. (b) Pseudo-Lyapunov-exponent $\lambda_p$ from exponential fits $D_\phi \sim \exp(\lambda_p t)$ before reaching saturation. (c) Collapse to characteristic exponential state separation with a rate of $\lambda_p \sim K\bar{\nu}$. (Distance has been rescaled with approximate perturbation strength $K \frac{\bar{\nu}}{\epsilon}$; the time has been rescaled with input rate $K\bar{\nu}$.) Inset: Different network sizes $N$ for $K = 100$. The parameters are as in Fig. 1 with $N = 100,000$. These are averages of ten initial conditions with 100 perturbations each.
D. Exponential decorrelation by an event cascade

The exponential decorrelation process after a perturbation is initiated and mediated by a different sequence of spikes in the perturbed trajectory compared to the reference trajectory. Taking a closer look at the individual distance measurements with finite perturbation size, one notices that every exponential separation of the trajectories appears to be initiated by a peculiar step of size $d_1$ [Fig. 6(a)]. This is even more visible in the histogram of all distance measurements. It reveals several of such steps $d_n$ between adjacent peaks in the histogram [Fig. 6(b)]. Even for different perturbation sizes $\varepsilon$, these steps are all approximately of the same size $d = 2J_0\sqrt{K}/N$ [Fig. 6(c)]. This indicates that at each step a spike in the reference trajectory is replaced by a different spike in the perturbed trajectory, because such a spike replacement results in a displacement of approximately $J_0\sqrt{K}$ in $K$ postsynaptic neurons of the failed spike in the reference trajectory and in the $K$ postsynaptic neurons of the new spike in the perturbed trajectory. Thus, the distance jumps by $d = 2J_0\sqrt{K}/N$. This expression is well fulfilled for a wide range of

![Diagram](image-url)

**FIG. 5.** Sensitivity to finite perturbations: (a) Distance $D_\varepsilon$ between perturbed and reference trajectories at spike times of the reference trajectory (projecting out possible time shifts) for perturbation strengths $\varepsilon = 10^{-5}, 10^{-3}, 10^{-1}$ in log-log plots. The gray lines indicate ten random directions perpendicular to the trajectory; the color lines indicate averages of exponentially separating or converging cases. (b) Probability $P_s(\varepsilon)$ of the exponential state separation versus perturbation strength $\varepsilon$ in the log-log plot. The dotted line indicates characteristic perturbation size $\varepsilon_{\text{FT}}$ separating stable from unstable dynamics; the shaded areas indicate single-synapse, single-spike failures. (c) Characteristic perturbation size $\varepsilon_{\text{FT}}$ versus network size $N$, connectivity $K$, and average firing rate $\bar{\nu}$ in log-log plots. The straight lines indicate fits to $\varepsilon_{\text{FT}} \sim 1/(\sqrt{NK}\bar{\nu})$. (d) Symbolic picture of dynamic flux tubes with radius $\varepsilon_{\text{FT}}$, the boxes indicate the different spike sequences assigned to different flux tubes. (Inside the flux tubes are stable dynamics, but adjacent flux tubes separate exponentially.) The parameters are as in Fig. 1 with $N = 100 000$. These are averages of ten initial conditions with 100 calculations and 100 random directions each.

**FIG. 6.** Rapid decorrelation through a cascade of new spikes replacing those in the original spike sequence. (a) Five examples of typical distance measurements $D_\varepsilon(t)$ for finite perturbations $\varepsilon = 0.01$. (b) Histogram of distances of all measurements. (c) Step sizes $d_n$ obtained from histograms as in panel (b) for different $\varepsilon$. The dashed line indicates distance due to one replaced spike $d = 2J_0\sqrt{K}/N$. (d) Step sizes $d_n$ versus connectivity $K$ for finite perturbations (solid lines) and single-spike perturbations (dotted lines). (e) Histograms of the individual rate activity $K$ versus connectivity $K$ for finite perturbations (solid lines) and single-spike perturbations (dotted lines). The $\lambda_\varepsilon$ (dashed lines) from Fig. 4 is plotted for comparison. The parameters are as in Fig. 1 with $N = 100 000$. These are averages of ten initial conditions with 100 calculations and 1000 random directions each.
connectivities $K$ in measurements with finite perturbation sizes and after single-spike failures [Fig. 6(d)]. The deviations for large $K$ and $n$ can be explained by an increased overlap of the sets of postsynaptic neurons; the steps $d_n$ are therefore smaller than $d$. However, these results strongly indicate that the decorrelation process is initiated and mediated by steps of constant height due to the emergence of a new spike sequence in the perturbed trajectory compared to the reference trajectory.

The fact that the exponential decorrelation process is mediated by discrete steps of constant height implies that the rate of these steps must increase exponentially. The single spike that initiates the rapid decorrelation thus triggers a cascade of new spikes replacing the reference spike sequence. This can be seen by assuming $D_{\phi}(t) \propto \exp(\lambda t)$ for the exponential state separation. From the discrete steps, we also know that $D_{\phi}(t_n) = nd$, where $t_n$ is the time of the $n$th step relative to the initializing one, and we obtain $\lambda = \ln(n)/t_n$ for each measurement. The histogram for all measurements is shown in Fig. 6(e). The observed convergence with increasing $n$ confirms the occurrence of an exponentially increasing rate of steps. The mean value of the fitted log-normal distributions (solid lines) provides an independent measure of the average rate of exponential separation $\lambda$ over all measurements [Fig. 6(f)]. It agrees very well with the pseudo-Lyapunov-exponent $\lambda_p$ (dashed line) previously obtained from exponential fits to the averaged distance after the single-spike failures depicted in Fig. 4. These results show that a perturbation leading to the change of one spike initiates a geometrically growing cascade of new spikes leading to a completely new spike sequence. Note that this high sensitivity was not visible in the Lyapunov spectra (Fig. 2) nor in the rate response after single-spike failures (Fig. 3). The perturbed trajectory belongs to a different dynamic flux tube with a different spike sequence but exhibits the same statistical properties and local stability.

To visualize the exotic phase-space structure generated by the dynamic flux tubes, we scanned a larger piece of phase-space volume. The cross sections of the phase space resemble stained glass [Fig. 7(a)]. For the axes, we chose two random $N$-dimensional vectors orthonormal to the direction of the trajectory $(1, 1, \ldots)$, spanning a two-dimensional cross section in the $N$-dimensional phase space. Each pixel in the picture represents a new initial condition. If trajectories of adjacent initial conditions converged to each other, the pixels were assigned identical colors; otherwise they were colored differently. Areas of one color can thus be thought of as the cross sections of the basins of attraction of unique stable trajectories, corresponding to individual cells in the stained-glass architecture. When a trajectory hits a wall (threshold boundary) of the phase space, a spike is emitted and the trajectory continues at a different point in the phase space in a different cell. The black lines visualize the boundaries of adjacent cells and are expected to represent trajectories that lead, at some point in time, to synchronous spikes in pairs of neurons; they hit the edges of the phase space. A symbolic concatenation of the stained-glass cells that are visited by one trajectory yield the flux-tube picture we have developed above. If a perturbation leads to a different flux tube, then the new trajectory diverges exponentially fast from the original one as described above.

The stained-glass cells composing the dynamic flux tubes differ in shape and size. If they, in fact, represent the flux tubes postulated before, their sizes should match the $e_{FT}$ derived from random perturbations. We thus, repeatedly, calculated different phase-space cross sections as presented in Fig. 7(a) for different parameter sets and determined the areas $A$ of all cross sections that were completely resolved. To obtain a characteristic scale for the radius, we used $r = \sqrt{A/\pi}$, which follows an exponential distribution [Fig. 7(b)]. Its average $r_{FT}$ confirms the previously obtained scaling relation of the typical flux-tube radius $\epsilon_{FT} \sim 1/(\sqrt{KNp})$ rather well [inset in Fig. 7(b)]. For this comparison, we used $\epsilon_{FT}(N = 100000, K = 1000, \dot{p} = 10 \text{ Hz}) = 8 \times 10^{-4}$. There is

![Image](332x332 to 544x442)

**FIG. 7.** Visualization of the phase-space structure. (a) Phase-space cross sections for $N = 200$ and $N = 2000$, spanned by two random $N$-dimensional vectors perpendicular to the trajectory. Flux-tube sections are drawn in one color and separated by solid lines. (b) Exponential distribution of radii of flux-tube cross sections for different parameter sets. The inset shows a comparison of average radius $r_{FT}$ with $\epsilon_{FT} \sim 1/(\sqrt{KNp})$ obtained in Fig. 5(c). Other parameters used are $K = 100$, $\dot{p} = 10 \text{ Hz}$, $J_0 = 1$, $\tau_m = 10 \text{ ms}$. 041007-8
some mismatch for low $K$ (open symbols). But this mismatch is actually also visible in Fig. 5(c) for low $K$, and using the measured value of $e_{FT}$ for low $K$ leads to a good agreement of $r_{FT}$ and $e_{FT}$ for all cases $N, K, \bar{p}$ (full symbols). We can thus conclude that the typical flux-tube radius scales essentially as $1/(\sqrt{KN\bar{p}})$.

This exotic phase-space structure embodies a good partition through which to estimate the entropy of spike sequences in the networks. Our results suggest that the phase space is composed of dynamic flux tubes. Trajectories within these flux tubes converge to a unique stable trajectory while trajectories of adjacent flux tubes separate exponentially fast. In the $N$-dimensional phase space, the total flux-tube length is $l_{FT} \approx 1^N/e_{FT}^{N-1} = (\sqrt{KN\bar{p}\tau_m})^{N-1}$. This in turn suggests that the entropy of distinct network-state sequences is $H = \ln(l_{FT}) \sim N\ln(\sqrt{KN\bar{p}\tau_m})$, growing faster than in the extensive case and increasing with a larger connectivity $K$ and firing rate $\bar{p}$. Expressed in terms the time needed by the network to display all possible sequences, this time would scale as $T = l_{FT}/(N\bar{p}) \sim (\sqrt{KN\bar{p}\tau_m})^{N-1}/(N\bar{p})$. Already for $N = 200$ neurons [Fig. 7(a)], it would take the network an astronomically large time, more than $10^{200}$ years, to display all distinct network-state sequences.

## IV. CONCLUSION

Our analysis reveals the coexistence of (i) dynamical stability to infinitesimal as well as small finite state perturbations, and (ii) a high sensitivity to single-spike and even single-synapse perturbations in networks of inhibitory LIF neurons in the balanced state. The networks exhibit negative-definite and extensive Lyapunov spectra that, at first sight, suggest a well-defined thermodynamic limit of the network dynamics, characterized by stable chaos, which has previously been proposed in Refs. [43–45]. In confirming the picture developed by London et al. [33], we found that single-spike perturbations induce a cascade of new spikes. Complementary to their proposition that this leads to an excess of extra spikes, we argue that in inhibitory networks, the new spike sequence replaces the original spike sequence without a considerable trace in the network statistics, e.g., the firing-rate response. The microstate of the network nevertheless follows a completely different dynamical path that diverges exponentially from the unperturbed one. We call the rate of exponential state separation in response to single-spike perturbations the pseudo-Lyapunov-exponent. It scales as $\lambda_p \sim KN\bar{p}$ and implies practically instantaneous decorrelation of network microstates.

The seemingly paradoxical coexistence of local stability and exponential state separation is reconciled by a unifying picture of dynamic flux tubes in the network’s phase space. States within a flux tube are attracted to a stable trajectory with a unique spike sequence. Different flux tubes, however, separate exponentially fast at a rate of $\lambda_p$. The radius of the dynamic flux tubes scales as $e_{FT} \sim 1/(\sqrt{KN\bar{p}\tau_m})$. This scaling can be used to estimate the entropy of the network’s repertoire of distinct state sequences $H \sim N\ln(\sqrt{KN\bar{p}\tau_m})$. Thus, the computation of the flux-tube radius presents a novel numerical approach to estimating the entropy of a neural network. With this approach, the entropy can be calculated with less effort than, e.g., with an approach that identifies all multiple-spike sequences numerically.

The coexistence of local stability and exponential state separation leads to an ambiguous nature of the network dynamics in the thermodynamic limit and shows that the notion of an edge of chaos is not applicable to the networks considered here. On the one hand, taking the large-system limit first, one obtains vanishingly thin flux tubes. Thus, the dynamics becomes unstable even for arbitrarily small perturbations. The resulting sensitivity to initial conditions is described by the positive pseudo-Lyapunov-exponent $\lambda_p \sim K\bar{p}$, that showed no sign of saturation. On the other hand, taking the small-perturbation limit first, one obtains a stable dynamics even when then taking the large-system limit. This dynamics is described by the negative Lyapunov exponent $\lambda = -1/\tau_m$. These findings suggest that the previously reported infinite Lyapunov exponent on the one hand [11] and the negative Lyapunov exponent on the other hand [43–45] resulted from the order in which the weak-perturbation limit and the thermodynamic limit were taken. Since there is no phase transition between order and chaos in this case, an edge of chaos does not exist in these networks.

We expect this form of state decorrelation to occur rather generally in similar classes of neural networks as a result of the noncommutativity of the spike order [59,60]. Trajectories in which at least two neurons would fire synchronous spikes at some time instant form the boundaries of the basins of attraction around stable trajectories with unique spike sequences. Infinitesimal perturbations to trajectories with such synchronous spikes lead to trajectories with different spike orders and the corresponding microstates diverge exponentially. Such a strong sensitivity, despite local stability due to the noncommutativity of $\delta$-like events, might be a general feature of dynamical systems displaying stable chaos [42]. There is evidence that stable chaos is not restricted to the idealized relatively simple networks studied here. In particular, it was shown that the local stability can be preserved when excitatory neurons are included [43], and also when the $\delta$-pulse coupling is replaced with temporally decaying postsynaptic currents [45]. We furthermore found that stable chaos can be preserved when the instantaneous action-potential generation of LIF neurons is replaced by a dynamic, yet extremely rapid, action-potential generation [61].

An important topic for future studies is the question of how the flux-tube structure of phase space is influenced by,
e.g., the network topology or synaptic plasticity. That the structure of network phase space is expected to strongly impact on the performance of a reservoir-computing architecture was already pointed out in studies of stable heteroclinic channels [62]. The novel phase-space structure of dynamic flux tubes, revealed here, may provide a basis for efficient stimulus categorization and segregation. If a perturbation leads to a switch of flux tubes, this can easily be distinguished by the completely different resulting spike sequence. In the context of reservoir computing, networks exhibiting flux tubes combine the fading memory and the separation property without the existence of an edge of chaos. The flux-tube radius can be thought of as the statistical border between the fading memory and the separation property. Applications of LIF neuron networks in reservoir computing may thus strongly benefit if the flux-tube structure of the network’s phase space is taken into account.

**ACKNOWLEDGMENTS**

We thank D. Battaglia, M. Bethge, E. Bodenschatz, T. Geisel, S. Jahnke, J. Jost, P. E. Latham, R. M. Memmesheimer, C. Kirst, A. Politi, H. Sompolinsky, M. Timme, and C. van Vreeswijk for fruitful discussions. This work was supported by BMBF (01GQ07113, 01GQ0811, 01GQ0922, 01GQ1005B), GIF (906-17.1/2006), DFG (SFB 889), VW-Stiftung (ZN2632), and the Max Planck Society.

[27] Prashant Joshi, From Memory-Based Decisions to Decision-Based Movements: A Model of Interval Discrimination Followed by Action Selection, Neural Netw. 20, 298 (2007).
[49] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevX.2.041007 for more details on the analytical derivations of the phase map and the single spike Jacobians, the random matrix approximation of the mean Lyapunov exponent, the numerical convergence and the extensivity of the Lyapunov spectra, and additional distance calculations.


