Linking structure and activity in nonlinear spiking networks

Gabriel Koch Ocker [1], Krešimir Josić [2,3], Eric Shea-Brown [4,1,5], Michael A. Buice* [1,4],

1 Allen Institute for Brain Science, Seattle, WA, USA
2 Department of Mathematics and Department of Biology and Biochemistry, University of Houston, Houston, TX, USA
3 Department of BioSciences, Rice University, Houston, TX, USA
4 Department of Applied Mathematics, University of Washington, Seattle, WA, USA
5 Department of Physiology and Biophysics, and UW Institute of Neuroengineering, University of Washington, Seattle, WA, USA

* michaelbu@alleninstitute.org

Abstract

Recent experimental advances are producing an avalanche of data on both neural connectivity and neural activity. To take full advantage of these two emerging datasets we need a framework that links them, revealing how collective neural activity arises from the structure of neural connectivity and intrinsic neural dynamics. This problem of *structure-driven activity* has drawn major interest in computational neuroscience. Existing methods for relating activity and architecture in spiking networks rely on linearizing activity around a central operating point and thus fail to capture the nonlinear responses of individual neurons that are the hallmark of neural information processing. Here, we overcome this limitation and present a new relationship between connectivity and activity in networks of nonlinear spiking neurons. We explicitly show how recurrent network structure produces pairwise and higher-order correlated activity, and how nonlinearities impact the networks' spiking activity. Finally, we demonstrate how correlations due to recurrent connectivity impact the fidelity with which populations encode simple signals and how nonlinear dynamics impose a new effect of correlations on coding: to translate response distributions in addition to stretching them. Our findings open new avenues to investigating how neural nonlinearities—including those expressed across multiple cell types—combine with connectivity to shape population activity and function.

Introduction

A fundamental goal in computational neuroscience is to understand how network connectivity and intrinsic neuronal dynamics relate to collective neural activity, and in turn drive neural computation. Experimental advances are vastly expanding both the scale and the resolution with which we can measure both neural connectivity and neural activity. Simultaneously, a wealth of new data suggests a possible partitioning of neurons into cell types with both distinct dynamical properties and distinct patterns of connectivity. What is needed is a way to link these three types of data: How is it that patterns of connectivity are translated into patterns of activity through neuronal dynamics?

Any model of neural activity should also capture the often-strong variability in spike trains across time or experimental trials. This variablity in spiking is often coordinated, or *correlated*, across cells which has a variety of implications. First, correlations play an essential role in plasticity of network structure [1–4]. Theories that describe spiking correlations allow for a self-consistent description of the coevolution of recurrent network structure and activity [5, 6]. Second, correlations between synaptic inputs control their effect on postsynaptic neurons: inputs that arrive simultaneously can produce stronger responses than those arriving separately. This has been referred to as "synergy" or "synchronous gain" in early work [7], and the magnitude of this synergy has been measured in the LGN by Usrey, Reppas & Reid [8] and cortex by Bruno & Sakmann [9] (but see [10]). Indeed, the level of correlation in an upstream population has been shown to act as a gain knob for firing rates downstream [11]. Finally, correlated fluctuations in activity can impact the fidelity with which populations can encode information [12, 13]. Importantly, the coding impact depends on a subtle interplay of how signals impact firing rates in a neural population and of how noise correlations occur across the population [14–18]. An accurate description of how network connectivity determines the individual and joint activity of neural populations is thus important for the understanding of neural activity, plasticity and coding.

Many studies of collective activity in spiking systems can be traced to the early work of Hawkes on self- or mutually-exciting point processes [19, 20]. The Hawkes model is also closely related to the linear response theory that can be used to describe correlations in integrate-and-fire networks [21, 22]. Here, each neuron and synapse is linearized around a central "operating point", and modes of collective activity are computed around that point [23–25]. Including a nonlinear transfer of inputs to rates in the Hawkes model gives a generalized linear model, which has been applied with considerable success to multineuron spike train data [26].

While analyses based on computing modes of collecting activity based on linearized dynamics have led to significant insights, they also impose a limitation. While shifts of the operating point can modulate the linearized dynamics of biophysical models [27], this approach cannot capture the impact of nonlinear neural dynamics at the operating point.

Here, we present a systematic method for computing correlations of any order for *nonlinear* networks of excitatory and inhibitory neurons. Nonlinear input-rate transfer couples higher-order spike train statistics to lower-order ones in a manner that depends on the order of the nonlinearity. In its simplest form, this coupling shows how pairwise–correlated inputs modulate firing rates. This generalizes the effects of pairwise correlations on neural gain in single-neurons [7, 8, 11] and feedforward circuits [28–31] to networks with high levels of recurrence and feedback.

We begin with simple models and progress to nonlinearly interacting networks of spiking neurons. Our method is *diagrammatic*, in the sense that the interplay of network connectivity and neural dynamics in determining network statistics is expressed and understood via a systematic series of graphical terms. We use this diagrammatic expansion to make and explain three main scientific points. First, we show how neural dynamics lead to spike correlations modulating firing rates in a recurrent, nonlinear network. Second, we illustrate an additional role of the prominent 'heavy-tailed" feature of neural connectivity where some neurons have many more connections than others, and some connections are much stronger than others. We show how this feature interacts with nonlinearities to control network activity. And third, we show how spike time correlations generated by recurrent connectivity can impact a simple population code in networks of linear and nonlinear neurons.

Results

Diagrammatic expansion for spike train statistics

We will show that any coupled point process model, even one with nonlinearities or negative interactions, has an associated expansion for all spike train cumulants organized by the strength of coupling of higher statistical moments with lower ones (e.g. the influence of the two-point correlation on the mean). The full model we describe is one where each neuron generates a spike train which is conditionally renewal with intensity:

$$r_i(t) = \phi\left(\sum_j \mathbf{g}_{ij} * \frac{dN_j}{dt}(t) + \lambda_i(t)\right) \tag{1}$$

Here $\mathbf{g}_{ij}(t)$ is a matrix of interaction filters, $*$ denotes convolution and $\lambda_i(t)$ is the baseline point process drive to neuron i. We will take the spike trains to be conditionally Poisson.

Current methods for the analysis of single- and joint spiking statistics rely on linear response techniques: using a self-consistent mean field theory to compute mean firing rates, and then linearizing the spiking around those rates to determine the stability and correlations. We begin with a simple example highlighting the need to account for nonlinear effects. We take an excitatory-inhibitory network where neurons have threshold-quadratic transfer functions ϕ. In this example we took network connectivity to be totally random (Erdös-Rènyi), with $p_{EE} = 0.2$ and $p_{EI} = p_{IE} = p_{II} = 0.5$. For simplicity, the magnitude of all connections of a given type ($E-E$, etc.) was taken to be the same. Furthermore, the time course of all network interactions is governed by the same filter $g(t) = \frac{t}{\tau^2}\exp(-t/\tau)$ (with $\tau = 10$ ms), so that $\mathbf{g}_{ij}(t) = \mathbf{W}_{ij}g(t)$. We set the strength of interactions such that the net inhibitory input weight on to a neuron was, on average, twice that of the net excitatory input weight so that for sufficiently strong interactions, the network was in an inhibitory-stabilized regime [32].

We examined the magnitude and stability of firing rates as we increase the strength of synaptic coupling. We used mean-field theory to predict the firing rates, and predicted their linear stability by the standard measure of the spectral radius of the stability matrix $\mathbf{\Psi}_{ij} = \phi'_i\mathbf{g}_{ij}$. ($\phi'_i$ denotes the first derivative of neuron i's transfer function with respect to the input.) As the strength of interactions increases, the mean field prediction for the firing rates loses accuracy (Fig. 1 A). This occurs well before the mean field theory crosses the stability boundary $|\mathbf{\Psi}| = 1$ (Fig. 1B). Examining simulations as the weights are increased reveals an even more fundamental failure of the theory: before the synaptic weights are strong enough for the mean field theory to be unstable, the simulations reveal divergent firing rates (Fig. 1C; the raster stops when the instantaneous rates diverge).

Rather than restricting theoretical analysis to regimes of extremely low coupling strength or linear models, we here develop a framework for activity cumulants that can apply to models with nonlinear input-spike transfer. This will allow us both to properly predict both the spiking cumulants and the location of the rate instability in the nonlinear network above. Thus, we develop a framework for activity cumulants that can apply to models with nonlinear input-spike transfer for strongly coupled networks. The mean field and linear response predictions for spiking cumulants correspond to the lowest order terms of this expansion, and are good whenever that lowest-order truncation is valid.

We will build this framework systematically: We begin with statistics of the drive $\lambda_i(t)$, then consider a filtered point process, $g * \lambda(t)$. In these simple models we will introduce the method and terminology which we will use to understand the more complicated models. We continue by considering the linearly self-exciting Hawkes process, taking a single neuron so $\mathbf{g} = g$ and $\phi(x) = x$, before proceeding to arbitrary nonlinearities ϕ. Finally, we introduce an arbitrary network structure $\mathbf{g}$. This model is intimately related to the popular generalized linear models (GLMs) for spiking activity, where the nonlinearity ϕ is commonly taken to be exponential, refractory dynamics can be embedded in the diagonal elements of $\mathbf{g}$, and λ corresponds to the filtered stimulus [26]. The common use of GLMs it to fit them to recorded spike trains, and then ask about the structure of the inferred functional connectivity $\mathbf{g}$. In contrast, we interpret $\mathbf{g}$ as reflecting the structural connectivity and synaptic and membrane dynamics of a specified model network and aim to compute statistics of its neurons' spike trains. The derivation given here will be heuristic. A more rigorous derivation of the expansion is given in Methods: Path integral representation.

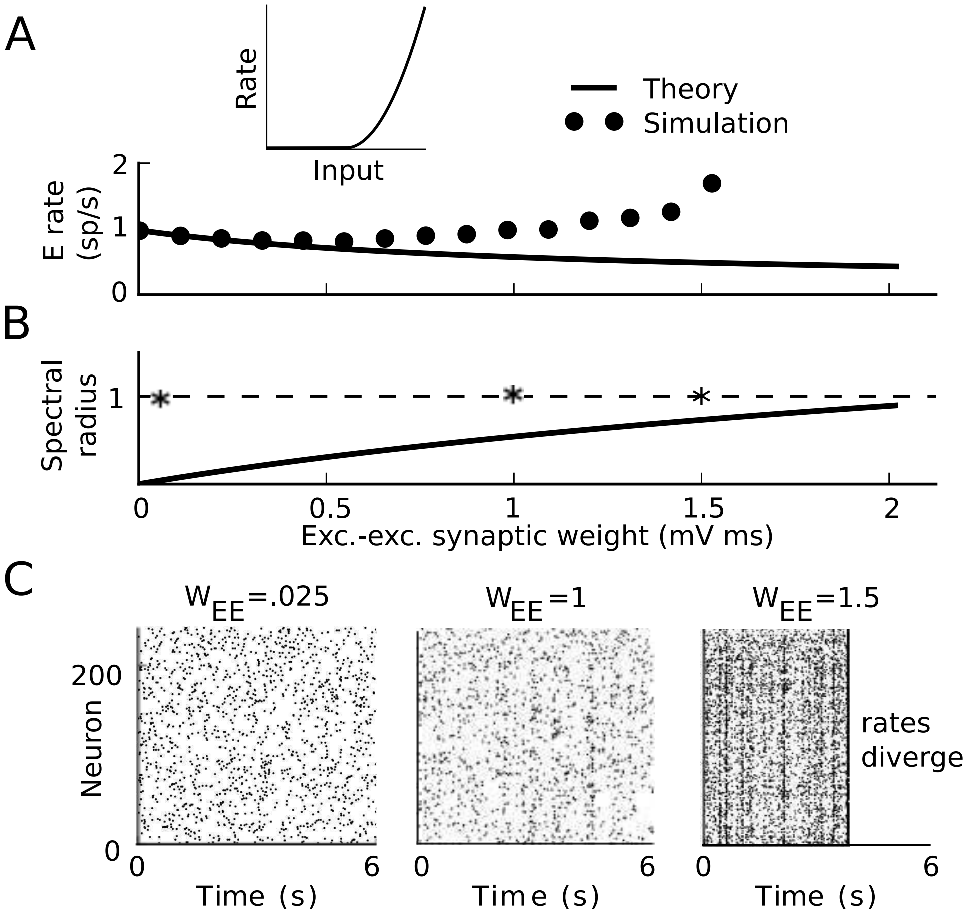


Fig 1. Dynamics approaching the firing-rate instability in threshold-quadratic networks. A) Average firing rate of the excitatory neurons as synaptic weights are scaled. Solid lines: prediction of mean field theory. Dots: result of simulation. Inset: threshold-quadratic transfer function. B) Spectral radius of the stability matrix of mean field theory as synaptic weights are scaled. Stars on indicate the weight values for the simulations below. C) Example realizations of activity for three different interaction strengths. As synapses become stronger, correlated activity becomes apparent. When synapses are strong enough the activity becomes unstable, even though the mean field theory is stable. Left: $(W_{EE}, W_{EI}, W_{IE}, W_{II})$ = (.025, -.1, .01, -.1) mV ms. C). Middle: $(W_{EE}, W_{EI}, W_{IE}, W_{II}) = (1, -4, .4, -4)$ mV ms. Right: $(W_{EE}, W_{EI}, W_{IE}, W_{II}) = (1.5, -6, .6, -6)$ mV ms.

Introduction to the general framework: Poisson process

An inhomogeneous Poisson process generates counts within a window dt independently with a Poisson distribution at rate $\lambda(t)$. A spike train produced by this process is

$$\frac{dN(t)}{dt} = \sum_k \delta(t - t_k) \tag{2}$$

where t_k is the kth spike time, and $N(t)$ is the spike count. The mean and autocovariance for this process are given by the familiar formulas:

$$\left\langle \frac{dN(t)}{dt} \right\rangle = \lambda(t) \tag{3}$$

$$\left\langle \frac{dN(t)}{dt} \frac{dN(t')}{dt'} \right\rangle_c = \lambda(t)\delta(t-t'), \tag{4}$$

where angular brackets denote the expectation across realizations and the subscript c denotes a cumulant, not the moment (i.e. we have subtracted all terms which factor into products of lower moments) [33]. The delta function arises because the process is independent at each time step, so that there is no correlation between events from one time t and any other time t'. In fact, because the events are generated independently at each time point, all of the cumulants of this process can be written as

$$\left\langle \prod_i \frac{dN(t_i)}{dt} \right\rangle_c = \int dt \lambda(t) \prod_i \delta(t_i - t) \tag{5}$$

where integrating out one of the delta functions puts the second cumulant in the above form. We can interpret this equation as describing a source of events appearing at rate $\lambda(t)$ at time t that propagate to times t_i. In this case, because the events are generated independently at each t, the events only propagate to the same time t. For a general point process, cumulants measured at the collection of times $\{t_i\}$ could be affected by events occurring at any past time, so that we would have to account for how events propagate across time.

The expansion for cumulants we will develop has a natural construction in terms of graphs, wherein components of the graph represent factors in each term. A set of defined rules dictate how each term in the expansion is formed from those graphs. While this graphical represenation is not necessary to understand the inhomogeneous Poisson process, we describe it in detail in this simple case to develop intuition and introduce terminology. We use cumulants in this construction because they provide the fundamental building blocks of moments; any n-th order moment can be constructed from up to n-th order cumulants). This also simplifies the developed expansion.

To begin, the nth cumulant has n arguments $\{t_i, i = 1, \ldots, n\}$, one for each factor of $\frac{dN}{dt}$ in Eq. (5). We represent each of these by an open white vertex in the graph labeled with the time point, t_i, and we represent the source term $\lambda(t)$ with a gray vertex. The white vertices are called "external" vertices whereas the gray vertex is called "internal".

The internal gray vertex represents the intensity of the underlying stochastic process generating events with rate $\lambda(t)$. The white external vertices represent the spike trains whose statistics we measure at times $\{t_i\}$. For each delta function, $\delta(t - t_i)$, in Eq. 5, we place an edge drawn as a dotted line from the gray vertex to the white vertex, i.e. from the event-generating process to the spike train's measurement times (Fig. 2). More generally, events generated by the source propagate through the graph to affect the external vertices.

In order to construct the term associated with each diagram, we multiply the factors corresponding to edges (delta functions linking t and t_i) or the internal vertex ($\lambda(t)$), and then integrate over the time t associated with the internal vertex. This links the generation of events by $\lambda(t)$ to their joint measurement at times $\{t_i\}$ through the propagator (here $\delta(t - t_i)$). For the diagrams shown in Fig. 2, these rules reproduce the cumulant terms in Eq. 5. Note that these graphs are directed, since we only consider causal systems where measured cumulants are only influenced by past events.

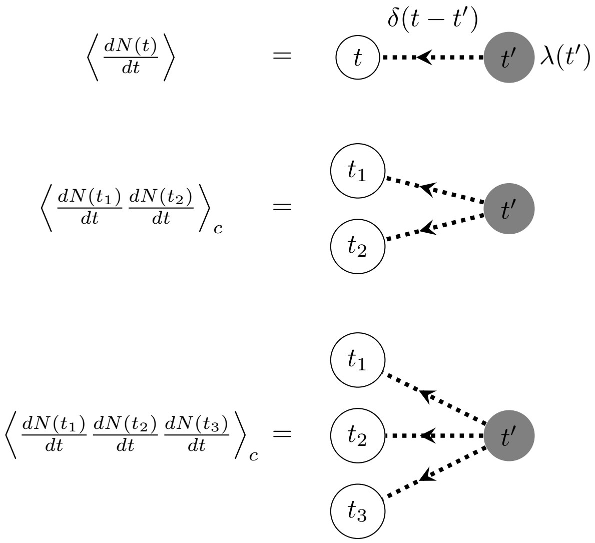

$\left\langle \frac{dN(t)}{dt} \right\rangle$ =
$\delta(t-t')$
t
t'
$\lambda(t')$
$\left\langle \frac{dN(t_1)}{dt}\frac{dN(t_2)}{dt} \right\rangle_c$ =
t_1
t_2
t'
$\left\langle \frac{dN(t_1)}{dt}\frac{dN(t_2)}{dt}\frac{dN(t_3)}{dt} \right\rangle_c$ =
t_1
t_2
t_3
t'

Fig 2. Feynman diagrams for the first three cumulants of the inhomogeneous Poisson process. Each dotted edge corresponds to a delta-function connecting the time indices of its two nodes. White nodes denote the measurement times, while gray nodes denote the times at which spikes are generated. The cumulants are constructed by convolving the term corresponding to the gray node with the product of all outgoing edges' terms (Eq. (5)).

In general, a given moment will be the sum of terms associated with many different graphs. For example the second moment is given by $\left\langle \frac{dN(t_1)}{dt}\frac{dN(t_2)}{dt} \right\rangle = \left\langle \frac{dN(t_1)}{dt}\frac{dN(t_2)}{dt} \right\rangle_c + \left\langle \frac{dN(t_1)}{dt} \right\rangle \left\langle \frac{dN(t_2)}{dt} \right\rangle$. Each term on the right hand side will have a corresponding graph. Moreover, the graph for the second term will include two disconnected components, one for each factor of the mean rate, which appears as in Figure 2. The graphs for the cumulant will always be described by connected graphs.

Filtered Poisson process

We proceed to a simple model of synaptic input: a presynaptic Poisson spike train, with count $N(t)$ and intensity $\lambda(t)$, drives postsynaptic potentials with shape $g(t)$:

$$\nu(t) = \epsilon g * \frac{dN}{dt}(t), \tag{6}$$

where $*$ denotes convolution: $g * f(t) = \int_{t_0}^{\infty} dt' g(t-t') f(t')$ (with the integral starting at the initial time for the realization). We assume g is normalized, $\int_{-\infty}^{\infty} g(t)dt = 1$, so that ϵ gives the magnitude of the filtering.

The cumulants of the postsynaptic potential $\nu(t)$ can be calculated directly. In general, they are given by:

$$\left\langle \prod_i \nu(t_i) \right\rangle_c = \int dt \lambda(t) \prod_i \epsilon g * \delta(t_i - t) \tag{7}$$

where the input spikes are generated at times t, arrive at times given by the delta functions and influence the cumulant measured at $\{t_i\}$ through g. Eq. 7 is the same as that for the inhomogeneous

Poisson process but with factors of $g*$. This provides a simple interpretation of Eq. 7: cumulants of the filtered Poisson process are given by taking the cumulants of the underlying Poisson process and examining how they can be filtered through the system at hand.

Similarly to the case for the Poisson process, we can represent the cumulants graphically. We again represent each measurement time on the left hand side of Eq. (7) by an external vertex (Fig. 3a). The convolution of δ and g in Eq. 7 corresponds to an internal time point which we integrate over (denoted by primes). We also represent these internal time points with white vertices that carry a factor of ϵ, the magnitude of the filter. We represent the source term $\lambda(t)$ with a gray vertex. All vertices that are not "external" are called internal. Every internal vertex also carries its own time index, t'.

The internal gray vertex again represents the intensity of the underlying stochastic process, $\lambda(t)$. The white external vertices represent the processes whose statistics we measure at times $\{t_i\}$. For each delta function, $\delta(t'-t)$, in Eq. 7, we place an edge drawn as a dotted line from the gray vertex to the white vertex, *i.e.* from the event-generating process to the arrival time of the event t'. In this example an event "arrives" as soon as it is generated. A wavy edge corresponds to the filter, g, and represents the effect of a spike arriving at time t' on the output process measured at time t_i (Fig. 3b). Events generated by the source thus propagate through the graph to affect the observed, external vertices.

In order to construct the expression associated with each diagram, we again multiply the factors corresponding to each edge in the diagram (e.g. $\delta(t'-t)$ or $g(t_i-t')$) or internal vertex (ϵ or $\lambda(t)$), and then integrate over the times associated with the internal vertices. Note that integration over the internal times t', t'', etc, results in the convolutions $\epsilon g * \delta(t_i - t)$. Integration over the time t associated with the source term corresponds to the outermost integral in Eq. (7) This links the generation of events by $\lambda(t)$ to their joint measurement at times $\{t_i\}$ through their arrival times (via $\delta(t-t')$), and temporal filtering ($g(t_i-t')$). For the diagrams shown in Fig. 3, these rules reproduce the cumulant terms in Eq. 7. Note that the graphs are directed, as for the expansion we describe the "propagator" term will be causal.

We can simplify the cumulants of this process (and the corresponding diagrammatic representations) by considering the propagator of $\nu(t)$ (also known as the linear response or impulse response). The propagator measures the change in $\langle \nu(t) \rangle$ in response to the addition of one input spike in $N(t)$. We can compute it by taking a functional derivative with respect to the input intensity $\lambda(t)$:

$$\begin{aligned}
\Delta(t,t') &= \frac{\delta}{\delta\lambda(t')} \langle \nu(t) \rangle \\
&= \frac{\delta}{\delta\lambda(t')} \left(\epsilon \int_{t_0}^{\infty} dt''\, g(t-t'') \left\langle \frac{dN(t'')}{dt} \right\rangle \right) \\
&= \epsilon \int_{t_0}^{\infty} dt''\, g(t-t'') \frac{\delta\lambda(t'')}{\delta\lambda(t')} \\
&= \epsilon g * \delta(t-t')
\end{aligned} \tag{8}$$

Since the dynamics are linear, this is also equivalent to the change of the expected rate with the addition of one spike to the input, *i.e.* taking $\lambda(t) \leftarrow \lambda(t) + \delta(t'-t)$ and $\langle \nu(t) \rangle_c \leftarrow \langle \nu(t) \rangle_c + \Delta * \delta(t)$ (or equivalently the Green's function of the expected rate). This allows us to rewrite the cumulants in terms of the input rate and the propagator:

$$\left\langle \prod_i \nu(t_i) \right\rangle_c = \int dt \lambda(t) \prod_i \Delta(t_i, t) \tag{9}$$

which can be represented graphically by introducing a solid, directed edge for $\Delta(t,t')$ (Fig. 3c). The propagator will be a central feature of the expansion for cumulants in more complicated models involving connections among neurons.

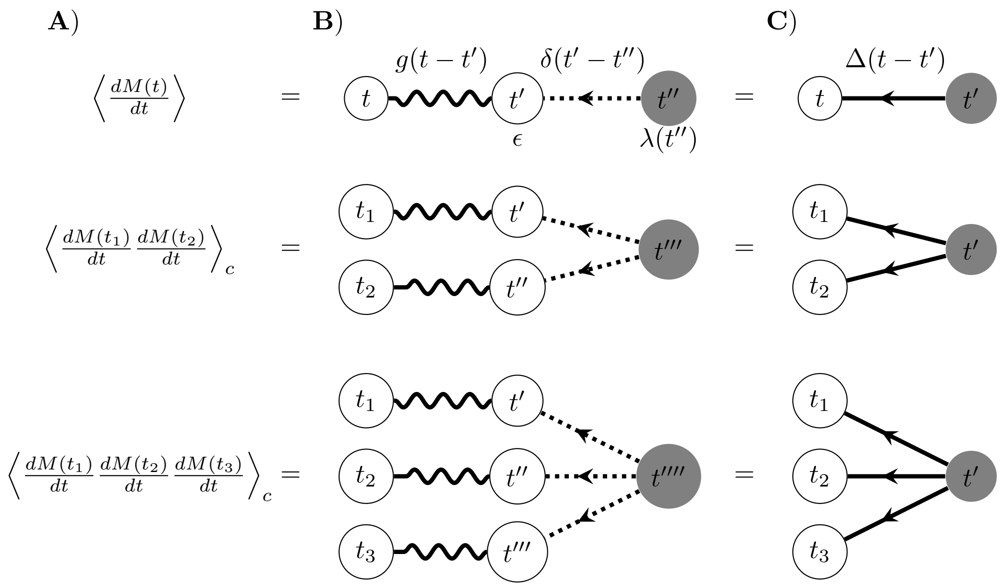

A)
B)
C)
$\left\langle \frac{dM(t)}{dt} \right\rangle$
=
$g(t-t')$
$\delta(t'-t'')$
t
t'
t''
ϵ
$\lambda(t'')$
=
$\Delta(t-t')$
t
t'
$\left\langle \frac{dM(t_1)}{dt} \frac{dM(t_2)}{dt} \right\rangle_c$
=
t_1
t'
t_2
t''
t'''
=
t_1
t_2
t'
$\left\langle \frac{dM(t_1)}{dt} \frac{dM(t_2)}{dt} \frac{dM(t_3)}{dt} \right\rangle_c$
=
t_1
t'
t_2
t''
t_3
t'''
t''''
=
t_1
t_2
t_3
t'

Fig 3. Feynman diagrams for the first three cumulants of the filtered inhomogeneous Poisson process. A) Cumulant corresponding to the graph. B) Diagrammatic expressions using the filter and the underlying Poisson process. Each dotted edge corresponds to a delta-function connecting the time indices of its two nodes. Each wavy edge corresponds to the filter g connecting the time indices of its two nodes. C) Diagrammatic expressions using the propagator. In all graphs, external white nodes (leaves of the graph) denote measurement times. Gray nodes denote the times at which spikes are generated in the input spike train. Internal white nodes (with time indices t') denote the times at which input spikes arrive at the postsynaptic neuron. The cumulants are constructed by convolving the term corresponding to the gray node with the product of all outgoing edges' terms (Eq. (7)).

Impact of self-excitation on activity statistics of any order: linearly self-exciting process

In order to generalize the graphical representation of Poisson cumulants, we begin with a linearly self-exciting process as considered by Hawkes [19]. Let the rate be a linear function of the instantaneous event rate (that is to say the firing rate conditioned on a particular realization of the event history)

$$r(t) = \epsilon g * \frac{dN}{dt}(t) + \lambda(t). \tag{10}$$

We assume that $g(\tau)$ and $\lambda(t)$ are such that $r(t) > 0$, and $\int_{-\infty}^{\infty} d\tau g(\tau) = 1$. If $\epsilon < 1$, then an event will generate less than one event on average, and the rate will not diverge. The history dependence of the firing rate will now enter into our calculations. We can compute the expected rate using the self-consistency equation:

$$\bar{r}(t) \equiv \left\langle \frac{dN}{dt}(t) \right\rangle = \epsilon g * \left\langle \frac{dN(t)}{dt} \right\rangle + \lambda(t) = \epsilon g * \bar{r}(t) + \lambda(t) \tag{11}$$

We provide an alternate derivation of this result that will prove useful below: We construct a perturbative expansion of the mean firing rate and show how this expansion can be re-summed to yield the full rate of the self-exciting process. This procedure can also be applied to obtain cumulants of arbitrary order for this process.

We will begin with a recursive formulation of the self-exciting process. In contrast to the filtered Poisson process of the previous section, here the process with count N generates events, which then influence its own rate, dN/dt. Each event can thus generate others in turn. In the case of a linear filter, g, the following approach is equivalent to the Poisson cluster expansion [34–36] and similar to the construction of previous linear response theories for spike train covariances [22]. Define the nth order self-exciting process, $N_n(t)$, to be the inhomogeneous Poisson process given by:

$$\frac{dN_n(t)}{dt} = \frac{dN_0(t)}{dt} + \frac{dM_{n-1}(t)}{dt}, \tag{12}$$

where $N_0(t)$ and $M_n(t)$ are inhomogeneous Poisson processes with rates $\lambda(t)$ and $\nu_n(t)$, respectively, where

$$\nu_n(t) = \epsilon g * \frac{dN_n}{dt}(t). \tag{13}$$

We can generate the processes recursively: To generate $N_n(t)$, we use a realization of $N_{n-1}(t)$ to compute the rate ν_{n-1} and generate a realization of $M_{n-1}(t)$. These are added to events generated from the original inhomogeneous Poisson process with rate $\lambda(t)$ to produce $N_n(t)$. We can use this recursive procedure to develop an expansion for the cumulants of the process at a given order in ϵ (thus a given order in the self-convolution of g).

Let us compute the value of $\left\langle \frac{dN(t)}{dt} \right\rangle$ in powers of ϵ using our recursive approach. The zeroth order solution, $\left\langle \frac{dN_0(t)}{dt} \right\rangle$, is the rate of the inhomogeneous Poisson process $\lambda(t)$. At order n, we compute $\left\langle \frac{dN_n(t)}{dt} \right\rangle$ using the $(n-1)$st order solution in the right hand side of Eq. (12). At first order, using the

Poisson solution for $\langle \frac{dN_0(t)}{dt} \rangle$ we get

$$\left\langle \frac{dN_1(t)}{dt} \right\rangle = \left\langle \frac{dN_0}{dt}(t) \right\rangle + \left\langle \frac{dM_0(t)}{dt} \right\rangle \tag{14}$$

$$= \lambda(t) + \epsilon g * \left\langle \frac{dN_0(t)}{dt} \right\rangle \tag{15}$$

$$= \lambda(t) + \epsilon g * \lambda(t) \tag{16}$$

$$= \lambda(t) + \epsilon \int_{t_0}^{\infty} g(t-t')\lambda(t')dt' \tag{17}$$

At second order we similarly arrive at

$$\left\langle \frac{dN_2(t)}{dt} \right\rangle = \left\langle \frac{dN_0(t)}{dt} \right\rangle + \left\langle \frac{dM_1(t)}{dt} \right\rangle \tag{18}$$

$$= \lambda(t) + \epsilon g * \left\langle \frac{dN_1(t)}{dt} \right\rangle \tag{19}$$

$$= \lambda(t) + \epsilon \int_{t_0}^{\infty} dt' g(t-t')\lambda(t') + \epsilon^2 \int_{-\infty}^{\infty} dt' g(t-t') \int_{t_0}^{\infty} dt'' g(t'-t'')\lambda(t'') \tag{20}$$

At higher orders we would obtain further terms with additional convolutions with g.

It will be useful to write these expansions in another way, which will allow their form to generalize to non-linear processes: we will construct the cumulants from the baseline rate and the propagator. We can always replace

$$\lambda(t) = \int_{t_0}^{\infty} dt' \delta(t-t')\lambda(t') \tag{21}$$

resulting in

$$\left\langle \frac{dN_1(t)}{dt} \right\rangle = \int_{t_0}^{\infty} dt' \delta(t-t')\lambda(t') + \epsilon \int_{t_0}^{\infty} dt' g(t-t') \int_{t_0}^{\infty} dt'' \delta(t'-t'')\lambda(t'') \tag{22}$$

Fig. 4a shows the graphical representation of this expansion. As before, the order of the moment is given by the number of external vertices and each external vertex carries a measurement time t_i. We have three types of internal vertices: two open white vertices that carry factors of ϵ (one type has one wavy incoming and one wavy outgoing line - the other has one incoming dotted line and one wavy outgoing line) and one gray vertex (that has one outgoing dotted line). As before, each gray internal vertex corresponds to the source term, and thus represents the factor $\lambda(t)$. The white internal vertices, and their edges represent how the events generated by the source are propagated through the filter $g(t)$. Each white vertex corresponds to a possible past event time, t'. To construct the cumulant corresponding to a diagram we integrate over all these possible internal times, weighting each by their influence on the current spiking rate. These weights are given by the filters, g, represented by the wavy edges. The graphical representation of $\left\langle \frac{dN_1(t)}{dt} \right\rangle$ (using the delta function as in Eq. 22) is shown in Figure 4a.

We can compute the firing rate of the self-exciting process $\bar{r}(t)$ as the limit of the nth order self-exciting processes, continuing the process outlined for Eq. 12:

$$\bar{r}(t) = \sum_{n=0}^{\infty} \epsilon^n g^{(n)} * \delta * \lambda(t), \tag{23}$$

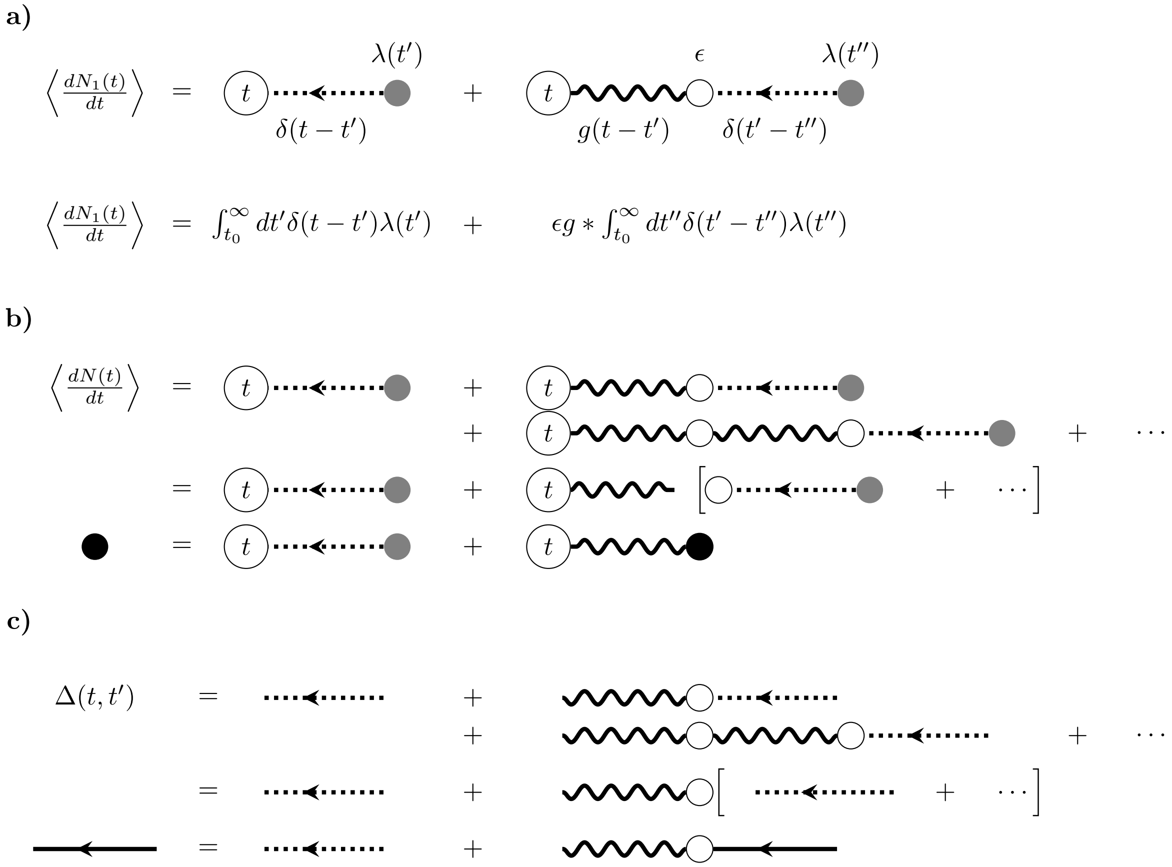

a)
$\left\langle \frac{dN_1(t)}{dt} \right\rangle$ =
$\lambda(t')$
t
$\delta(t-t')$
+
ϵ
$\lambda(t'')$
$g(t-t')$
$\delta(t'-t'')$
$\left\langle \frac{dN_1(t)}{dt} \right\rangle = \int_{t_0}^{\infty} dt' \delta(t-t')\lambda(t') + \epsilon g * \int_{t_0}^{\infty} dt'' \delta(t'-t'')\lambda(t'')$
b)
$\left\langle \frac{dN(t)}{dt} \right\rangle$ =
+ ⋯
c)
$\Delta(t,t')$ =
+ ⋯

Fig 4. Diagrammatic expansion for the mean firing rate and linear response of the self-exciting process. A) First-order approximation of the firing rate. B) Diagrams corresponding to the re-summing of the expansion of the mean field rate (Eq. 24), which is represented by the black dot. C) Diagrams corresponding to the re-summing calculation of the propagator (Eq. 27), which is represented by the solid edge. In all diagrams, time indices associated with internal vertices have been suppressed.

where $g^{(n)}$ is the n-fold convolution of g with itself and $g^{(0)}(t) = \delta(t)$. Indeed, we can see that this expression for $\bar{r}(t)$ yields the same recursive self-consistency condition as above:

$$\begin{aligned}
\bar{r}(t) &= \epsilon^0 g^{(0)} * \delta * \lambda(t) + \sum_{n=1}^{\infty} \epsilon^n g^{(n)} * \delta * \lambda(t) \\
&= \lambda(t) + \epsilon g * \sum_{n=1}^{\infty} \epsilon^{n-1} g^{(n-1)} * \delta * \lambda(t) \\
&= \lambda(t) + \epsilon g * \bar{r}(t). \qquad (24)
\end{aligned}$$

We can also represent this recursive relation graphically as in Fig. 4b. The infinite sum defined by Eq. (12) has a specific graphical representation: Notice that the leftmost vertex and wavy line in the right-hand side of Figure 4b (top) can be detached and factored, with the remaining series of diagrams corresponding exactly to those of the mean. This series of subgraphs on the right hand side sums to $\left\langle \frac{dN(t)}{dt} \right\rangle$, leading to the recursion relation in Eq. 24 (Fig. 4b). The graphical representation is equivalent to the recursion relation.

The propagator, $\Delta(t,t')$, measures the fluctuation in the expected rate (around the mean-field value) in response to the addition of one spike at time t' to the drive $\lambda(t)$. Setting $\lambda(t) \leftarrow \lambda(t) + \delta(t-t')$ and $\bar{r}(t) \leftarrow \bar{r}(t) + \Delta * \delta(t)$ in Eq. 11 gives:

$$\begin{aligned}\bar{r}(t) + \Delta * \delta(t) &= \epsilon g * \big(\bar{r}(t) + \Delta * \delta(t)\big) + \lambda(t) + \delta(t-t') \\ \Delta(t,t'') &= \epsilon g * \Delta(t-t') + \delta(t-t')\end{aligned} \tag{25}$$

As with the expected rate $\bar{r}(t)$, we can examine the propagators of the n-th order self-exciting processes. For the first-order process $N_1(t)$,

$$\Delta_1(t,t') = \delta(t-t') + \epsilon g * \delta(t-t') \tag{26}$$

The first term is the propagator of the inhomogeneous Poisson process. The second term of Δ_1 is the propagator of the filtered Poisson process, Eq. 8. This equation can be represented by the same type of graphs as for the expected rate (Fig. 4c top), but stand for functions between two time points: the initial time t' of the perturbation, and the later time t, at which we are computing the rate of the process. We don't represent these initial and final points as vertices, because the propagator is a function that connects two vertices. However, we still integrate over the times corresponding to the internal vertices since the propagator accounts for the total effect of a perturbation of the source on the observed activity.

In general, the propagator for the nth-order self-exciting process can be computed by taking a functional derivate of the rate with respect to the input rate λ:

$$\begin{aligned}\Delta_n(t,t') &= \frac{\delta}{\delta\lambda(t')}\left(\lambda(t) + \epsilon g * \bar{r}_{n-1}(t)\right) \\ &= \delta(t-t') + \epsilon g * \frac{\delta}{\delta\lambda(t')}\bar{r}_{n-1}(t) \\ &= \delta(t-t') + \epsilon g * \sum_{k=0}^{n-1} \epsilon^k g^{(k)} * \Delta_k(t,t')\end{aligned} \tag{27}$$

This recursion relation can be expressed graphically just as for the mean rate (Fig. 4c, top). Factoring out $\epsilon g*$ corresponds to popping off an edge and vertex from the series (Fig. 4c, middle). Taking the limit $n \to \infty$ in Eq. 27 yields the self-consistency condition for the full propagator $\Delta(t,t')$ given by Eq. 25, and indicated by the solid black line in Fig. 4c (bottom).

These diagrammatic expansions may seem cumbersome for so simple a model. Even for the self-exciting Hawkes process, however, they allow the fast calculation of any order of spike train cumulant. Let us begin with the second cumulant of the instantaneous firing rate. Again we will construct an expansion in ϵ, i.e. powers of $g(\tau)$. To zeroth order, this is the inhomogeneous Poisson solution. To first order in ϵ we have

$$\begin{aligned}\left\langle \frac{dN_1(t)}{dt}\frac{dN_1(t')}{dt}\right\rangle_c &= \left\langle \left(\frac{dN_0(t)}{dt}+\frac{dM_0(t)}{dt}\right)\left(\frac{dN_0(t')}{dt}+\frac{dM_0(t')}{dt}\right)\right\rangle_c \\ &= \left\langle \frac{dN_0(t)}{dt}\frac{dN_0(t')}{dt}\right\rangle_c + \left\langle \frac{dN_0(t)}{dt}\frac{dM_0(t')}{dt}\right\rangle_c \\ &+ \left\langle \frac{dM_0(t)}{dt}\frac{dN_0(t')}{dt}\right\rangle_c + \left\langle \frac{dM_0(t)}{dt}\frac{dM_0(t')}{dt}\right\rangle_c \\ &= \int_{t_0}^{\infty} ds\delta(t-s)\delta(t'-s)\lambda(s) + \epsilon\int_{t_0}^{\infty} ds\delta(t-s)g*\delta(t'-s)\lambda(s) \\ &+ \epsilon\int_{t_0}^{\infty} ds\delta(t'-s)g*\delta(t-s)\lambda(s) \\ &+ \epsilon\int_{t_0}^{\infty} ds\delta(t-s)\delta(t'-s)g*\int_{t_0}^{\infty} ds'\delta(s-s')\lambda(s')\end{aligned} \tag{28}$$

The first term on the second line is the second cumulant of the inhomogenous Poisson process. The other terms arise from the dependency of the processes $M_0(t)$ and $N_0(t)$: The process $M_0(t)$ is conditionally dependent on the rate of $N_0(t)$; this gives rise to the second and third terms on the right-hand side. The dependencies of the stochastic process $M_0(t)$ at two points in time are captured by the fourth term. This decomposition relies on the distributivity of the expectation operator.

We can construct diagrams for each of these terms using the same rules as before, with the addition of two new internal vertices (Fig. 5a). These new vertices are distinguished by their edges. The first has two outgoing dotted lines representing the zeroth-order propagator $\delta(t-t')$, as in the second cumulant of the inhomogeneous Poisson process. It represents events that are generated by the drive $\lambda(t)$ and propagate jointly to the two measurement time points, The second new vertex has the same two outgoing lines and one incoming wavy line for the filter $g(t,t')$ - it represents the fourth term on the right hand side of Eq. (28). This vertex carries a factor of ϵ and represents the filtering of past events that then propagate to the two measurement time points.

Continuing the computation of the second cumulant to any order in ϵ will result in higher order terms of the linear response and expected rate being added to the corresponding legs of the graph. At a given order n, one leg of each diagram will be associated with a particular term in the expansion, to order n, of the expected rate or the linear response. The second cumulant of dN_2/dt would thus add diagrams with two filter edges to the diagrams of Fig. 5a, either both on the same leg of the graph or distributed amongst the graph's three legs.

As with the filtered Poisson process, we can simplify this sum of four diagrams for the second cumulant of the first-order self-exciting process. Examining subgraphs of each term on the right-hand side of Fig. 5A reveals a connection to the linear response and mean rate of the first-order self-exciting processes. On each leg emanating from the internal branching vertex, the four terms sum to the product of two copies of the linear responses of the first-order self-exciting process $N_1(t)$ (compare subgraphs on the top branch of the diagrams in Fig. 5a with Fig. 4a). Similarly, the sum of the legs coming in to the branching vertex is the firing rate of $N_1(t)$ (compare to Fig. 4b). So, we will group the terms on the legs of the graph into contributions to the linear response and the mean (Fig. 5b middle).

When we add the diagrams of up to order n together, we can separately re-sum each of these expansions because of the distributivity of the expectation. So, we can replace the entire series *to all orders in* ϵ with simpler diagrams using the full representations for the linear response and expected rate

(Fig. 5b). This can be proved inductively, or by rigorously deriving the Feynman rules from the cumulant generating functional (Methods). This yields the following result for the second cumulant, which corresponds to the final graph at the bottom far right of Fig. 5b:

$$\left\langle \frac{dN(t)}{dt}\frac{dN(t')}{dt}\right\rangle_c = \int_{t_0}^{\infty} \Delta(t-s)\Delta(t'-s)\bar{r}(s)ds \tag{29}$$

This is the complete analytic result for the second cumulant of the self-exciting process for fluctuations around the mean field solution $\bar{r}(t)$ [19]. It can be represented by the single term on the right-hand side of Eq. 29 and the corresponding single diagram (Fig. 5b, right). Compare this with the filtered Poisson process, which has a diagram of the same topology but with different constituent factors (Fig. 3C, middle row). The Feynman diagrams capture the form of the re-summed perturbative expansions for the cumulants, while the definitions of the vertices and edges capture the model-specific rate, $\bar{r}(t)$, and propagator, $\Delta(t,t')$.

One might think that the higher cumulants are generated as simply by replacing each leg of the filtered inhomogeneous Poisson process with the correct propagator, along with the rate $\bar{r}(t)$. This means that the general cumulant term would be given by:

$$\left\langle \prod_i \frac{dN(t_i)}{dt}\right\rangle_c = \int_{t_0}^{\infty} dt\bar{r}(t)\prod_i \Delta(t-t_i) \tag{30}$$

This would be incorrect, as many important terms arising from the self-interaction would be lost.

The reason this naive generalization fails is that it neglects the higher-order responses to perturbations in the event rate. For example, the second cumulant responds to perturbations in the rate; this quadratic response impacts the third cumulant. We can see this in the third cumulant of the first-order self-exciting process:

$$\begin{aligned}\left\langle \frac{dN_1(t)}{dt}\frac{dN_1(t')}{dt}\frac{dN_1(t'')}{dt}\right\rangle_c &= \left\langle \left(\frac{dN_0(t)}{dt}+\frac{dM_0(t)}{dt}\right)\left(\frac{dN_0(t')}{dt}+\frac{dM_0(t')}{dt}\right)\left(\frac{dN_0(t'')}{dt}+\frac{dM_0(t'')}{dt}\right)\right\rangle_c \\ &= \int_{t_0}^{\infty} ds\delta(t-s)\delta(t'-s)\delta(t''-s)\lambda(s) \\ &+\epsilon\int_{t_0}^{\infty} ds\delta(t-s)\delta(t'-s)g*\delta(t''-s)\lambda(s)+(t\leftrightarrow t'\leftrightarrow t'') \\ &+\epsilon\int_{t_0}^{\infty} ds\delta(t-s)\delta(t'-s)\delta(t''-s)g*\int_{t_0}^{\infty} ds'\delta(s-s')\lambda(s') \\ &+\epsilon\int_{t_0}^{\infty} ds'\delta(t''-s')\int_{t_0}^{\infty} ds\delta(t-s)\delta(t'-s)g*\delta(s-s')\lambda(s')+(t\leftrightarrow t'\leftrightarrow t'').\end{aligned} \tag{31}$$

The first term is the third cumulant of the inhomogeneous Poisson process. The second and third are generalizations of the terms found in the second cumulant (we have used $(t\leftrightarrow t'\leftrightarrow t'')$ to denote "all other permutations of t,t',t''"). These terms are part of the naive expression in (30). The last term is the novel one that arises due to the "quadratic response". It appears when we compute

$$\left\langle \frac{dN_0(t)}{dt}\frac{dM_0(t')}{dt}\frac{dM_0(t'')}{dt}\right\rangle_c = \int_{t_0}^{\infty} ds\delta(t'-s)\delta(t''-s)\left\langle \frac{dN_0(t)}{dt}g*\frac{dN_0}{dt}(s)\right\rangle_c \tag{32}$$

We have to take into account that the process $\frac{dN_0(t)}{dt}$ is correlated with the rate of the process $\frac{dM_0(t)}{dt}$ (since one is a linear function of the other!). This produces a "cascade" effect that results in the

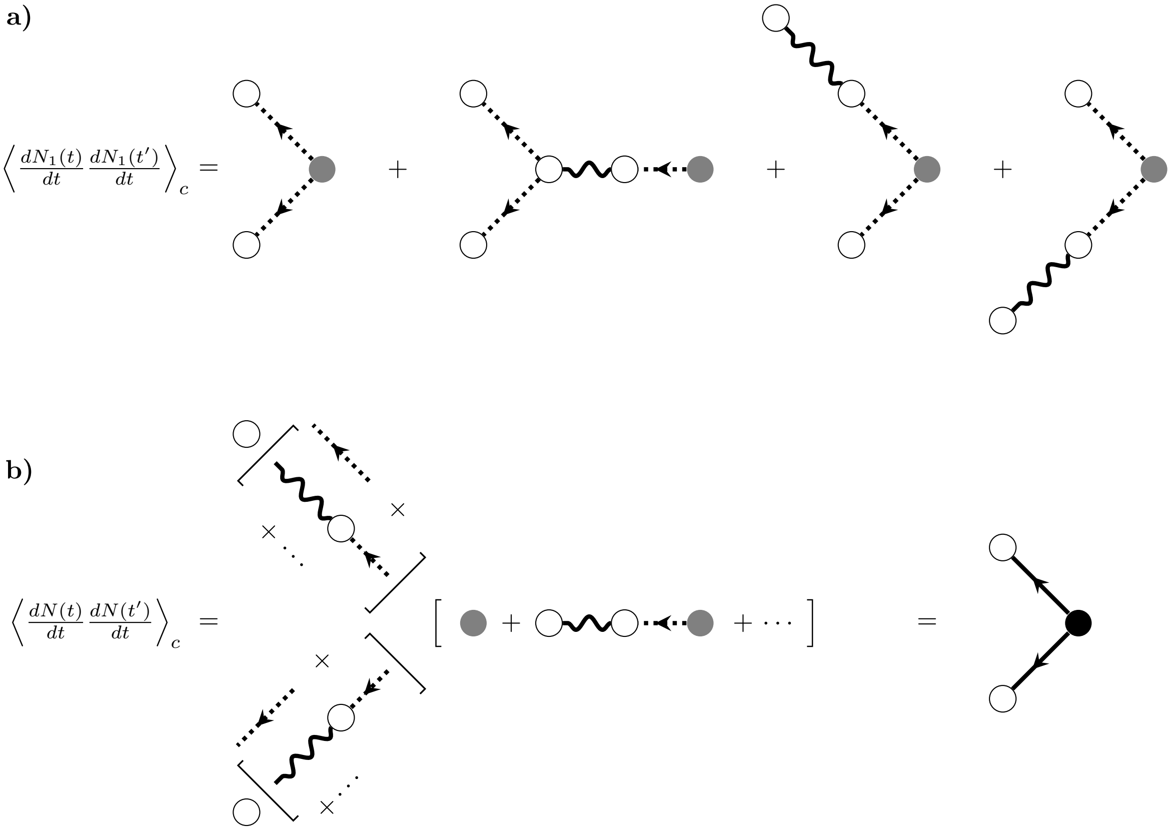

a)
$\left\langle \frac{dN_1(t)}{dt} \frac{dN_1(t')}{dt} \right\rangle_c =$
+
+
+
b)
$\left\langle \frac{dN(t)}{dt} \frac{dN(t')}{dt} \right\rangle_c =$
×
×
…
×
×
…
[+ + ⋯]
=

Fig 5. Diagrammatic expansion for the second cumulant for the self-exciting process. A) First-order approximation of the second cumulant. B) Re-summing to obtain the full second cumulant. Compare the expansions within the square brackets adjacent to external vertices to the expansion of the propagator, Fig. 4c, and compare the the expansion of the source term to that of the mean field rate, Fig.4b.

quadratic response. For the first-order process, only one step in the cascade is allowed. By introducing branching internal vertices similar to those in Fig. 5, we can express these somewhat unwieldy terms with diagrams. These are shown in Figure 6. The cascade of one source spike producing three spikes in the first-order process is represented by the second diagram of Fig. 6a and the cascade of one source spike producing two spikes, one of which then produces another two spikes in the first-order process, is represented by the last diagram of Fig. 6a.

As before, continuing to higher orders in the recursively self-exciting process would add diagrams with additional filter edges along the legs of the graphs in Fig. 6a, corresponding to additional steps in the potential cascades of induced spikes. For example, the third cumulant of the second-order process, $\left\langle \frac{dN_2(t)}{dt} \frac{dN_2(t')}{dt} \frac{dN_2(t'')}{dt} \right\rangle_c$, would add diagrams with two filter edges to those of Fig. 6a, with those two filter edges appearing either sequentially on the same leg of the graph or distributed amongst the legs of the graph. We can then use the same ideas that allowed us to resum the graphs representing the second cumulant. As before, we identify the expansions of the mean-field rate, $\bar{r}$, and the linear response, Δ, along individual legs of the graph and use the multilinearity of cumulants to resum those expansions to give the diagrams at the bottom of Fig. 6.

Considering the resummed graphs, we have the following result for the third cumulant:

$$\begin{aligned}\left\langle \frac{dN_1(t)}{dt} \frac{dN_1(t')}{dt} \frac{dN_1(t'')}{dt} \right\rangle_c &= \int_{t_0}^{\infty} ds \Delta(t-s)\Delta(t'-s)\Delta(t''-s)\bar{r}(s) \\ &+ \int_{t_0}^{\infty} ds' \Delta(t''-s') \int_{t_0}^{\infty} ds \Delta(t-s)\Delta(t'-s) g * \Delta(s-s')\bar{r}(s') \\ &+ (t \leftrightarrow t' \leftrightarrow t'')\end{aligned} \tag{33}$$

The types of diagram developed for up to the third cumulant encompass all the features that occur in the diagrammatic computations of cumulants of linearly self-exciting processes. The general rules for diagrammatically computing cumulants of this process are given in Fig. 7; they are derived in general in Methods: Path integral representation. The graphs generated with this algorithm correspond to the re-summed diagrams we computed above.

For the nth cumulant, $\left\langle \prod_i \frac{dN(t_i)}{dt} \right\rangle_c$, begin with n white external vertices labelled t_i for each i. Construct all fully connected, directed graphs with the vertex and edge elements shown in Figure 7. For each such fully, connected directed graph constructed with the component vertices and edges, the associated mathematical term is constructed by taking the product of each associated factor, then integrating over the time points of internal vertices. The nth cumulant is the sum of these terms. This produces cumulants of up to third order, as recently shown by Jovanović, Hertz & Rotter [36], as well as cumulants of any order. As we show next, this procedure can also be generalized to calculate cumulants in the presence of a nonlinearity, including both thresholds enforcing positive activity (as commonly disregarded in studies of the Hawkes process) and any nonlinear input-rate transfer function.

Nonlinearities impose bidirectional coupling between different orders of activity: nonlinearly self-exciting process

Now we include a nonlinearity in the firing rate, so that the process produces events dN/dt with a rate given by

$$r(t) = \phi\left(g * \frac{dN(t)}{dt} + \lambda(t)\right) \tag{34}$$

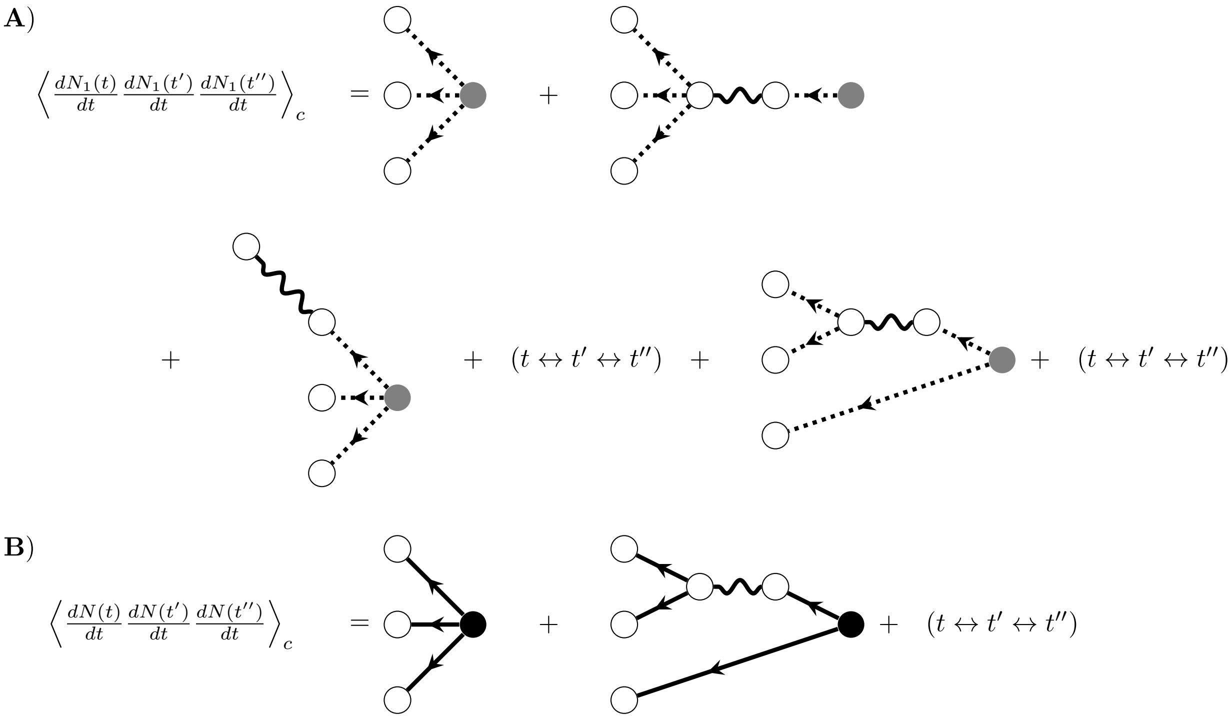

A)
$\left\langle \frac{dN_1(t)}{dt} \frac{dN_1(t')}{dt} \frac{dN_1(t'')}{dt} \right\rangle_c$
=
+
+
+
$(t \leftrightarrow t' \leftrightarrow t'')$
+
+
$(t \leftrightarrow t' \leftrightarrow t'')$
B)
$\left\langle \frac{dN(t)}{dt} \frac{dN(t')}{dt} \frac{dN(t'')}{dt} \right\rangle_c$
=
+
+
$(t \leftrightarrow t' \leftrightarrow t'')$

Fig 6. Diagrams corresponding to third order cumulants. A) Diagrams corresponding to the third cumulant of the first-order self-exciting process. B) Diagrams corresponding to the third cumulant of the self-exciting process, after resumming the perturbative expansion. Nodes and edges correspond to the same terms as in Fig. 5.

Feynman Rules for the Self-Exciting Process

For the nth cumulant, $\left\langle \prod_i \frac{dN(t_i)}{dt} \right\rangle_c$:

I. For each i, introduce a vertex labelled t_i. These are *external* vertices.

III. Construct all directed, fully connected graphs such that each vertex from I. has a single incoming propagator (Δ) edge, using the available vertices and edges below.

Filter edges can only impact internal vertices. Each internal vertex has an associated time variable

IV. To construct the cumulant: for each graph, multiply the associated vertex or edge factors together and integrate over the times t for all internal vertices. Add the terms so obtained for each graph together.

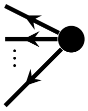
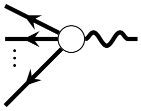
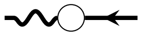
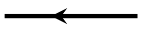
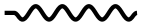
Vertex or Edge	Factor
	$\bar{r}(t)$
	1
	1
	$\Delta(t, t')$
	$g(t, t')$

Fig 7. Feynman rules for the self exciting process. These rules provide an algorithm for computing the expansion of the cumulants around the mean field solution $\bar{r}(t)$. The dots between the legs of the first two vertices indicate that there are such vertices with any number of outgoing legs greater than or equal to two.

We begin by considering the mean-field solution $\bar{r}$ which, if it exists, is self-consistently given by $\bar{r}(t) = \phi\left(g * \bar{r}(t) + \lambda(t)\right)$. Thus, as always the mean-field solution is given by neglecting second and higher-order cumulants of the spiking process. Next, we consider the propagator, which as above is the linear response of the rate around the mean field, given by taking the linearization of Eq. 34 around the mean-field solution $\bar{r}(t)$ and differentiating it with respect to the drive $\lambda(t)$. This propagator obeys:

$$\Delta(t,t') = \phi^1(g * \bar{r}(t) + \lambda(t))\left(g * \Delta(t,t') + \delta(t-t')\right) \tag{35}$$

where ϕ^1 is the first derivative of ϕ with respect to the input, evaluated at $g * \bar{r} + \lambda$. We will first develop a recursive formulation of the mean-field rate and propagator, which will be required for calculating cumulants of the full process. First, we Taylor expand ϕ about 0. For simplicity, consider a quadratic ϕ so that:

$$r(t) = \epsilon_1 \left(g * \frac{dN}{dt}(t) + \lambda(t)\right) + \epsilon_2 \left(g * \frac{dN}{dt}(t) + \lambda(t)\right)^2 \tag{36}$$

where ϵ_m is the mth Taylor coefficient, evaluated at 0. We now develop the point process dN/dt recursively at each order of the nonlinearity:

$$\frac{dN_{m,n}(t)}{dt} = \frac{dN_{0,0}(t)}{dt} + \frac{dM_{m-1,n}(t)}{dt} + \frac{dP_{m,n-1}(t)}{dt} \tag{37}$$

where $M_{m,n}$ is an inhomogeneous Poisson process with rate $\epsilon_1 g * \frac{dN_{m,n}}{dt}(t)$ and $P_{m,n}$ is an inhomogeneous Poisson process with rate $\epsilon_2 \left(g * \frac{dN_{m,n}(t)}{dt}\right)^2$ and $N_{0,0}(t)$ is an inhomogeneous Poisson process with rate $\lambda(t)$. To generate a set of events in N at order m in the linear term of ϕ and order n in the quadratic term of ϕ, we take events at each previous order, $(m-1,n)$ and $(m,n-1)$ and use those to develop $M_{m-1,n}(t)$ and $P_{m,n-1}(t)$. These, together with $N_{0,0}(t)$, give $\frac{dN_{m,n}}{dt}(t)$. In contrast to the linear self-exciting process, the quadratic process here is recursively defined on a lattice.

Similar to the case of the linearly self-exciting process, we can use this recursive definition to develop an expansion for the mean-field rate and propagator in powers of ϵ_1 and ϵ_2. When we calculate higher-order cumulants, we will identify the expansions of the mean-field firing rate and propagator which will allow us to use them to simplify the resulting diagrams. The mean-field rate to finite order in m, n is once again given by neglecting second and higher-order cumulants of $N_{n,m}$ which allows us to take an expectation inside the quadratic term of Eq. (37). Taking the expectation of both sides of this equation in the mean field approach then yields:

$$\bar{r}_{m,n}(t) = \lambda(t) + \epsilon_1 g * \bar{r}_{m-1,n}(t) + \epsilon_2 \left(g * \bar{r}_{m,n-1}(t)\right)^2 \tag{38}$$

For example,

$$\bar{r}_{1,1} = \lambda(t) + \epsilon g * \bar{r}_{0,1} + \epsilon_2 (g * \bar{r}_{1,0})^2 \tag{39}$$

where

$$\bar{r}_{1,0}(t) = \lambda(t) + \epsilon_1 g * \lambda(t) \tag{40}$$

$$\bar{r}_{0,1}(t) = \lambda(t) + \epsilon_2 \left(g * \lambda(t)\right)^2 . \tag{41}$$

Similarly, the propagator (for the dynamics of the recursive process, linearized around zero) is, to finite order in m, n:

$$\Delta_{m,n}(t,t') = \delta(t-t') + \epsilon_1 g * \Delta_{m-1,n}(t,t') + 2\epsilon_2 \left(g * \bar{r}_{m,n-1}(t)\right) g * \Delta_{m,n-1}(t,t') \tag{42}$$

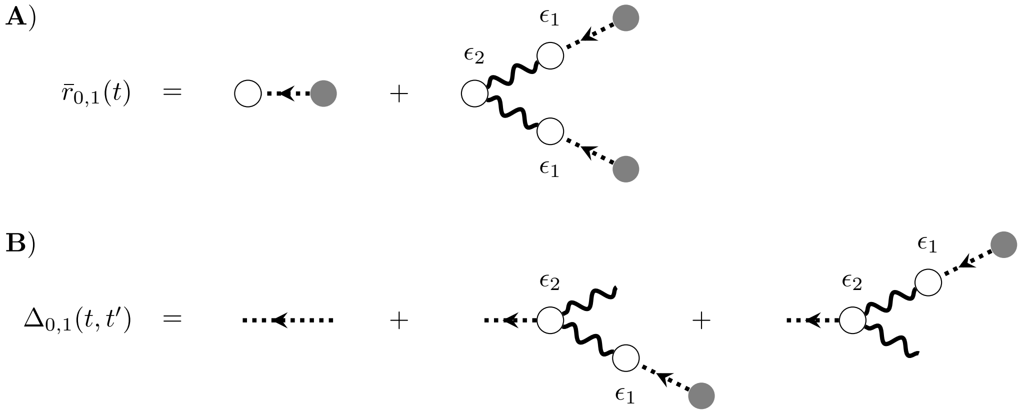


Fig 8. Expansion of the mean firing rate and propagator for the nonlinearly self-exciting process. A) One of the first nonlinear terms of the expansion of the mean-field firing rate, to first order in the quadratic term of the nonlinearity and zeroth order in the linear term. The two diagrams shown correspond to the two terms in (41). B) First nonlinear terms of the expansion of the propagator around the mean-field firing rate.

To zeroth order in ϵ_2, this yields an expansion of the mean-field rate $\bar{r}(t)$ which takes the same form as the expansion of the rate of the linearly self-exciting process, Eq. 12 and admits the same graphical representation (Fig. 4b). Similarly, a perturbative expansion of the linear response about the mean-field rate to zeroth order in ϵ_2 takes the same form as for the linearly self-exciting process (Eq. 27) and admits the same graphical representation (Fig. 4c).

To account for the nonlinear terms arising at first order and greater in ϵ_2, we will need to add another type of internal vertex in diagrammatic descriptions of the cumulants. These vertices, carrying factors of ϵ_2, will have two incoming edges and any number of outgoing edges. Each incoming edge carries the operator $g*$ and the number of incoming edges corresponds to the order in the Taylor expansion of the nonlinearity. (It also corresponds to the order of cumulant influencing that vertex's activity. The number of outgoing edges corresponds to the order of cumulant being influenced, locally in that subgraph.) The factor of $\bar{r}(t)$ that appears in other vertices is modified to be consistent with the mean firing rate under the quadratic nonlinearity, and will thus obey Eq.(34) above.

The mean-field rate and propagator, to first order and greater in ϵ_2, can be represented diagrammatically using the new vertex (e.g. Fig. 8a, b). Notice that these directed graphs are tree-like, but with their leaves in the past. Repeating these calculations to the next order in ϵ_2 can be accomplished by taking the basic structure of e.g. Fig. 8 and, along each leg entering the new vertex for ϵ_2, inserting the previous-order graphs (Figs. 4a, 8a). Including higher-order terms in ϵ_1 would require inserting those graphs along the ϵ-carrying vertices of Fig. 4a.

We next consider the fluctuations of the firing rate around the mean-field:

$$\left\langle \frac{dN}{dt} \right\rangle = r + \bar{r}. \tag{43}$$

Again taking a quadratic nonlinearity, we have:

$$r(t) = \epsilon_1 g * r(t) + \epsilon_2 \left(g * r(t)\right)^2 \quad . \tag{44}$$

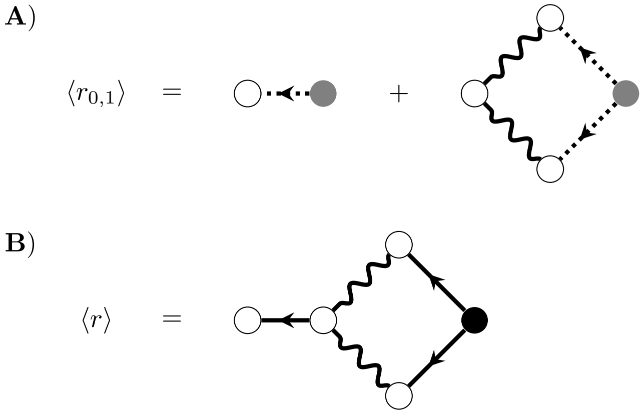


Fig 9. Corrections to the mean-field firing rate of the nonlinearly self-exciting process, to quadratic order in the nonlinearity ϕ. A) The correction to mean-field theory for the firing rate to first order in the quadratic term of ϕ. B) The full one-loop correction to the mean-field rate.

where $\epsilon_m = \phi^m/m!$, evaluated at $g * \bar{r}(t) + \lambda(t)$. Similarly to before, we recursively define dN/dt:

$$\frac{dN_{m,n}(t)}{dt} = \frac{dN_{0,0}(t)}{dt} + \frac{dM_{m-1,n}(t)}{dt} + \frac{dP_{m,n-1}(t)}{dt} \tag{45}$$

Once again, $M_{m,n}$ is an inhomogeneous Poisson process with rate $\epsilon_1 g * \frac{dN_{m,n}(t)}{dt}$ and $P_{m,n}$ is an inhomogeneous Poisson process with rate $\epsilon_2 \left(g * \frac{dN_{m,n}}{dt}\right)^2(t)$. $N_{0,0}(t)$ now has rate $\delta\lambda(t) = \lambda(t) - \bar{r}(t)$. We will begin calculating cumulants of dN/dt to finite order in m, n. The first nonlinear correction to the firing rate appears at first order in ϵ_2:

$$\left\langle \frac{dN_{0,1}(t)}{dt} \right\rangle_c = \delta\lambda(t) + \epsilon_2 \left\langle \left(g * \frac{dN_{0,0}}{dt}(t)\right)^2 \right\rangle_c \tag{46}$$

which can be represented diagrammatically using the new vertex (Fig. 9a). Notice that in contrast to the corresponding graph for the mean-field expansion (Fig. 8a), this diagram has a “loop” (a cycle were it an undirected graph). This reflects the dependence of the rate on the second cumulant of the baseline process $N_{0,0}$. This dependence of the firing rate on higher-order spiking cumulants is a fundamental feature of nonlinearities.

Proceeding beyond the first order in both ϵ_1 and ϵ_2, we see that the expansion of each term of the nonlinearity depends on the other:

$$\left\langle \frac{dN_{1,1}(t)}{dt} \right\rangle_c = \delta\lambda(t) + \epsilon_1 g * \left\langle \frac{dN_{0,1}(t)}{dt} \right\rangle_c + \epsilon_2 \left\langle \left(g * \frac{dN_{1,0}(t)}{dt}\right)^2 \right\rangle_c \tag{47}$$

so that at each order in ϵ_2 we must insert the solution at the previous order in ϵ_2 and the same order in ϵ_1 (and vice versa). This recursive mixing of expansions between the linear and nonlinear terms of ϕ seems intractable. However, this joint expansion can be re-summed to obtain the full correction to the firing rates [37]. The Feynman rules for the re-summed diagrams of the nonlinearly self-exciting process are given in Fig. 10. For a quadratic nonlinearity, this yields the one-loop correction to the firing rate:

$$r_1 = \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \Delta(t, t_1) \left(g * \Delta(t_1, t_2)\right)^2 \bar{r}(t_2) \tag{48}$$

where we relabel r as r_1 to denote that this is the one-loop correction.

The appearance of a loop in the Feynman diagram for the mean rate of the quadratically self-exciting process is a general feature of nonlinearities. It represents the influence of higher-order spike train cumulants on lower order ones. In order to measure that dependency, we can count the number of loops in the graphs. To do this, we add a bookkeeping variable, h. We count the number of loops by multiplying factors of h and $1/h$. Each internal vertex adds a factor of $1/h$ and each outgoing edge a factor of h. In this way every vertex with more than one outgoing edge will contribute a factor of h for every edge beyond the first. h thus effectively counts the order of fluctuations contributed by the vertex. For example, the mean for the linear self-exciting process has a graph with a single internal vertex and a single internal edge, so it is zeroth order in h (Fig. 4b). The two point function, however, having two edges and one internal vertex (Fig. 5b), is first order in h. Similarly, the tree-level diagrams will always contribute a total of h^{n-1}, where n is the order of the cumulant.

In terms of powers of h, a graph for a nth order cumulant with one loop will be equivalent to a graph for a $n+1$st order cumulant with one less loop. Consider cutting one of the lines that form the loop in Fig. 9b at the internal vertex and leaving it hanging. Now the graph for the one-loop correction to the mean rate appears to be a graph for a second cumulant - it has two endpoints. The power counting in terms of h, however, has not changed. The one-loop correction to the mean is of the same order in h as the tree-level second cumulant. In general, we will have that the order h^m will be given by

$$m = n + l - 1 \tag{49}$$

where n is the number of external vertices and l is the number of internal loops. The number of loops thus tracks the successive contributions of the higher order fluctuations. This expansion is called the “loop” expansion and is equivalent to a small-fluctuation expansion. If one can control the size of the fluctuations, one can truncate the loop expansion as an approximation for the statistics of the system. One way of doing this with networks is to insure that the interactions are $\mathcal{O}(1/N)$ so that $h \propto 1/N$ and the expansion becomes a system size expansion.

The Taylor expansion of an arbitrary nonlinearity ϕ could have infinitely many terms. This would lead, in the recursive formulation, to infinitely many processes $\{M, P, \ldots\}$. Even after re-summing the recursive formulation, this would leave an infinite number of diagrams corresponding to any given cumulant. There are two ways to approach this. The first is to insist on a perturbative expansion in the non-linear terms, e.g. only consider terms up to a fixed order in the Taylor expansion of the nonlinearity ϕ.

The second approach to controlling the order of the loop expansion is to consider a regime in which mean field theory is stable as this will also control the fluctuations, limiting the magnitude of the loop contributions [37]. The expansion then breaks down in the regime of a bifurcation or “critical point”. In this case, the linear response diverges, causing all loop diagrams to similarly diverge. This is a fluctuation-dominated regime in which mean field theory, along with the fluctuation expansion around it, fails.

Interaction between single-neuron nonlinearities and network structure

No new concepts are required in moving from a nonlinear self-exciting process to a network of interacting units. Each external and internal vertex must now be associated with a unique neuron index i and the integrations over time for the internal vertices must now be accompanied by summations over the indices of the internal vertices. In addition, the filter $g(\tau)$ must be expanded to include coupling across units. In general, this is given by $g_{ij}(\tau)$ for the coupling from neuron j to neuron i. We will

Feynman Rules for the Nonlinear Self-Exciting Process

For the nth cumulant, $\left\langle \prod_i \frac{dN(t_i)}{dt} \right\rangle_c$:

I. For each i, introduce a vertex labelled t_i. These are *external* vertices.

II. Construct all directed, fully connected graphs such that each vertex from I. has a single incoming propagator (Δ) edge using the available vertex factors and edges below.

III. For each graph, multiply the associated vertex or edge factor and integrate over the times t for all internal vertices.

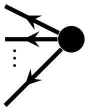
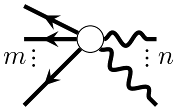
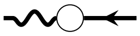
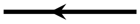
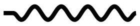
Vertex or Edge	Factor
	$\bar{r}(t)$
	$\phi^n(g * \bar{r} + \lambda(t))/n!$
	1
	$\Delta(t, t')$
	$g(t, t')$

Fig 10. Feynman rules for the nonlinear self exciting process. These rules provide an algorithm for computing the expansion of the cumulants around the mean field solution $\bar{r}(t)$. The dots between the outgoing legs of the first vertex indicate that there any number of outgoing legs greater than or equal to two. The number n of incoming edges of the second vertex correspond to it's factor containing the nth derivative of ϕ, evaluated at the mean field input. The m dots between the outgoing edges of the second vertex indicate that it can have any number of outgoing edges such that $m + n \geq 3$.

Feynman Rules for Nonlinear Networks of Poisson Neurons

For the nth cumulant, $\left\langle \prod_i \frac{dN_i(t_i)}{dt} \right\rangle_c$:

I. For each i, introduce a vertex labelled t_i, i. These are *external* vertices.

II. Construct all directed, fully connected graphs such that each vertex from I. has a single incoming propagator (Δ) edge using the available vertex factors and edges below.

III. For each graph, multiply the associated vertex or edge factors, integrate over the time t_i, and sum over the neuron index i for all internal vertices.

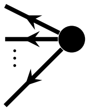
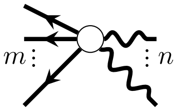
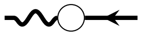
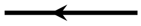
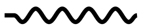
Vertex or Edge	Factor
	$\bar{r}_i(t)$
	$\phi^n \left(\sum_j g_{ij} * \bar{r}_j + \lambda_i(t) \right) /n!$
	1
	$\Delta_{ij}(t,t')$
	$g_{ij}(t,t')$

Fig 11. Feynman rules for networks of stochastically spiking neurons with nonlinear input-rate transfer ϕ. These rules provide an algorithm for computing the expansion of the cumulants around the mean field solution $\bar{r}(t)$. The dots between the outgoing legs of the first vertex indicate that there are any number of outgoing legs greater than or equal to two. The number n of incoming edges of the second vertex correspond to it's factor containing the nth derivative of ϕ, evaluated at the mean field input. The m dots between the outgoing edges of the second vertex indicate that it can have any number of outgoing edges such that $m + n \geq 3$.

consider the general model of a network of units that generate conditionally Poisson-distributed events, given an input variable. The conditional rate for unit i is given by

$$r_i(t) = \phi\left(\sum_j g_{ij} * \frac{dN_j}{dt} + \lambda_i(t)\right) \tag{50}$$

Similarly, the propagator now obeys

$$\Delta_{ij}(t,t') = \phi_1\left(\sum_s g_{is} * \bar{r}_s(t) + \lambda_s(t)\right)\left(\sum_k g_{ik} * \Delta_{kj}(t,t') + \delta_{ij}\delta(t-t')\right) \tag{51}$$

These dynamics are qualitatively the same as those of the nonlinearly self-exciting process (Eq. 34 but replace the neuron's own rate with the sum over its presynaptic inputs). Introducing these sums over neuron indices yields the complete set of rules for generating Feynman diagrams for the cumulants of this model, shown in Figure 11.

Mid-course summary: diagrammatic expansion reveals interplay of network structure and neural nonlinearities in driving neural activity

To summarize, we now have in hand a set of tools—a fluctuation expansion—to compute spike train cumulants of arbitrary order in networks of linear-nonlinear-Poisson neurons. This expansion provides a systematic way to account for the synergistic dependence of lower-order activity on higher-order activity through the spiking nonlinearity, and naturally incorporates the full microstructure of the neuronal network. The order of the nonlinearity (for non-polynomials, the order of its Taylor expansion) determines the single-neuron transfer gain (Fig. 12 left). It *also* determines how activity propagates through the network (Fig. 12).

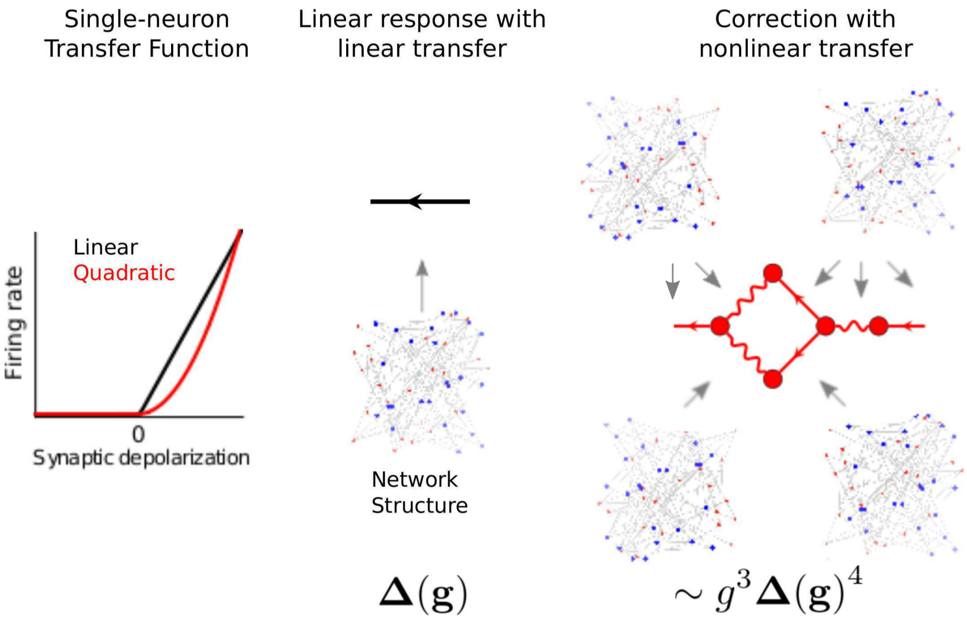

Single-neuron Transfer Function
Linear response with linear transfer
Correction with nonlinear transfer
Linear
Quadratic
Firing rate
0
Synaptic depolarization
Network Structure
$\mathbf{\Delta}(\mathbf{g})$
$\sim g^3 \mathbf{\Delta}(\mathbf{g})^4$

Fig 12. Fluctuation expansion links single-neuron nonlinearities and network structure to determine network activity. The linear response for linear neurons depends on the network structure both explicitly and implicitly, through the mean-field rates. The first nonlinear correction brings in additional explicit and implicit dependencies on the connectivity.

We next provide an example of using this fluctuation expansion to compute a cumulant of spiking activity: the first nonlinear correction to the second cumulant. We will compute these using the Feynman rules (Fig. 11) to construct the corresponding diagrams, from which we will write the corresponding equation.

We begin by placing external vertices corresponding to the measurement times, each with one propagator edge coming in to it. For the two-point cumulant there are two external vertices (Fig. 13a). The propagators coming into those external vertices can, according to the Feynman rules, arise from either a source vertex or from an internal vertex with incoming filter edges (Fig. 11). If both propagators arise from a source vertex, we arrive at the tree-level diagram of Fig. 5b, which provides the linear prediction for the two-point cumulant. To obtain the first nonlinear correction, we will begin by adding an internal vertex. There are two ways we can do this: with one internal vertex providing the propagators for both external vertices or with the internal vertex providing the propagator of just one external vertex (Fig. 13b, top and bottom respectively).

We next proceed to add another layer of vertices. The internal vertices added in each diagram of Fig. 13b have incoming propagator edges. Those edges could emanate from other internal vertices or from source verticies. We will start by finishing the diagrams where they emanate from source vertices; placing these yields the top two, final diagrams of Fig. 13c. We then continue constructing the diagram with an internal vertex providing the two propagators (Fig. 13c, bottom). Note that if we added an internal vertex to the propagator hitting the t_2 external vertex, it would require at least two incoming filter edges (in order to obey $n + m \geq 3$ per the rules of Fig. 11) which would give rise to a second loop in the graph.

This last diagram has a hanging filter edge, which must arise from an internal vertex with one incoming propagator edge. We finish the diagram with that internal vertex and the source vertex providing its propagator (Fig. 13d). We could not add additional internal vertices along that path, since they would either violate $n + m \geq 3$ or give rise to more than one loop in the diagram (and thus belong at a higher order in the loop expansion).

Following the rules of Fig. 11, while restricting ourselves to graphs with a certain number of loops (here one) thus allowed us to construct the diagrams correspond to the first nonlinear correction to the two-point cumulant. We next write the equation corresponding to these diagrams. For each of the complete diagrams, we begin at the external vertices, and proceeding from left to right in the graph, multiply the factors corresponding to each edge and vertex together. We finish by summing over indices for all internal vertices and integrating over all internal time points. The contributions from each diagram are then added together. This yields:

$$\begin{aligned}
\left\langle \frac{dN_i(t_1)}{dt} \frac{dN_i(t_2}{dt} \right\rangle_c^{1\text{loop}} = {} & \sum_{k,l,m,n} \int dt' \int dt'' \, \Delta_{ik}(t_1 - t') \Delta jk(t_2 - t') \frac{\phi_k''}{2} \left(g_{kl} * \Delta_{ln}(t' - t'')\right) \left(g_{km} * \Delta_{mn}(t' - t'')\right) \bar{r}_l(t'') \\
& + \sum_{k,l,m,n} \int dt' \int dt'' \, \Delta_{ik}(t_1 - t') \frac{\phi_k'''}{2} \left(g_{kl} * \Delta_{ln}(t' - t'')\right) \left(g_{km} * \Delta_{mn}(t' - t'')\right) \Delta_{jn}(t_2 - t'') \bar{r}_n(t'') \\
& + \sum_{k,l,m,n} \int dt' \int dt'' \, \Delta_{ik}(t_2 - t') \frac{\phi_k'''}{2} \left(g_{kl} * \Delta_{ln}(t' - t'')\right) \left(g_{km} * \Delta_{mn}(t' - t'')\right) \Delta_{jn}(t_1 - t'') \bar{r}_n(t'') \\
& + \sum_{k,l,m,n,o,p} \int dt' \int dt'' \int dt''' \Big[\Delta_{ik}(t_1 - t') \frac{\phi_k''}{2} \left(g_{kl} * \Delta_{ln}(t' - t'')\right) \left(g_{km} * \Delta_{mn}(t' - t'')\right) \\
& \quad g_{no} * \Delta_{op}(t'' - t''') \Delta_{jp}(t_2 - t''') \bar{r}_p(t''') \Big] \\
& + \sum_{k,l,m,n,o,p} \int dt' \int dt'' \int dt''' \Big[\Delta_{ik}(t_2 - t') \frac{\phi_k''}{2} \left(g_{kl} * \Delta_{ln}(t' - t'')\right) \left(g_{km} * \Delta_{mn}(t' - t'')\right) \\
& \quad g_{no} * \Delta_{op}(t'' - t''') \Delta_{jp}(t_1 - t''') \bar{r}_p(t''') \Big] \qquad (52)
\end{aligned}$$

Demonstrating the interplay between single-neuron transfer, connectivity structure, and network dynamics

Our methods predict how spike time statistics of all orders emerge from the interplay of single-neuron input-output properties and the structure of network connectivity. Here we demonstrate how these methods can be used to predict key phenomena in recurrent spiking networks: the fluctuations and stability of population activity, and stimulus coding. First, we isolate the contributions of nonlinearities in single-neuron dynamics to network activity and coding as a whole. We do so by computing "one-loop" correction terms; these correspond to the first structures in our diagrammatic expansion that arise from nonlinear neural transfer. The one-loop corrections provide for the dependence of nth order spiking

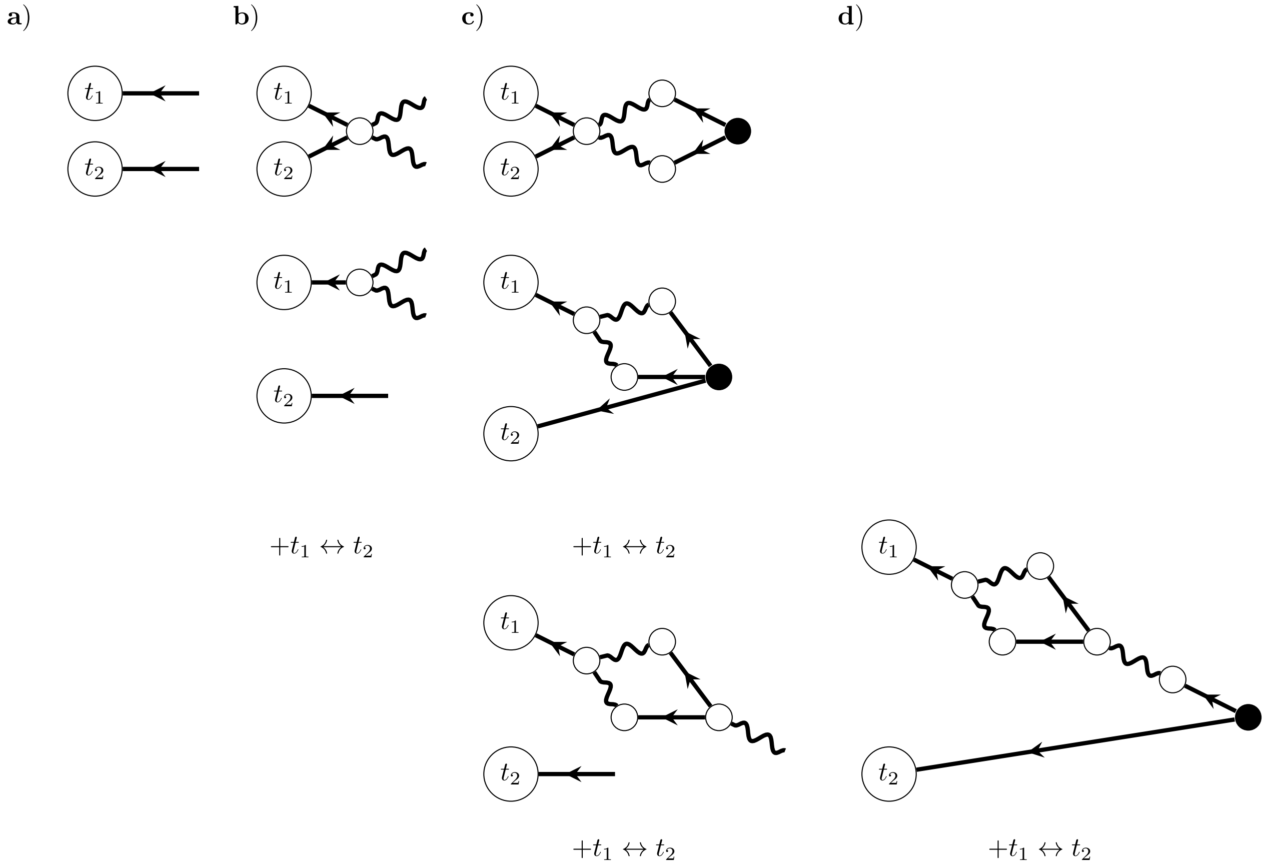

a)
b)
c)
d)
t_1
t_2
$+t_1 \leftrightarrow t_2$
$+t_1 \leftrightarrow t_2$
$+t_1 \leftrightarrow t_2$
$+t_1 \leftrightarrow t_2$

Fig 13. Construction of Feynman diagrams for the first nonlinear correction to the two-point cumulant (graphs containing one loop). In each panel, we add a new layer of vertices to the diagrams, until we arrive at a source vertex. When there are multiple potential ways to add vertices, we add diagrams to account for each of those constructions. A) External vertices corresponding to the two measurement times, with incoming propagator (Δ) edges. B) Diagrams with one internal vertex added. $t_1 \leftrightarrow t_2$ corresponds to switching the two external vertices in the bottom diagram; the top diagram is symmetric with respect to that switch. C) Diagrams with two layers of vertices. The top diagram finishes that of B, top. The second two arise from the second diagram of B, and each also have copies with $t_1 \leftrightarrow t_2$. D) Last diagrams containing one loop. The final diagrams corresponding to the one-loop correction to the second cumulant are the top two of C) and that of D).

cumulants on $n+1$st order cumulants. Predictions that would be made by linearizing neural dynamics, as in classic approaches for predicting pairwise correlations [19, 22, 38] and recent ones for higher-order correlations [36], are described as "tree-level." We show how these one-loop corrections, which give new, explicit links between network structure and dynamics (Fig. 12), predict spiking statistics, stability, and the accuracy of coding in recurrent networks.

2. Recurrent spike-train correlations drive firing statistics and stability in nonlinear networks

In our analysis of the impact of nonlinear neural transfer on network dynamics, a principal finding was that spike correlations could affect firing rates, as described by the one-loop correction to the mean-field firing rates. In this section we illustrate the importance of this effect in a class of networks under intensive study in neuroscience: randomly connected networks of excitatory and inhibitory cells. We began with a network for which we expect classical theoretical tools to work well, taking the neurons to have threshold-linear transfer functions. Here, as long as the neurons do not receive input fluctuations that push their rates below this threshold, the "tree-level" theory that takes transfer to be entirely linear should work well. We then move on to consider nonlinear effects.

As in our original motivational example, we took network connectivity to be totally random (Erdös-Rènyi), with $p_{EE} = 0.2$ and $p_{EI} = p_{IE} = p_{II} = 0.5$. The magnitude of all connections of a given type ($E-E$, etc.) was taken to be the same and the time course of all network interactions was governed by the same filter $g(t) = \frac{t}{\tau^2}\exp\left(-t/\tau\right)$ (with $\tau = 10$ ms), so that $\mathbf{g}_{ij}(t) = \mathbf{W}_{ij} g(t)$. The net inhibitory input weight on to a neuron was, on average, twice that of the net excitatory input weight.

We examined the spiking dynamics as the strength of synaptic weights proportionally increased (Fig. 14A), and studied network activity with using both theory and direct simulation. Due to the high relative strength of inhibitory synapses in the network, firing rates decreased with synaptic weight (Fig. 14D). The magnitude of spike train covariances (reflected by the integrated autocovariance of the summed excitatory population spike train) increased (Fig. 14E). These changes were also visible in raster plots of the network's activity (Fig. 14B,C).

At a critical value of the synaptic weights, the mean-field theory for the firing rates loses stability (Fig. 14F). The location of this critical point is predicted by the linear stability of the dynamics around the mean-field rate; the spectrum of the propagator $\mathbf{\Delta}(\omega) = \left(\mathbf{I} - \phi'\mathbf{g}(\omega)\right)^{-1}$ diverges when the spectral radius of $\phi'\mathbf{g}$ is ≥ 1. This is the point where the spectral radius of the inverse propagator crosses zero. Until that critical point, however, the "tree-level" predictions for both firing rates and spike train covariances (i.e. mean-field theory and linear response theory) provided accurate predictions (Fig. 14D,E). This combination of mean-field theory for firing rates and a linearization around it to predict spike train covariances has a long history in theoretical neuroscience (e.g. [21–23, 38, 39]).

We next give a simple example of how nonlinearities in neural transfer cause this standard tree-level theory (mean-field formulas for rates and linear response theory for covariances) to fail – and how tractable "one-loop" corrections from our theory give a much-improved description of network dynamics. We take the same network as above, but replace neurons' threshold-linear transfer functions with a rectified power law (Fig. 15A). This has been suggested as a good description of neural transfer near threshold [40–42]. For simplicity, we take the power law to be quadratic. As we increased synaptic weights, the tree-level theory qualitatively failed to predict the magnitude of spike train covariances and firing rates (Fig. 15D,E black curve vs dots). This occurred well before the mean-field firing rates lost stability (Fig. 15F, black).

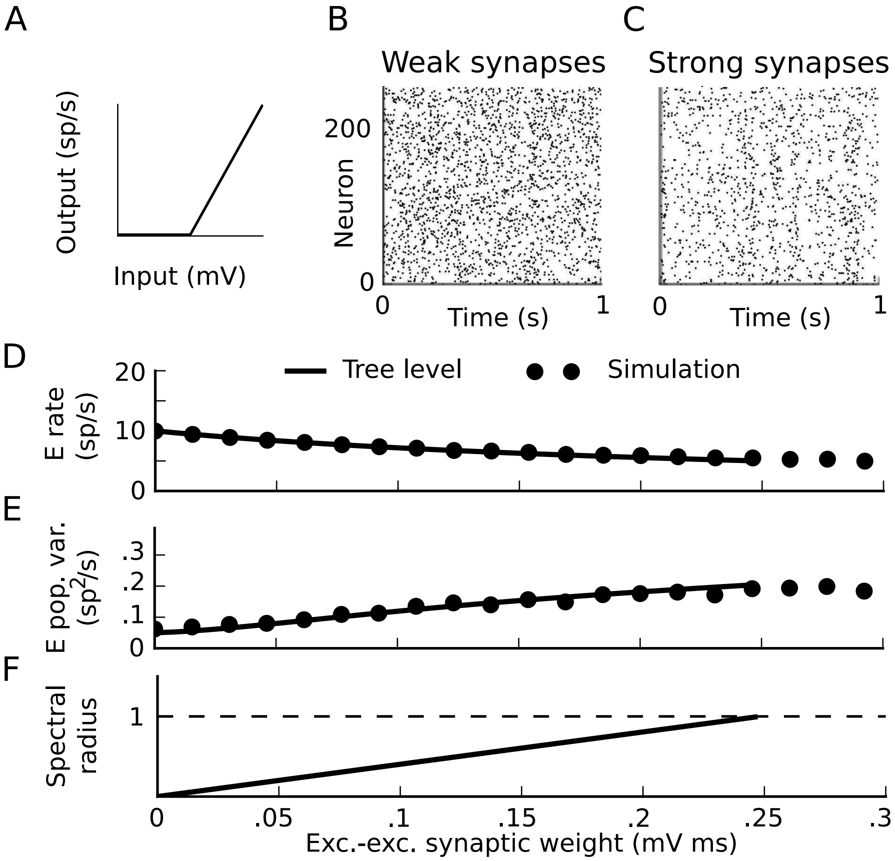


Fig 14. Dynamics approaching the firing-rate instability in threshold-linear networks. A) Threshold-linear input-rate transfer function. B,C) Raster plots of 1 second realizations of activity for weak and strong synaptic weights. Neurons 0-199 are excitatory and 200-240 are inhibitory. B) $(W_{EE}, W_{EI}, W_{IE}, W_{II}) = (.025, -.1, .01, -.1)$ mV. C) $(W_{EE}, W_{EI}, W_{IE}, W_{II}) = (.2, -.8, .08, -0.8)$ mV. D-F) Average firing rate of the excitatory neurons (D), integral of the auto-covariance function of the summed population spike train (E), and spectral radius of the stability matrix of mean-field theory. (F) vs excitatory-excitatory synaptic weight. While excitatory-excitatory weight is plotted on the horizontal axis, all other synaptic weights increase proportionally with it. Black lines: tree-level theory: mean-field firing rates and covariance computed by linearizing dynamics around it, for each value of synaptic weights. Dots: simulation.

Higher-order terms of the loop expansion described above (Nonlinearities impose bidirectional coupling between different orders of activity) provide corrections to mean-field theory for both firing rates and spike train correlations. These corrections represent coupling of higher-order spike train cumulants to lower order cumulants. In the presence of an input-rate nonlinearity, for example, synchronous (correlated) presynaptic spike trains will more effectively drive postsynaptic activity [7, 43]. This effect is described by the one-loop correction to the firing rates (Fig. 9).

Similarly to the nonlinearly self-exciting process, the one-loop correction to the mean-field rate of neuron i is:

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \sum_{j,k} \Delta_{ij}(t, t_1) \frac{1}{2} \phi_j'' \left(\sum_l g_{jl} * \Delta_{lk}(t_1, t_2) \right) \left(\sum_l g_{jl} * \Delta_{lk}(t_1, t_2) \right) \bar{r}_k(t_2), \tag{53}$$

where t_0 is the initial time. This correction was, on average, positive (Fig. 15D for excitatory neurons; also true for inhibitory neurons). Similarly to firing rates, the loop expansion provides corrections to higher-order spike train cumulants. The one-loop correction to the spike train covariances accounts for the impact of triplet correlations (third joint spike train cumulants) on pairwise correlations and provided an improved prediction of the variance of the population spike train as synaptic weights increased (Fig. 15E).

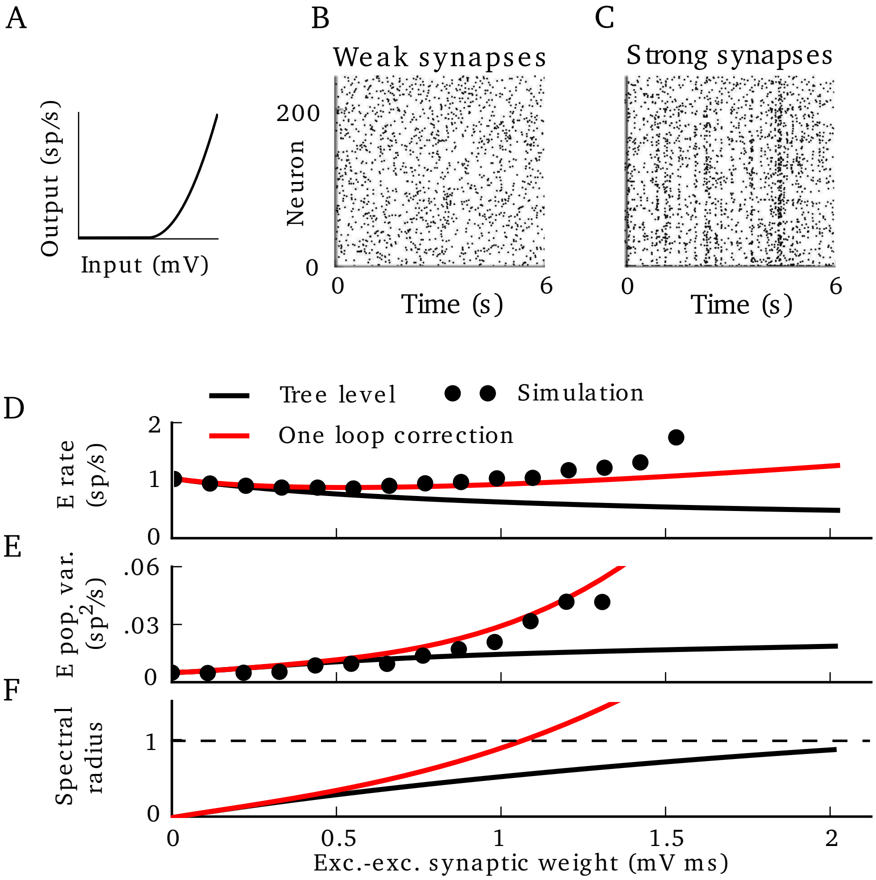


Fig 15. Correlation-driven instability in nonlinear networks. A) Threshold-quadratic input-rate transfer function. B,C) Raster plots of 6 second realizations of activity for weak and strong synaptic weights. Neurons 0-199 are excitatory and 200-240 are inhibitory. B) $(W_{EE}, W_{EI}, W_{IE}, W_{II}) = (.025, -.1, .01, -.1)$ mV. C) $(W_{EE}, W_{EI}, W_{IE}, W_{II}) = (1.5, -6, .6, -6)$ mV. D-F) Average firing rate of the excitatory neurons (D), integral of the auto-covariance function of the summed population spike train (E), and spectral radius of the stability matrix of mean-field theory (F) vs excitatory-excitatory synaptic weight. While excitatory-excitatory weight is plotted on the horizontal axis, all other synaptic weights increase proportionally with it. Black line: tree-level theory. Red line: one-loop correction accounting for impact of the next order (pairwise correlations' influence on mean and triplet correlations' influence on pairwise). Dots: simulation. All dots after the one-loop spectral radius crosses 1 represent results averaged over the time period before the activity diverges.

Since the one-loop correction to the firing rates could be large, we also asked whether it could impact the stability of the firing rates - that is, whether pairwise correlations could, through their influence on firing rates through the nonlinear transfer function, induce an instability. This is a question of when the eigenvalues of the propagator diverge—or equivalently, when the eigenvalues of the inverse propagator cross zero. The inverse of the one-loop propagator is given by the "proper vertex" obtained by amputating the outside propagator edges of the one-loop correction to the propagator; or equivalently, calculated from the Legendre transform of the cumulant-generating functional [37]. This yields::

$$\mathbf{\Gamma} = -\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \sum_k \frac{1}{2}\phi''_j \left(\sum_l g_{jl} * \Delta_{lk}(t_1,t_2)\right)\left(\sum_l g_{jl} * \Delta_{lk}(t_1,t_2)\right)\phi'_k g_{kl} \tag{54}$$

where ϕ''_j denotes the second derivative of the transfer function of neuron j, evaluated at its mean-field input (and similar for ϕ'_k). This is the correction to the inverse propagator $\mathbf{\Delta}^{-1}$, so the correction to the steady-state stability matrix $\int dt g(t)\phi'\mathbf{W}$ is given by $-\mathbf{\Gamma}$. This predicted that the firing rates should lose stability significantly before the bifurcation of the mean-field theory (Fig. 15F, red vs black). Indeed, we saw in extended simulations that the spiking network could exhibit divergent activity even with synaptic weights which the mean-field theory predicted should be stable (Fig. 1C). In summary, mean-field theory can mis-predict the bifurcation of the rate of spiking models since it fails to capture the impact of correlations on firing rates through nonlinear transfer functions.

3. Impact of connectivity structure on correlation-driven instabilities in nonlinear networks

Recent work has shown that cortical networks are more structured than simple Erdős-Rènyi networks. One feature of cortical networks is a broad spread of neurons' in- and out-degree distributions (i.e., the distributions of the number of synaptic inputs that each neuron receives or sends); another is broadly spread synaptic weights. These network properties, in turn, can have a strong impact on population activity [44–48]. Here, we illustrate the link between network structure and activity in the presence of nonlinear neural transfer. To generate structured networks, we began with the type of excitatory-inhibitory networks discussed in the previous section, but took the excitatory-excitatory coupling to have both heavy-tailed degree and weight distributions. Specifically, we took it to have truncated, correlated power law in- and out-degree distributions. We then took to the synaptic weights to be log-normally distributed [49, 50]. For simplicity, we took the location and scale parameters of the weight distribution to be the same.

We then examined the network dynamics as the location and scale of the excitatory-excitatory synaptic weights increased. For each mean weight, we sampled the excitatory-excitatory weights from a lognormal distribution with that mean and variance. The excitatory-inhibitory, inhibitory-excitatory and inhibitory-inhibitory weights remained delta-distributed. For strong and broadly distributed synaptic weights (Fig. 16B, the network exhibited a similar correlation-induced instability as observed in the Erdős-Rènyi network (Fig. 16C) even though mean-field theory predicted that the firing rates should be stable (Fig. 16F, black vs. red curves). As synaptic weights increased from zero, the mean-field theory for firing rates provided a misprediction (Fig. 16D, black line vs dots) and the linear response prediction for the variance of the population spike train also broke down (Fig. 16E, black line vs dots). The one-loop corrections, accounting for the impact of pairwise correlations on mean rates and of triplet correlations on pairwise correlations, yielded improved predictions (Fig. 16D, E red lines) and a much more accurate prediction for when firing rates would lose stability (Fig. 16C, F). These effects were similar to those seen in Erős-Rènyi networks (Fig. 15), but the transition of the firing rates occurred sooner, both for the mean field (because of the effect of the weight and degree distributions on the eigenvalues of the weight matrix) and one loop theories (because of the impact of the correlations on the firing rates).

4. Stimulus coding and discrimination in nonlinear, recurrent networks

We next examined how internally-generated spike train correlations impacted a network's ability to perform a simple coding task: discriminating two different inputs. The possible impact of correlations on

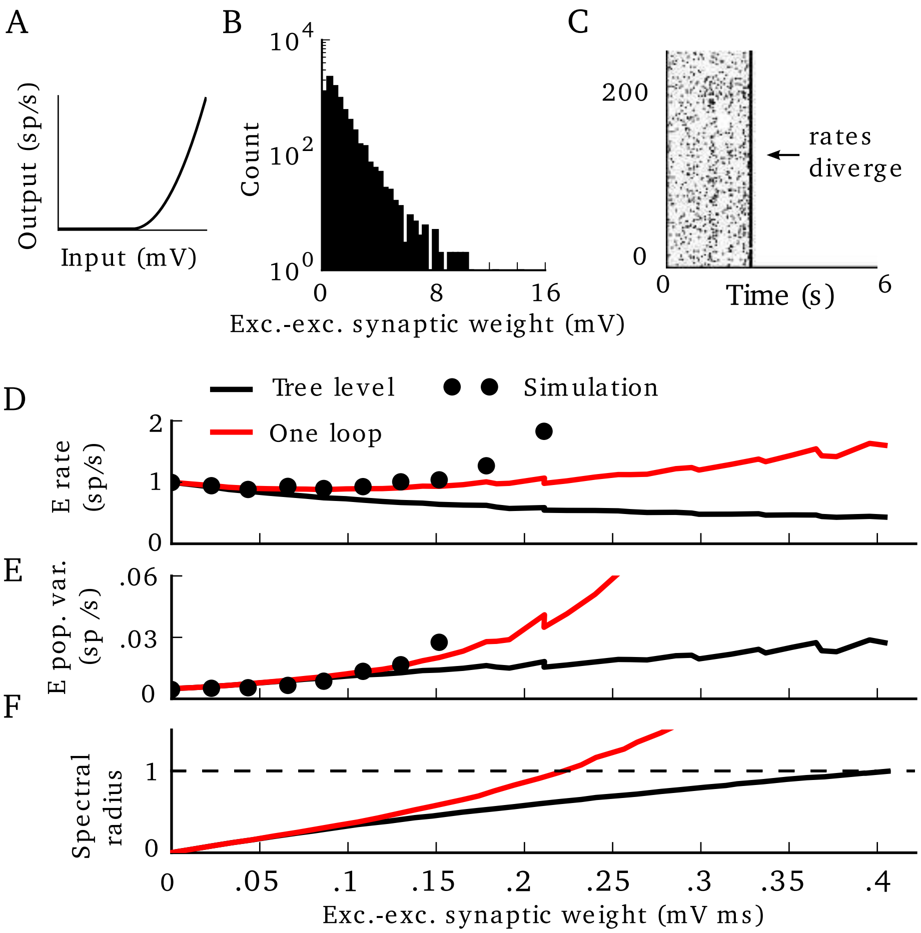


Fig 16. Correlation-driven instability in a non-Erdős-Rényi network with broadly distributed excitatory-excitatory weights. A) Threshold-quadratic input-rate transfer function. B) Histogram of excitatory-excitatory synaptic weights with location parameter of 1.42 (mean of .29 mV), corresponding to the simulation in panel C. C) Raster plots of 6 second realizations of activity. Neurons 0-199 are excitatory and 200-240 are inhibitory. $(W_{EE}, W_{EI}, W_{IE}, W_{II}) = (1.125, -4.5, .45, -4.5)$ mV. D-F) Average firing rate of the excitatory neurons (D), integral of the auto-covariance function of the summed population spike train (E), and spectral radius of the stability matrix of mean-field theory (F) vs excitatory-excitatory synaptic weight. While the mean excitatory-excitatory weight is plotted on the horizontal axis, all other synaptic weights increase proportionally with it. Black line: tree-level theory. Red line: one-loop correction. Dots: simulation. Simulation results in D, E are from realizations in which the rates did not diverge.

this task depended on the neurons' transfer function. For linear transfer (or threshold-linear, away from the threshold), we examined how the discriminability of the response distributions and the performance of decoder of the stimulus depended on the presence or absence of correlations: in other words, how assuming independent Poisson activity impacted the coding performance. Correlations stretch multi-neuron activity distributions [12], which can make the two response distributions either easier or harder to discriminate depending on the relation between the direction of the correlations and the stimulus structure (Fig. 17, middle). In the presence of a nonlinear transfer function, a further potential impact of correlations is to change neurons' mean activities, translating the multi-neuron response distributions. This could, in principle, either increase or decrease the discriminability of the response distributions (Fig. 17, bottom). Our goal was to determine which of the effects above occurs in sparse, random neural networks with distinct sets of neurons receiving inputs for each of two stimuli.

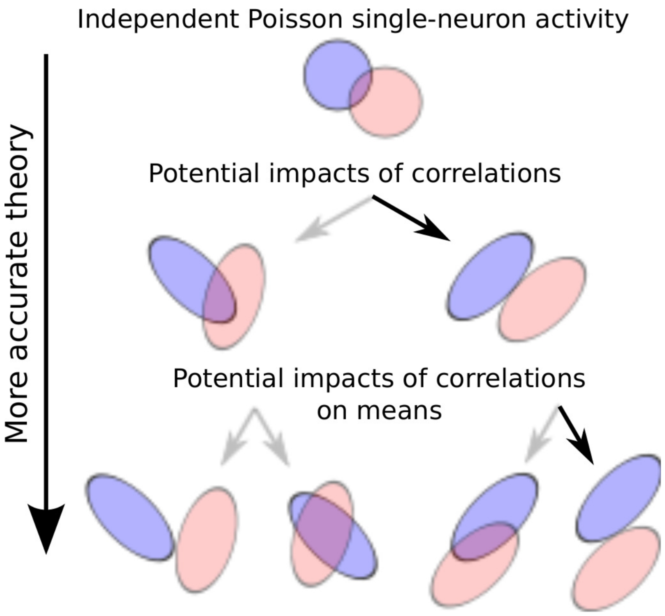

Independent Poisson single-neuron activity
Potential impacts of correlations
Potential impacts of correlations on means
More accurate theory

Fig 17. Potential impacts of correlations on coding in a presence of a nonlinearity. The independent Poisson assumption for neurons gives rise to uncorrelated distributions of population activity (top layer). Correlations could increase or decrease the overlap of those distributions, decreasing or increasing the decoding performance (middle layer left and right, respectively). The impact of correlations on the mean responses can shift those distributions, again either decreasing (bottom layer, middle) or increasing the decoding performance (bottom layer, left and right). For the network, nonlinearity and coding task we used, correlations improved the decoding performance but the shift of the distributions due to the nonlinear effect did not significantly affect the decoding performance (solid arrows).

We used the same type of excitatory-inhibitory network as before, with Erdős-Rènyi connectivity ($p_{EE} = 0.5, p_{EI} = p_{IE} = p_{II} = 0.5$), and provided it with two different inputs from a feedforward layer of neurons. Input A came from one half of the feedforward neurons and targeted one half of the excitatory neurons in the recurrent network, while input B came from the other half of the sensory layer neurons and targeted the other half of the recurrent excitatory neurons. We called input A's targets “assembly A” and likewise for B. The two inputs targeted non-overlapping assemblies. Feedforward neurons of input A (B) projected to assembly A (B) with $p = 0.5$ (Fig. 18A). There was no connectivity within the feedforward layer and assemblies A and B in the recurrent layer differed only in their feedforward inputs.

We next quantify how well the activity of assembly B could discriminate activation of the two inputs, measuring it as the spike count in a 200 ms window (and examining separately the population-averaged spike count or the population vector of spike counts). Activation of an input meant setting its neurons' rates to 1 sp/s. We used this weak stimulus activation in order to maintain a range of synaptic weights before the network became unstable. We began by examining the mean spike count of assembly B in response to activation of either of the two inputs. For a given strength of interactions, we approximated the distribution of the mean spike count of assembly B as a Gaussian distribution with mean and variance given by the theoretical predictions (Methods: Population coding). For threshold-linear transfer functions, the linear predictions corresponding to the tree-level diagrams an excellent match with simulation (Fig. 18). The distributions of activity in response to the two inputs appeared barely

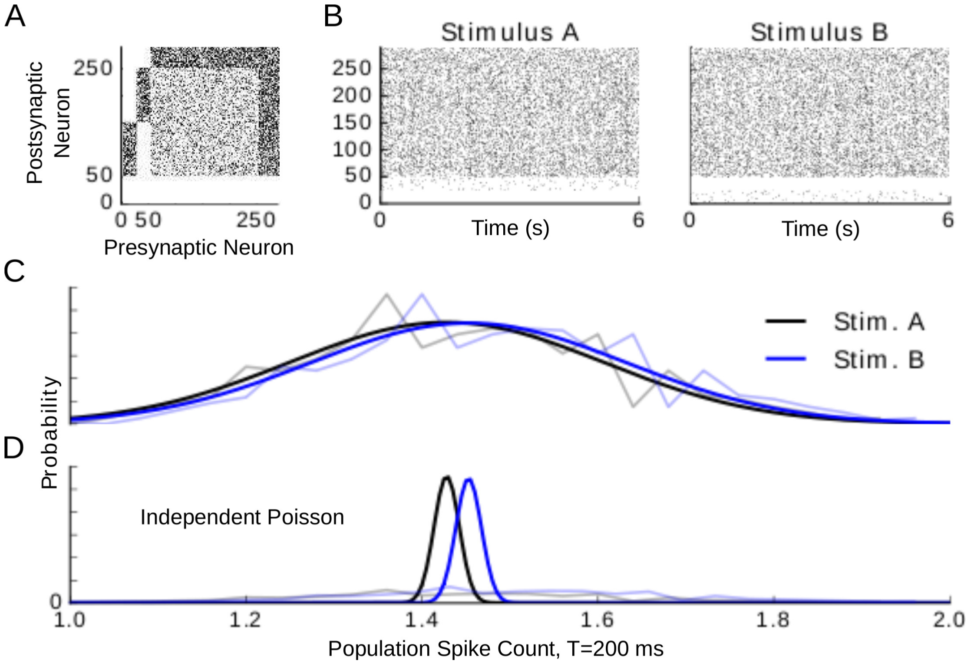


Fig 18. Stimulus-driven population activity. A) Connectivity of the network with the feedforward layer. Neurons 1-50 are the feedforward layer which projects to two subsets of the recurrent layer's excitatory neurons. B,C) Rasters of the network activity with each stimulus; the stimuli correspond to activation of feedforward neurons 1-25 (Stimulus A) or 26-50 (Stimulus B). The neurons whose activity we examine to decode the stimulus (neurons of "assembly B") are highlighted in blue. D) distributions of activity in response to stimulus A (black) and B (blue). Solid lines: theory (mean-field firing rates and tree-level correlations). E) Same as in D) but assuming that neurons spike independently, so that the variance of the population mean spike count is proportional to the sum of the variance of each neurons' spike counts.

distinguishable (Fig. 18C). Ignoring correlations between the activity of neurons in assembly B yielded theoretical distributions of the mean activity that appeared much more distinguishable due to their lower variance, however these theoretical distributions were a much worse match for the actual activity (Fig. 18D).

We next examined how the response distributions and the performance of a simple decoder of the stimulus were affected by the strength of interactions. The Kullback-Leibler divergence $D_{KL}\big(p(x|B)||p(x|A)\big)$ measured the loss of information when the distribution of activity in response to stimulus A is used to approximate the distribution of activity in response to stimulus B (Methods: Population coding). The D_{KL} increased with synaptic weights for both the population mean response and the population vector response (Fig. 19A, D). In general, neglecting correlations (assuming that individual neurons in assembly B spiked as independent Poisson processes) increased the D_{KL}, although this was much stronger for the population mean response than the population vector.

We used two decoders to predict the stimulus from realizations of assembly B's activity. The first was a naive Bayesian decoder, using the theoretical response distributions under the independent Poisson assumption. The second was the quadratic discriminant, using the theoretical distributions obtained with the tree-level prediction for the covariances. Both of these discriminate a stimulus based on the ratio of the probabilities of each stimulus given a particular response $\frac{p(B|x)}{p(A|x)}$, with the logarithm of the probability ratio being zero providing a convenient decision boundary for exponential family distributions. We saw that the independent Poisson assumption did not significantly affect the average of the probability ratio for the population mean response (Fig. 19B). The fraction of correctly decoded stimuli was, however, significantly impacted. Indeed, while increasing synaptic weights improved the decoding performance under the tree level theory, neglecting correlations led to a decoding performance that decreased with synaptic weights (Fig. 19C).

As expected, the decoding performance for the population vector was significantly greater than that of the population mean, both in terms of the average log probability ratio (Fig. 19E and the fraction of correct responses. Indeed, under the tree level theory the stimuli were almost perfectly decodable, with approximately 98% correctly decoded trials across all synaptic weights (Fig. 19F). For the population vector response, the independent Poisson assumption significantly reduced the decoding performance (Fig. 19E,F) although the fraction of correctly decoded trials still increased with synaptic weights.

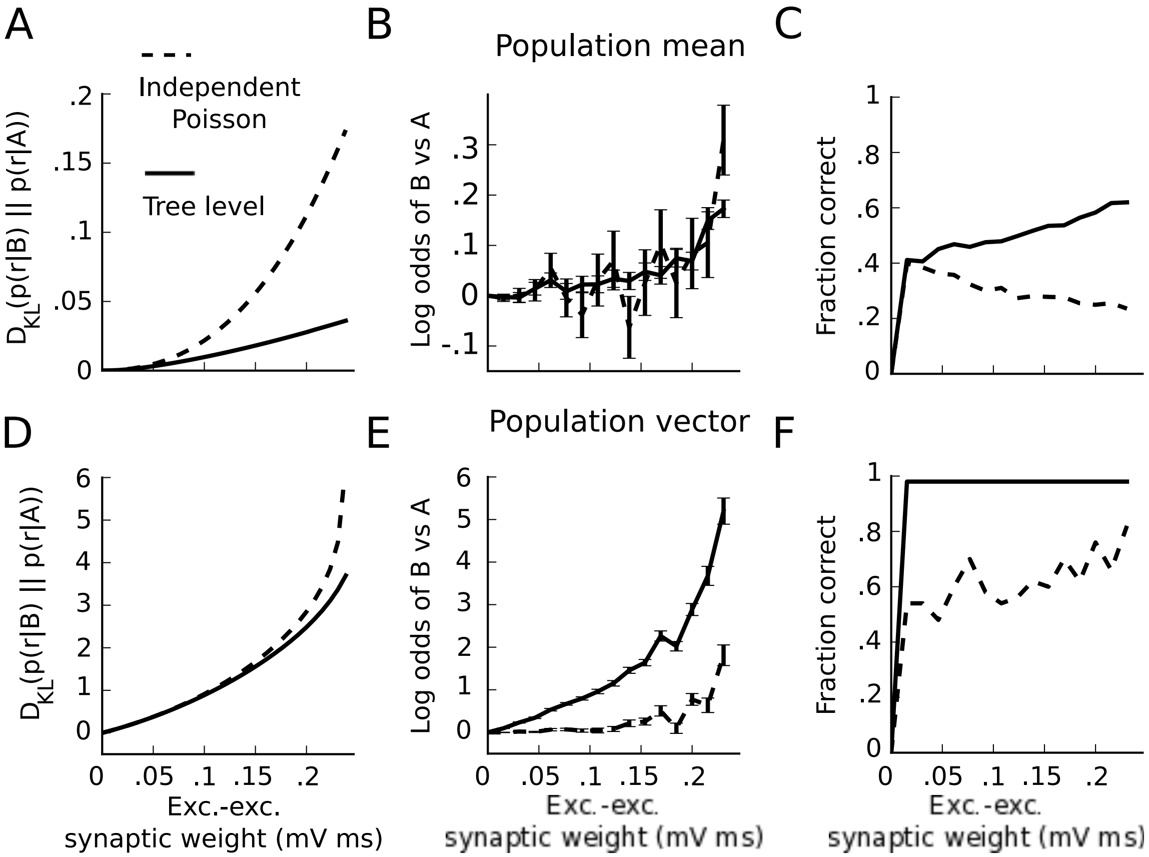


Fig 19. Stimulus distrimination with threshold-linear transfer functions. A) Kullback-Leibler divergence between the two conditional response distributions of the population mean spike count, assuming that they are Gaussian. B) Logarithm of the posterior probability of stimulus B vs A, $\ln \frac{p(x|B)}{p(x|A)}$ for the population mean spike count of assembly B. C) Fraction correct decoding performance (fraction of trials with log probability ratio > 0). D-F) Same as A-C), but for the distributions of the vector of spike counts from assembly B. Solid black lines: tree level theory, with the population spike count's statistics predicted with mean-field firing rates and pairwise covariances predicted by the linear response about the mean field. Dashed lines: independent Poisson assumption, with the population spike count statistics estimated with the mean-field rates and assuming neurons spike as independent Poisson processes with autocovariance functions $r_i\delta(s)$. Error bars: standard error of the mean.

We next turned to networks with nonlinear neurons having threshold-quadratic transfer functions. We examined the same coding task and the same effects of neglecting correlations (assuming independent Poisson) activity, as well as the contribution of the nonlinear correction for the impact of pairwise correlations on firing rates (the one-loop correction). As for the linear network, neglecting correlations increased the Kullback-Leibler divergence of the conditional response distributions to a large degree for the population mean response and a small degree for the population vector (Fig. 20A, D). The nonlinear correction, while it did impact the firing rates (e.g. as in the absence of a feedforward stimulus, Fig. 15D), did not affect the D_{KL} of the two conditional response distributions.

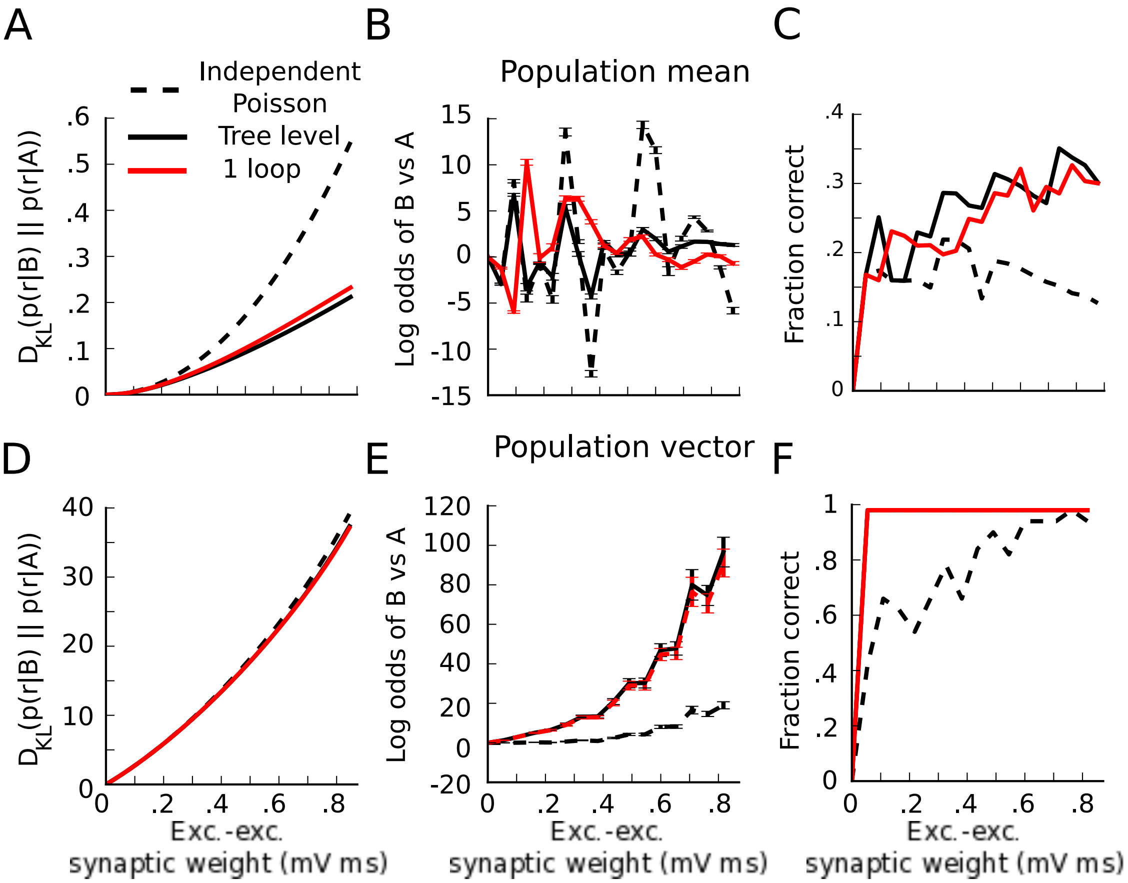


Fig 20. Stimulus discrimination with threshold-quadratic transfer functions. A) Distributions of population spike count of assembly B, $p(x|B)$ and $p(x|A)$, where x is the mean spike count of assembly B in a 200 ms window. Red: quadratic noise approximation, taking the distribution of the population spike count to be Gaussian and using the one-loop correction for both the first and second cumulants. Black: Independent Poisson approximation. C) Logarithm of the odds ratio, $\log \frac{p(x|B)}{p(x|A)}$, as a function of synaptic weight. D) Kullback-Leibler divergerence between the distributions of activity in response to each stimulus. E,F) Same as A,B but with the population vector response rather than the mean.

The nonlinear correction for the firing rates also had little effect on the average decoding performance of the population mean or vector response, compared to the effect of the the independent Poisson assumption (Fig. 20B,C,E,F). Thus while the nonlinearity affected firing rates, for this network and stimuli it did not have the effect of separating the response distributions. This highlights the potential diversity of nonlinear effects: while correlations can stretch response distributions to help or hurt coding performance, their impact on rates can translate the response distributions to help or hurt their discrimination, or to leave it unaffected (Fig. 17).

Discussion

Joint spiking activity between groups of neurons can control population coding and controls the evolution of network structure through plasticity. Theories for predicting the joint statistics of activity in model networks have been locally linear so far. We present a systematic and diagrammatic fluctuation expansion (or, in reference to those diagrams, loop expansion) for spike-train cumulants, which relies on a stochastic field theory for networks of stochastically spiking neurons. It allows the computation of arbitrary order joint cumulant functions of spiking activity and dynamical response functions which provide a window into the linear stability of the activity, as well as nonlinear corrections to all of those quantities.

Using this expansion, we investigated how nonlinear transfer can affect firing rates and fluctuations in population activity, imposing a dependence of rates on higher-order spiking cumulants. This coupling could significantly limit how strong synaptic interactions could become before firing rates lost stability. We then investigated the impact of correlations on a simple coding task, both directly through stretching the distributions of population activity and indirectly, by shifting the distributions through the impact of correlations on mean rates.

Relationship to other theoretical methods

A classic and highly influential tool for analyzing the dynamics of neural rate models with Gaussian-distributed synaptic weights is dynamical mean field theory, which reveals a transition to chaotic rate fluctuations in networks of rate units with Gaussian connectivity [51]. Dynamical mean field theory proceeds, briefly, by taking the limit of large networks and replacing interactions through the quenched heterogeneity of the synaptic weights by an effective Gaussian process mimicking their statistics. Recent extensions of dynamical mean field theory have incorporated a number of simple biological constraints, including positive-valued firing rates [52–54] and certain forms of cell type-specific connectivity [52, 55, 56]. In this framework, spiking is described in the limit of slow synapses as additive noise in the rates which can shift the transition to chaotic rate fluctuations to higher coupling strengths and smooth the dynamics near the transition [52, 57].

An alternative approach to dynamical mean-field theory is to start from the bottom up: to posit an inherently stochastic dynamics of single neurons and from it derive a set of equations for statistics of their activity [58, 59]. This approach provides a rigorous derivation of the popular rate models as the mean field of the underlying stochastic activity, as well as the opportunity to calculate higher-order activity statistics for the activity of a particular network [60–63]. This is the approach taken here with the popular and biologically motivated class of linear-nonlinear-Poisson models. A similar approach underlies linear response theory for computing spike train covariances [21, 64], which corresponds to the tree level of the loop expansion presented here. For integrate-and-fire neuron models receiving Poisson inputs, Fokker-Planck theory for the membrane potential distributions can be used to calculate the linear response function of an isolated neuron [65], which together with the synaptic filter and weight matrix determines the propagator.

Dynamics and stability in spiking networks

Fluctuations in large spiking networks Networks of excitatory and inhibitory neurons with instantaneous synapses have been shown, depending on their connectivity strengths and external drive, to exhibit a variety of dynamics, including the “classical” asynchronous state, oscillatory population activity, and strong, uncorrelated rate fluctuations [24, 66, 67]. The classical asynchronous state and oscillatory regimes exist in the presence of Poisson-like single-neuron activity, either due to external white noise or to internally generated high-dimensional chaotic fluctuations [68, 69]. Transitions between these modes correspond to bifurcations in which a given state loses stability. The present results allow one to compute these transition points with greater accuracy, by explicitly computing correlations of arbitrary order and, crucially, how these correlations "feed back" to impact firing rates and the stability of states with different rates.

Inhibitory-stabilized and supralinear-stabilized networks Beyond the overall stability of network states, an important general question is how firing rates depend on inputs in recurrent neural networks. Recent theoretical work shows how "inhibitory-stabilized" networks can have surprising dependencies on inputs, with rates trending in opposite directions from what one would at first expect [32]. Supralinear input-rate transfer in inhibitory-stabilized networks can explain a variety of physiological observations [70, 71]. Our results are therefore useful in predicting how correlations emerge and couple to firing rates in these networks. The impact of cell type-specific dynamics on dynamics and coding remains to be fully elucidated [72].

An open door to future work on structure-driven activity

An important next step is to develop explicit, tractable relations between the statistics of localized network structures involving only a few cells at once (connectivity motifs) and the statistics of the global features of network activity that we study here [38, 45, 46, 73]. The interplay of neural nonlinearities and synaptic motifs to control network activity is now ripe for investigation.

In summary, we revealed new ways that network connectivity and nonlinear neural dynamics combine to determine activity in neural populations. This work has developed a framework that will allow us to meet several goals applied to modern datasets on neural activity, cell types, and connectivity. In particular, it opens the door to studies revealing how local circuit architecture and the particular input-rate transfer of different electrophysiological cell types combine to control both the fidelity of neural coding and the structure of neural representations in different functional cell types.

Materials and Methods

non-Erdős-Rènyi network model

For Fig. 16, we generated the excitatory-excitatory connectivity with a truncated power-law degree distribution. The marginal distributions of the number of excitatory-excitatory synaptic inputs (in-degree) or outputs (out-degree) obeyed:

$$p(d) = \begin{cases} C_1 d^{\gamma_1}, & 0 \leq d \leq L_1 \\ C_2 d^{\gamma_2}, & L_1 \leq L_2 \\ 0, & \text{else} \end{cases} \tag{55}$$

where d is the in- or out-degree. The in- and out-degree distributions were then coupled by a Gaussian copula with correlation coefficient ρ to generate in- and out-degree lists. These lists generated likelihoods for each possible connection proportional to the in-degree of the postsynaptic neuron and the out-degree of the presynaptic neuron. We then sampled the excitatory-excitatory connectivity according to those likelihoods.

Path integral representation

Here we outline the derivation of a path integral formalism for a network of processes with nonlinear input-rate transfer, following methods developed in nonequilibrium statistical mechanics [74–78]. We will begin by developing the formalism for a simple model, where a spike train is generated stochastically with intensity given by some input $\nu(t)$. We will specify the cumulant generating functional of the spike train given $\nu(t)$ below.

The general strategy is to introduce an auxiliary variable, called the "response variable", whose dynamics will determine how information from a given configuration (e.g. the spike counts $n(t)$) at one time will effect future configurations of the dynamics. Introducing the response variable allows us to write the probability density functional for the process in an exponential form. The integrand of that exponential is called the "action" for the model, which can then be split into a "free" action (the piece bilinear in the configuration and response variables) and an "interacting" one (the remainder). Cumulants of the process can then be computed, in brief, by taking expectations of the configuration and response variables and the interacting action against the probability distribution given by the free action.

Let $n(t)$ be the number of spike events recorded since some fiducial time t_0. In a time bin dt, Δn events are generated with some distribution $P(\Delta n)$ and added to $n(t)$. Let the events generated in any two time bins be conditionally independent given some inhomogeneous rate $\nu(t)$, so that $P(\Delta n) = P(\Delta n|\nu)$. So, assuming that initially $n(t_0) = 0$, the probability density functional of the vector of events over M time bins is:

$$P[\Delta n(s) : s \leq t] = \prod_{i=1}^{M} P(\Delta n_i|\nu_i) = \prod_{i=1}^{j} \int \frac{d\tilde{n}_i}{2\pi i} e^{-\tilde{n}_i \Delta n_i} P(\tilde{n}_i|\nu_i) = \prod_{i=1}^{j} \int \frac{d\tilde{n}_i}{2\pi i} e^{-\tilde{n}_i \Delta n_i + W[\tilde{n}_i|\nu_i]} \tag{56}$$

where $P(\tilde{n}_i|\nu_i)$ is the Laplace transform of $P(\Delta n_i|\nu_i)$ and $W[\tilde{n}_i|\nu_i]$ is the cumulant generating functional for the auxiliary variable. In the third step we have written the distribution of $P(\Delta n_i)$ as the inverse Laplace transform of the Laplace transform. The Laplace transform variable $\tilde{n}_i$ is our auxiliary response variable. In the fourth step we identified the Laplace transform of the probability density functional as the moment generating functional, so that $W[\tilde{n}_i|\nu_i]$ is the cumulant generating functional of the spike count. Note that these are complex integrals. The contour for the integration over $\tilde{n}_i$ is parallel to the imaginary axis.

Taking the continuum limit $M \to \infty, dt \to 0$ then yields the probability density functional of the spike train process $\dot{n}$:

$$P[\dot{n}] = \int \mathcal{D}\tilde{n}(t) e^{-\int dt(\tilde{n}(t)\dot{n}(t) - W[\tilde{n}(t)])} \tag{57}$$

where $\mathcal{D}\tilde{n}(t) = \lim_{M\to\infty} \prod_{i=1}^{M} \frac{d\tilde{n}_i}{2\pi i}$ and $\dot{n} = \frac{dn}{dt}$ and we suppress the conditional dependence of $\tilde{n}(t)$ on $\nu(t)$. In the continuum limit the integral is a functional or path integral over realizations of $\tilde{n}(t)$. We will call the integrand of the exponential in Eq. 57 the action:

$$S[\tilde{n}, n] = \int dt \big(\tilde{n}\dot{n} - W[\tilde{n}]\big) \tag{58}$$

We have slightly abused notation here in that a factor of $1/dt$ has been absorbed into $W[\tilde{n}]$. We will justify this below.

We have not yet specified the conditional distribution of the events given the input $\nu(t)$, leaving $W[\tilde{n}(t)]$ unspecified. Here, we will take the events to be conditionally Poisson, so that

$$W[\tilde{n}] = \big(e^{\tilde{n}} - 1\big)\, \nu(t) \tag{59}$$

(In the continuum limit, the rate $\nu(t)$ allowed us to absorb the factor of $1/dt$ into W. A finite size time bin would produce $\nu(t)dt$ events in bin dt.)

This representation of the probability density functional yields the joint cumulant generating functional (CGF) of n and $\tilde{n}$:

$$Z[J,\tilde{J}] = \int \mathcal{D}n(t) \int \mathcal{D}\tilde{n}(t)\, e^{-S[\tilde{n},n]+J\tilde{n}+\tilde{J}n} \tag{60}$$

and the cumulant generating functional of $\dot{n}$:

$$Z[\tilde{J}] = \int \mathcal{D}n(t) \int \mathcal{D}\tilde{n}(t)\, e^{-S[\tilde{n},n]+\tilde{J}\dot{n}} \tag{61}$$

The above strictly applies only to the inhomogeneous Poisson process. This formalism is adapted to the self-exciting process by introducing conditional dependence of the rate $\nu(t)$ on the *previous* spiking history. In the discrete case, before taking the limit $M \to \infty$, we say that the rate $\nu_i = \phi[n_{i-}]$, where ϕ is some positive function and n_{i-} indicates all spiking activity up to but not including bin i. This requirement is equivalent to an Ito interpretation for the measure on the stochastic process $\dot{n}(t)$. Because of this assumption, the previous derivation holds and we can write

$$W[\tilde{n}] = \left(e^{\tilde{n}(t)} - 1\right)\phi(n(<t)) \tag{62}$$

In the continuum limit, there is an ambiguity introduced by the appearance of the time variable t in both $\tilde{n}(t)$ and $n(t)$. This is resolved in the definition of the measure for the functional integral, and affects the definition of the linear response (below). Again, this is a manifestation of the Ito assumption for our process.

The specific model used in this paper assumes a particular form for the argument of ϕ. We assume that the input is given by

$$\nu(t) = \phi(g * \dot{n}(t) + \lambda(t)) \tag{63}$$

where $g(t)$ is a filter that defines the dynamics of the process in question and $\lambda(t)$ is an inhomogeneous rate function.

The result is that the action for non-linearly self-exciting process is given by

$$S[\tilde{n},n] = \int dt\big(\tilde{n}\dot{n} - \left(e^{\tilde{n}(t)} - 1\right)\phi(g * \dot{n}(t) + \lambda(t))\big) \tag{64}$$

The only extension required to move from the above action to the network model is to introduce indices labelling the neurons and couplings specific for each neuron pair. Nothing of substance is altered in the above derivation and we are left with

$$S[\tilde{n},n] = \sum_i \int dt\big(\tilde{n}_i\dot{n}_i - \left(e^{\tilde{n}_i(t)} - 1\right)\phi(\sum_j g_{ij} * \dot{n}_j(t) + \lambda_i(t))\big) \tag{65}$$

Mean field expansion and derivation of Feynman rules.

We could use the above action in order to derive Feynman rules for these processes. The expansions so described would be equivalent to our initial expansions before resumming (the sets of diagrams that use dashed lines). These would describe an expansion about $\dot{n}(t) = 0$. We can arrive at this expansion by separating the action into two pieces, called the "free" action and the "interacting" action: $S[\tilde{n},n] = S_0[\tilde{n},n] + S_V[\tilde{n},n]$. The free action, $S_0[\tilde{n},n]$ is defined by the bilinear component of $S[\tilde{n},n]$ in an expansion around 0, i.e.

$$S_0[\tilde{n},n] = \sum_{i,j} \int dt dt' \tilde{n}_i(t) K_{ij}(t,t') n_j(t') \tag{66}$$

for some operator $K_{ij}(t,t')$. Define

$$\langle n_i(t)\tilde{n}_j(t')\rangle = \Delta_{ij}(t,t') \tag{67}$$

Taking the expectation with respect to the probability density given by the free action yields, by completion of squares,

$$\int ds K_{ik}(t,s)\Delta_{kj}(s,t') = \delta(t-t')\delta_{ij} \tag{68}$$

so that K is the operator inverse of Δ.

Computing moments requires functional integrals of the form

$$\langle\prod_i n_i(t_i)\prod_j \tilde{n}_j(t_j)\rangle = \int \mathcal{D}n(t)\mathcal{D}\tilde{n}(t)\prod_i n_i(t_i)\prod_j \tilde{n}_j(t_j)e^{-S[\tilde{n},n]} \tag{69}$$

We Taylor expand the term e^{-S_V}. This leaves a series of moments with respect to the free action S_0, each of which can be computed using the formula above. Feynman rules are simply a recipe for the terms of this expansion.

The “resummed” rules that we present in the Results arise from first performing a slight change of variables in the integrand. Instead of considering the fluctuations about $n(t) = 0$, we shift the configuration and response variables by their mean field solutions. Defining $r_i(t) = \langle\dot{n}_i(t)\rangle$ and $\tilde{r}_i(t) = \langle\tilde{n}_i(t)\rangle$, these are determined by

$$\begin{aligned}\tilde{r}_i(t) &= 0\\ r_i(t) &= \phi\left(\sum_j g_{ij} * r_j(t) + \lambda_i(t)\right)\end{aligned} \tag{70}$$

We shift by these solutions by defining

$$\delta\dot{n}_i(t) = \dot{n}_i(t) - r_i(t) \tag{71}$$

This leaves us with the action

$$S[\tilde{n}_i, n_i] = \sum_i \int dt \big(\tilde{n}_i\delta\dot{n}_i - \left(e^{\tilde{n}_i(t)} - 1\right)\phi\left(\sum_j g_{ij} * (\dot{n}_j(t) + r_j(t)) + \lambda_i(t)\right)\big) + \tilde{n}_i(t)r_i(t) \tag{72}$$

We introduce diagrams for the expansion about mean field theory by using solid lines for the propagator, $\Delta_{ij}(t,t')$ and wavy edges for the filter g_{ij}, open circles for the “external vertices” that represent the moment being computed, and solid dots for the vertices. Each term in the interacting action defined according to the mean field expansion above will produce a separate vertex for the Feynman rules. The approach here is a standard device in field theory and can be found many texts, for one example see [37]. The rules so derived are shown in Figure 11.

Population coding

We measured the activity of one assembly (“B”) of excitatory neurons in the recurrent excitatory-inhibitory network, in response to activation input A or input B. We measured the activity of assembly B as the spike count in a 200 ms window x, and approximated the conditional response distributions $p(x|A)$ and $p(x|B)$ as Gaussians with theoretically predicted mean μ_α and variance σ_α^2 for $\alpha = A$ or B. The variance of the population spike count is given by the integral of the autocovariance function of the population spike train against the self-convolution of the box counting window [79]. For large counting windows this approaches the integral of the autocovariance function; we used this approximation, which provided a good measure of variability over $T = 200$ ms (Fig. 18).

We examined the coding performance with both the population-averaged spike count of assembly B, x and the population vector of spike counts in assembly B, $\vec{x}$. When using the vector of spike counts, we used the theoretically predicted covariance matrices. The Kullback-Leibler divergence between the distributions of the population spike counts is:

$$D_{KL}\big(p(x|B)||p(x|A)\big) = \ln\frac{\sigma_A}{\sigma_B} + \frac{\sigma_B + (\mu_B - \mu_A)^2}{2\sigma_A} - \frac{1}{2} \tag{73}$$

and that between two multivariate Gaussian distributions by:

$$D_{KL}\big(p(\vec{x}|B)||p(\vec{x}|A)\big) = \frac{1}{2}\left(\mathrm{Tr}\left(\Sigma_A^{-1}\Sigma_B\right) + \left(\vec{\mu}_A - \vec{\mu}_B\right)^T \Sigma_A^{-1}\left(\vec{\mu}_A - \vec{\mu}_B\right) - N_B + \ln\left(\frac{\det\Sigma_A}{\det\Sigma_B}\right)\right) \tag{74}$$

where Σ_α is the covariance matrix of the neurons in assembly α and N_B the size of assembly B (100 neurons here).

Table 1. Model parameters

Parameter	Description	Value
N_E	Number of excitatory neurons	200
N_I	Number of inhibitory neurons	40
N_{FF}	Number of feed-forward neurons	50
p_{EE}	Excitatory-excitatory connection probability	0.2
p_{EI}	Inhibitory-excitatory connection probability	0.2
p_{IE}	Excitatory-excitatory connection probability	0.2
p_{II}	Excitatory-excitatory connection probability	0.2
τ	Time constant for postsynaptic potentials	10 ms
$g(t)$	Shape of postsynaptic potentatials	$\frac{t}{\tau^2}e^{-t/\tau}$

Acknowledgments

We thank Brent Doiron and Robert Rosenbaum for helpful comments on the manuscript. We acknowledge the support of the NSF through grants DMS-1514743 and DMS-1056125 (ESB), DMS-1517629 (KJ), and NSF/NIGMS-R01GM104974 (KJ), as well as the Simons Fellowships in Mathematics (ESB and KJ). The authors wish to thank the Allen Institute founders, Paul G. Allen and Jody Allen, for their vision, encouragement and support.

References

1. Sejnowski TJ. Storing covariance with nonlinearly interacting neurons. Journal of Mathematical Biology. 1977;4(4):303–321. doi:10.1007/BF00275079.
2. Bienenstock EL, Cooper LN, Munro PW. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. The Journal of Neuroscience. 1982;2(1):32–48.
3. Gerstner W, Kempter R, van Hemmen JL, Wagner H. A neuronal learning rule for sub-millisecond temporal coding. Nature. 1996;383(6595):76–78. doi:10.1038/383076a0.
4. Pfister JP, Gerstner W. Triplets of Spikes in a Model of Spike Timing-Dependent Plasticity. The Journal of Neuroscience. 2006;26(38):9673–9682. doi:10.1523/JNEUROSCI.1425-06.2006.

5. Ocker GK, Litwin-Kumar A, Doiron B. Self-Organization of Microcircuits in Networks of Spiking Neurons with Plastic Synapses. PLoS Comput Biol. 2015;11(8):e1004458. doi:10.1371/journal.pcbi.1004458.
6. Tannenbaum NR, Burak Y. Shaping Neural Circuits by High Order Synaptic Interactions. PLOS Comput Biol. 2016;12(8):e1005056. doi:10.1371/journal.pcbi.1005056.
7. Abeles M. Role of the cortical neuron: integrator or coincidence detector? Israel Journal of Medical Sciences. 1982;18(1):83–92.
8. Usrey WM, Reppas JB, Reid RC. Paired-spike interactions and synaptic efficacy of retinal inputs to the thalamus. Nature. 1998;395(6700):384–387. doi:10.1038/26487.
9. Bruno RM, Sakmann B. Cortex is driven by weak but synchronously active thalamocortical synapses. Science (New York, NY). 2006;312(5780):1622–1627. doi:10.1126/science.1124593.
10. Histed MH, Maunsell JHR. Cortical neural populations can guide behavior by integrating inputs linearly, independent of synchrony. Proceedings of the National Academy of Sciences. 2014;111(1):E178–E187. doi:10.1073/pnas.1318750111.
11. Salinas E, Sejnowski TJ. Impact of correlated synaptic input on output firing rate and variability in simple neuronal models. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience. 2000;20(16):6193–6209.
12. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. Nat Rev Neurosci. 2006;7(5):358–66. doi:10.1038/nrn1888.
13. Panzeri S, Macke JH, Gross J, Kayser C. Neural population coding: combining insights from microscopic and mass signals. Trends in Cognitive Sciences. 2015;19(3):162–172. doi:10.1016/j.tics.2015.01.002.
14. Seriès P, Latham PE, Pouget A. Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. Nature Neuroscience. 2004;7(10):1129–1135. doi:10.1038/nn1321.
15. Josić K, Shea-Brown E, Doiron B, de la Rocha J. Stimulus-dependent correlations and population codes. Neural Comp. 2009;21:2774–2804.
16. Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A. Information-limiting correlations. Nature Neuroscience. 2014;17(10):1410–1417. doi:10.1038/nn.3807.
17. Zylberberg J, Cafaro J, Turner MH, Shea-Brown E, Rieke F. Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code. Neuron. 2016;89(2):369–383. doi:10.1016/j.neuron.2015.11.019.
18. Franke F, Fiscella M, Sevelev M, Roska B, Hierlemann A, Azeredo da Silveira R. Structures of Neural Correlation and How They Favor Coding. Neuron. 2016;89(2):409–422. doi:10.1016/j.neuron.2015.12.037.
19. Hawkes AG. Spectra of some self-exciting and mutually exciting point processes. Biometrika. 1971;58(1):83–90. doi:10.1093/biomet/58.1.83.
20. Brillinger D. Estimation of the second-order intensities of a bivariate stationary point process. Journal of the Royal Statistical Society Series B (Methodological). 1976;38(1):60–66.
21. Doiron B, Lindner B, Longtin A, Maler L, Bastian J. Oscillatory activity in electrosensory neurons increases with the spatial correlation of the stochastic input stimulus. Phys Rev Let. 2004;93(4).

22. Trousdale J, Hu Y, Shea-Brown E, Josić K. Impact of Network Structure and Cellular Response on Spike Time Correlations. PLoS Computational Biology. 2012;8(3):e1002408. doi:10.1371/journal.pcbi.1002408.
23. Ginzburg I, Sompolinsky H. Theory of correlations in stochastic neural networks. Physical Review E. 1994;50(4):3171–3191. doi:10.1103/PhysRevE.50.3171.
24. Brunel N. Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. Journal of Computational Neuroscience. 2000;8(3):183–208. doi:10.1023/A:1008925309027.
25. Mattia M, Del Giudice P. Population dynamics of interacting spiking neurons. Physical Review E. 2002;66(5):051917. doi:10.1103/PhysRevE.66.051917.
26. Kass RE, Eden U, Brown E. Analysis of neural data. 1st ed. Springer series in statistics. Springer-Verlag; 2014.
27. Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K. The mechanics of state-dependent neural correlations. Nature Neuroscience. 2016;19(3):383–393. doi:10.1038/nn.4242.
28. Abeles M. Corticonics: Neural circuits of the cerebral cortex. Cambridge University Press; 1991.
29. Diesmann M, Gewaltig MO, Aertsen A. Stable propagation of synchronous spiking in cortical neural networks. Nature. 1999;402(6761):529–533. doi:10.1038/990101.
30. Reyes AD. Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. Nature Neuroscience. 2003;6(6):593–599. doi:10.1038/nn1056.
31. Kumar A, Rotter S, Aertsen A. Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. Nat Rev Neurosci. 2010;11:615–627.
32. Tsodyks MV, Skaggs WE, Sejnowski TJ, McNaughton BL. Paradoxical Effects of External Modulation of Inhibitory Interneurons. The Journal of Neuroscience. 1997;17(11):4382–4388.
33. Risken H. The Fokker-Planck equation: methods of solution and applications. 3rd ed. Haken H, editor. Springer-Verlag; 1996.
34. Hawkes AG, Oakes D. A cluster process representation of a self-exciting process. Journal of Applied Probability. 1974;11:493–503.
35. Saichev AI, Sornette D. Generating functions and stability study of multivariate self-excited epidemic processes. The European Physical Journal B. 2011;83(2):271–282. doi:10.1140/epjb/e2011-20298-3.
36. Jovanović S, Hertz J, Rotter S. Cumulants of Hawkes point processes. Physical Review E. 2015;91(4):042802. doi:10.1103/PhysRevE.91.042802.
37. Zinn-Justin J. Quantum Field Theory and Critical Phenomena. Clarendon Press; 2002.
38. Pernice V, Staude B, Cardanobile S, Rotter S. How structure determines correlations in neuronal networks. PLoS Comput Biol. 2011;7(5):e1002059. doi:10.1371/journal.pcbi.1002059.
39. Ostojic S, Brunel N, Hakim V. How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. J Neurosci. 2009;29(333):10234–10253.
40. Heeger DJ. Normalization of cell responses in cat striate cortex. Visual Neuroscience. 1992;9(2):181–197.

41. Troyer TW, Krukowski AE, Priebe NJ, Miller KD. Contrast-Invariant Orientation Tuning in Cat Visual Cortex: Thalamocortical Input Tuning and Correlation-Based Intracortical Connectivity. The Journal of Neuroscience. 1998;18(15):5908–5927.

42. Miller KD, Troyer TW. Neural noise can explain expansive, power-law nonlinearities in neural response functions. J Neurophysiol. 2002;87(2):653–659.

43. Reid RC. Divergence and reconvergence: multielectrode analysis of feedforward connections in the visual system. In: Progress in Brain Research. vol. 130. Elsevier; 2001. p. 141–154.

44. Iyer R, Menon V, Buice M, Koch C, Mihalas S. The Influence of Synaptic Weight Distribution on Neuronal Population Dynamics. PLOS Comput Biol. 2013;9(10):e1003248. doi:10.1371/journal.pcbi.1003248.

45. Hu Y, Trousdale J, Josić K, Shea-Brown E. Motif statistics and spike correlations in neuronal networks. Journal of Statistical Mechanics: Theory and Experiment. 2013;2013(03):P03012. doi:10.1088/1742-5468/2013/03/P03012.

46. Hu Y, Trousdale J, Josić K, Shea-Brown E. Local paths to global coherence: Cutting networks down to size. Physical Review E. 2014;89(3):032802. doi:10.1103/PhysRevE.89.032802.

47. Roxin A. The role of degree distribution in shaping the dynamics in networks of sparsely connected spiking neurons. Frontiers in Computational Neuroscience. 2011;5:8. doi:10.3389/fncom.2011.00008.

48. Zhao L, Beverlin BI, Netoff T, Nykamp DQ. Synchronization from second order network connectivity statistics. Frontiers in Computational Neuroscience. 2011;5:28. doi:10.3389/fncom.2011.00028.

49. Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB. Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol. 2005;3(3):e68. doi:10.1371/journal.pbio.0030068.

50. Cossell L, Iacaruso MF, Muir DR, Houlton R, Sader EN, Ko H, et al. Functional organization of excitatory synaptic strength in primary visual cortex. Nature. 2015;518(7539):399–403. doi:10.1038/nature14182.

51. Sompolinsky H, Crisanti A, Sommers HJ. Chaos in Random Neural Networks. Physical Review Letters. 1988;61(3):259–262. doi:10.1103/PhysRevLett.61.259.

52. Kadmon J, Sompolinsky H. Transition to Chaos in Random Neuronal Networks. Physical Review X. 2015;5(4):041030. doi:10.1103/PhysRevX.5.041030.

53. Harish O, Hansel D. Asynchronous Rate Chaos in Spiking Neuronal Circuits. PLOS Comput Biol. 2015;11(7):e1004266. doi:10.1371/journal.pcbi.1004266.

54. Mastrogiuseppe F, Ostojic S. Intrinsically-generated fluctuating activity in excitatory-inhibitory networks. arXiv:160504221 [q-bio]. 2016;.

55. Aljadeff J, Stern M, Sharpee T. Transition to Chaos in Random Networks with Cell-Type-Specific Connectivity. Physical Review Letters. 2015;114(8):088101. doi:10.1103/PhysRevLett.114.088101.

56. Aljadeff J, Renfrew D, Vegué M, Sharpee TO. Low-dimensional dynamics of structured random networks. Physical Review E. 2016;93(2):022302. doi:10.1103/PhysRevE.93.022302.

57. Goedeke S, Schuecker J, Helias M. Noise dynamically suppresses chaos in random neural networks. arXiv:160301880 [nlin, q-bio]. 2016;.

58. Buice MA, Chow CC. Beyond mean field theory: statistical field theory for neural networks. Journal of Statistical Mechanics (Online). 2013;2013:P03003. doi:10.1088/1742-5468/2013/03/P03003.

59. Bressloff PC. Path-Integral Methods for Analyzing the Effects of Fluctuations in Stochastic Hybrid Neural Networks. The Journal of Mathematical Neuroscience. 2015;5(1). doi:10.1186/s13408-014-0016-z.

60. Ohira T, Cowan JD. Master-equation approach to stochastic neurodynamics. Physical Review E. 1993;48(3):2259–2266. doi:10.1103/PhysRevE.48.2259.

61. Buice MA, Cowan JD. Field-theoretic approach to fluctuation effects in neural networks. Physical Review E. 2007;75(5):051919. doi:10.1103/PhysRevE.75.051919.

62. Bressloff P. Stochastic Neural Field Theory and the System-Size Expansion. SIAM Journal on Applied Mathematics. 2009;70(5):1488–1521. doi:10.1137/090756971.

63. Buice MA, Chow CC, Cowan JD. Systematic fluctuation expansion for neural network activity equations. Neural Comp. 2010;22(377-426).

64. de la Rocha J, Doiron B, Shea-Brown E, Josić K, Reyes A. Correlation between neural spike trains increases with firing rate. Nature. 2007;448(7155):802–6. doi:10.1038/nature06028.

65. Fourcaud N, Brunel N. Dynamics of the Firing Probability of Noisy Integrate-and-Fire Neurons. Neural Computation. 2002;14(9):2057–2110. doi:10.1162/089976602320264015.

66. Ledoux E, Brunel N. Dynamics of networks of excitatory and inhibitory neurons in response to time-dependent inputs. Frontiers in Computational Neuroscience. 2011;5:25. doi:10.3389/fncom.2011.00025.

67. Ostojic S. Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. Nature Neuroscience. 2014;17(4):594–600. doi:10.1038/nn.3658.

68. van Vreeswijk C, Sompolinsky H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. Science (New York, NY). 1996;274(5293):1724–1726.

69. Renart A, Rocha Jdl, Bartho P, Hollender L, Parga N, Reyes A, et al. The Asynchronous State in Cortical Circuits. Science. 2010;327(5965):587–590. doi:10.1126/science.1179850.

70. Ozeki H, Finn IM, Schaffer ES, Miller KD, Ferster D. Inhibitory Stabilization of the Cortical Network Underlies Visual Surround Suppression. Neuron. 2009;62(4):578–592. doi:10.1016/j.neuron.2009.03.028.

71. Rubin DB, Van Hooser SD, Miller KD. The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. Neuron. 2015;85(2):402–417. doi:10.1016/j.neuron.2014.12.026.

72. Litwin-Kumar A, Rosenbaum R, Doiron B. Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. Journal of Neurophysiology. 2016;115(3):1399–1409. doi:10.1152/jn.00732.2015.

73. Jovanović S, Rotter S. Interplay between Graph Topology and Correlations of Third Order in Spiking Neuronal Networks. PLOS Comput Biol. 2016;12(6):e1004963. doi:10.1371/journal.pcbi.1004963.

74. Martin PC, Siggia ED, Rose HA. Statistical Dynamics of Classical Systems. Physical Review A. 1973;8(1):423–437. doi:10.1103/PhysRevA.8.423.

75. Doi M. Second quantization representation for classical many-particle system. Journal of Physics A: Mathematical and General. 1976;9(9):1465. doi:10.1088/0305-4470/9/9/008.

76. Doi M. Stochastic theory of diffusion-controlled reaction. Journal of Physics A: Mathematical and General. 1976;9(9):1479. doi:10.1088/0305-4470/9/9/009.

77. Peliti L. Path integral approach to birth-death processes on a lattice. Journal de Physique. 1985;46(9):1469–1483. doi:10.1051/jphys:019850046090146900.

78. Täuber UC. Field-Theory Approaches to Nonequilibrium Dynamics. In: Henkel M, Pleimling M, Sanctuary R, editors. Ageing and the Glass Transition. No. 716 in Lecture Notes in Physics. Springer Berlin Heidelberg; 2007. p. 295–348. Available from: `http://link.springer.com/chapter/10.1007/3-540-69684-9_7`.

79. Cox D, Isham V. Point Processes. Monographs on Statistics and Applied Probability. CRC Press; 1980.