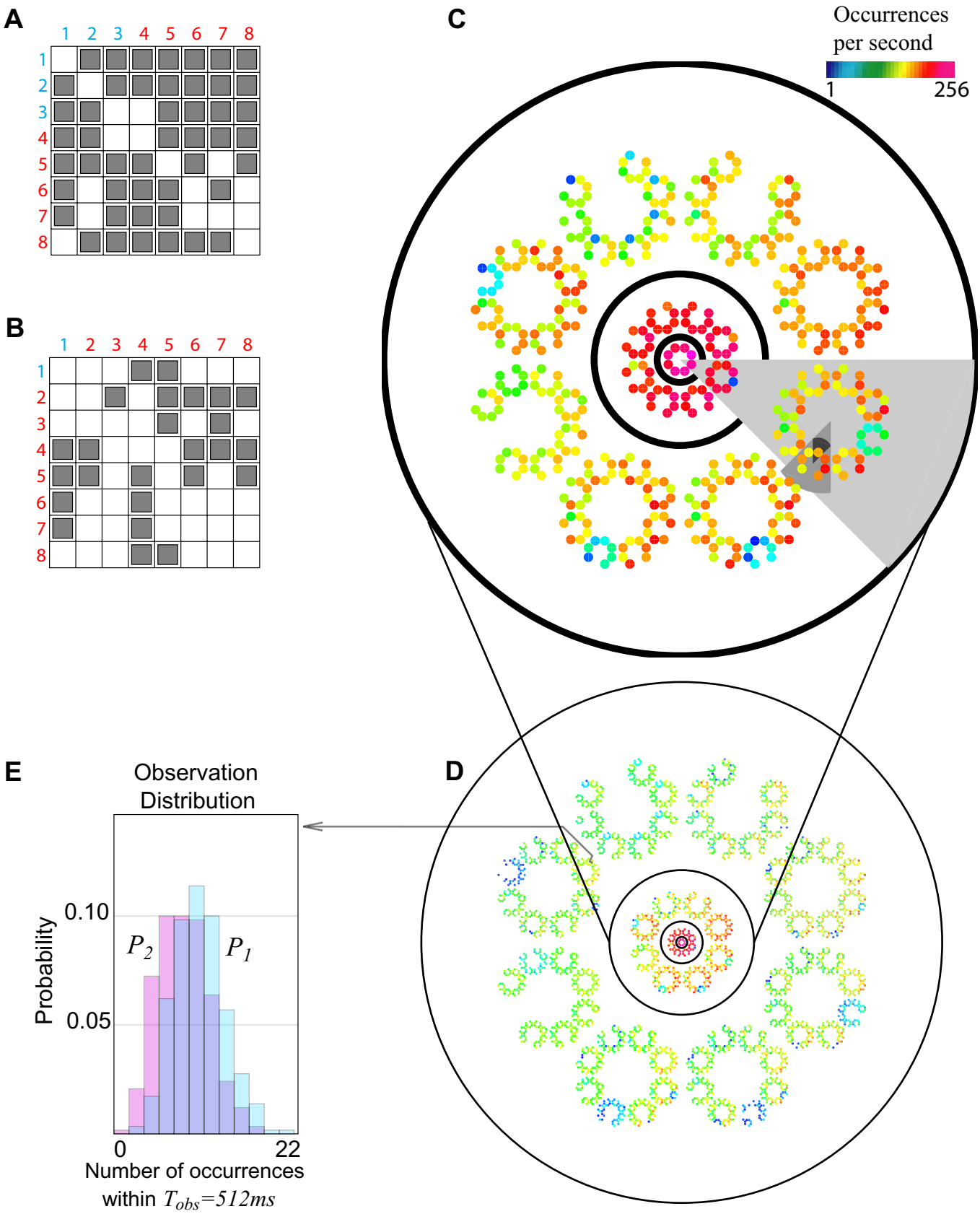

A
1 2 3 4 5 6 7 8
B
1 2 3 4 5 6 7 8
C
Occurrences
per second
1
256
D
E
Observation
Distribution
Probability
0.10
0.05
P_2
P_1
0
22
Number of occurrences
within T_{obs}=512ms

Supplementary Figure S1 (Best viewed electronically): **The methods and the utility of our event-tree analysis** within model networks of conductance-based I&F neurons, driven by independent Poisson input. For reference, we briefly describe the standard conductance-based integrate-and-fire (I&F) point neuron (Koch, 1999). Given a network of N I&F neurons with label i, each with voltage V_i, conductances $G_i^{\text{input}}, G_i^{\text{ex}}, G_i^{\text{in}}$, and type $\tau_i = \{\text{ex}, \text{in}\}$, its dynamics is governed by

$$\frac{d}{dt}V_i(t) = -G^L(V_i(t) - \epsilon^L) - G_i^{\text{input}}(t)\,(V_i(t) - \epsilon^{\text{ex}}) - G_i^{\text{ex}}(t)\,(V_i(t) - \epsilon^{\text{ex}}) - G_i^{\text{in}}(t)\,\left(V_i(t) - \epsilon^{\text{in}}\right)$$

$$G_i^{\text{ex}}(t) = \sum_j \sum_k S^{\text{ex}}_{\tau_i,\tau_j}\Delta_{i,j}e^{-\frac{t-T_{j,k}}{\sigma^{\text{ex}}}}\theta(t - T_{j,k}),$$

$$G_i^{\text{in}}(t) = \sum_j \sum_k S^{\text{in}}_{\tau_i,\tau_j}\Delta_{i,j}e^{-\frac{t-T_{j,k}}{\sigma^{\text{in}}}}\theta(t - T_{j,k}),$$

$$G_i^{\text{input}}(t) = f\sum_k F_{\tau_i}e^{-\frac{t-T^F_{i,k}}{\sigma^{\text{ex}}}}\theta(t - T^F_{i,k}),$$

where $\theta\,(t)$ is the Heaviside function. The voltage $V_i(t)$ evolves continuously until it reaches a threshold $V_i(t) = \epsilon^T$. At this point in time the i^{th} neuron produces a spike (the k^{th} spike of neuron i is recorded as $T_{i,k}$), and the voltage V_i is reset to ϵ^R, and held there for an absolute refractory period of τ_{ref} ms. Here, G^L is the leak conductance and ϵ^L is leakage potential. The various synaptic conductances G^Q ($Q = \text{ex}, \text{in}$) are characterized by their different decay time scales σ^Q and reversal potentials ϵ^Q. Each spike from the j^{th} neuron gives rise to an instantaneous increase in the Q-type conductance of neuron i of magnitude $S^Q_{\tau_i,\tau_j}\Delta_{i,j}$. The coupling matrix $\Delta_{i,j}$ (with entries either 0 or 1) indicates the network connectivity. The coupling strengths $S^Q_{\tau_i,\tau_j}$ (indicating the amount of Q-type conductance added to τ_i-type neurons by τ_j-type firing events) only depend on the types of neurons i and j, and not directly on their index. The system is also driven by feedforward input. The k^{th} input spike from the feedforward input to the i^{th} neuron is denoted by $T^F_{i,k}$, and instantaneously increases that neuron's G^{input} conductance by magnitude fF_{τ_i}. For the simulations presented in Fig. 2 in the main text, we use the following parameters (with conductance in units of $[\text{ms}]^{-1}$): $G^L = 0.00667\text{ms}^{-1}$, $\epsilon^R = \epsilon^L = -60.95\text{mV}$, $\epsilon^T = -48\text{mV}$, $\epsilon^{\text{ex}} = 0\text{mV}$, $\epsilon^{\text{in}} = -70\text{mV}$, $\tau_{\text{ref}} = 2\text{ms}$, $\sigma^{\text{ex}} = 2\text{ms}$, $\sigma^{\text{in}} = 10\text{ms}$ with input strength factors $F_{\text{ex}} = 1$, $F_{\text{in}} = 0.4$. For the three regimes described in Fig. 2, the inputs to each neuron in the system are independent Poisson processes with rate $\nu \sim 0.5\text{kHz}$ and strength $f \sim 0.005$.

(i) For the **phase-oscillator regime** in Fig. 2*A*, the coupling strengths are given by

$$S^{\text{ex}}_{\text{ex,ex}} = 0.350,\ S^{\text{ex}}_{\text{in,ex}} = 0.165,\ S^{\text{in}}_{\text{ex,in}} = 0.361,\ S^{\text{in}}_{\text{in,in}} = 0.0424\,,$$

and the network connectivity illustrated in Panel *A*. Panel *A* indicates the coupling matrix $\Delta_{i,j}$, with grey squares indicating entries of 1, and blank squares indicating entries of 0. The neurons labeled by blue numerals denote inhibitory neurons, and the red denote excitatory neurons.

(ii) For the **bursty-oscillator regime** in Fig. 2*B*, the coupling strengths are given by

$$S^{\text{ex}}_{\text{ex,ex}} = 0.350,\ S^{\text{ex}}_{\text{in,ex}} = 0.165,\ S^{\text{in}}_{\text{ex,in}} = 0.148,\ S^{\text{in}}_{\text{in,in}} = 0.0424.$$

The network connectivity is the same as in Panel A.

(iii) For the **sustained-firing regime** in Fig. $2C$, the coupling strengths are given by

$$S^{\rm ex}_{\rm ex,ex} = 0.118,\ S^{\rm ex}_{\rm in,ex} = 0.165,\ S^{\rm in}_{\rm ex,in} = 0.0856,\ S^{\rm in}_{\rm in,in} = 0.0751,$$

and the network connectivity is illustrated in Panel B.

We emphasize that there is a broad range of parameter values of coupling strengths and connectivity matrices that produce similar dynamics with the same qualitative coding properties. Specifically, the general phenomena of (a) *robust event-tree discriminability* and (b) the ability of event-tree analysis to distinguish between multiple stimuli (which may differ along distinct stimulus dimensions) persist for a large class of parameter values, and for a wide variety of dynamical regimes similar to the bursty-oscillator and sustained-firing regime.

Event trees are extended space-time entities. In Panel C, we display a representation of the 3-event-tree collected over $T_{obs} = \infty$ (with $\alpha = 2$ms) of the sustained-firing regime (Fig. $2C$) under stimulus I_1 (with input rate $\nu_1 = 0.5$kHz and strength $f_1 = 0.005$). The colors, in general, correspond to occurrence rate, plotted logarithmically from 1 occurrence per second to 256 occurrences per second. Event-chains with an occurrence rate less than 1 occurrence per second have intentionally not been plotted. The representation of an event tree begins by fixing the number N of neurons (or the number N of coarse-grained sets of neurons), the maximum length $m_{\max}$ of the event chains in the tree, the time T_{obs} of observation, and the coarse-grained time interval α. The tree is then represented by a collection of rings (circles). On each ring, the discrete set of angles $\theta_j = 2\pi\,(j - 0.5)\,/N, j = 1, \cdots N$ labels the N neurons (or coarse-grained sets of neurons).

The collection of rings which represents the 3-event tree in panel C is organized as follows. In panel C, the three black circles indicate the organization of event-chains by the depth of events in the chain. The central black circle contains 1-event-chains, the interior annulus contains 2-event-chains, and the outer annulus contains 3-event-chains. The location of each m-event-chain is specified by m angles, which correspond to the indices of the m neuronal events defining that event-chain. For example, consider the 3-event-chain $\sigma^8 \to \sigma^6 \to \sigma^2$ indicated by the three grey pie-slices. The smallest darkest pie slice contains a single circle. This single circle lies within a hierarchical structure — a ring of rings of rings. The location of this single circle can be defined in terms of the following three angles: the angle of the major ring in which it lies (the largest light grey pie slice corresponds to the orientation angle $2\pi\,(j - 0.5)\,/8$ with $j = 8$, since the first event is σ^8), followed by the angle of the medium ring in which it lies (the medium grey pie slice corresponds to the orientation angle $2\pi\,(j - 0.5)\,/8$ with $j = 6$, since the second event is σ^6), followed again by the angle of the smallest ring in which it lies (the smallest darkest grey pie slice corresponds to the orientation angle $2\pi\,(j - 0.5)\,/8$ with $j = 2$, as the third event is σ^2). Thus, the color of this single circle corresponds to the occurrence rate of the 3-event-chain $\sigma^8 \to \sigma^6 \to \sigma^2$. All of the 3-event-chains are organized this way within the outer annulus. All the 2-event-chains can be organized in a similar hierarchical structure — namely, a ring of rings. This smaller ring of rings has been scaled, and placed in the middle annulus. Finally, the single ring corresponding to the set of 1-event-chains is also appropriately scaled and placed in the

center of the diagram. Note that the color coding of the 1-event chain merely labels the firing rates of each of the N neurons on the ring.

More generally, in this hierarchial organization, the m-event-chain $\sigma^{j_1} \to \cdots \to \sigma^{j_m}$ is plotted at complex vector location $\Sigma_{k=1}^{m} a^{k-1} e^{2\pi i (j_k - 0.5)/8}$ with respect to the center of the diagram, where the scale factor $a \sim 3$ is chosen to place the chain in the appropriate annulus for clarity of visualization.

In summary, Panel C depicts a 3-event-tree collected over a very long observation window ($T_{obs} = \infty$). Similar diagrams can be used to depict event-trees of different lengths m and shorter observation windows T_{obs}. This event-tree representation provides a succinct means of organizing the data associated with an entire collection of event-chains up to a specified length, and can also serve to pinpoint relevant features within a network's dynamics. For another example, see *Supplementary Fig. S3*. For short observation windows, different presentations of the same stimulus will produce different $m_{\max}$-event-trees. As event-trees may differ only slightly for different but similar stimuli, the probability distribution of event chains over many independent trials can be used to disentangle the dynamical differences caused by distinct stimuli from the fluctuations caused by short T_{obs} observation windows. These distributions can be a sensitive function of the stimulus and thus can provide a statistically accurate measurement that can distinguish between distinct stimuli.

Note that although the event $\sigma^7 \to \sigma^7$ does not occur very often (the location is blank), the event $\sigma^7 \to \sigma^7 \to \sigma^j$ (for $j \neq 7$) occurs more often. This peculiar phenomenon stems from our definition of event-chains associated with the particular choice of $\alpha = 2$ms for temporal coarse-graining. For example, in this case of $\alpha = 2$ms, the set of events $\{\sigma^7_{t=0.1}, \sigma^7_{t=2.6}, \sigma^1_{t=2.7}\}$ would contain both the 2-event-chain $\sigma^7 \to \sigma^1$ and the 3-event-chain $\sigma^7 \to \sigma^7 \to \sigma^1$, but not the 2-event-chain $\sigma^7 \to \sigma^7$. There are many alternative definitions of event-chains which produce event-trees capable of robust discriminability, but we will not fully discuss those here. As an example, given a set of time scales $\alpha_1, \ldots \alpha_{m-1}$, we can easily define an m-event-chain, denoted by $\left\{\sigma^{j_1} \overset{\alpha_1}{\to} \sigma^{j_2} \overset{\alpha_2}{\to} \cdots \overset{\alpha_{m-1}}{\to} \sigma^{j_m}\right\}$, as the event $\sigma_t^{j_m}$, preceded by the events $\sigma^{j_{m-1}}_{I_{m-1}}$, $\ldots$, $\sigma^{j_1}_{I_1}$ (with $I_k = [t - \Sigma_{l=k}^{l=m-1}\alpha_l, t - \Sigma_{l=k+1}^{l=m-1}\alpha_l)$). Such an event-chain can capture correlations within and across multiple synchronous activities within a system. For clarity, we have chosen uniform $\alpha_k = \alpha$, but this is in no way essential to our general results. Our definition of event-chains is structured to allow for efficient data collection techniques which (a) treat every spike computationally equally, and (b) produce event-trees which are invariant to time-translation of the system dynamics. This can be contrasted with a definition of event-chains which employs a global time discretization, and therefore produces event-trees which depend on the particular choice of time bins. Our techniques will be discussed in future work.

In Panel D, we display a representation of the 4-event-tree collected over $T_{obs} = \infty$ (with $\alpha = 2$ms) of the sustained-firing regime under stimulus I_1 (where $I_1 = (\nu_1 = 0.5, f_1 = 0.005)$ and $I_2 = (\nu_2 = 0.525, f_2 = 0.005)$ are the same as those in Fig. 2). The diagram of panel C is nested in the center of this diagram.

Panels C and D display measurements of the $T_{obs} = \infty$ event-tree of the system under stimulus I_1. However, in general (see text), each different trial presentation of a particular stimulus over a shorter T_{obs}, say $T_{obs} = 512$ms, will produce a different $m_{\max}$-event-tree. The

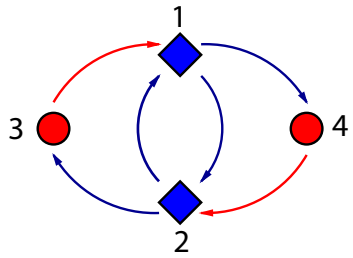
distribution of $m_{\max}$-event-trees over $T_{obs} = 512$ms, measured by collecting multiple trials, can be a sensitive function of the stimulus. Thus, in order to estimate the functional power of this network, we may attempt to estimate the difference between the T_{obs}-distribution of the $m_{\max}$-event-trees produced by this system under stimulus I_1 and the T_{obs}-distribution produced under I_2, with $T_{obs} = 512$ms, and $m_{\max} = 4$. If the difference between these two T_{obs}-distributions is significant, then the event-tree-projected network dynamics observed within a typical single trial of $T_{obs} = 512$ms can be used to discriminate between the stimuli I_1 and I_2.

In practice, one usually cannot measure the full multi-dimensional T_{obs}-distribution of the 4-event-tree for a system, as such a joint probability distribution of occurrence counts of multiple different event-chains is defined on a space of dimension $N + N^2 + N^3 + N^4 \sim N^5$. For stationary inputs, as in our example here, such joint distributions are only measurable for T_{obs} either extremely small (i.e., only a few short event-chains occur once within each trial, and most event-chains do not occur at all, thus the T_{obs}-distribution reduces to the joint distribution of singleton events) or extremely large (i.e., the T_{obs} distribution is well represented by the $T_{obs} = \infty$ event-tree, which captures the mean occurrence count, or observation rate of each event-chain). To circumvent this curse of dimension, instead of constructing the T_{obs}-distribution of the 4-event-tree for this system, we measure the T_{obs}-distribution of each 4-event-chain within the 4-event-tree separately. In other words, we first record many independent samples of this network's $m_{\max}$-event-tree (over multiple independent $T_{obs} = 512$ms observations) under both stimulus I_1 and I_2. With this collection of data, we obtain the empirical observation distributions of each m-event-chain's occurrence count (for $m \leq m_{\max}$). Thus, for each event-chain $\{\sigma^{j_1} \to \sigma^{j_2} \to \cdots \to \sigma^{j_m}\}$, we obtain the collection of probabilities $P\left(\{\sigma^{j_1} \to \sigma^{j_2} \to \cdots \to \sigma^{j_m}\} \text{ occurs } k \text{ times within a given } T_{obs} = 512\text{ms} \mid \text{stimulus } I_l\right)$ for each integer $k \geq 0$ and each stimulus $l = 1, 2$.

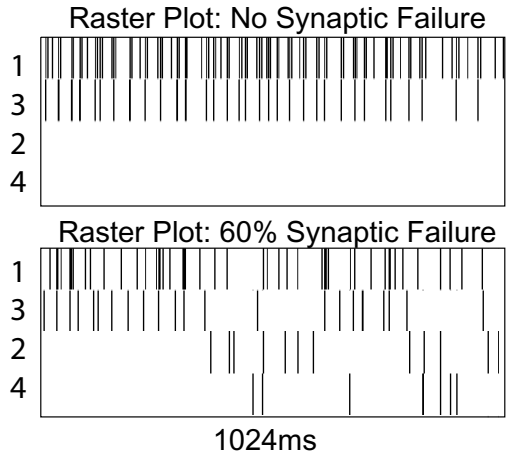
Some event-chains are not indicators of the stimulus (i.e., the $T_{obs} = 512$ms distribution of occurrence count is very similar for I_1 and I_2). However, there are other event-chains that can be used to discriminate between the stimuli. For example, the $T_{obs} = 512$ms distribution of occurrence counts of the 4-event-chain $\sigma^4 \to \sigma^1 \to \sigma^2 \to \sigma^8$ is quite different under stimulus I_1 than under stimulus I_2. The corresponding observation distributions are plotted in Panel E, with the blue and red histograms, P_1, and P_2, corresponding to I_1 and I_2 respectively (the overlapping region is purple). Note that the mean occurrence count for this 4-event-chain actually decreases when the rate of the input increases. The occurrence count of this event-chain alone can be used to discriminate the inputs I_1 and I_2. For example, we may take many sample $T_{obs} = 512$ms observations of the network's dynamics under different randomly chosen stimuli (either I_1 or I_2, with equal probability). For each of these samples, we examine the occurrence count ρ of the $\{\sigma^4 \to \sigma^1 \to \sigma^2 \to \sigma^8\}$ event-chain, and determine a possible candidate stimulus. We choose I_1 if $P_1(\rho) > P_2(\rho)$, otherwise we choose I_2 (e.g., in this case in Panel E, we guess I_2 if $\sigma^4 \to \sigma^1 \to \sigma^2 \to \sigma^8$ occurs 9 or fewer times, otherwise we guess I_1). Applying this procedure to many different independent sample trials results in a hit rate $A = \frac{1}{2}\int_0^\infty \max(P_1, P_2)\, dn$, and false alarm rate $B = 1 - A$, and the "information ratio" $\mathcal{I}_{I_1,I_2}^{\sigma^{j_1} \to \cdots \to \sigma^{j_m}} \equiv A/B$. For the case in Panel E, we have $A = 63\%$, $B = 37\%$, and $\mathcal{I}_{I_1,I_2}^{\sigma^4 \to \sigma^1 \to \sigma^2 \to \sigma^8} = A/B = 1.7$

The procedure described above classifies the stimulus underlying a single sample T_{obs}-observation by considering only the occurrence count of a single event-chain (i.e., a single element of the event-tree) associated with that T_{obs}-observation. We can easily extend this procedure to incorporate every event-chain within the observed event-tree. For example, given a sample $T_{obs} = 512$ms observation, and its associated event-tree, we can independently use the occurrence count of each event-chain within that event-tree to identify a possible candidate stimulus (which will be different, in general, for each event-chain). Thus, in this procedure, each event-chain 'votes' for either stimulus I_1 or I_2 (i.e., a vote of ± 1, respectively). We then sum up the votes across the entire event-tree to determine a single candidate stimulus underlying the sample T_{obs}-observation. It is natural to weight each vote with the factor $\log\left(\mathcal{I}_{I_1,I_2}^{\sigma^{j1}\to\cdots\to\sigma^{jm}}\right)$, which is a function of the information ratio of the contributing event-chain. This weighting implies that N votes of all 1's with error rate B have the same weight as a single vote with a far smaller error rate $\frac{B^N}{(1-B)^N+B^N}$. We define the 'discriminability' of the $m_{\max}$-event-tree (for this 2-way discriminability task) to be the percentage of sample observations which were correctly classified under our voting procedure. To perform 3-way discriminability tasks, we go through an analogous procedure, performing all three pairwise discriminability tasks for each sample observation, and ultimately selecting the candidate stimulus corresponding to the majority (with triple dead heats or other exceptional cases automatically counted as incorrect). Note that the discriminability is a function of α, T_{obs} and $m_{\max}$. For most of the systems we have observed, the discriminability increases as $m_{\max}$ and T_{obs} increase.

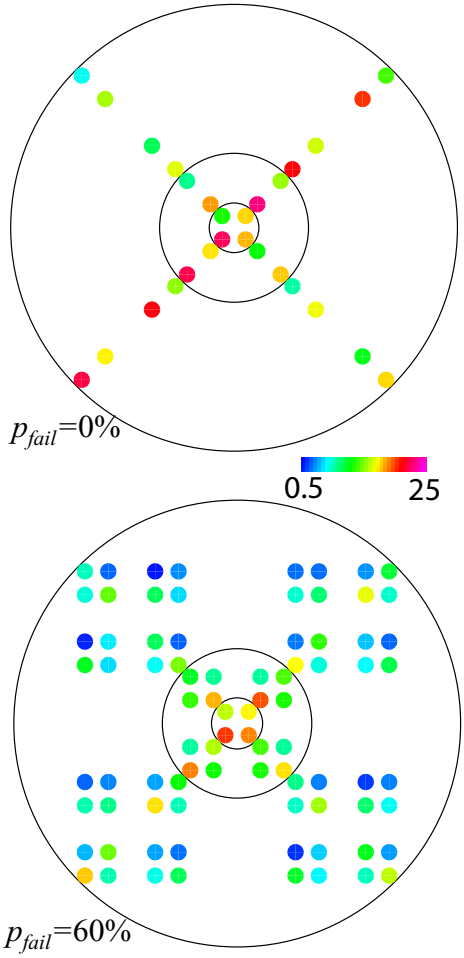
A


1
3
4
2

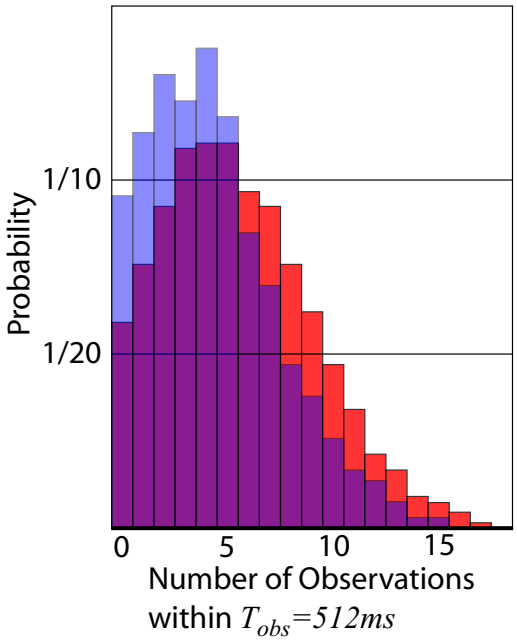
B


Raster Plot: No Synaptic Failure
1
3
2
4
Raster Plot: 60% Synaptic Failure
1
3
2
4
1024ms

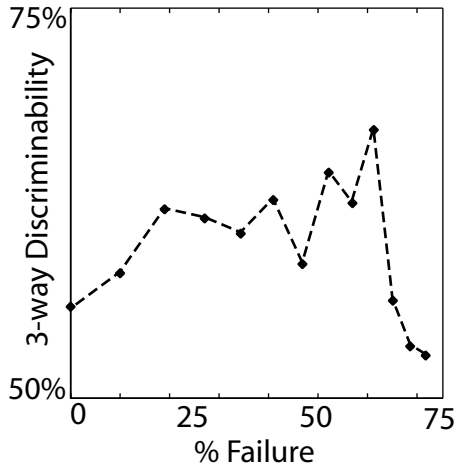
C


p_{fail}=0%
0.5
25
p_{fail}=60%

D

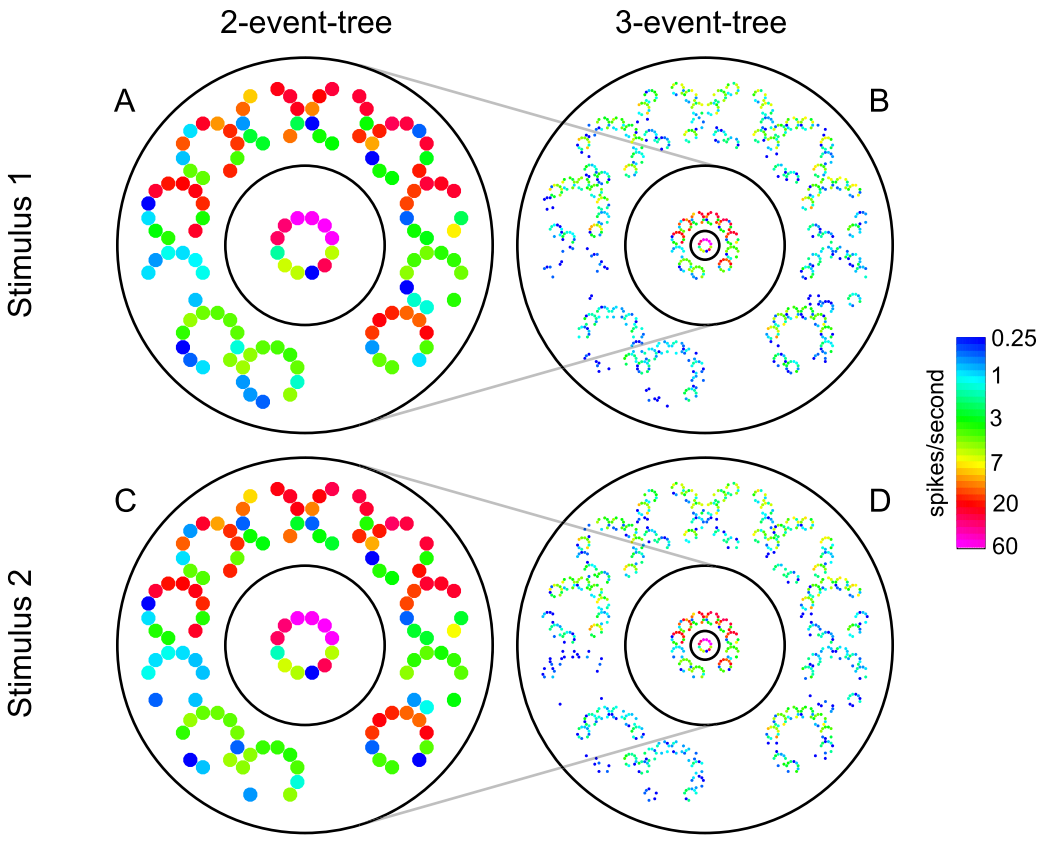

Probability
1/10
1/20
0
5
10
15
Number of Observations
within T_{obs}=512ms

E


75%
3-way Discriminability
50%
0
25
50
75
% Failure

Supplementary Figure S2: **Functional power vs. reliability of neurons.** Here we illustrate that the functional power of a network is not simply an increasing function of the synaptic reliability of neurons within that network. For clarity, we focus on a specific I&F neuronal network with $N = 4$ neurons. We model synaptic failure by randomly determining if each neuron in the network is affected by any presynaptic firing event (If the probability of spike transmission is $(1 - p_{fail})$, we scale the synaptic strength by $1/(1 - p_{fail})$, in order to keep the mean network properties the same). *(A):* The network is composed of 2 excitatory neurons (red), and 2 inhibitory neurons (blue), with connectivity indicated by the arrows. We examine this network's dynamics under drive with each neuron receiving independent Poisson input. If there is no synaptic failure in the network (i.e., the inter-neuronal connections are 100% reliable, and the synaptic failure probability $p_{fail} = 0\%$), then the strong recurrent connectivity within the network forces the system into one of two states: Either neurons 1 and 3 repeatedly fire (suppressing neurons 2 and 4), or neurons 2 and 4 repeatedly fire (suppressing neurons 1 and 3). However, if the synapses within the network have a probability of failure, then it is possible for the system to escape from either of these locked states, and all four neurons may fire. *(B):* Upper Panel: A typical 1024ms raster plot indicating locked activity when $p_{fail} = 0\%$. Lower Panel: A typical 1024ms raster plot illustrating the type of unlocked activity observed when $p_{fail} = 60\%$. *(C):* Representations of the 3-event-trees associated with these two dynamics, with the upper and lower panels corresponding to the $p_{fail} = 0\%$ and $p_{fail} = 60\%$ dynamics, respectively (These event trees were collected over $T_{obs} = \infty$, with $\alpha = 32$ms). The colorscale stands for the occurrence rate of any particular event-chain, ranging logarithmically from 0.5Hz to 25Hz. The observation rate of the event $\sigma^{j_1} \to \cdots \to \sigma^{j_m}$ is plotted at complex vector location $\Sigma_{k=1}^{m} 3^{k-1} e^{2\pi i (j_k - 0.5)/4}$ with respect to the center of each panel. Events with an observation rate of less than 0.01Hz are not plotted, with their locations intentionally left blank. The concentric circles indicate the separation between 1-event-chains (central circle), 2-event-chains (interior annulus) and 3-event-chains (exterior annulus). Clearly, the event-tree structure is far richer in the case of failure, than in the case without failure. *(D):* As a test of the network's functional power, we drive the $p_{fail} = 60\%$ system with two similar inputs I_1 and I_2, (where I_1 and I_2 are the same as those in Fig. 2) and record the $T_{obs} = 512$ms 3-event-trees. The red (blue) histogram illustrates the distribution of $T_{obs} = 512$ms occurrence counts, as observed under stimulus I_1 (I_2), for a typical 2-event-chain. These histograms indicate that the occurrence count of individual event-chains can be used to distinguish between these two stimuli, even within $T_{obs} = 512$ms (discrimination based solely on the single event-chain corresponding to the histograms shown is accurate 58% of the time). Moreover, the entire 3-event-tree (i.e, the collection of event-chains, along with their occurrence counts) can be used to discriminate these two stimuli (over $T_{obs} = 512$ms) with $\sim 70\%$ accuracy. *(E):* We estimate the functional power of this network as a function of p_{fail}. We drive the system with three similar inputs, I_1, I_2, I_3, (where I_1, I_2 and I_3 are the same as those in Fig. 2) and again record the $T_{obs} = 512$ms 3-event-trees. We use the 3-event-trees to perform a 3-way discrimination task (33% would be chance level). The discriminability is plotted for a set of different p_{fail} (the dots indicate data points, the dashed line is to guide the eye only), which clearly demonstrates that the event-trees associated with the network's dynamics are richer, and more capable of fine discrimination, when $p_{fail} \sim 60\%$, than when $p_{fail} = 0\%$. Fig. 4 in

the main text is reproduced as Panel E here for completeness.


2-event-tree
3-event-tree
A
B
C
D
Stimulus 1
Stimulus 2
spikes/second
0.25
1
3
7
20
60

Supplementary Figure S3 (Best viewed electronically): **Illustration of the utility of the graphical representation of event-trees**. We consider a specific sparsely coupled I&F neuronal network with $N = 12$ neurons operating in a sustained firing regime. We drive the network with two similar inputs, I_1 and I_2. Shown in the first row of the figure are *(A)* the 2-event-tree collected over $T_{obs} \gg 1$ for stimulus I_1, and *(B)* the 3-event-tree collected over $T_{obs} \gg 1$ for stimulus I_1. The occurrence counts are normalized by T_{obs} and plotted as occurrence rates with colors corresponding to the colorbar on the right side of the figure (occurrence rates below 0.25 Hz are not plotted). The set of occurence rates are organized as explained in Figure $S1$. Panels *(C)* and *(D)* display analagous event-trees collected for stimulus I_2. It can readily be seen by comparing the inner rings of panels A and C that the observed firing rates under I_1 are nearly indistinguishable from the firing rates observed under I_2. There are, however, apparent differences between the occurrence rates of 2-chains and 3-chains under I_1 and under I_2, which can be seen by comparing the outer rings of panels A and C (for 2-chains), or the outer rings of panels B and D (for 3-chains).