frontiers in
COMPUTATIONAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 07 January 2014
doi: 10.3389/fncom.2013.00187

How pattern formation in ring networks of excitatory and inhibitory spiking neurons depends on the input current regime

Birgit Kriener[1]*†, Moritz Helias[2†], Stefan Rotter[3,4], Markus Diesmann[2,5] and Gaute T. Einevoll[1]

[1] *Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, Ås, Norway*
[2] *Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6), Jülich Research Centre and JARA, Jülich, Germany*
[3] *Faculty of Biology, University of Freiburg, Freiburg, Germany*
[4] *Bernstein Center Freiburg, University of Freiburg, Freiburg, Germany*
[5] *Medical Faculty, RWTH Aachen University, Aachen, Germany*

Edited by:
Klaus R. Pawelzik, University Bremen, Germany

Reviewed by:
Klaus R. Pawelzik, University Bremen, Germany
Brent Doiron, University of Pittsburgh, USA

****Correspondence:***
Birgit Kriener, Computational Neuroscience Group, Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, PO Box 5003, 1432 Ås, Norway
e-mail: birgit.kriener@umb.no
† These authors have contributed equally to this work.

Pattern formation, i.e., the generation of an inhomogeneous spatial activity distribution in a dynamical system with translation invariant structure, is a well-studied phenomenon in neuronal network dynamics, specifically in neural field models. These are population models to describe the spatio-temporal dynamics of large groups of neurons in terms of macroscopic variables such as population firing rates. Though neural field models are often deduced from and equipped with biophysically meaningful properties, a direct mapping to simulations of individual spiking neuron populations is rarely considered. Neurons have a distinct identity defined by their action on their postsynaptic targets. In its simplest form they act either excitatorily or inhibitorily. When the distribution of neuron identities is assumed to be periodic, pattern formation can be observed, given the coupling strength is supracritical, i.e., larger than a critical weight. We find that this critical weight is strongly dependent on the characteristics of the neuronal input, i.e., depends on whether neurons are mean- or fluctuation driven, and different limits in linearizing the full non-linear system apply in order to assess stability. In particular, if neurons are mean-driven, the linearization has a very simple form and becomes independent of both the fixed point firing rate and the variance of the input current, while in the very strongly fluctuation-driven regime the fixed point rate, as well as the input mean and variance are important parameters in the determination of the critical weight. We demonstrate that interestingly even in "intermediate" regimes, when the system is technically fluctuation-driven, the simple linearization neglecting the variance of the input can yield the better prediction of the critical coupling strength. We moreover analyze the effects of structural randomness by rewiring individual synapses or redistributing weights, as well as coarse-graining on the formation of inhomogeneous activity patterns.

Keywords: pattern formation, spiking neurons, linear model, mean-driven, fluctuation driven, ring networks, small-world networks

1. INTRODUCTION

Understanding the dynamics of neuronal networks and its dependence on connection topology, weight distribution or input statistics is essential to understand network function. One very successful field in uncovering structure-dynamics relationships are so-called neural field models (see e.g., Beurle, 1956; Wilson and Cowan, 1972; Amari, 1977; Ermentrout and Cowan, 1979a; Ben-Yishai et al., 1995; Ermentrout, 1998; Coombes, 2005). In these models the positions of neurons in space are substituted by densities and their respective coupling is described by coupling kernels depending on pairwise spatial distance. Given certain symmetries such as translation invariance, and homogeneous input the resulting non-linear dynamics is often reduced to a low-dimensional system that can be studied analytically. Such systems can produce various spatio-temporal phenomena, such as traveling waves, activity bumps and formation of periodic patterns, which can be linked to activity in biological systems, e.g., visual hallucinations, feature selectivity, short term memory, or EEG rhythms (see e.g., Ermentrout, 1998; Coombes, 2005 and references therein).

However, when thinking of actual neuronal tissue, the symmetry requirements and continuity assumptions apply on a rather macroscopic scale, when neuron densities are high and heterogeneities are negligible. Moreover, such rate-based models cannot fully resolve the statistics of synaptic input currents, but it is well-known that neurons and neuronal populations react quite strongly to changes in the statistics of incoming currents in terms of mean, auto- and cross-covariance (see e.g., Mainen and Sejnowsky, 1996; Silberberg et al., 2004; Fourcaud-Trocmé and Brunel, 2005; De la Rocha et al., 2007; Boucsein et al., 2009;

Tchumatchenko and Wolf, 2011). Finally, it is often unclear how to quantitatively interpret the parameters of abstract field models, especially when compared to spiking network simulations.

Along these lines, Usher et al. (1995) demonstrated numerically how periodic pattern-formation takes place in a two-dimensional toroidal network of excitatory and inhibitory pulse-coupled integrate-and-fire oscillators. The bifurcation occurs when the relative surround inhibition reaches a certain threshold that is determined from the Fourier-transform of the effective coupling matrix (dispersion relation). In their study, Usher et al. (1995) moreover observed that the dynamics of the emerging pattern depends on the average external input current, with slowly diffusing hexagonally arranged activity bumps for low input current, over stationary bumps for intermediate input strength, to coherently moving bumps for strong external input.

In a later study, Bressloff and Coombes (1998, 2000) showed rigorously how periodic orbits in networks of pulse-coupled inhibitory and excitatory neurons with Mexican-hat topology indeed undergo a discrete Turing-Hopf-bifurcation, if the system is coupled strongly enough.

Roxin et al. (2005) studied a neuronal field model with particular emphasis on the role of delays. Similar to earlier studies of neuronal field models, (see e.g., Ben-Yishai et al., 1995; Compte et al., 2000; Shriki et al., 2003) neurons were distributed along a ring and coupling was distance-dependent. In particular, the coupling kernel was of the form $J(|x-y|) = J_0 + J_1\cos(x-y)$, where x and y are neuron positions in space and J_0 and J_1 are coupling strength coefficients. Depending on the choice of J_0, J_1 and the delay the system can show a plethora of activity states such as traveling waves, stationary or oscillatory bumps, and also aperiodic behavior. However, this choice of interaction between two neurons is rather qualitative. It is usually symmetric and changes sign with distance, properties that are not in line with the biology of individual neurons and synapses.

Thus, to compare the field-model results to simulations of networks of excitatory and inhibitory spiking neurons, Roxin et al. (2005) set up a model where the same number of excitatory and inhibitory neurons were uniformly distributed across two rings and were coupled by sinusoidally modulated connection probability. This set-up resulted in an effectively one-dimensional model where at each position excitation and inhibition combine to a net-coupling comparable to the field model, and spatio-temporal patterns qualitatively matched the field-model predictions. The parameters in the spiking model were chosen with regard to biophysically plausible values, and the non-linearity of the current-to-rate transduction in the field model was chosen as threshold linear in an *ad-hoc* manner, however a direct quantitative mapping between neuronal network and field model was not tried.

Here, we investigate the dynamics of identical leaky integrate-and-fire neurons that are arranged on ring and grid networks where both inhibitory and excitatory neurons are uniquely assigned positions in space rather than representing a density. No two neurons can be at the same position, and this is thus a step toward spatially embedded networks of spiking neurons instead of neuronal populations with possibly ambiguous coupling weights. Instead, individual neurons have a distinctive identity in that they act either excitatorily or inhibitorily on all their post-synaptic targets (Dale's principle) and the coupling between any two neurons is usually not symmetrical. This is important since neglecting the identity of a neuron can lead to dramatically different dynamics (Kriener et al., 2008, 2009). If the system is translation-invariant, a Turing-instability occurs for a critical synaptic coupling strength J_c, such that there is a bistability between the spatially homogeneous firing rate distribution and an inhomogeneous periodic spatial pattern, similar to what was described before (Usher et al., 1995; Bressloff and Coombes, 1998, 2000).

We find that the value of J_c depends strongly on the statistics of the neuronal input: if neurons are mean-driven, i.e., receive suprathreshold input, the instability occurs for much smaller coupling strength than when neurons are in the strongly fluctuation-driven regime, where the mean input is subthreshold and spikes are evoked by spontaneous suprathreshold fluctuations. The latter is the key mechanism in explaining asynchronous-irregular firing in balanced neuronal networks.

In assessing stability of an invariant network state to perturbations the usual *modus operandi* is to linearize around the fixed point (see e.g., Ermentrout and Cowan, 1979a,b, 1980; Usher et al., 1995; Bressloff and Coombes, 1998, 2000; Coombes, 2005), implying crucially the derivative of the neuronal input-output function. In the fluctuation-driven regime the input-output function depends on both the mean and variance of the input current and thus both respective derivatives are expected to influence the quantitative prediction of linear fixed point stability (Amit and Brunel, 1997; Tetzlaff et al., 2012; Helias et al., 2013). Indeed, in the strongly fluctuation-driven regime, such that the standard-deviation of the input current is much larger than the average distance to firing threshold, the correct critical coupling strength is derived by linearization with respect to both mean and variance around the fixed point. In the mean-driven regime, on the other hand, the linearization becomes independent on the input current variance, and the derivative with respect to the mean is constant over wide ranges of working points implying independence on the exact location of the rate fixed point. In this regime pattern formation typically occurs for considerably smaller coupling strength than in the fluctuation-driven regime. Interestingly, if the input current is in an intermediate regime with both moderate subthreshold input mean, yet considerable variance the linearization independent of the input variance can give a better estimate than the one taking it into account. We will see that usually, the true critical coupling strength will lie inbetween both estimates, with J_c generally closer to the mean-driven regime prediction.

Moreover, in order to relate our ring model to the earlier findings of pattern formation in the classical Ermentrout-Cowan networks (Ermentrout and Cowan, 1979a,b, 1980) we also introduce a coarse-grained network that maps to a two-dimensional system which is formally identical to Ermentrout-Cowan networks, but lacks some of the details of the full, effectively five-dimensional ring network considered here.

Finally, we study the role of robustness of pattern formation to randomness: first, structural randomness is caused by the introduction of short-cuts which move the ring topology into the

so-called small-world regime (Watts and Strogatz, 1998); secondly, we also discuss the effect of randomness in the weight distribution introduced by lifting the constraint that neurons are either fully excitatory in their action or fully inhibitory (Dale's principle).

2. MATERIALS AND METHODS

2.1. NEURON AND NETWORK MODEL

We study ring networks of N leaky integrate-and-fire (LIF) neurons with current-based synapses. $N_\mathrm{E} = \beta N$, $\beta \in [0, 1]$, of all neurons are excitatory, the residual $N_\mathrm{I} = N - N_\mathrm{E}$ neurons are inhibitory. The membrane potential $V_i(t)$ of neuron i is governed by the differential equation

$$\tau_\mathrm{m} \frac{dV_i(t)}{dt} = -V_i(t) + RI_{\mathrm{s,i}}(t-d) + RI_{\mathrm{x,i}}(t)\,, \qquad (1)$$

where the membrane potential $V_i(t)$ is reset to V_res whenever it reaches the threshold potential V_thr and a spike is emitted. The neuron is then refractory for a time τ_ref. τ_m denotes the membrane time constant, $RI_{\mathrm{s,i}}(t-d) = \tau_\mathrm{m} \sum_{j=1,\, i\neq j}^{N} W_{ij} s_j(t-d)$, where I_s is the input current received from the local network, and $s_j(t) = \sum_k \delta(t - t_{j,k})$ is the spike train produced by some neuron j. d is the transmission delay, and W is the coupling matrix, with entries that are either zero (no synapse present), $J_\mathrm{E} = J$ (excitatory synapse) or $J_\mathrm{I} = -gJ$ (inhibitory synapse). To keep spiking activity from entering a high rate state, we assume that the recurrent input I_s is net inhibitory by choosing $g > N_\mathrm{E}/N_\mathrm{I}$.

I_x denotes the external input current that is modeled as stationary Poisson noise injected via current-based synapses of strength J_x (for details, see below), and R is the membrane resistance. Such leaky integrate-and-fire neurons were extensively studied in various input (see e.g., Amit and Tsodyks, 1991; Amit and Brunel, 1997; Lindner, 2004; Fourcaud-Trocmé and Brunel, 2005; De la Rocha et al., 2007; Vilela and Lindner, 2009; Helias et al., 2010) and network settings, especially in the context of balanced random networks (Brunel, 2000; Tetzlaff et al., 2012). Here, we consider ring networks of LIF neurons, such that each fifth neuron is inhibitory and the coupling matrix W is given by (cf. **Figure 3A**)

$$W_{ij} = \begin{cases} J & \text{if } j \text{ excitatory, } |i-j| \bmod N \leq \kappa/2 \text{ and } i \neq j \\ -gJ & \text{if } j \text{ inhibitory, } |i-j| \bmod N \leq \kappa/2 \text{ and } i \neq j \\ 0 & \text{otherwise} \end{cases}. \qquad (2)$$

Thus, each neuron is connected to its κ nearest neighbors, where we assume $\kappa = 0.1N$. We note that the network layout chosen here is motivated by an actual embedding of individual neurons into some space, in contrast to earlier studies of pattern formation in spiking neuron networks where inhibitory and excitatory neurons are often distributed in equal numbers along two identical rings, and where the main difference lies in the spatial interaction ranges. These systems are usually motivated by population-density descriptions and can often be reduced to a two-dimensional model (see e.g., Ben-Yishai et al., 1995; Compte et al., 2000; Shriki et al., 2003; Roxin et al., 2005). As we will demonstrate in section 3.2, the network studied here leads to an effectively five-dimensional model. In section 3.4 we will discuss a reduction to a two-dimensional model that however already deviates noticeably in its details from the full five-dimensional network.

2.2. CONSTANT EXTERNAL INPUT IN THE MEAN-DRIVEN AND INTERMEDIATE REGIME

In the case of the mean-driven ($\mu[RI] \geq V_\mathrm{thr} - V_\mathrm{res}$) and intermediate regimes, which will be discussed later, the external current is purely excitatory, such that $RI_\mathrm{x}(t) = \tau_\mathrm{m} J_\mathrm{x} s_\mathrm{x}(t)$ with $J_\mathrm{x} > 0$ and rate $\nu_\mathrm{X} = \mathrm{E}[s_\mathrm{x}(t)]$, where E[.] denotes the expectation value. To quantify the effective strength of the external input, we express $\nu_\mathrm{X} = \eta\theta/J_\mathrm{x}\tau_\mathrm{m}$, i.e., the parameter η gives the strength of the external drive in terms of the excitatory rate necessary to reach threshold. Assuming stationary spiking activity with rate ν_o for both the excitatory and inhibitory population, i.e., $\nu_\mathrm{E} = \mathrm{E}[s_\mathrm{E}(t)] = \nu_o$ and $\nu_\mathrm{I} = \mathrm{E}[s_\mathrm{I}(t)] = \nu_o$ respectively, the mean and variance of the total input $RI = R(I_\mathrm{s} + I_\mathrm{x})$ are thus given by

$$\begin{aligned} \mu\left[R\left(I_\mathrm{s}+I_\mathrm{x}\right)\right] &= \sum_{j\in\{\mathrm{E,I,X}\}} \tau_\mathrm{m}\nu_j J_j = \kappa\tau_\mathrm{m}\nu_o J(\beta - g(1-\beta)) \\ &\quad + \tau_\mathrm{m}\nu_\mathrm{X} J_\mathrm{x} \qquad (3) \\ \sigma^2\left[R\left(I_\mathrm{s}+I_\mathrm{x}\right)\right] &= \sum_{j\in\{\mathrm{E,I,X}\}} \tau_\mathrm{m}\nu_j J_j^2 = \kappa\tau_\mathrm{m}\nu_o J^2(\beta + g^2(1-\beta)) \\ &\quad + \tau_\mathrm{m}\nu_\mathrm{X} J_\mathrm{x}^2. \end{aligned}$$

For stronger local coupling strength J the negative mean recurrent contribution $\mu[RI_\mathrm{s}]$ to the total input is stronger, while the variance $\sigma^2[RI_\mathrm{s}]$ increases. For the networks we consider here, for small J the network activity is thus typically driven by the mean $\mu[RI]$ of the total input, while for larger J the variance $\sigma^2[RI]$ has a notable impact. Moreover, the network firing rate ν_o depends on J.

2.3. EXTERNAL INPUT ADJUSTED TO ENSURE CONSTANT TOTAL CURRENT IN THE STRONGLY FLUCTUATION-DRIVEN REGIME

The strongly fluctuation-driven regime is characterized by subthreshold mean total input, $\mu[RI] < V_\mathrm{thr} - V_\mathrm{res}$, but high variance $\sigma^2[RI]$, such that spikes are initiated by transient input fluctuations. To guarantee that we are in the high variance regime for all J, we inject excitatory and inhibitory external currents with rates $\nu_\mathrm{Ex}, \nu_\mathrm{Ix}$, such that the total input current mean and variance, i.e., $\mu[RI]$ and $\sigma^2[RI]$, stay the same when varying the local coupling strength J. This is achieved by adapting $\mu[RI_\mathrm{x}] = \mu[RI] - \mu[RI_\mathrm{s}]$ and $\sigma^2[RI_\mathrm{x}] = \sigma^2[RI] - \sigma^2[RI_\mathrm{s}]$ as a function of $\nu_\mathrm{Ex}, \nu_\mathrm{Ix}$. The external current injected into each local neuron by current-based synapses of strengths $J_\mathrm{Ex} = J_\mathrm{x}$ and $J_\mathrm{Ix} = -gJ_\mathrm{x}$ is thus given by

$$RI_x(t) = \tau_m J_x (s_{Ex}(t) - g s_{Ix}(t)) \tag{4a}$$

$$\text{with} \quad \nu_{Ex} := \mathrm{E}[s_{Ex}(t)] = \frac{1}{\tau_m J_x (1+g)} \left(\frac{\sigma^2[RI_x]}{J_x} + g\mu[RI_x] \right) \tag{4b}$$

$$\text{and} \quad \nu_{Ix} := \mathrm{E}[s_{Ix}(t)] = \frac{1}{\tau_m J_x g (1+g)} \left(\frac{\sigma^2[RI_x]}{J_x} - \mu[RI_x] \right). \tag{4c}$$

Because the variance of the recurrent input $\sigma^2[RI_s]$ increases quickly with increasing J, the total variance $\sigma^2[RI]$– in order to keep it fixed for all J– usually needs to be chosen quite large in order to avoid non-sensical rates $\nu_{Ix} < 0$. This moreover implies that this approach of choosing the external input is not useful for the mean-driven scenario.

For all neurons we set $\theta = V_{thr} - V_{res} = 20\,\mathrm{mV}$, $\tau_m = 20\,\mathrm{ms}$, and $R = 80\,\mathrm{M\Omega}$. The excitatory synaptic coupling strengths J, J_x, the relative strength of inhibition g, as well as τ_{ref} and d will be specified individually. The model and parameters are listed in **Table A1** in Appendix A1. All simulations were carried out with NEST (Gewaltig and Diesmann, 2007).

3. RESULTS

For small coupling strength J and large mean current input $\mu[RI] = \mathrm{E}[R(I_x + I_s)]$ the spiking dynamics of the ring networks is characterized by locally clustered activity on the spatial scale of the footprint κ, cf. **Figures 1A,G** (Kriener et al., 2009). For increasing J, however, a clear spatial pattern emerges,

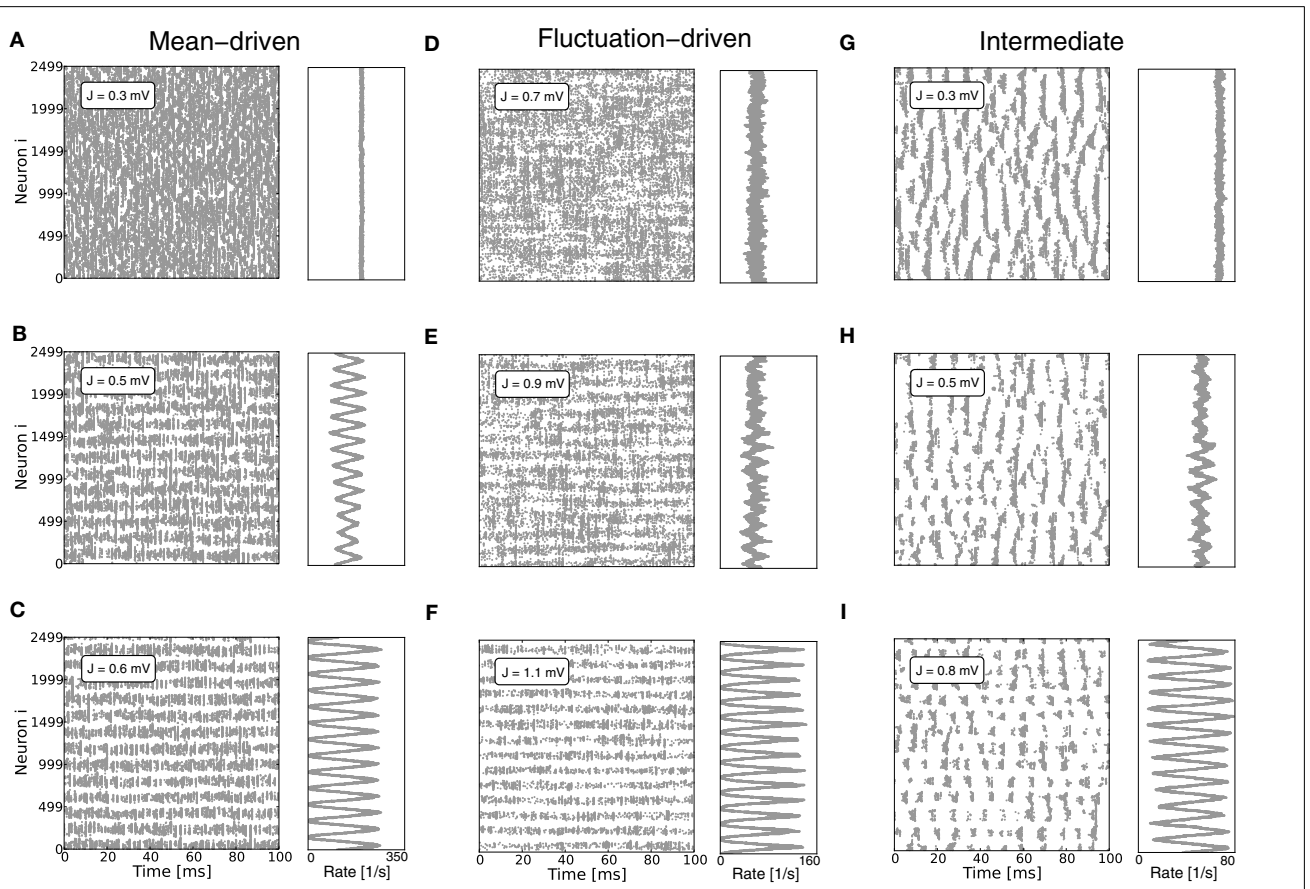


FIGURE 1 | Pattern formation in a ring network of 2500 neurons with $g = 6$ and varying absolute coupling strength parameter J. **(A–C)** Show the mean-driven case, where the external input is given by Poissonian input spikes with rate $\nu_x = 10\theta/J_x\tau_m$ and amplitude $J_x = 0.1\,\mathrm{mV}$, resulting in an excitatory input current of amplitude $I_x = 2500\,\mathrm{pA}$. The critical coupling strength J_c for the spatially homogeneous rate distribution to become bistable with a spatially modulated pattern is around 0.5 mV. **(D–F)** Show the same network in the fluctuation driven regime, where mean and variance of the external input current I_x were chosen such that mean and standard deviation of the total input current $I = I_x + I_s$ were fixed at $\mu[RI] = 5\,\mathrm{mV}$ and $\sigma[RI] = 60\,\mathrm{mV}$, cf. Equation (4). **(G–I)** Show the corresponding plots for an intermediate input regime, where the external input current was Poissonian with rate $3.5\theta/J_x\tau_m$ and amplitude $J_x = 0.1\,\mathrm{mV}$, resulting in an input current of amplitude $I_x = 875\,\mathrm{pA}$. Here, refractory time constant τ_{ref} and delay d were chosen to be 0.1 ms to minimize loss of input current and avoid pronounced delay oscillations in the mean-driven case (Helias et al., 2013). As we will demonstrate in section 3.2, eigenvalue analysis predicts an instability of an eigenmode with non-zero wavenumber $\Lambda_c > 0$ (here, $\Lambda_c = 13$) for coupling $J_c^{md} \approx 0.5\,\mathrm{mV}$ in the mean-driven case, and $J_c^{fd} \approx 0.9\,\mathrm{mV}$ in the fluctuation-driven case. We see that already for $J \lesssim J_c$ there can be deviations from the spatially homogeneous mode $\Lambda = 0$ due to the fluctuating nature of the activity and selective amplification (Dayan and Abbott, 2001), especially in the fluctuation-driven case. For larger J the pattern becomes more prominent and finally some neurons fall completely silent **(C,F,I)**.

cf. **Figures 1C,F**. Since all neurons receive by construction statistically identical input, all neurons should fire at the same rate. But there is a distinct critical coupling strength J_c beyond which the homogeneous distribution of firing rates becomes unstable and the system enters a spatially inhomogeneous state after sufficient perturbation (Bressloff and Coombes, 1998). We note that the spatio-temporal spiking activity can be quite rich in networks with translation-invariant symmetry, as discussed for example in (Ben-Yishai et al., 1995; Usher et al., 1995; Bressloff and Coombes, 1998, 2000; Shriki et al., 2003; Roxin et al., 2005; Marti and Rinzel, 2013). Here, we do not systematically analyze the temporal aspects of the activity in terms of traveling waves or oscillations, but focus on the onset of periodic pattern formation in space as a function of coupling strength.

What determines this critical coupling strength J_c where the bistability of the homogeneous spatial rate distribution and the spatially modulated distribution occurs? We find that J_c depends strongly on the input regime the neurons operate in: if neurons are driven predominantly by the mean of the input current, i.e., if $\mu[RI] \geq \theta$, we find that J_c is much smaller than if neurons are strongly fluctuation-driven, i.e., $\mu[RI] < \theta$ with large enough variance of the input current to occasionally drive the membrane potential to threshold, cf. **Figure 1**.

In the following we present two different linear rate models of the integrate-and-fire dynamics that explain and predict the respective occurrence of the pattern formation.

3.1. TWO LINEARIZATIONS FOR THE SELF-CONSISTENT RATE IN NETWORKS OF LIF-NEURONS

The analysis of the stability of spatially homogeneous activity dynamics to spatial perturbations follows a simple concept: first, the stationary, homogeneous state is determined in a mean-field approximation. In particular, in the diffusion limit, i.e., under the assumption that the coupling strength J is small compared to the distance between reset and threshold θ and that all neurons receive statistically identical input, the stationary firing rate ν_o of the neurons can be derived self-consistently as a function of mean and variance of the input current (a solution of the mean first-passage-time problem first derived by Siegert (Siegert, 1951), but also see e.g., Amit and Tsodyks, 1991; Brunel, 2000), such that

$$\nu_o^{-1} = \tau_{\mathrm{ref}} + \tau_{\mathrm{m}}\sqrt{\pi}\int_{\frac{V_{\mathrm{res}}-\mu_o}{\sigma_o}}^{\frac{V_{\mathrm{thr}}-\mu_o}{\sigma_o}} \exp\left[x^2\right]\,(1+\mathrm{erf}[x])\,dx, \quad (5)$$

where $\mu_o = \sum_{i\in\{\mathrm{E,I,X}\}} J_i \nu_{i,o} \tau_{\mathrm{m}}$ and $\sigma_o^2 = \sum_{i\in\{\mathrm{E,I,X}\}} J_i^2 \nu_{i,o} \tau_{\mathrm{m}}$ are the self-consistent input mean and variance. Because all neurons receive statistically identical input, in absence of symmetry-breaking the firing rate distribution is spatially homogeneous, i.e., $\nu_{o,i} \equiv \nu_o$ for all i. To assess simple stability of the homogeneous state to spatial perturbations, we need to perform linear perturbation analysis. The critical eigenvector of the resulting linear stability operator then determines the spatial pattern that develops.

We will show that the exact nature of the appropriate linearization depends critically on the input current regime, and derive the respective linearization in the limits of subthreshold mean input with fluctuation-driven spiking in section 3.1.1, and of dominant suprathreshold mean total input current $\mu[R(I_x + I_s)] > \theta$ in section 3.1.2. In the fluctuation-driven regime a quantitative approximation is obtained from the modulation of both mean and fluctuations of the synaptic input to the neurons, while in the mean-driven regime we recover the intuitive result that the system is governed by the noise-free input-output relation given by the f-I-curve of the individual neuron. In general, the full spiking dynamics will lie somewhere in between these two limiting cases, the case we call "intermediate."

3.1.1. Linearization of the input-output relation in the fluctuation-driven regime

We first discuss the case when the system is fluctuation-driven, i.e., the average input current is subthreshold ($\mu[RI] < \theta$) and spikes are induced by membrane-potential fluctuations.

To assess linear stability in this regime we need to derive a linear model that self-consistently describes the dynamics of the fluctuations $u_i(t) = \langle s_i(t) - \nu_o \rangle = \nu_i(t) - \nu_o$ of the activity of neuron i around its mean firing rate or working point ν_o. Treating the response of integrate-and-fire neurons by first order perturbation theory is a well-established method first introduced in Abbott and van Vreeswijk (1993) for the perfect integrator and in Amit and Brunel (1997); Brunel and Hakim (1999) for the leaky integrate-and-fire neuron.

The dynamics of $u(t)$ can then be written in the form

$$u(t) \simeq h(t) * [u(t) + x(t)]\,, \quad (6)$$

where $h(t)$ is a linear filter and $x(t)$ is a white noise imitating the spiking nature of the signal. In the Fourier-domain Equation (6) becomes the simple product $U(\omega) = H(\omega)[U(\omega) + X(\omega)]$, where the capitalizations indicate the Fourier-transforms of $h(t)$, $u(t)$ and $x(t)$. Pattern formation is typically a slow process and the transfer function $H(\omega)$ of leaky integrate-and-fire neurons is dominated by low-pass behavior[1]. In order to assess stability, we therefore estimate $H(\omega)$ in the low-frequency limit, i.e., $H(0) = \int_0^\infty h(t)\,dt$. With $\frac{d\mu_{o,i}}{ds_j} = \tau_{\mathrm{m}} W_{ij}$ and $\frac{d\sigma^2_{o,i}}{ds_j} = \tau_{\mathrm{m}} W_{ij}^2$ as well as $\frac{d\sigma^2_{o,i}}{ds_j} = 2\sigma_{o,i}\frac{d\sigma_{o,i}}{ds_j}$, we obtain (see also Amit and Brunel, 1997; Tetzlaff et al., 2012; Helias et al., 2013 for details) the effective coupling matrix $\tilde{W}$ of the system as the local derivative of the input-output-curve given by Equation (5):

$$\tilde{W}_{ij}(W_{ij}) = \int_0^\infty h_{ij}(t)\,dt = \frac{\partial \nu_{o,i}}{\partial s_j} \quad (7a)$$

$$= \frac{\partial \nu_{o,i}}{\partial \mu_{o,i}}\tau_{\mathrm{m}} W_{ij} + \frac{\partial \nu_{o,i}}{\partial \sigma_{o,i}}\frac{\tau_{\mathrm{m}}}{2\sigma_{o,i}} W_{ij}^2 \quad (7b)$$

[1]Note, that for exponential response kernels $h(t) = \tau \exp[-t/\tau]$ the response is structurally equivalent to a linear first order differential equation, though it has been demonstrated that in ring networks this response kernel can be substantially different from exponential and even contain oscillatory components (Lindner and Schimansky-Geier, 2001; Pernice et al., 2012).

$$= (\nu_o\tau_m)^2\sqrt{\pi}\frac{W_{ij}}{\sigma_{o,i}}\left(f(y_{\theta,i})\left(1+\frac{W_{ij}}{2\sigma_{o,i}}y_{\theta,i}\right)\right.$$
$$\left.-f(y_{r,i})\left(1+\frac{W_{ij}}{2\sigma_{o,i}}y_{r,i}\right)\right), \quad (7c)$$

with

$$f(y)=e^{y^2}(\mathrm{erf}[y]+1),\quad y_\theta=\frac{\theta-\mu_o}{\sigma_o},\quad \text{and}\quad y_r=\frac{V_{\mathrm{res}}-\mu_o}{\sigma_o}. \quad (8)$$

The effective coupling matrix $\tilde{W}(W_{ij})$ is thus structurally identical to W given by Equation (2), with $\tilde{W}_{ij}=\tilde{W}(J)$ if j is excitatory, and $\tilde{W}_{ij}=\tilde{W}(-gJ)$ if j is inhibitory, yielding an effective relative inhibitory strength $\tilde{g}:=\left|\tilde{W}(-gJ)/\tilde{W}(J)\right|$. The zero frequency mode $U(0)$ therefore fullfills the equation

$$U(0)=\left[\mathbb{I}-\tilde{W}\right]^{-1}X(0). \quad (9)$$

We expect linearity to break down if $\tilde{W}$ has any eigenvalue with real part larger than unity. The critical coupling strength J_c^{fd} for which this is the case can be determined implicitly from Equation (7).

3.1.2. Linearization in the low input current noise limit

Here, we discuss the linearization in the case that neurons are mean-driven, i.e., when the mean of the input current $\mu[RI]$ is suprathreshold, while the variance of the input is negligible [see also Amit and Tsodyks (1991) for analogous arguments]. In this case, the input-output-relation of the neuron becomes effectively linear in $\mu[RI]$ until it reaches saturation at $\nu_{\max}=1/\tau_{\mathrm{ref}}$ (Amit and Tsodyks, 1991; Salinas and Sejnowsky, 2000). This can be seen by considering the input-output relation of an individual neuron in response to a constant current, i.e., the f-I-curve

$$\nu(I)=\left[\tau_{\mathrm{ref}}-\tau_m\mathrm{Log}\left[1-\frac{\theta}{RI}\right]\right]^{-1}. \quad (10)$$

For $RI\gg\theta$, i.e., $\frac{\theta}{RI}\ll 1$ and negligible τ_{ref}, it holds (see Appendix A2 for details)

$$\nu(I)\doteq\frac{RI}{\tau_m\theta}-\frac{1}{2\tau_m}. \quad (11)$$

This linear-threshold-type relationship $\nu(I)=\left[\frac{RI}{\tau_m\theta}-\frac{1}{2\tau_m}\right]_+$ fits that observed in simulations very well, cf. **Figures 2A,B**.

To derive a rate equation for the full spiking network dynamics, we first define the instantaneous rate of neuron i as

$$\nu_i(t)=\mathrm{E}_{\Delta t}[s_i(t)]:=\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{-\Delta t/2}^{\Delta t/2}\mathrm{E}[s_i(t)]\,dt, \quad (12)$$

where E[.] is the average over realizations.

Substituting the spike trains by their respective firing rates, the input to neuron i takes the form $RI_i=\tau_m\sum_j W_{ij}\nu_j+RI_x$. If we only consider networks with rates $\nu_i\geq 0$ and approximate the firing rate of neuron i by the linear expression (11), we arrive at the linear equation

$$\nu_i=-\frac{1}{2\tau_m}+\frac{1}{\theta\tau_m}\left(\tau_m\sum_j W_{ij}\nu_j+RI_x\right)\quad\Leftrightarrow$$
$$\tau_m\nu=[\theta\mathbb{I}-W]^{-1}(RI_x-\theta/2). \quad (13)$$

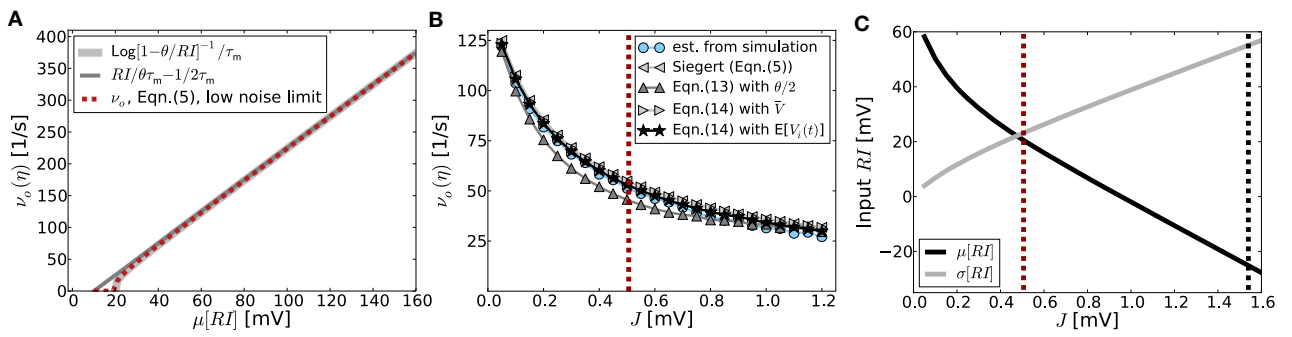


FIGURE 2 | (A) Demonstrates that the input-output-relation of the LIF neuron Equation 5 indeed gets linear in the strongly mean-driven regime. The light gray line shows the f-I-curve in the noise-less case [Equation (10)], dark gray corresponds to its linear approximation Equation (11), while the red curve is the corresponding self-consistent rate given by Equation (5) in the low-noise limit $\sigma[RI]\to 0$. **(B)** Shows the output rates as a function of network coupling strength J as they are actually obtained from network simulations averaged over 10 trials (blue dotted curve) vs. the prediction from Equation (5) (gray left-pointing triangles), as well as the linear model prediction Equation (13) with $\mathrm{E}[V(t)]=\theta/2$ (dark gray triangles). Also shown is the prediction from solving Equation (14) self-consistently with $\tilde{V}=\int_{-\infty}^{\theta}VP(V)\,dV$ [with stationary membrane potential probability density function $P(V)$ as derived in Brunel (2000); gray right-pointing triangles], as well as from using the estimated average membrane potentials $\mathrm{E}[V_i(t)]$ obtained from simulations (black asterisks). The vertical red line represents the expected critical coupling strength J_c^{md} for the mean-driven regime. The external drive is constant Poissonian noise ($\eta=3.5$), while the local network coupling strength J, and hence $\mu[RI_s]$ and $\sigma[RI_s]$, vary. In this setting the system undergoes a regime change from the mean-driven to the fluctuation-driven scenario. This is demonstrated in **(C)**, where the total mean and standard deviation of $RI=R(I_x+I_s)$ are shown as function of J in black and gray, respectively. The red dashed line corresponds to that in **(B)**, while the black dashed line indicates the critical J_c^{fd} expected from Equation (9). The corresponding spike activity is shown for three exemplary cases in **Figures 1G–I**. Other parameters in **(B,C)** are $N=2500$, $\kappa=250$, $\theta=20$ mV, $\tau_m=20$ ms, $\tau_{\mathrm{ref}}=0.1$ ms, and $g=6$.

Indeed, if τ_{ref} is negligible, Equation (1) can be rewritten as (Kriener et al., 2008)

$$\tau_{\text{m}} \frac{dV_i(t)}{dt} = -V_i(t) - \tau_{\text{m}}\theta s_i(t) + \tau_{\text{m}} \sum_{j=1,\, i\neq j}^{N} W_{ij} s_j(t-d) + RI_{\text{x,i}} \,. \tag{14}$$

Equation (13) can then be derived analogously by temporal averaging, assuming stationarity $d\,\text{E}_{\Delta t}[V(t)]/dt \equiv 0$, i.e., $\text{E}_{\Delta t}[V(t)] \equiv \text{const}$, and thus $\text{E}_{\Delta t}[\nu(t)] = \text{const} = \nu_o$. For high rates in response to large $\mu[RI]$ we expect $\text{E}[V(t)] \approx (V_{\text{thr}} - V_{\text{res}})/2 = \theta/2$, as the voltage passes through the interval between reset and threshold with approximately constant velocity, leading to

$$0 = -\theta/2 - \tau_{\text{m}}\theta\nu_i(t) + \tau_{\text{m}} \sum_j W_{ij}\nu_j + RI_{\text{x,i}}, \tag{15}$$

which is identical to (13). We note, that the self-consistent solution of Equation (15) yields the correct quantitative rates over a wide range of relative input magnitude, if instead of $\theta/2$ the actual $\text{E}[V_i(t)]$-values measured in simulations are inserted. If all neurons are identical and receive statistically identical input, this mean value can be obtained from $\forall_i\, \text{E}[V_i(t)] \equiv \bar{V} := \int_{-\infty}^{\theta} VP(V)\, dV$, where $P(V)$ denotes the stationary membrane potential probability density function as e.g., derived in Brunel (2000). All output-rate predictions are compared to the outcome from simulations in **Figure 2B**. Aside from the prediction that assumes $\text{E}[V(t)] = \theta/2$ (dark gray triangles) and underestimates the true rate, all predictions fit the simulation results (blue dots) very well. In particular, the Siegert equation (5) coincides perfectly with the linearized rate Equation (14) if we assume $\text{E}[V(t)] = \bar{V}$ (light gray triangles), while the best fit is obtained with Equation (14) and the actual measured $\text{E}[V_i(t)]$ (black asterisks).

The stability of the fixed point rate to small perturbations is determined by the largest real part of all eigenvalues of W/θ. If one eigenvalue λ_c has real part larger than unity, we expect the corresponding eigenvector ν_c to grow exponentially, only limited by the non-linearities due to rate-rectification for small rates, and the saturation because of neuronal refractoriness for high rates. This determines the critical coupling strength

$$J_c^{\text{md}} = \frac{1}{\text{Re}[\lambda_c]} \tag{16}$$

that in this limit is independent of the exact working point $\nu_o > 0$. Even though for coupling strengths $J > J_c^{\text{md}}$ the rate instability can lead to pattern formation such that individual neurons fire at quite different rates, the population rate is still captured well by the firing rate predictions assuming homogeneous rates across neurons, see **Figure 2B**.

Finally, we note that for the constant external drive scenario considered here the mean and variance of the total input current vary with varying J. In particular, for increasing J the system undergoes a cross-over from the mean-driven to the fluctuation-driven regime, see **Figure 2C**.

3.2. EIGENSYSTEM OF DALE-CONFORM TRANSLATION-INVARIANT RING NETWORKS

In this section we will analytically derive the eigenvalue spectrum of the ring networks under consideration that yields the critical eigenvalue and thus the critical coupling strength J_c with respect to the two linearizations Equation (9) and Equation (13). As described in section 2 we consider ring networks of size N with regular connectivity, in that each neuron is connected to its κ nearest neighbors, $\kappa/2$ on each side of the neuron but not to itself. Inhibitory neurons shall be distributed periodically across the network as illustrated in **Figure 3A**, where the ratio between excitation and inhibition is $4:1$ ($\beta = 0.8$). This is in line with the observed frequencies of excitatory and inhibitory cells

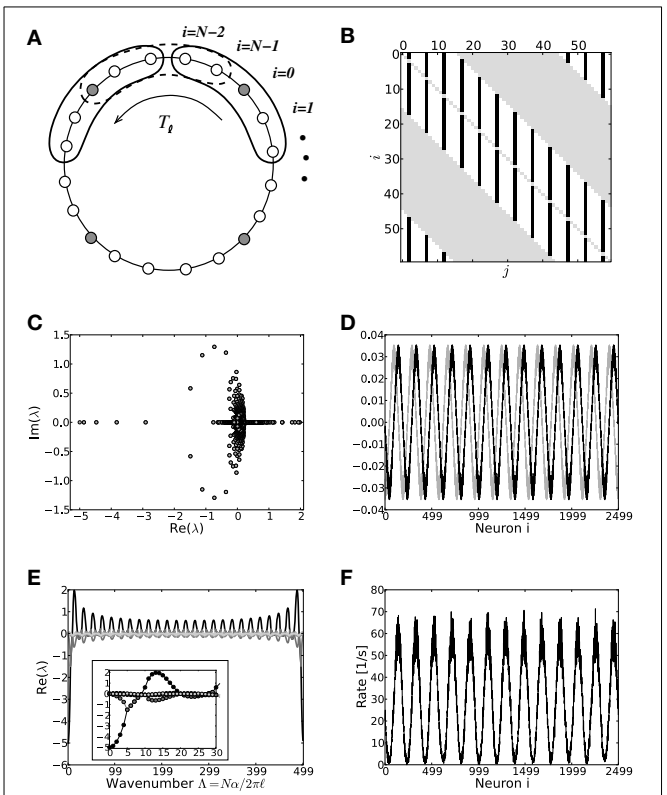


FIGURE 3 | (A) Sketch of the neuron layout on the ring topology. Each neuron is connected to its κ nearest neighbors. A shift in neuron label by five yields the same ring as before and is formally expressed by the shift operator T_ℓ with $\ell = 5$. **(B)** Shows the coupling matrix W of the ring: black squares depict inhibitory, white squares excitatory, and gray squares zero coupling from j to i ($N = 60$, $\kappa = 30$). **(C)** Eigenvalue spectrum of a ring matrix W/θ with $N = 2500$, $\kappa = 250$, $J = 1$ mV, and $g = 6$. **(D)** The two eigenvectors belonging to the twice degenerated eigenvalue from **(C)** with largest real part λ_c. **(E)** The eigenvalues form five bands and are shown as a function of the wavenumber. The black curve depicts the band containing the critical (maximal) eigenvalue λ_c which here corresponds to wavenumber $\Lambda_c = 13$, so the two crictical eigenvectors have 13 major peaks (cf. **D,F**). **(F)** The rate as a function of the neuron index in a simulation of $N = 2500$ integrate-and-fire neurons with $J = 1$ mV and $I_x = 750$ pA. Note that the spectra (shown in **C**) look slightly different in the fluctuation-driven case, because the absolute amplitude of the entries $|\tilde{W}_{ij}| \neq |W_{ij}|$, and also the balance between positive and negative entries $\tilde{g} \neq g$. The maximum of the dispersion relation and the eigenvectors are, however, the same.

in cortex (Schüz and Braitenberg, 1994). The ring can thus be divided into structurally identical elementary cells of four excitatory neurons and one inhibitory neuron. We choose κ as an integer multiple of the size ℓ of an elementary cell (here $\ell = 5$, $\kappa \bmod \ell = 0$), such that it moreover divides the total number of neurons N ($N \bmod \kappa = 0$). The resulting coupling matrix is sketched in **Figure 3B**, where black squares depict inhibitory weights, white depict excitatory weights and gray denotes the lack of a connection. In this way all neurons receive the exact same amount of inhibition and excitation, and the translation invariant mode $(1, 1, \cdots, 1)^\top$ is an eigenvector of the coupling matrix with eigenvalue $\mu = \kappa J(\beta - g(1-\beta))$. To keep network activity balanced, inhibitory synapses are assumed to be g-times stronger, with $g > 4$, cf. Equation (2). Hence, the local network is inhibition dominated.

To compute the eigensystem of the N-dimensional system we make use of its symmetry properties. The coupling matrix commutes with the unitary operator that shifts all neuron indices by multiples of $\ell = 5$. Hence, we can diagonalize both in the same eigenbasis and moreover reduce the problem to an effectively five-dimensional one as outlined in Appendix A3. **Figure 3C** shows the exemplary eigenvalue spectrum of in this case the rescaled coupling matrix W/θ of a network of size $N = 2500$ with absolute coupling strength $J = 1$ mV and relative strength of inhibition $g = 6$. Thus, the eigenvalue λ_c of W/θ with largest real part will exceed unity, if the amplitude of the absolute coupling strength J exceeds $J_c^{\mathrm{md}} = 1/\mathrm{Re}[\lambda_c]$, cf. Equation (16), here $J_c^{\mathrm{md}} \approx 0.5$.

The corresponding critical weight J_c^{fd} for the fluctuation-driven case is obtained by the identical eigenvalue decomposition of $\tilde{W}$ as defined by Equation (7). Since $\tilde{W}$ depends non-trivially on the self-consistent working point ν_o of the system, the critical coupling strength J_c^{fd} is the solution of

$$J_c^{\mathrm{fd}} := J \;\text{ such that }\; \max_i\left[\mathrm{Re}[\lambda_i(\tilde{W}(\nu_o, J))]\right] \equiv 1\,, \qquad (17)$$

where λ_i, $i \in \{1, \ldots, N\}$ are the N eigenvalues of $\tilde{W}$. If $J > J_c^{\mathrm{md/fd}}$ (depending on the regime), the spatially homogeneous rate distribution becomes unstable to perturbations in favor of the fastest growing eigenmode ν_c, which only depends on the symmetry properties of the matrix as long as g and $\tilde{g}$ are larger than four, and it is hence the same irrespective of the input regime, cf. **Figures 1C,F,I**. The spatial frequency of the emerging spatial pattern (see **Figure 3**), i.e., the number of maxima of the rate distribution along the ring, is thus given by the wavenumber Λ (see Appendix A3) that corresponds to the respective $\lambda_c := \max_i\left[\mathrm{Re}[\lambda_i]\right]$. We note, that the eigenvalue spectra in the fluctuation-driven case look slightly different from those in the mean-driven case (cf. **Figure 3C**) because of the in general different scaling $\tilde{g} \neq g$ between positive and negative entries in $\tilde{W}$. The maximum of the dispersion relation determining Λ_c and the eigenvectors are, however, usually the same.

Figure 4 shows the critical eigenvalue λ_c as a function of J for the linearizations Equation (9) (gray lines) and Equation (13) (red line) in the constant external drive regime, parameterized by η. The critical eigenvalue J_c is given by the intersection with the horizontal line at one. As was pointed out in section 3.1.2, in the noiseless approximation Equation (13) the eigenvalues of the linear operator do not depend on the working point, and thus the critical weight J_c^{md} is unique for all η such that $\nu_o > 0$ (here, critical $J_c^{\mathrm{md}} = 0.506$ indicated by red circle).

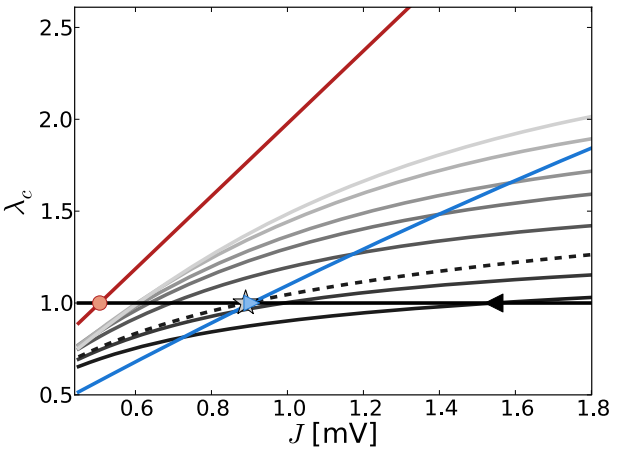


FIGURE 4 | Eigenvalue of W/θ (red) and $\tilde{W}$ (varying gray levels) with largest real part λ_c in dependence of J and external drive η (increasing $\eta = \{3.5, 5, 10, 15, 20, 30, 40\}$ from dark to light gray). The eigenvalues of W/θ do not depend on η and thus the J_c value such that $\lambda_c = 1$ [indicated by red circle, cf. Equation (13)] holds uniquely for $J_c^{\mathrm{md}} = 0.506$. The gray lines show the dependence of λ_c of $\tilde{W}$ as obtained from Equation (7) with both contributions from $d\nu/d\mu$, $d\nu/d\sigma$ (solid lines) and with the $d\nu/d\mu$ contribution only (dashed, here for clarity we just plot the curve for $\eta = 3.5$). The black triangle and the gray asterisk indicate the respective predictions of $J_c^{\mathrm{fd}} = 1.54$ and $J_c^{\mathrm{md},d\nu/d\mu} = 0.89$ for $\eta = 3.5$. For comparison, the blue line depicts the J-dependence in the input scenario with fixed total input current mean and variance, i.e., $\mu[RI] = 5$ mV and $\sigma[RI] = 60$ mV, independent of changes in J. In that case, the self-consistent population firing rate is expected to be constant and thus Equation (7) is of the form $\tilde{W} = c_1 W_{ij} + c_2 W_{ij}^2$, with constants c_1, c_2. For large σ and $c_1 > c_2$ the curve is dominated by the linear weight dependence. For the parameters shown here, the expected $J_c^{\mathrm{fd}} = 0.905$ mV (cf. **Figure 1**).

For the linearization Equation (9), and taking into account both μ and σ in Equation (7b) (solid lines), the critical weight J_c^{fd} is strongly dependent on the value of $\eta = \{3.5, 5, 10, 15, 20, 30, 40\}$ (increasing from dark to light gray), and never comes close to the prediction of the noiseless linearization J_c^{md}. In particular, for larger η J_c^{fd} increases again.

If we neglect the σ-dependence in Equation (7b), and just take into account $d\nu_o/d\mu_o = d\nu_o^{\mathrm{md}}/dRI = 1/\tau_{\mathrm{m}}\theta$ as given by Equation (11), Equation (7) is linear in W_{ij} and we recover the linear noiseless approximation Equation (13) for Equation (9). We might thus expect that Equation (9) in general will give a J_c^{fd}-prediction close to J_c^{md} if we neglect the σ-dependent quadratic term in Equation (7b). Although the prediction of this $J_c^{\mathrm{fd}|d\nu_o/d\mu_o}$ becomes smaller (cf. black triangle for $J_c^{\mathrm{fd}} = 1.54$ mV vs. the gray asterisk for $J_c^{\mathrm{fd}|d\nu_o/d\mu_o} = 0.89$ mV for $\eta = 3.5$) it never becomes as small as J_c^{md}. This is because every change in μ_o will also influence σ_o via its impact on ν_o, and thus $d\nu_o(\mu_o, \sigma_o)/d\mu_o \neq d\nu_o^{\mathrm{md}}(RI)/dRI$.

We note that the observation $J_c^{\mathrm{fd}} > J_c^{\mathrm{fd},d\nu_o/d\mu_o}$ is explained by the fact that $\tilde{W}(W_{ij})|_{d\nu_o/d\mu_o}$ is linear in W_{ij}, and thus the ratio

$\tilde{g} = |\tilde{W}(-gJ)/\tilde{W}(J)| = g$, while for the full Equation (7), $\tilde{W}(W_{ij})$ is quadratic in W_{ij}, yielding a $\tilde{g} < g$. Evaluating Equation (A20) in Appendix A3 shows that the critical eigenvalue λ_c in both linearizations is a linearly increasing function of $g, \tilde{g}$, respectively, and J_c is thus decreasing as a function of $g, \tilde{g}$.

For comparison, **Figure 4** also shows the J-dependence of the real part of the critical eigenvalue λ_c in the strongly fluctuation-driven regime ($\mu[RI] = 5\,\text{mV}$, $\sigma[RI] = 60\,\text{mV}$, blue line), with the critical J_c indicated by the crossing of the unity-line (blue triangle).

3.3. LINEAR STABILITY IN DEPENDENCE OF THE INPUT CURRENT REGIME

We will now compare the general performance of the linearizations Equations (9) and (13) and in particular the resulting predictions of the critical coupling strength J_c with the actual onset of pattern formation as observed in simulations. **Figure 1** demonstrates how the identical ring network structure undergoes pattern formation at very different values of network coupling strength J depending on the input current regime. In particular, pattern formation sets in later if the system is in a fluctuation-driven regime. This can be understood in the two limit-cases that were analyzed in sections 3.1.1 and 3.1.2.

But what determines pattern formation onset in intermediate cases that comprise the most interesting and relevant cases? As we saw in section 3.1.2, the network will even undergo a crossover from the mean-driven to a more fluctuation-driven regime, if the external input is kept constant and not adapted to counterbalance the changes in recurrent input structure for changing J, see **Figure 2C**. Also, due to the spiking nature of the activity, even in a strongly mean-driven scenario there will always be a considerable amount of variance which might change the onset of pattern formation with respect to J_c^{md}.

To give more insight into the actual onset of pattern formation, useful reduced measures are the variance and kurtosis of the distribution of firing rates over neurons, cf. **Figure 5**. **Figures 5A–C** show the histogram of rates, and the variance and kurtosis of these histograms for various J-values in the mean-driven regime averaged over 10 trials each. **Figures 5D–F** show the same for the fluctuation-driven case and **Figures 5G–I** for the intermediate case, all for the same parameters as in **Figure 1**.

When the system is still well in the linear regime $J < J_c$ the distribution is approximately Gaussian with small variance and a kurtosis close to zero (cf. **Figures 5A,D,G**, light gray curves), while once the critical weight J_c is surpassed the variance increases strongly (cf. **Figures 5B,E,H**) and the kurtosis becomes negative (platycurtic), indicating the broadening of the distribution due to the inhomogeneous rate per neuron (cf. **Figures 5C,F,G** and **5A,D,G,** dark gray curves). The respective critical weights are indicated by vertical dashed lines in **Figures 5B,E,H** and **C,F,I** for visual guidance (red: J_c^{md}, black: J_c^{fd}, gray: $J_c^{\text{fd}|d\nu_o/d\mu_o}$).

The onset of pattern formation in the strongly fluctuation-driven regime at J_c^{fd} is more shallow than in the mean-driven regime at J_c^{md}. This might be due to deviations of the input spike statistics from the asynchronous-irregular activity assumption underlying the linear response derivation of J_c^{fd}: the standard-deviation of the input is extremely large at $\sigma = 60\,\text{mV}$. This implies that the membrane-potential will be hyperpolarized for long times, while at other times it is highly depolarized, and several spikes are emitted in short succession. This typically leads to higher coefficients of variation, i.e., more irregular firing than expected for Poisson spike statistics [see also Brunel (2000) and supplementary material section S1]. Also, even in the highly fluctuation-driven regime there may be residual correlations between spike trains. Lastly, the increased network coupling strength might lead to deviations from the Gaussian white noise approximation of input currents underlying Equation (5) and the linear response theory yielding Equation (7) (see supplementary material section S1 for an analysis). How such deviations of the input statistics can indeed change the spike response behavior of LIF neurons was studied, e.g., in Moreno et al. (2002); Renart et al. (2007); Moreno-Bote et al. (2008); Helias et al. (2010).

However, another contributing factor, and more likely the reason for the shallow transition at J_c^{fd}, is the degeneracy of the maximal eigenvalue, see **Figures 3D,E**. In the fluctuation-driven regime the system is subject to strong perpetual perturbations that lead to a switching between the two dominant eigenvectors depicted in **Figure 3D** in the transition regime $J_c^{\text{fd}} \pm \Delta J$. Indeed, as can be seen from **Figure 1E**, the periodic pattern is already clearly distinguishable at J_c^{fd} over several hundred milliseconds, but the average activity depicted in the histogram is still quite flat, yielding smaller variance and kurtosis of the respective rate distribution.

Finally, **Figures 5D–F** demonstrate clearly how pattern formation occurs for intermediate J-values $J_c^{\text{md}} \lesssim J_c^{\text{inter}} < J_c^{\text{fd}}$ if the system is neither clearly in the mean- nor strongly fluctuation-driven regime, and even changes from one to the other regime with increasing coupling strength, cf. **Figure 2C**. As we demonstrate in the supplementary material section S1, most input and network settings have pattern formation onsets that usually agree much better with the J_c^{md} prediction than with J_c^{fd}. This is a very consistent finding, even in cases where neurons are well in the fluctuation-driven regime in terms of subthreshold mean and pronounced input variance.

As we discuss in the supplementary material section S1 the reason for this finding lies in an asymmetry between the excitatory and inhibitory compound input current statistics. The excitatory subpopulation tends to be synchronized, even in the presence of strong balanced external noise, while inhibition actively decorrelates itself. The explanation for this effect lies in the local connectivity of ring networks, leading to a population-specific pattern of correlations already below the critical coupling: excess synchrony of the excitatory population will reinforce further spiking, while spikes from the inhibitory population decreases the instantaneous firing probability [see also Helias et al. (2013) for a related discussion of this effect in the framework of balanced random networks]. Near synchrony is thereby effectively enhanced in the excitatory population and suppressed in the inhibitory one. Volleys of excitatory input spikes act like compound pulses with large amplitude (on the order of up to θ) in an otherwise balanced asynchronous background activity. At these amplitudes the effective gain Equation (7) derived from linear response theory becomes linear in J with a slope proportional to $1/\theta$, i.e., the slope of the linear model Equation (11), see supplementary

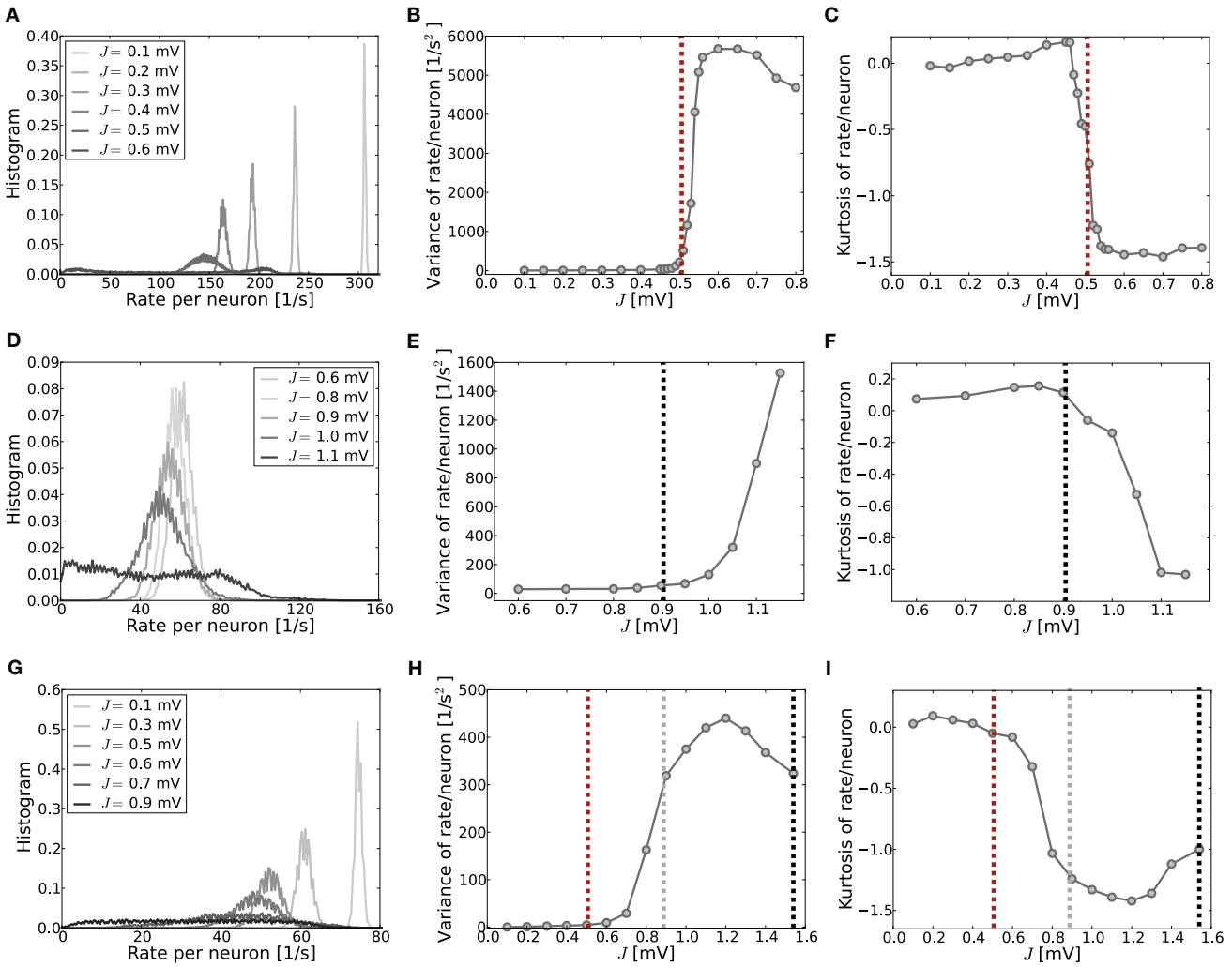


FIGURE 5 | The histograms of rates, variance and kurtosis for various weights J in a network of $N = 2500$, $\kappa = 250$, and $g = 6$ averaged over 10 trials. (A–C) Shows the mean-driven case ($I_x = 2500$ pA delivered as Poisson noise with external synaptic strength $J_x = 0.1$ mV), **(D–F)** the fluctuation-driven case (Poisson noise adapted s. t. $\mu[RI] = 5$ mV and $\sigma[RI] = 60$ mV) and **(G–I)** an intermediate case, where the external input was Poisson noise (current-amplitude 875 pA, $J_x = 0.1$ mV). For subcritical $J < J_c$ the distributions are approximately Gaussian with small variance and kurtosis close to zero. For supracritical weight the distributions become broader and platykurtic, indicated by the negative kurtosis. The vertical dashed lines in **(B,C)** and **(E,F)** mark the respective critical weights, i.e., $J_c = 0.506$ mV [red, using Equation (13)] for the mean-driven, and $J_c = 0.905$ mV for the strongly fluctuation-driven case [black, using Equation (9)]. The dashed vertical lines in **(H,I)** mark the respective predictions for the critical weight by Equation (13) (red, independent of η), and by Equation (9) with only the contribution of $\partial\nu/\partial\mu$ (gray) and both linear and quadratic contributions $\partial\nu/\partial\mu$, $\partial\nu/\partial\sigma$ (black) for the estimation of the effective coupling strengths, cf. Equation (7b). The increase in the kurtosis and decrease in the variance for very large J in **(B,C)** and **(H,I)** is due to the rectification of rates that leads to increasing mass at zero. Note that sampling is denser around the expected J_c-value in the first two rows.

material section S1. This explains the fact that Equation (13) has better predictive power also in the intermediate or—comparably weakly—fluctuation-driven scenarios. Moreover, the fluctuation-driven prediction only becomes valid if the additional external noise sufficiently dilutes residual synchrony, such as it is the case for $\mu[RI] = 5$ mV and $\sigma[RI] = 60$ mV.

3.4. COARSE-GRAINED RING NETWORK

In section 3.2 we saw that the full system can effectively be reduced to a five-dimensional system. However, the computation of eigensystems (cf. Appendix A3) is still quite involved. In this section we study how well a further simplified system that is formally identical to the well-known Ermentrout-Cowan networks [Ermentrout and Cowan (1979a,b, 1980) and supplementary material section S2] predicts the dynamics of the full network.

We coarse-grain the system and combine groups of $\ell = N/N_I$ neurons, such that they contain four excitatory and one inhibitory neuron (ℓ needs to be odd in the following), as indicated in **Figure 6**. The neurons $i_c \in \{0, \ldots, \ell - 1\}$ within each cell $c \in \{0, \ldots, C - 1\}$, $C = N/\ell$, are connected to every other neuron within their cell c but not to themselves. Moreover, all

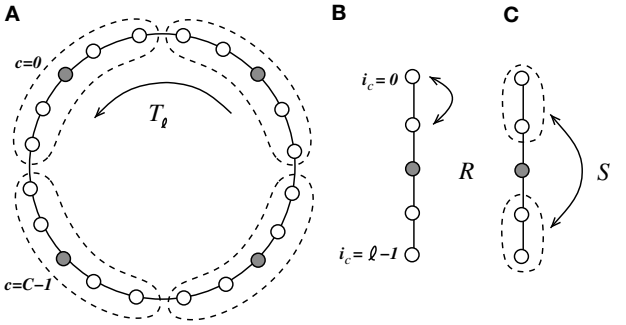


FIGURE 6 | We coarse-grain the system by introduction of cells containing four excitatory neurons and one inhibitory neuron [indicated in (A) by the dashed line]. The coarse-grained system is invariant to three symmetry operations. **(A)** T_ℓ is a translation of the cell by $\ell = 5$, **(B)** R is the mirror symmetry of two neighboring excitatory neurons within each cell, and **(C)** S is the mirror symmetry of the two pairs of neighboring excitatory neurons within each cell.

neurons within one local cell c are connected to all neurons within the $K < C/2$ neighboring cells to the left and to the right, i.e., $\{(c-K) \bmod C, \ldots, (c-1) \bmod C, (c+1) \bmod C, \ldots, (c+K) \bmod C\}$. The computation of the eigensystem is outlined in Appendix A4. Due to the two mirror symmetries corresponding to exchange of neurons within the unit cell (with corresponding eigenvalues $r, s \in \{-1, 1\}$) and the translational symmetry with the eigenvalue determined by the wavenumber α, the system can be reduced to effectively two dimensions.

3.4.1. Coarse-grained network dynamics

In the coarse-grained system we only need knowledge about two elements of the ring, one excitatory and one inhibitory neuron within the same cell (arbitrarily chosen as ν_0, $\nu_{(\ell-1)/2}$ in $c = 0$, without loss of generality). For a fixed set of eigenvalues r, s, α, the activities of all remaining neurons are then uniquely determined. The homogeneous mode $\nu_o(t)$ couples to the symmetric subspace $r, s = 1$ only and we can write $\nu_o(t)$ thus as a linear combination of the two eigenvectors $\nu^{\mathrm{I}}_{\alpha_0}$ and $\nu^{\mathrm{E}}_{1,1,\alpha_0}$, where $\alpha_0 = 0$ (cf. Appendix A4).

To analyze the stability of this mode with respect to spatial perturbations we write the population activity as a linear combination of eigenvectors with non-zero wavenumbers α_i, $\nu^{\mathrm{E}}_{1,1,\alpha_i}$, $\nu^{\mathrm{I}}_{\alpha_i}$, such that

$$\nu(t) = \sum_{i=0}^{C-1} a_i(t) \left(\nu^{\mathrm{E}}_{1,1,\alpha_i} + \nu^{\mathrm{I}}_{\alpha_i}\right) . \tag{18}$$

Then, for all neurons $n \in \{0, \ldots, N-1\}$ the input, $(W\nu(t))[n]$, in the reduced system becomes

$$(W\nu(t))[n] = \sum_{c=(\mathcal{J}[n]-K) \bmod C}^{(\mathcal{J}[n]+K) \bmod C} \sum_{j=0}^{\ell-1} \left(W_{n,(c\ell \bmod N+j)}\, \nu(t)\right) \left[\mathcal{J}[n]\ell \bmod N + j\right] \tag{19}$$

$$\begin{aligned}
&= J\left(\sum_{c=-K}^{K} T^c_\ell (1+R)(1+S) \sum_{i=0}^{C-1} a_i(t)\, \nu^{\mathrm{E}}_{1,1,\alpha_i}\right) [\mathcal{J}[n]\ell \bmod N] \\
&\quad - gJ\left(\sum_{c=-K}^{K} T^c_\ell \sum_{i=0}^{C-1} a_i(t) \nu^{\mathrm{I}}_{\alpha_i}\right) [\mathcal{J}[n]\ell \bmod N + (\ell-1)/2] \\
&= 4J\left(\sum_{i=0}^{C-1} \sum_{c=-K}^{K} e^{i\alpha_i c}\, a_i(t)\, \nu^{\mathrm{E}}_{1,1,\alpha_i}\right) [\mathcal{J}[n]\ell \bmod N] \\
&\quad - gJ\left(\sum_{i=0}^{C-1} \sum_{c=-K}^{K} e^{i\alpha_i c}\, a_i(t) \nu^{\mathrm{I}}_{\alpha_i}\right) [\mathcal{J}[n]\ell \bmod N + (\ell-1)/2] \\
&= 4J\left(\sum_{i=0}^{C-1} \Omega[\alpha_i]\, a_i(t)\, \nu^{\mathrm{E}}_{1,1,\alpha_i}\right) [\mathcal{J}[n]\ell \bmod N] \\
&\quad - gJ\left(\sum_{i=0}^{C-1} \Omega[\alpha_i]\, a_i(t) \nu^{\mathrm{I}}_{\alpha_i}\right) [\mathcal{J}[n]\ell \bmod N + (\ell-1)/2]
\end{aligned}$$

with $\mathcal{J}[n] := \lfloor n/\ell \rfloor$. In the first identity we made use of the operators R, S and T_ℓ to express the input-connectivity of the network. The second identity follows from the linearity of the sum over eigenmodes and the respective eigenvalues of the operators R, S and T_ℓ. We excluded synaptic self-coupling, so we need to subtract that input of the cell to itself, i.e.

$$\begin{aligned}
&J\left(\sum_{i=1}^{C-1} a_i(t) \nu^{\mathrm{E}}_{1,1,\alpha_i}\right)[n] \cdot (1 - \delta_{n,\mathcal{J}[n]\ell \bmod N + (\ell-1)/2}) \\
&\quad -gJ\left(\sum_{i=1}^{C-1} a_i(t) \nu^{\mathrm{I}}_{\alpha_i}\right)[n] \cdot \delta_{n,\mathcal{J}[n]\ell \bmod N + (\ell-1)/2}.
\end{aligned}$$

Leaving out one inhibitory input if the receiving neuron is inhibitory, and one excitatory input if the receiver is excitatory leads to an asymmetry in the total input of excitatory and inhibitory neurons. This implies that the spatially homogeneous mode is not an eigenmode of the system anymore. To compensate for this the weights are rescaled such that the input per neuron equals that in the full (non-coarse-grained) system, i.e., to $\mu = \kappa J(\beta - g(1-\beta))$. Hence,

$$W'_{ij|i\in\mathcal{E}} = W_{ij|i\in\mathcal{E}} \frac{\mu}{J((3+8K) - g(2K+1))} =: \mu_{\mathcal{E}} W_{ij|i\in\mathcal{E}} \tag{20}$$

$$W'_{ij|i\in\mathcal{I}} = W_{ij|i\in\mathcal{I}} \frac{\mu}{J(4+(4-g)2K)} =: \mu_{\mathcal{I}} W_{ij|i\in\mathcal{I}}$$

Note, that if we do not exclude synaptic self-coupling in the coarse-grained system the eigensystem looks dramatically different from that of the full system. In particular the eigenvalues

are nearly fully real and the critical eigenvalue is much smaller implying a higher critical weight.

In the homogeneous mode the input to each neuron $n \in \{0, \ldots, N-1\}$ is completely identical, all neurons fire at the same rate ν_o given by the self-consistent solution

$$\begin{pmatrix} \tau_m \nu_o[0] \\ \tau_m \nu_o[2] \end{pmatrix} = \left[\begin{pmatrix} -\theta & 0 \\ 0 & -\theta \end{pmatrix} - J \begin{pmatrix} 4\mu_{\mathcal{E}}\Omega[\alpha] - 1 & -g\mu_{\mathcal{E}}\Omega[\alpha] \\ 4\mu_{\mathcal{I}}\Omega[\alpha] & -g\mu_{\mathcal{I}}(\Omega[\alpha]-1) \end{pmatrix} \right]^{-1} \begin{pmatrix} RI_x - \theta/2 \\ RI_x - \theta/2 \end{pmatrix}$$

where $\Omega(\alpha) := \frac{\sin(\alpha(K+1/2))}{\sin(\alpha/2)}$. The eigenvalues of the reduced rescaled coupling matrix

$$\frac{W'}{\theta} = \frac{J}{\theta} \begin{pmatrix} 4\mu_{\mathcal{E}}\Omega[\alpha] - 1 & -g\mu_{\mathcal{E}}\Omega[\alpha] \\ 4\mu_{\mathcal{I}}\Omega[\alpha] & -g\mu_{\mathcal{I}}(\Omega[\alpha]-1) \end{pmatrix} \tag{21}$$

then determine for which parameters g, J the spatially homogeneous state becomes unstable and which wavenumber $\Lambda_i(\alpha_i)$ will grow fastest and define the spatial pattern. The critical eigenvalue for the coarse-grained system is thus given by the eigenvalue of $\frac{W'}{\theta}$ with largest real part that will cross unity from below first, i.e.

$$\lambda_c(g,J) = \max_{i \in \{0,\ldots,C-1\}} \left\{ \frac{\mu}{2\theta\mu_{\mathcal{I}}\mu_{\mathcal{E}}} \Big(\mu_{\mathcal{I}}(4\Omega[\alpha_i]-1) - g\mu_{\mathcal{E}}(\Omega[\alpha_i]-1) + \sqrt{4g\mu_{\mathcal{I}}\mu_{\mathcal{E}}(1-5\Omega[\alpha_i]) + (\mu_{\mathcal{I}}(4\Omega[\alpha_i]-1) - g\mu_{\mathcal{E}}(\Omega[\alpha_i]-1))^2} \Big) \right\} \tag{22}$$

In the fluctuation-driven regime J and g in Equation (21) need to be substituted by $\tilde{J} := \tilde{W}(J)$ and $\tilde{g}$, such that the critical eigenvalue is given by $\lambda_c^{\mathrm{fd}}(\tilde{g}, \tilde{J})$.

On a mesoscopic level the coarse-grained and full system give rise to very similar activity patterns—differences become apparent in the fine-scale features, e.g., the full analysis of a network of 2500 neurons with 10% connectivity and $g = 6$ predicts destabilization of wavenumber 14 in both mean- and fluctuation-driven regime, while the coarse-grained model predicts 13, see also **Figure 7**. We note that the coarse-grained model presented in this section is formally identical to the linearization derived in Ermentrout and Cowan (1980) in the context of a neural field ring model if one assumes a boxcar-footprint. We summarize the respective derivation in the supplementary material section S2.

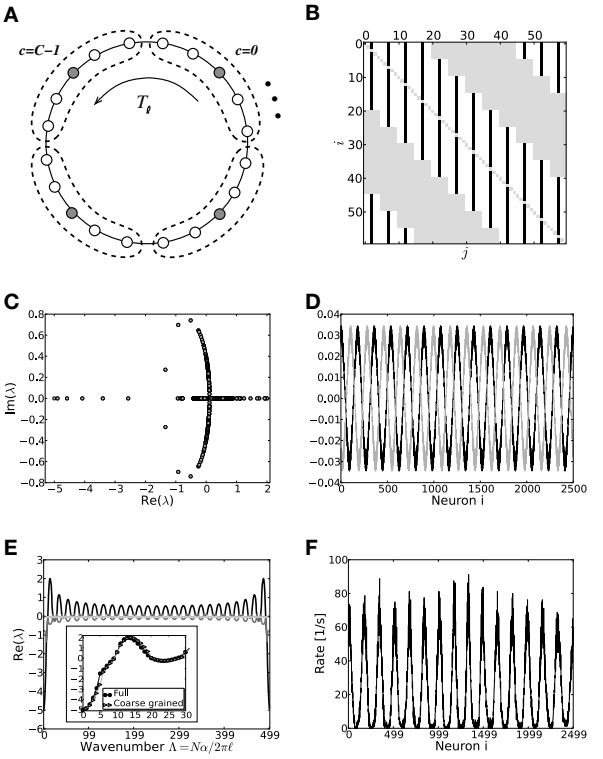


FIGURE 7 | (A) Sketch of the layout of the coarse-grained system. Each neuron is connected to the $L-1$ other cells within its cell and to all neurons in its $2K$ nearest cell neighbors. A shift in neuron label by five yields the same ring as before and is formally expressed by the shift operator T_ℓ with $\ell = 5$. **(B)** Shows the coupling matrix W of the ring: black squares depict inhibitory, white squares excitatory, and gray squares zero coupling from j to i ($N = 60$, $\kappa = 30$). **(C)** Eigenvalue spectrum of a scaled ring matrix W/θ with $N = 2500$, $\kappa = 250$, and $g = 6$. **(D)** The two eigenvectors belonging to the twice degenerated eigenvalue from **(C)** with largest real part, λ_c^{md}. **(E)** Five bands of eigenvalues as a function of the wavenumber. The black line depicts the band containing λ_c^{md}. The inset shows the critical bands of the system in section 3.2 (black circles, cf. also **Figure 3**) vs. the critical band in the coarse-grained system (gray triangles). The critical wavenumber differs by one. **(F)** The corresponding rate per neuron in a simulation of $N = 2500$ integrate-and-fire neurons with $J = 1$ mV and $I_x = 750$ pA.

3.5. SENSITIVITY TO RANDOMNESS

To conclude, we study how robust the findings of the previous sections are to effects of randomness in either structure or weight distribution. Though we again only show the eigenvalue spectra for the synaptic coupling matrix W/θ, that is the relevant linear operator in the mean-driven case, all results map straight-forwardly to the fluctuation-driven regime with appropriate rescaling of J and g.

3.5.1. Small-world networks: structural noise

First, we will study the effect of the introduction of structural randomness by rewiring individual connections with probability p_r [for details see Kriener et al. (2009)] such that the input composition from the network to every neuron stays constant at μ. The random rewiring procedure means that each realization of a network will be different and moreover lacks invariance to T_ℓ such that the eigensystems of individual coupling matrices need to be computed numerically. For low p_r the rate model for the translation-invariant network still gives very good results and the patterns that form for large J are indeed strongly correlated with the critical eigenvector v_c, see **Figure 8**. For higher p_r predictions become worse and more than one vector can contribute to the pattern (e.g., **Figures 8C2,C3**). For low rewiring probability $p_r < 0.02$ we can still analytically estimate the critical eigenvalue by employing a mean field model (Grabow et al., 2012) for small-world networks in which we rewire "on average," cf. Appendix A5 and **Figures 8A2–C2,A4–C4**. Hence, also for small-world networks of excitatory and inhibitory spiking neurons, rates and linear

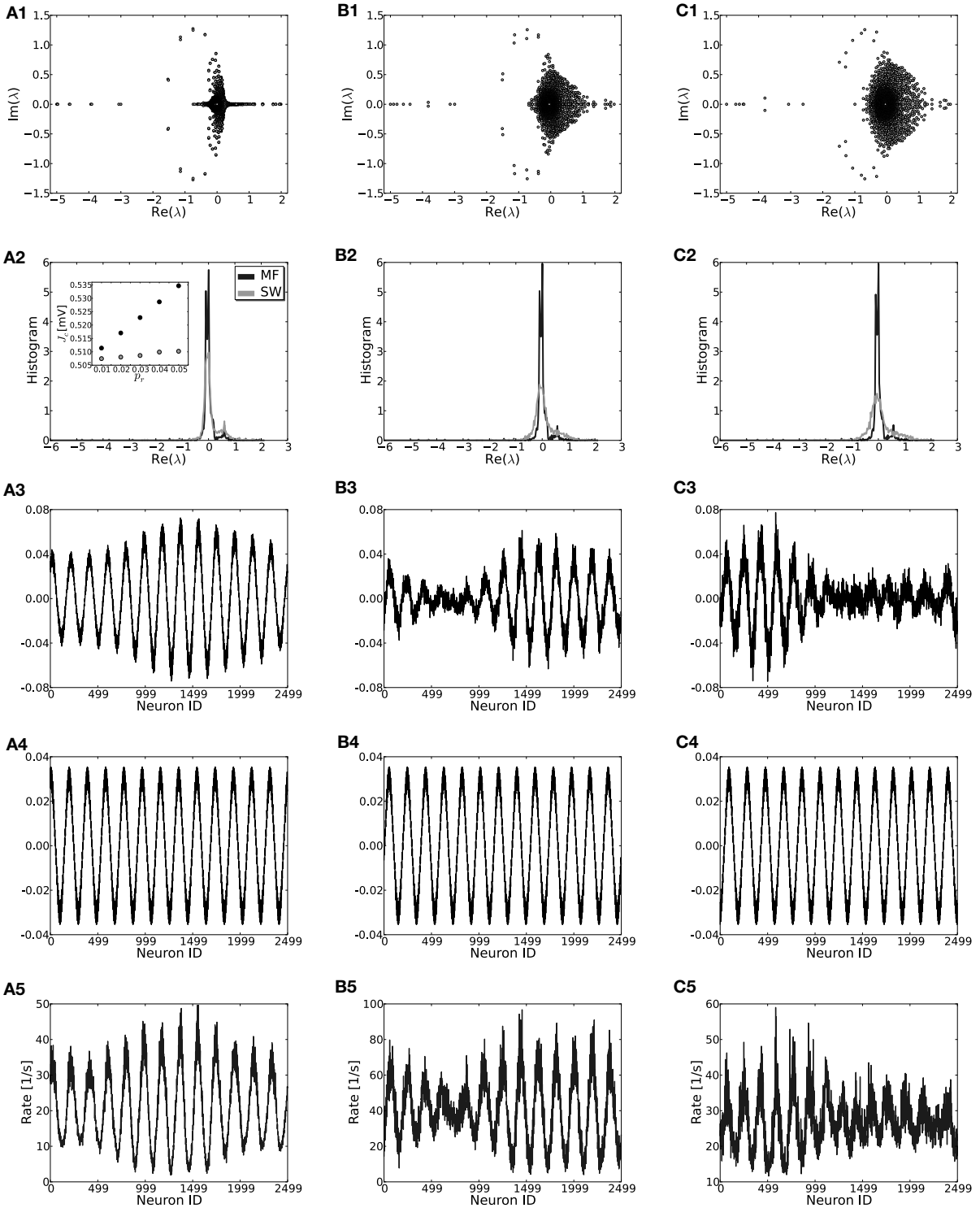


FIGURE 8 | Effect of rewiring of edges on the eigenvalue distributions (row 1), the mean-field density of the real parts of eigenvalues (row 2), the crictical eigenvectors of the specific realization (row 3) and the mean-field system (row 4) and the rate per neuron (row 5) in small-world networks for three different rewiring probabilities $p_r = 0.01, 0.03, 0.05$ in (A,B,C), respectively (cf. Watts and Strogatz, 1998; Kriener et al., 2009). While in **(A3,B3)** it is the dominant eigenmode v_c that correlates strongest with the actual rate distribution [**(A5,B5)**, Pearson correlation coefficients of $c = 0.967$ and $c = 0.945$, respectively], in **(C3)** the linear combination of the two leading eigenvectors correlates best with the actual rates [**(C5)**, v_c has $c = 0.68$ while the combination of the two leading modes gives $c = 0.91$]. The inset in **(A2)** shows the estimate for the critical weight J_c as a function of p_r in the mean field model (MF, black) and as an average of 50 network realizations (SW, gray). Parameters: $N = 2500$, $J = 1$ mV and $I_x = 750$ pA. During rewiring the number of excitatory and inhibitory inputs was kept constant.

stability can be estimated from the translation-invariant network rate model if p_r is small. The results of the mean-field model are slightly worse than for unweighted networks (cf. Grabow et al., 2012) because of the two qualitatively different types of edges in the networks considered here (excitatory and inhibitory). For the small number of edges κ we assume in **Figure 8** it can occur that for very small p_r no inhibitory edges are rewired at all, while the mean-field model still reassigns inhibitory connection density. We find the mean field to work better for denser or larger networks, and if the excitatory and inhibitory input per neuron is kept the same during rewiring (not shown).

If the constant input per neuron constraint is dropped, the homogeneous mode $(1, \ldots, 1)^\top$ is not an eigenmode anymore and also the prediction of stability fails in large parts. On the one hand there are activity inhomogeneities forming out simply due to the fact that some areas on the ring will by chance get more or less net inhibitory input, and hence some neuronal subpopulation will be driven to zero rates, while others—lacking net inhibitory input from these rectified neurons—will have very high rates. Similar effects arise when the distribution of inhibitory neurons is not periodic anymore, but, e.g., distributed randomly along the ring.

3.5.2. Ring networks not conform with Dale's principle

Finally, if Dale's principle—i.e., the biological fact that each neuron can only either depolarize or hyperpolarize all its postsynaptic targets, but never both at the same time—is violated, both the eigensystem (**Figure 9A**), as well as the spiking dynamics looks akin to that of a random network, even if the underlying topology is a ring, cf. **Figures 9B,C**. Hence, the identity of neurons in terms of being excitatory or inhibitory is a necessary condition for the formation of structured patterns in the case of identical footprints for inhibition and excitation considered here.

Even though in all cases pattern formation is either strongly or completely impaired, the prediction Equation (14) of individual rates in the low coupling regime is excellent, also in inhomogeneous networks, where different neurons receive very different inputs, given the individual mean membrane potentials $\langle V_i(t)\rangle_t$ are used in Equation (15).

4. DISCUSSION

In this paper we analyzed the dynamics of pattern formation in networks of excitatory and inhibitory spiking neurons that are embedded on a ring architecture. In particular, we studied the dependence of the underlying bifurcation on the input current statistics and structural noise, and derived a coarse-grained system that is formally analogous to the classical Ermentrout-Cowan networks. To conclude, we will summarize and discuss our findings.

4.1. IMPACT OF INPUT CURRENT STATISTICS

The very nature of spiking neuron networks leads to considerable input current fluctuations, and hence in general both mean $\mu[RI]$ and standard deviation $\sigma[RI]$ of the input will shape the output rate of the neuron. For leaky integrate-and-fire neurons the output rate as a function of $\mu[RI]$ and $\sigma[RI]$, assuming stationary Poisson-like uncorrelated input spike trains and weak synaptic coupling J, is given by the Siegert equation (5) (Amit and Tsodyks, 1991; Amit and Brunel, 1997). Thus, in order to analyze the linear stability of this self-consistent solution, in general the derivatives with respect to both $\mu[RI]$ and $\sigma[RI]$ need to be taken into account, cf. Equation (7).

We analyzed two different ways to drive the networks: in the first all neurons in the network were driven with the same excitatory Poisson-noise ν_X, independent on the recurrent coupling strength J. In this case the total input RI the neurons receive, i.e., the external input RI_x together with the recurrent input RI_s from the network, changes with increasing J, such that the mean $\mu[RI]$ decreases, while the standard deviation $\sigma[RI]$ increases.

In the second input scenario, we corrected for the change in recurrent input contributions with changing J by administering compensating inhibitory and excitatory external Poisson input ν_{Ix}, ν_{Ex}. Thus, $\mu[RI]$ and $\sigma[RI]$, and hence also ν_o are approximately constant with increasing J.

We find that in the first input scenario the critical J_c as observed in simulations deviates strongly from the one predicted by Equations (7),(9), which are derived from the self-consistent solution Equation (5) by linear response theory. In particular,

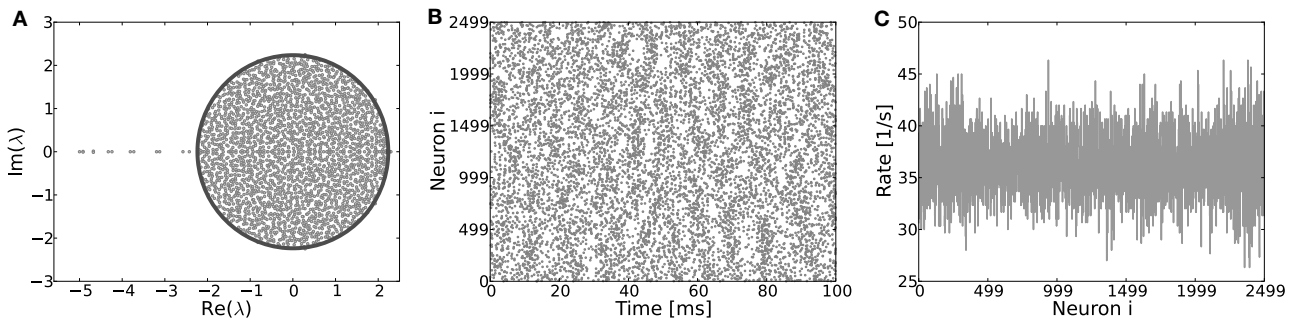


FIGURE 9 | (A) Eigenvalue distribution in the complex plane of a ring network with random assignment of weights, irrespective of the identity of a neuron, hence each existing synapse has weight J with probability β and $-gJ$ with probability $(1-\beta)$. These ring coupling matrices have eigenvalue spectra very akin to those of corresponding random networks and most of the eigenvalues adhere to the circle law prediction (Girko, 1984) for random networks $r = \sqrt{Np(\beta + g^2(1-\beta)) - \mu^2/N}$ (dark line). Only a few singular eigenvalues on the left of the center indicate the underlying ring topology. The eigenvectors of such "hybrid" ring matrices have no apparent structure and no pattern formation occurs when the system becomes supracritical, as clearly visible from the spiking activity **(B)** and the rates per neuron **(C)**. Here, $J_c \approx 0.44$ mV, simulation parameters $J = 0.6$, $g = 6$ and $N = 2500$.

for the mean-driven regime, where $\mu[RI] > V_{\text{thr}}$, J_c^{md} becomes basically independent of the rate ν_o as well as of $\sigma[RI]$, irrespective of its yet considerable magnitude. Instead, a very simple linearization derived from the noiseless f-I-curve Equation (10) applies. The J_c^{md} derived from this linearization is always considerably smaller than the prediction by Equations (7), (9). For intermediate cases where neurons are neither mean- nor strongly fluctuation-driven, the critical coupling strength lies in between the two predictions, and often closer to J_c^{md}.

Possible reasons for the failure of the prediction of Equations (7), (9) in the constant external drive scenario (parameterized by η, cf. section 2) are the break-down of the weak-coupling assumption, as well as the uncorrelated Poisson assumption for the input spike trains in the recurrent contribution. In line with earlier findings (Tetzlaff et al., 2012), we could not identify deviations with respect to coupling strength, however, we observed clear deviations from the Poisson assumption.

If spiking is Poissonian we expect the interspike intervals to be distributed exponentially, and the coefficient of variation (CV) of the interspike intervals (ISI) to be unity. A smaller CV indicates more regular, a larger one more irregular spiking. As can be seen from **Figures 1A–C** and **1G–I**, activity is spatio-temporally structured, with locally clustered spiking that can induce significant pairwise correlations, and low coefficients of variation of the interspike intervals for small J (cf. also supplementary material section S1, Figures S1B,C). In effect, these deviations from uncorrelated Poisson spiking lead to deviations from the Gaussian white noise approximation underlying the derivation of Equations (5) and (9), especially for increasing J.

We observe that in particular the excitatory subpopulation tends to synchronize because of the highly recurrent local structure and positive feedback, while the inhibitory population actively desynchronizes itself due to negative feedback (see supplementary material section S1, Figures S3, S4, and Helias et al. (2013) for a related discussion). Thus, high amplitude volleys of excitatory input in a balanced background comprised of asynchronous inhibition and external Poisson-drive move the network in a regime that is better captured by the mean-driven low-noise approximation Equation (11).

Activity in the strongly fluctuation-driven regime, on the other hand, is much more asynchronous and irregular by construction, cf. **Figures 1D–F** with CV[ISI] of unity and larger (see supplementary material, section S1, Figures S1B,C). When in this scenario $\sigma[RI]$ is very large, while $\mu[RI]$ is sufficiently subthreshold, the prediction for the critical coupling strength J_c^{fd} from Equations (7), (9) for linearity to break down indeed appears to be correct, or even seems to underestimate the actual onset of pattern formation.

This shallow transition around J_c^{fd} can be explained by slow noise-induced transitions between the two degenerate eigenvectors which grow quickest once pattern formation takes place. However, the fact that the CV is larger than unity will also lead to deviations from the prediction. In a series of papers Moreno et al. (2002); Renart et al. (2007) and Moreno-Bote et al. (2008) studied the impact of deviations of the Poisson-assumption of input spike trains on the firing rates of neurons. They studied the effect of positive and negative spike train auto- and cross-correlations parametrized by the Fano-factor $F = \sigma^2[\text{counts}]/\mu[\text{counts}]$ of spike counts on the output firing rate of current-based LIF neurons in both the fluctuation- and the mean-driven regime. For renewal processes the CV of interspike intervals is directly related to the Fano-factor of counts by $F = \text{CV}^2$ (Cox, 1962). In Moreno et al. (2002); Renart et al. (2007); Moreno-Bote et al. (2008) it was shown that in the fluctuation-driven regime output rates are quite sensitive to input noise correlations (parameterized by the Fano factor), while in the mean-driven regime sensitivity is less pronounced.

In simulations of ring networks in the strongly fluctuation-driven case we observe that rates decrease with increasing correlations, while in otherwise equivalent random networks, rates increase (not shown). So the different recurrent input structure plays a crucial role for the effective input current and thus network dynamics properties. These differences are often, as here for ring and grid networks, a direct consequence of complex network topology, and it is thus important to understand how connectivity structure translates to single neuron activity, as well as collective network states. A more thorough analysis of such effects in the context of firing rate stability in random and spatially structured networks will be subject of subsequent research.

4.2. IMPACT OF FINITE TIME SCALES

We assumed throughout the manuscript that the system has a negligible refractory time constant, small delays and instantaneous post-synaptic currents (δ-synapses). If these assumptions are dropped we observe differences especially for the mean-driven case: for larger delays d in the order of several milli-seconds, pronounced oscillations can occur that tend to synchronize the system locally on the order of the footprint (Kriener et al., 2009), or for very strong external drive also on a population-wide level (Brunel, 2000). A systematic study of the complex effects of delays on pattern formation in neural-field-type ring networks of spiking neurons was presented in Roxin et al. (2005). In the network we analyzed here, noisy delay oscillations mainly perturb developing patterns and can thus shift the onset of clear pattern formation to larger J_c.

Similar effects arise in the case of finite refractory time τ_{ref}. A larger τ_{ref}, especially in the presence of oscillations, i.e., highly coincidental input spiking, decreases the effective network input, in particular if $\tau_{\text{ref}} > d$, and thus also shifts the input regime. This can be compensated for by explicitly taking into account τ_{ref} in the derivation of the noiseless linearization.

Finally, if synaptic time constants are finite, while total input currents stay the same, delay oscillations can be smoothened out, and the effect of absolute refractoriness is decreased. Increasing synaptic time constants can thus counteract the effect of finite τ_{ref} and d.

4.3. SPIKING NETWORKS vs. NEURAL FIELD MODELS

Pattern formation, such as activity bump formation or periodic patterns, is a well-studied phenomenon in ring and toroidal networks (see e.g., Ermentrout and Cowan, 1979a,b; Ben-Yishai et al., 1995; Usher et al., 1995; Bressloff and Coombes, 1998, 2000; Shriki et al., 2003; Marti and Rinzel, 2013). Earlier studies of periodic pattern formation in spiking neuron networks already

showed that ring- and grid-networks of mean-driven neurons can undergo a Turing-bifurcation (Usher et al., 1995; Bressloff and Coombes, 1998, 2000). In particular, Usher et al. (1995) observed a dependence on the input level. For weaker, yet suprathreshold mean input current, activity patches diffuse chaotically across the spatial extend of the system, while for stronger inputs a stable spatial pattern develops that relates to the critical eigenmode. We observe similar dynamics in the networks considered here as well, if the system is in the intermediate regime. This can be understood as follows: if the noise component is relatively strong, it shifts the actual critical coupling strength J_c^{md} to higher values J_c^{inter} inbetween J_c^{md} and J_c^{fd}. If the system is thus effectively slightly sub-critical in terms of J_c^{inter}, input-noise can transiently excite eigenmodes close to effective criticality, due to Hebbian (Dayan and Abbott, 2001) or non-normal amplification (Murphy and Miller, 2009).

To compare our results to the classical work on neural field models on ring topologies (Ermentrout and Cowan, 1979a,b, 1980) we moreover studied a coarse-grained model that reduces the N-dimensional system to an effectively two-dimensional one that is structurally equivalent to the linearized two-population model presented, e.g., in (Ermentrout and Cowan, 1980) (see supplementary material section S2) which, however, allows for a direct quantitative mapping to the spiking network dynamics.

Because of the coarse-graining some of the connectivity details get lost and thus the critical eigenmode is not identical with that of the original ring network. Moreover, we observe the importance of excluding self-coupling in the coarse-grained model: if that is not excluded the resulting eigensystem looks dramatically different with nearly completely real-valued eigenvalues.

4.4. IMPACT OF STRUCTURAL NOISE

For the translation-invariant distribution of inhibitory neurons across a ring topology the eigensystem, and hence the respective linear stability, can be computed analytically. If inhibitory neurons are however randomly distributed, some parts of the ring will have a higher, others a lower density of inhibitory neurons, such that these networks form an ensemble of possible realizations, most of which are not invariant to translations. The resulting inhomogeneous rate per neuron in the subcritical regime can still be predicted well from the linear rate model. However, even when knowing the dominant eigenvector, it will usually not be periodic, and due to the inhomogeneous input per neuron different neurons will be at quite different working points. There is interference between the rate distribution and the dominant mode. Thus, the activity pattern that evolves in the supracritical regime is not straightforward to predict.

If the weights are distributed fully randomly across the whole ring topology irrespective of the identities of the neurons, i.e., in violation of Dale's principle, the resulting eigensystems look very similar to those of corresponding networks with random topologies. Only a few surviving singular eigenvalues serve to distinguish the underlying ring from the random topology. This directly translates to the spiking dynamics: the activity is much more asynchronous than that of Dale-conform networks and most of all pattern formation can not take place due to the effective lack of lateral inhibition.

4.5. IMPACT OF NETWORK SIZE

If κ/N is kept constant when varying the network size N the critical coupling strength in both the mean- and the fluctuation-driven regime decreases. For example, for $N = 10000$ and $\kappa = 1000$ the critical coupling strength becomes $J_c^{\mathrm{md}} \approx 0.2\,\mathrm{mV}$ for the mean-driven, and $J_c^{\mathrm{fd}} \approx 0.32\,\mathrm{mV}$ for the fluctuation-driven regime. If the number of synapses per neuron κ is fixed when N increases, both J_c^{md} and J_c^{fd} hardly change because the maximal eigenvalue is approximately unaffected by network dilution.

We emphasize that the analysis presented here can be extended to two-dimensional grid networks in a straightforward manner, see supplementary material section S3. We conclude by remarking that the rate model Equation (13) is also a valuable tool to study the rate dynamics of spatially embedded spiking neuron networks with e.g., inhomogeneous neuron distribution and distance-dependent connectivity in the mean-driven regime, allowing for a treatment of more general networks than commonly studied random networks or neural field models. The presented model is in conclusion a further step in the efforts to find and understand appropriate descriptions of the high-dimensional spiking activity in structured networks.

ACKNOWLEDGMENTS

We want to thank the reviewers for their helpful comments. We gratefully acknowledge funding by the eScience program of the Research Council of Norway under grant 178892/V30 (eNeuro), the Helmholtz Association: HASB and portfolio theme SMHB, the Next-Generation Supercomputer Project of MEXT, EU Grant 269921 (BrainScaleS), and by the German Federal Ministry for Education and Research (BMBF, grant 01GQ0420 to BCCN Freiburg, and grant 01GQ0830 to BFNT Freiburg*Tübingen). All network simulations were carried out with NEST (http://www.nest-initiative.org).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fncom.2013.00187/abstract

REFERENCES

Abbott, L. F., and van Vreeswijk, C. (1993). Asynchronous states in networks of pulse-coupled oscillators. *Phys. Rev. E* 48, 1483–1490. doi: 10.1103/PhysRevE.48.1483

Amari, S.-i. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87. doi: 10.1007/BF00337259

Amit, D., and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* 7, 237–252. doi: 10.1093/cercor/7.3.237

Amit, D. J., and Tsodyks, M. V. (1991). Quantitative study of attractor neural network retrieving at low spike rates I: substrate – spikes, rates and neuronal gain. *Network* 2, 259. doi: 10.1088/0954-898X/2/3/003

Ben-Yishai, R., Bar-Or, R. L., Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3844–3848. doi: 10.1073/pnas.92.9.3844

Beurle, R. L. (1956). Properties of a mass of cells capable of regenerating pulses. *Philos. Trans. R. Soc. Lond.* B240, 55–94. doi: 10.1098/rstb.1956.0012

Boucsein, C., Tetzlaff, T., Meier, R., Aertsen, A., and Naundorf, B. (2009). Dynamical response properties of neocortical neuron ensembles: multiplicative versus additive noise. *J. Neurosci.* 29 1006–1010. doi: 10.1523/JNEUROSCI.3424-08.2009

Bressloff, P. C., and Coombes, S. (1998). Spike train dynamics underlying pattern formation in integrate-and-fire oscillator networks. *Phys. Rev. Lett.* 81, 2384–2387. doi: 10.1103/PhysRevLett.81.2384

Bressloff, P. C., and Coombes, S. (2000). Dynamics of strongly coupled spiking neurons. *Neural Comput.* 12, 91–129. doi: 10.1162/089976600300015907

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* 8, 183–208. doi: 10.1023/A:1008925309027

Brunel, N., and Hakim, V. (1999). Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural Comput.* 11, 1621–1671. doi: 10.1162/089976699300016179

Compte, A., Brunel, N., Goldman-Rakic, P. S., and Wang, X. J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* 10, 910–923. doi: 10.1093/cercor/10.9.910

Coombes, S. (2005). Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* 93, 91–108. doi: 10.1007/s00422-005-0574-y

Cox, D. R. (1962). *Renewal Theory. (Science Paperbacks).* London: Chapman and Hall.

Dayan, P., and Abbott, L. F. (2001) *Theoretical Neuroscience*. Cambridge, MA: MIT Press.

De la Rocha, J., Doiron, B., Shea-Brown, E., Kresimir, J., and Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature* 448 802–807. doi: 10.1038/nature06028

Ermentrout, B. (1998). Neural networks as spatio-temporal pattern-forming systems. *Rep. Prog. Phys.* 61, 353–430. doi: 10.1088/0034-4885/61/4/002

Ermentrout, G. B., and Cowan, J. D. (1979a). A mathematical theory of visual hallucination patterns. *Biol. Cybern.* 34, 137–150. doi: 10.1007/BF00336965

Ermentrout, G. B., and Cowan, J. D. (1979b). Temporal oscillations in neuronal nets. *J. Math. Biol.* 7, 265–280. doi: 10.1007/BF00275728

Ermentrout, B., and Cowan, J. (1980). Large scale spatially organized activity in neural nets. *SIAM J. Appl. Math.* 38, 1–21. doi: 10.1137/0138001

Fourcaud-Trocmé, N., and Brunel, N. (2005). Dynamics of the instantaneous firing rate in response to changes in input statistics. *J. Comput. Neurosci.* 18, 311–321. doi: 10.1007/s10827-005-0337-8

Gewaltig, M.-O., and Diesmann, M. (2007) NEST (NEural Simulation Tool). *Scholarpedia* 2, 1430.

Girko, V. (1984). Circular law. *Theory Probab. Appl.* 29, 694–706. doi: 10.1137/1129095

Grabow, C., Grosskinsky, S., and Timme, M. (2012). Small-world spectra in mean field theory. *Phys. Rev. Lett.* 106, 218701. doi: 10.1103/PhysRevLett.108.218701

Hanuschkin, A., Kunkel, S., Helias, M., Morrison, A., and Diesmann, M. (2010). A general and efficient method for incorporating precise spike times in globally time-driven simulations. *Front. Neuroinform.* 4:113. doi: 10.3389/fninf.2010.00113

Helias, M., Deger, M., Rotter, S., and Diesmann, M. (2010). Instantaneous non-linear processing by pulse-coupled threshold units. *PLoS Comput. Biol.* 6:e1000929.

Helias, M., Tetzlaff, T., and Diesmann, M. (2013). Echoes in correlated neural systems. *New J. Phys.* 15, 023002. doi: 10.1088/1367-2630/15/2/023002

Kriener, B., Helias, M., Aertsen, A., and Rotter, S. (2009). Correlations in spiking neuronal networks with distance dependent connections. *J. Comput. Neurosci.* 27, 177–200. doi: 10.1007/s10827-008-0135-1

Kriener, B., Tetzlaff, T., Aertsen, A., Diesmann, M., and Rotter, S. (2008). Correlations and population dynamics in cortical networks. *Neural Comput.* 20, 2185–2226. doi: 10.1162/neco.2008.02-07-474

Lindner, B. (2004). Interspike interval statistics of neurons driven by colored noise. *Phys. Rev. E* 69, 0229011–0229014. doi: 10.1103/PhysRevE.69.022901

Lindner, B., and Schimansky-Geier, L. (2001). Transmission of noise coded versus additive signals through a neuronal ensemble. *Phys. Rev. Lett.* 86, 2934–2937. doi: 10.1103/PhysRevLett.86.2934

Mainen, Z. F., and Sejnowski, T. J. (1996). Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature* 382, 363–366. doi: 10.1038/382363a0

Marti, D., and Rinzel, J. (2013). Dynamics of feature categorization. *Neural Comput.* 25, 1–45. doi: 10.1162/NECO_a_00383

Moreno, R., de la Rocha, J., Renart, A., and Parga, N. (2002). Response of spiking neurons to correlated inputs. *Phys. Rev. Lett.* 89, 288101. doi: 10.1103/PhysRevLett.89.288101

Moreno-Bote, R., Renart, A., and Parga, N. (2008). Theory of input spike auto- and cross-correlations and their effect on the response of spiking neurons. *Neural Comput.* 20, 1651–1705. doi: 10.1162/neco.2008.03-07-497

Murphy, B. K., and Miller, K. D. (2009). Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* 61, 635–648. doi: 10.1016/j.neuron.2009.02.005

Pernice, V., Staude, B., Cardanobile, S., and Rotter, S. (2012). Recurrent interactions in spiking networks with arbitrary topology. *Phys. Rev. E* 85, 031916. doi: 10.1103/PhysRevE.85.031916

Renart, A., Moreno-Bote, R., Wang, X. J., and Parga, N. (2007). Mean-driven and fluctuation-driven persistent activty in recurrent networks. *Neural Comput.* 19, 1–46. doi: 10.1162/neco.2007.19.1.1

Roxin, A., Brunel, N., and Hansel, D. (2005). The role of delays in shaping spatio-temporal dynamics of neuronal activity in large networks. *Phys. Rev. Lett.* 94, 238103. doi: 10.1103/PhysRevLett.94.238103

Salinas, E., and Sejnowski, T. J. (2000). Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. *J. Neurosci.* 20, 6193–6209.

Schüz, A., and Braitenberg, V. (1994). "Constraints to a random plan of cortical connectivity," in *Structural and Functional Organization of the Neocortex*, eds B. Albowitz, K. Albus, U. Kuhnt, H.-C. Nothdurft, and P. Wahle (Heidelberg: Springer-Verlag), 161–169.

Shriki, O., Hansel, D., and Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural Comput.* 15, 1809–1841. doi: 10.1162/08997660360675053

Siegert, A. J. (1951). On the first passage time probability problem. *Phys. Rev.* 81, 617–623. doi: 10.1103/PhysRev.81.617

Silberberg, G., Bethge, M., Markram, H., Pawelzik, K., and Tsodyks, M. (2004). Dynamics of population codes in ensembles of neocortical neurons. *J. Neurophysiol.* 91, 704–709. doi: 10.1152/jn.00415.2003

Tchumatchenko, T., and Wolf, F. (2011). Representation of dynamical stimuli in populations of threshold neurons. *PLoS Comput. Biol.* 7:e1002239. doi: 10.1371/journal.pcbi.1002239

Tetzlaff, T., Helias, M., Einevoll, G., and Diesmann, M. (2012). Decorrelation of neural-network activity by inhibitory feedback. *PLoS Comput. Biol.* 8: e1002596. doi: 10.1371/journal.pcbi.1002596

Usher, M., Stemmler, M., and Olami, Z. (1995). Dynamic pattern formation leads to 1/f noise in neural populations. *Phys. Rev. Lett.* 74, 326–329. doi: 10.1103/PhysRevLett.74.326

Vilela, R. D., and Lindner, B. (2009). Comparative study of different integrate-and-fire neurons: spontaneous activity, dynamical response, and stimulus-induced correlation. *Phys. Rev. E* 80, 031909. doi: 10.1103/PhysRevE.80.031909

Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature* 393, 440–442. doi: 10.1038/30918

Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24. doi: 10.1016/S0006-3495(72)86068-5

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 April 2013; accepted: 09 December 2013; published online: 07 January 2014.

Citation: Kriener B, Helias M, Rotter S, Diesmann M and Einevoll GT (2014) How pattern formation in ring networks of excitatory and inhibitory spiking neurons depends on the input current regime. Front. Comput. Neurosci. 7:187. doi: 10.3389/fncom.2013.00187

This article was submitted to the journal Frontiers in Computational Neuroscience.

Copyright © 2014 Kriener, Helias, Rotter, Diesmann and Einevoll. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A. APPENDIX

A1. MODEL AND SIMULATION DESCRIPTION

Table A1 | Model description, simulation paradigm and parameters.

A	**MODEL SUMMARY**
Populations	Three: excitatory, inhibitory, external input
Connectivity	Ring connectivity with excitatory and inhibitory neurons arranged on the same ring, such that each fifth neuron is inhibitory; same connection footprint for both populations
Neuron model	Leaky integrate-and-fire (LIF), fixed voltage threshold precise spike timing (Hanuschkin et al., 2010)
Synapse model	δ-current inputs (discontinuous voltage jumps)
Input	Independent Poisson spike trains

B	**POPULATIONS**	
Name	**Elements**	**Size**
E	LIF neuron	$N_E = 4N_I$
I	LIF neuron	N_I
E_x	excitatory Poisson input generator	one realization per neuron
I_x	inhibitory Poisson input generator (only for the fluctuation-driven case)	one realization per neuron

C	**CONNECTIVITY**		
Name	**Source**	**Target**	**Pattern**
EE	E	E	connect to all excitatory neurons within a given distance on the ring, weight J, delay d
IE	E	I	connect to all excitatory neurons within a given distance on the ring, weight J, delay d
EI	I	E	connect to all excitatory neurons within a given distance on the ring, weight $-gJ$, delay d
II	I	I	connect to all excitatory neurons within a given distance on the ring, weight $-gJ$, delay d
E_x	E_x	E ∪ I	independent Poisson spike trains, weight J_x, rate ν_{Ex}
I_x	I_x	E ∪ I	independent Poisson spike trains, weight $-gJ_x$, rate ν_{Ix}

(Continued)

Table A1 | Continued

D	NEURON AND SYNAPSE MODEL	
Name	LIF neuron	
Type	Leaky integrate-and-fire, δ-current input	
Subthreshold dynamics	$\tau_m \dot{V}(t) = -V(t) + RI(t)$	if $t > t^* + \tau_{ref}$
	$V(t) = V_{res}$	else
	$RI(t) = \sum_{j=1, i\neq j}^{N+N_{ext}} \tau_m W_{ij} \sum_k \delta(t - t_{i,k} - d) + RI_x(t)\,,$	$W_{ij} \in \{-gJ, 0, J\}$
Spiking	If $V(t-) < \theta \wedge V(t+) \geq \theta$	
	1. set $t^* = t$	
	2. emit spike with time-stamp t^*	

E	INPUT
Type	**Description**
Poisson generators	Rate ν_{Ex}, ν_{Ix}, each generator projects independent realizations to all neurons

A2. APPROXIMATION OF THE SIEGERT FORMULA BY AN AFFINE-LINEAR INPUT-OUTPUT RELATION

For strong mean drive $\mu[RI] \gg \theta$ we can approximate the input-output relation (10) by an affine linear relationship. In the following we assume the firing rate $0 < \nu \ll 1/\tau_{ref}$, which is the case for all relevant regimes, so we can neglect the effect of the refractory time τ_{ref}. We use the Ansatz

$$\nu(I) = -\left(\tau_m \mathrm{Log}\left[1 - \theta/(RI)\right]\right)^{-1} \doteq a + bRI \qquad \text{(A1)}$$
$$\Leftrightarrow -\tau_m \mathrm{Log}\left[1 - \theta/(RI)\right] \doteq (a + bRI)^{-1}$$

By taking the derivative with respect to RI on both sides of the last expression we obtain

$$-\frac{\tau_m}{1 - \theta/(RI)}\theta/(RI)^2 = -b\,(a + bRI)^{-2} \qquad \text{(A2)}$$
$$\frac{\theta\tau_m}{\theta RI - (RI)^2} = \frac{-1}{a^2/b + 2aRI + b(RI)^2}$$

Equating the terms in proportion to (RI) and in proportion to $(RI)^2$ in the denominator fixes the parameters $a = -\frac{1}{2\tau_m}$ and $b = \frac{1}{\theta\tau_m}$, leading to the approximation

$$\nu(RI) \doteq -\frac{1}{2\tau_m} + \frac{1}{\theta\tau_m}RI \qquad \text{(A3)}$$

Indeed, to show that for $\theta \gg RI$

$$-\frac{1}{\tau_m \mathrm{Log}\left[1 - \frac{\theta}{RI}\right]} \doteq -\frac{1}{2\tau_m} + \frac{1}{\theta\tau_m}RI\,. \qquad \text{(A4)}$$

we introduce the auxiliary variable $\theta/RI =: (1 - x)$ and compute the limit

$$\lim_{x\nearrow 1}\left[\frac{1}{\mathrm{Log}[x]} + \frac{1}{1-x}\right] = \lim_{x\nearrow 1}\left[\frac{1 - x + \mathrm{Log}[x]}{\mathrm{Log}[x](1-x)}\right]. \qquad \text{(A5)}$$

Both numerator and denominator of this expression tend to zero, thus we apply L'Hôpital's rule and obtain for (A5)

$$\lim_{x\nearrow 1}\left[\frac{1/x - 1}{1/x - 1 - \mathrm{Log}[x]}\right]. \qquad \text{(A6)}$$

Again, numerator and denominator tend to zero, thus we apply L'Hôpital's rule again and obtain for (A6)

$$\lim_{x\nearrow 1}\left[\frac{1/x^2}{(1+x)/x^2}\right] = \lim_{x\nearrow 1}\left[\frac{1}{1+x}\right] = \frac{1}{2}\,. \qquad \text{(A7)}$$

This yields (A4).

As a heuristic observation, we note moreover, that in the large-σ_o-limit , i.e., $\sigma_o \gg \mu_o$, the integrand of Equation (5) becomes basically unity, and the integral thus scales as $\nu_o^{-1} \sim \sqrt{\pi}\tau_m\theta/\sigma_o$. In particular we observe that the input-output-relation Equation (5) for $V_{res} = 0$ takes the following linear form in σ_o:

$$\nu_o(\sigma_o, \mu_o) = \frac{1}{\tau_m\sqrt{\pi}}\left(\frac{\sigma_o + \mu_o}{\theta} - \frac{1}{\pi}\right). \qquad \text{(A8)}$$

The input-output-relation of the LIF is thus also linear in σ_o for $\sigma_o \gg \mu_o$, see supplementary material section S4.

A3. EIGENSPECTRUM OF THE RING GRAPH WITH EQUIDISTANT INHIBITORY NEURONS

We consider graphs, in which each neuron i is connected to neurons $i - \kappa/2$ to $i + \kappa/2$, κ even, but not to itself. To keep analysis simple, we assume that each ℓ-th neuron, $N \bmod \ell = 0$ is inhibitory.

Define T_ℓ as the translation operator that shifts all neuron indices by ℓ such that for every vector v we have

$$T_\ell\, v = T_\ell(v_0, v_1, v_2, ..., v_{N-1})^\top$$
$$= (v_\ell, v_{\ell+1}, v_{\ell+2}, ..., v_{N-1}, v_0, ..., v_{\ell-1})^\top \qquad \text{(A9)}$$

i.e., for the node index n

$$T_\ell[n] = (n+\ell) \bmod N\,, \quad n \in \{0, ..., N-1\}\,. \tag{A10}$$

T_ℓ is a unitary operator and has eigenvalues φ_l and eigenvectors w_l, such that

$$T_\ell w_l = e^{i\,\alpha_l} w_l =: \varphi_l w_l\,, \tag{A11}$$

with

$$T_\ell^{N/\ell} w_l = w_l \;\Rightarrow\; \alpha_l = 2\pi l\ell/N\,,\; l \in \{0, ..., N/\ell - 1\}\,. \tag{A12}$$

Hence the N/ℓ ℓ-fold degenerated eigenvalues of T_ℓ are $\varphi_l = e^{i2\pi l\ell/N}$, $l \in \{0, ..., N/\ell - 1\}$. The eigenvectors w_l are discrete harmonics and fullfil

$$w_{l_{(i+\ell) \bmod N}} = \varphi_l\, w_{l_{i \bmod N}}\,. \tag{A13}$$

We define

$$\mathbb{C}^\ell \ni \tilde{w} = (w_0, w_1, w_2, w_3, ..., w_{\ell-1})^T \tag{A14}$$

and the map η that embeds the $\mathbb{C}^\ell$ as a subspace into the $\mathbb{C}^N$

$$\eta : \mathbb{C}^\ell \to \mathbb{C}^\ell \subset \mathbb{C}^N,\;\; \tilde{w} \mapsto \sum_{i=0}^{\ell-1} \tilde{w}_i \mathrm{e}_i =: \tilde{v}\,, \tag{A15}$$

where e_i, $i \in \{0, ..., N-1\}$ are the N canonical basis-vectors of the $\mathbb{R}^N$ and $\tilde{w}_i \in \mathbb{C}$ is the i-th vector component of $\tilde{w}$. We moreover define the operator family P_l, $l \in \{0, ..., N/\ell - 1\}$

$$P_l : \mathbb{C}^\ell \subset \mathbb{C}^N \to \mathbb{C}^N\,,\;\; \tilde{v} \mapsto \sum_{j=0}^{N/\ell-1} T_\ell^j\, e^{-i2\pi jl\ell/N}\, \tilde{v}\,. \tag{A16}$$

We see that $(P_l \circ \eta)$ indeed projects into the eigenspaces and is an isomorphism

$$P_l \circ \eta : \mathbb{C}^\ell \to \mathrm{Eig}(T_\ell, \varphi_l) \subset \mathbb{C}^N \tag{A17}$$

as can be readily checked:
We see that for each $l \in \{0, ..., N/\ell - 1\}$

$$\begin{aligned} T_\ell(P_l \circ \eta) \;&=\; T_\ell \sum_{j=0}^{N/\ell-1} T_\ell^j\, e^{-i2\pi jl\ell/N} = \sum_{j=0}^{N/\ell-1} T_\ell^{j+1}\, e^{-i2\pi jl\ell/N} \\ &\overset{(*)}{=} e^{i2\pi l\ell/N} \sum_{k=0}^{N/\ell-1} T_\ell^k\, e^{-i2\pi kl\ell/N} = \varphi_l(P_l \circ \eta), \end{aligned} \tag{A18}$$

where in the equation marked by (*) we performed an index shift and used $(e^{-i\alpha}T_\ell)^{N/\ell} = (e^{-i\alpha}T_\ell)^0$.
Linearity:

$$\begin{aligned} \forall_{v,\, w \in \mathbb{C}^\ell,\, a \in \mathbb{C}}\;\; (P_l \circ \eta)(a\,w) &= a\,(P_l \circ \eta)w \;\text{ and }\; (P_l \circ \eta)(w+v) \\ &= (P_l \circ \eta)w + (P_l \circ \eta)v \end{aligned} \tag{A19}$$

Injectivity:

$$\mathrm{Ker}\,(P_l \circ \eta) = \{0\}$$

Surjectivity:

$$\dim\,[(P_l \circ \eta)(\mathbb{C}^\ell)] = \dim\,[\mathrm{Eig}(T_\ell, \varphi_j)] = \ell\,.$$

Now, we can calculate the spectrum of W. Due to $[W, T_\ell] = 0$, i.e., W and T_ℓ commute, we can diagonalize both matrices, such that they have a common system of eigenvectors. Hence, it is sufficient to reduce the problem to the ℓ-dimensional vector-space embedded in $\mathbb{C}^N$, spanned by eigenvectors of $\mathrm{Eig}(T_\ell, e^{i\varphi_l})|_{\mathbb{C}^\ell}$, i.e., $\tilde{w}_l$, $l \in \{0, ..., N/\ell - 1\}$. To obtain the eigenspaces $\mathrm{Eig}(W,\lambda_{li})$ we only need to solve the diagonalization problem for the corresponding $(\ell \times \ell)$-matrices $\eta^{-1} \circ W \circ P_l \circ \eta = (P_l\, W)|_{\mathbb{C}^{\ell\times\ell}}$:

$$(\eta^{-1} \circ W \circ P_l \circ \eta)\tilde{w} = \lambda\tilde{w} \tag{A20}$$

Once we got the ℓ eigenvalues and eigenvectors $\{\lambda_{l,i}, \tilde{w}_{l,i}\}$ of W in the respective sub-spaces, the operator $P_l \circ \eta$ generates the corresponding N-dimensional eigenvector w_l.

A4. EIGENSYSTEM OF THE REDUCED RING SYSTEM

We now consider a coarse-grained network model where connectivity is defined between $C = N/\ell$ cells of ℓ neurons each as depicted in **Figure 6A**. Each neuron is connected to all other $\ell - 1$ neurons within the cell it resides in, as well as to all neurons in the neighboring $2K = pC$ cells, yielding block-wise connectivity matrices as indicated in **Figure 7B**. We introduce a new global neuron labeling such that

$$n = c\,\ell + i_c, \quad c = \left\lfloor \frac{n}{\ell} \right\rfloor, \quad \text{and} \quad i_c = n \bmod \ell \tag{A21}$$

Symmetries:
The system in invariant to shifts about elementary cells, i.e., for the cell index n we have again invariance to the operation

$$T_\ell[n] = (n+\ell) \bmod N\,. \tag{A22}$$

Moreover, there are two mirror symmetries R and S within each cell (cf. **Figures 6B,C**), such that

$$Rw_{[n]} = Rw_{[(c\ell+i_c) \bmod N]} = \begin{cases} w_{[(c\ell+(1-i_c)) \bmod N]} & \text{if } i_c \le 1 \\ \\ w_{[n]} & \text{if } i_c > 1 \end{cases} \tag{A23}$$

$$Sw_{[n]} = Sw_{[(c\ell+i_c) \bmod N]} = w_{[(c\ell+\ell-1-i_c) \bmod N]} \tag{A24}$$

R and S are operators, such that for the eigenvalues r_j and s_j, and for the eigenvectors v_j and w_j we have $R^2 v_j = r_j^2 v_j = v_j$ and $S^2 w_j = s_j^2 w_j = w_j$. Hence, $r_j, s_j \in \{-1, 1\}$. We have three commuting operators T_ℓ, R, S, i.e.,

$$[T_\ell, R] = [T_\ell, S] = [S, R] = 0\,.$$

So, we can find a common basis of eigenvectors to span the complete system. For each wavenumber $\Lambda_i = \alpha_i N/2\pi\ell$ one eigenvector is obviously given by

$$v^{\mathrm{I}}_{\alpha_i} := \sum_{c=0}^{N/\ell-1} (e^{-i\alpha_i} T_\ell)^c v_{(\ell-1)/2}\, e_{(\ell-1)/2} \tag{A25}$$

where $e_{(\ell-1)/2}$ is the $(\ell-1)/2$-th canonical basisvector of the $\mathbb{R}^N$ and $(\ell-1)/2$ is the index of the first inhibitory neuron on the ring enumerated as in **Figure 6**, since this dimension does not couple to either R or S. To get the residual $N - N_{\mathrm{I}}$ eigenvectors, we choose

$$v^{\mathrm{E}}_{r,s,\alpha_i} := \sum_{c=0}^{C-1} (e^{-i\alpha_i} T_\ell)^c (1+sS)(1+rR) v_0\, e_0 \tag{A26}$$

with first canonical basis vector e_0.

Indeed, we see again that for each $c \in \{0, ..., N/\ell - 1\}$

$$\begin{aligned} T_\ell (e^{-i\alpha} T_\ell)^c &= e^{-i\alpha c}\, T_\ell^{c+1} = e^{i\alpha} e^{-i\alpha(c+1)}\, T_\ell^{c+1} \\ &= e^{i\alpha} (e^{-i\alpha} T_\ell)^{c+1} \\ \text{and} \quad (e^{-i\alpha} T_\ell)^C &= 1\,, \end{aligned} \tag{A27}$$

and

$$\begin{aligned} \mathcal{X}(1 + x\mathcal{X}) &= \mathcal{X} + x\mathcal{X}^2 = \mathcal{X} + x1 = x(x\mathcal{X} + 1) \\ &= x(1 + x\mathcal{X}) \end{aligned} \tag{A28}$$

with $x \in \{r, s\}$ and $\mathcal{X} \in \{R, S\}$, so $v_{r,s,\alpha}$ are eigenvectors with eigenvalues r, s, α. Evaluation of the eigenvectors for the two possible eigenvalues of R, S each yields

$$\begin{aligned} v^{\mathrm{E}}_{1,1,\alpha} &= \sum_{c=0}^{C-1} (e^{-i\alpha} T_\ell)^c \big((v_0, v_0, 0, v_0, v_0, 0, \ldots, 0)^\top\big) \\ v^{\mathrm{E}}_{-1,1,\alpha} &= \sum_{c=0}^{C-1} (e^{-i\alpha} T_\ell)^c \big((v_0, -v_0, 0, -v_0, v_0, 0, \ldots, 0)^\top\big) \\ v^{\mathrm{E}}_{1,-1,\alpha} &= \sum_{c=0}^{C-1} (e^{-i\alpha} T_\ell)^c \big((v_0, v_0, 0, -v_0, -v_0, 0, \ldots, 0)^\top\big) \\ v^{\mathrm{E}}_{-1,-1,\alpha} &= \sum_{c=0}^{C-1} (e^{-i\alpha} T_\ell)^c \big((v_0, -v_0, 0, v_0, -v_0, 0, \ldots, 0)^\top\big) \end{aligned} \tag{A29}$$

A5. MEAN-FIELD APPROXIMATION OF SMALL-WORLD NETWORKS OF EXCITATORY AND INHIBITORY NEURONS

Analogous to the mean-field approach developed in (Grabow et al., 2012) for unweighted networks we will now compute the eigensystem of the mean-field small world network derived from the ring network with equidistant inhibition by random rewiring of edges (Kriener et al., 2009). In this case instead of assigning explicit weights $W_{ij} \in \{0, J, -gJ\}$ to individual entries of the coupling matrix, all possible connections carry their expected weight as a function of the rewiring probability p_r such that all weights within the original ring neighborhood are scaled by $p_1(p_r)$, and all weights that were 0 in the original ring are set to $W_{ij} = p_2(p_r)\, J$ if j is excitatory, and $W_{ij} = -p_2(p_r)\, gJ$ if j is inhibitory, where (cf. Kriener et al., 2009; Grabow et al., 2012)

$$\begin{aligned} p_1(p_r) &= (1 - p_r) + \frac{p_r^2 \kappa}{(N - 1 - \kappa(1 - p_r))} \\ p_2(p_r) &= \frac{p_r}{(N - 1 - \kappa(1 - p_r))} \end{aligned} \tag{A30}$$

The resulting average coupling matrix is again invariant with respect to T_ℓ and the eigensystem can be determined exactly as described in Appendix A3.