

RESEARCH ARTICLE

Cooperative coding of continuous variables in networks with sparsity constraint

Paul Züge *, Natalie Schieferstein*, Raoul-Martin Memmesheimer *

Institute for Genetics, University of Bonn, Bonn, Germany

* pzuege@uni-bonn.de (PZ), nschliefe@uni-bonn.de (NS), rm.memmesheimer@uni-bonn.de (RMM)



OPEN ACCESS

Citation: Züge P, Schieferstein N, Memmesheimer R-M (2025) Cooperative coding of continuous variables in networks with sparsity constraint. PLoS Comput Biol 21(7): e1012156.

<https://doi.org/10.1371/journal.pcbi.1012156>

Editor: Stefano Panzeri, Universitätsklinikum Hamburg-Eppendorf, GERMANY

Received: May 10, 2024

Accepted: May 16, 2025

Published: July 3, 2025

Copyright: © 2025 Züge et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The code written for the spiking network implementations of cooperative coding is publicly available at <https://github.com/NatalieSchieferstein/CoopCodeSpiking> (<https://doi.org/10.5281/zenodo.15553757>). The code written for the implementation of cooperative coding in rate networks is publicly available at <https://github.com/PaulZOnGit/CoopCodingRate> (<https://doi.org/10.5281/zenodo.15555348>).

Abstract

A hallmark of biological and artificial neural networks is that neurons tile the range of continuous sensory inputs and intrinsic variables with overlapping responses. It is characteristic for the underlying recurrent connectivity in the cortex that neurons with similar tuning predominantly excite each other. The reason for such an architecture is not clear. Using an analytically tractable model as well as spiking neural networks, we show that it can naturally arise from a cooperative coding scheme. In this scheme neurons with similar responses specifically support each other by sharing their computations to obtain the desired population code. This sharing allows each neuron to effectively respond to a broad variety of inputs, while only receiving few feedforward and recurrent connections. Few strong, specific recurrent connections then replace many feedforward and less specific recurrent connections, such that the resulting connectivity optimizes the number of required synapses. This suggests that the number of required synapses may be a crucial constraining factor in biological neural networks. Synaptic savings increase with the dimensionality of the encoded variables. We find a trade-off between saving synapses and response speed. The response speed improves by orders of magnitude when utilizing the window of opportunity between excitatory and delayed inhibitory currents that arises if, as found in experiments, spike frequency adaptation is present or strong recurrent excitation is balanced by strong, shortly-lagged inhibition.

Author summary

Neurons represent continuous sensory or intrinsic variables in their joint activity, with rather broad and overlapping individual response profiles. In particular there are often many neurons with highly similar tuning. In the cortex, these neurons predominantly excite each other. We provide a new explanation for this type of recurrent excitation, showing that it can arise in a novel cooperative coding scheme that minimizes the number of required synapses. This suggests the number of required synapses as a crucial

Funding: RMM received a grant supporting this research from the German Federal Ministry of Education and Research (BMBF, <https://www.bmbf.de>) via the Bernstein Network (<https://bernhard-network.de>), Bernstein Award 2014, grant number 01GQ1710), as well as an uncommitted grant from Huawei Technologies Co., Ltd. (<https://www.huawei.com>). PZ received a salary from these grants. RMM further received a grant as part of the iBehave Network: Sponsored by the Ministry of Culture and Science of the State of North Rhine-Westphalia (<https://ibehave.nrw/>, grant number NW21-049 A-F). NS received a salary from this grant. This publication was supported by the Open Access Publication Fund of the University of Bonn. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

constraining factor in biological neural networks. In our cooperative coding scheme, neurons use few strong and specific excitatory connections to share their computations with those neurons that also need it. This way, neurons can generate a large part of their response by leveraging inputs from neurons with similar responses, requiring fewer feedforward inputs. We find a trade-off between saving synapses and response speed. Theoretical estimates and numerical simulations show that specific features of biological single neurons and neural networks can drastically increase the response speed, improving the trade-off.

Introduction

The brain encodes continuous sensory or intrinsic variables in the coordinated activity of populations of neurons. The tuning curves (response profiles) of individual neurons in such populations are rather broad, leading to large overlaps between them [1,2]. Further, there are often many neurons with highly similar tuning. Neuron populations with such features include simple cells in the primary visual cortex (V1) [3,4], head direction cells in the anterior thalamic nucleus [5], tactile neurons in primary somatosensory cortex [6], place cells in the hippocampus [7] and grid cells in the medial entorhinal cortex [8,9]. In machine learning, convolutional networks have overlapping receptive fields (RFs) that tile the input space [10]. RFs similar to those in visual cortex emerge by learning a sparse code for natural images [11], and RFs similar to grid cells emerge through training on navigation tasks [12,13].

Neurobiological data show that neurons with strongly overlapping RFs are predominantly excitatorily coupled: Synaptic connections between similarly tuned excitatory principal neurons are more likely [14], stronger and more often bidirectional [15,16]. In line with this, the strongest incoming synapses provide excitation that matches a neuron's RF [16,17]. Furthermore, highly similarly tuned principal neurons have overall, i.e., including indirect, polysynaptic connections, a net excitatory effect on each other [18,19]. In contrast, if the tuning is barely similar or dissimilar, the net effect is inhibitory.

Such recurrent excitatory connectivity may seem unintuitive from a normative standpoint, as it amplifies noise [20] and can increase response times [21,22]. Previous studies suggested that it may support persistent activity and thus working memory [23,24] or sampling-based inference [25].

Neural networks, however, evolved subject to physiological and physical constraints [26–29], including metabolic cost and available space. Optimizing for specific features can largely determine the neural network and lead to solutions that are in other aspects sub-optimal. A prominent example for this is a recent version of the efficient coding hypothesis [30–34]. It posits that neural networks greedily minimize the number of used spikes or the rate activity, which contribute to metabolic cost. The network connectivity obtained from the optimization is, however, very dense, which is not found in experiments. Further, the coding scheme is “competitive”, in the sense that similarly tuned neurons compete for the opportunity to generate spikes. In other words, such neurons take away spikes and activity from each other. This predicts inhibitory couplings between very similarly tuned neurons, contrary to the experimentally observed physiological and effective excitatory interconnectivity between them.

Here we explore the implications of “cooperative coding” in a neural network. In this newly proposed scheme, neurons avoid replicating computations through feedforward weights whose results are already accessible from the activity of other feature neurons. Instead, each feature neuron performs only a non-redundant feedforward computation. It

then achieves the required response by additionally incorporating the results obtained by similarly tuned feature neurons through recurrent connections. In other words, feature neurons do not independently replicate shared parts of the computations through feedforward weights, but they transmit them through recurrent connections to each other. The resulting connectivity is like-to-like, i.e., strong and effectively excitatory between similarly tuned principal neurons, as observed in experiments. Interestingly the scheme can optimize the number of synapses in a network, while maintaining the required neural network dynamics. Such an optimization differs from the common focus on saving dynamical quantities such as spikes and may be imposed by space restrictions and the cost of maintaining synapses [28,35].

Results

To demonstrate the concept of cooperative coding, we consider a layer of feature neurons (output neurons), which receive feedforward input from an input layer as well as recurrent input. The task of the feature neurons is to generate a weighted sum of the inputs with weight strengths that decay exponentially with the distance of an input from the preferred input. We assume that the functionally relevant network response, representing the desired features (outputs), is the steady state activity. The desired outputs are linear functions of the inputs. Neural responses can hence be characterized by linear RFs and implemented by feedforward connectivity alone. Importantly, they can also be implemented using mixtures of feedforward and recurrent input.

We will compare the different network implementations in terms of the space requirement, approximated by the number of required synapses, and in terms of the metabolic cost to keep up the stationary state. Furthermore, we will compare the response times and demonstrate how they can be substantially decreased in networks with spike frequency adaptation (SFA) or balancing inhibition. Finally we will verify that our findings translate to cooperative coding in spiking neural networks.

We will analyze three concrete examples of cooperative coding: (i) encoding a one-dimensional stimulus, (ii) simultaneously encoding two one-dimensional stimuli with linear mixed selectivity (MS) and (iii) encoding a two-dimensional stimulus. For ease of description, we focus on translationally invariant RFs. (Approximate) translational invariance, meaning that offset RFs have similar shapes, is a common characteristic of experimentally encountered RFs [2,4,5,8,9]. Further, it is a common characteristic of RFs that emerge in machine learning [10,11]. Although the RFs that we consider do not have the precise shape of measured RFs, for example those of simple cells in V1 [36], they share the key properties of localized, overlapping and broadening RFs that tile the represented space. Indeed, RFs of neurons in hierarchically higher layers are often broader and constructed from those in lower layers [37,38].

Encoding a 1D stimulus

As a concrete, analytically tractable model that illustrates how cooperative coding works and can save synapses, we consider RFs that tile the one-dimensional parameter space of a stimulus (see Fig 1A). An input neuron j , $j = 1, \dots, N$, signals the presence and strength of a stimulus with a specific parameter j by nonzero activity $r_j > 0$. The task of the feature layer is to generate a response that is maximal at the preferred stimulus parameter and then decays exponentially the more different the stimulus becomes from the preferred one. This behavior is qualitatively similar to commonly observed tuning curves such as orientation tuning curves or place fields. We further assume that if multiple stimuli are present, the feature layer responses to their different parameters superpose linearly.

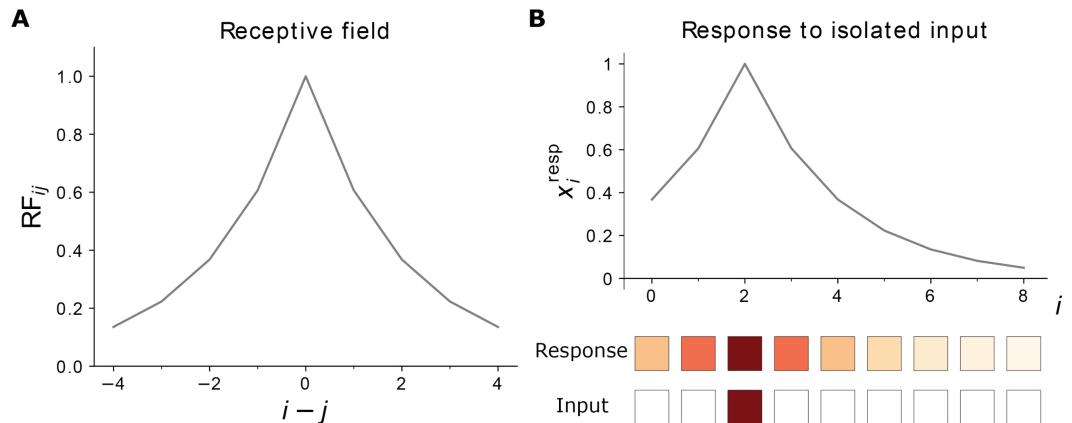


Fig 1. Receptive field and response of 1D network. (A) The RF of neuron i is peaked at $j = i$ and decays exponentially with $|i - j|$. The RF width parameter is here $d = 2$. (B) Top: The network response to an isolated unit input, here located at $j_0 = 2$, has the same shape and amplitude as a neuron's RF. It peaks at $i = j_0$. Bottom: Input and feature neurons are shown with color-coded activities r_j and x_i , respectively. Increasingly dark red color represents higher activity; white squares indicate inactive neurons.

<https://doi.org/10.1371/journal.pcbi.1012156.g001>

As an example, the inputs may be interpreted as a simple model for the representation of the orientation of a bar in the early visual system. $r_j > 0$ then means that the orientation is within the j th bin of the total orientation range $[0, 180^\circ]$. The transformation from input to features in our model describes the combination of responses from hierarchically lower visual areas to hierarchically higher ones [39]. As another example, the neurons may model the activity of place cells on a periodic, closed track. The transformation then models the transformation from input neurons with smaller place fields to neurons with larger place fields. Such a transformation may take place from the hippocampal dentate gyrus to the downstream area CA3 [40]. In our model, the input generates a simple encoding of the current location, where input neuron j is active if the animal is in the j th location.

The desired stationary feature layer activity can be expressed as

$$x_i^{\text{resp}} = \sum_{j=1}^N \text{RF}_{ij} r_j, \quad \text{RF}_{ij} = e^{-\frac{|i-j|}{d}} = \gamma^{|i-j|}. \quad (1)$$

Here r_j is the activity of the j th input neuron, $\gamma = \exp(-1/d)$ and d defines the width of the RF. We use periodic boundary conditions. For computations with neuron indices, this means that $|i - j|$ means $\min_{n \in \{-1, 0, 1\}} |i - j + nN|$. There are as many input as feature neurons. We note that, because of the symmetry $\text{RF}_{ij} = \text{RF}_{ji}$, the vector RF_k , describing the RF of feature neuron k , is the same as RF_k , the network response when only input neuron k is active, compare Fig 1A and 1B. Summarized in a formula, we have $x_k^{\text{resp}}|_{r_j=\delta_{jk}} = \text{RF}_{kl} = \text{RF}_{lk} = x_l^{\text{resp}}|_{r_j=\delta_{jk}}$, where k is fixed and l variable.

Feedforward implementation. To model the temporal dynamics of the neurons, we choose a standard simple linear rate network model [41,42]. The purely feedforward network that generates the response Eq 1 as stationary state is then given by

$$\tau \dot{x}_i(t) = -x_i(t) + \sum_{j=1}^N W_{ij}^{\text{ff}} r_j(t), \quad (2)$$

with $W_{ij}^{\text{ff}} = \text{RF}_{ij}$ and a time constant τ . In the stationary state, we have $\dot{x}_i = 0$ for all i and thus

$$x_i = \sum_{j=1}^N W_{ij}^{\text{ff}} r_j = \sum_{j=1}^N \text{RF}_{ij} r_j = x_i^{\text{resp}}. \quad (3)$$

This state is asymptotically stable and globally attracting; the flow is a contraction to it. These properties follow immediately from the fact that the system is linear and has a unique fixed point, which is asymptotically stable because all eigenvalues of the matrix specifying the homogeneous differential equation are negative, equal to $-1/\tau$ [43,44]. To approximate the network with a characteristic number of feedforward weights that is smaller than N , we require synapses $W_{ij}^{\text{ff}} = \gamma^{|i-j|}$ only where $|i - j| \leq d$. This defines the RF size $n_{\text{RF}} = 2d + 1$ as the number of feedforward synapses per neuron needed to implement the RF within a distance d around its center.

Cooperative implementation. The same stationary neuronal responses can be obtained as the steady state of a recurrent network that uses cooperative coding. It requires only three synapses per feature neuron, two recurrent and one feedforward synapse. This network's dynamics are given by

$$\tau \dot{x}_i(t) = -x_i(t) + \sum_{j=1}^N W_{ij}^{\text{rec}} x_j(t) + \sum_{j=1}^N W_{ij}^{\text{ff}} r_j(t) \quad (4)$$

$$= -x_i(t) + w^{\text{rec}} (x_{i+1}(t) + x_{i-1}(t)) + w^{\text{ff}} r_i(t), \quad (5)$$

with weights $w^{\text{rec}} = \frac{1}{\gamma + \gamma^{-1}} = \frac{\gamma}{1 + \gamma^2}$ and $w^{\text{ff}} = 1 - 2\gamma w^{\text{rec}} = \frac{1 - \gamma^2}{1 + \gamma^2}$. If the RFs are not narrow (d is not small against 1), the two recurrent connections are strong, in the sense that w^{rec} is not small against 1. Thus the network features strong like-to-like excitation and is driven by feedforward input. One can straightforwardly verify that $x_i = x_i^{\text{resp}}$ is indeed a stationary state of the network, by inserting Eq 1 into Eq 5, see Eq S2 in [S1 Appendix](#). The reason for this is ultimately that the desired response of a neuron i can be largely generated by summing the responses of the two neurons $i \pm 1$ neighboring i , see [Eq 11](#) and [Fig 2B](#). This is achieved by the recurrent connections. The missing part is contributed by the feedforward input. This state is asymptotically stable as all real parts of the eigenvalues of the matrix defining the homogeneous system are negative, see [Eq S12](#) in [S1 Appendix](#). For broad RFs (where $\gamma \lesssim 1$), the recurrent connections are nearly as strong as possible: their sum $2w^{\text{rec}}$ is close to 1, the value beyond which the network becomes unstable. The stationary state is also the only stationary state. Since the system is linear, the state is therefore a global attractor as for the feedforward network [43,44]. Thus, for constant input the network forms this stable response pattern.

Cooperative coding. Cooperative coding can be understood as sharing of the information that an individual neuron obtains from external input specifically with those neurons that also need it. This allows to generate most of the neuronal responses from sparse recurrent connectivity. Especially very similarly tuned neurons will project strongly excitatorily onto each other; oppositely tuned neurons would inhibit each other.

As a concrete example, we introduced the networks Eq 5, where it suffices that each neuron receives input from only one input and two feature neurons. Still, each neuron effectively responds to $\mathcal{O}(d)$ input neurons. This is possible because the feature neurons recurrently share their activity, and hence their access to feedforward input, with their neighbors. These in turn share it with their neighbors, thus propagating it through the network. The network

response then forms dynamically through the interplay of feedforward input and recurrent interactions.

The coding harnesses the fact that despite very few feedforward and recurrent synapses, poly-synaptic connectivity can still be far-reaching [45,46]: To show this, we consider the formal steady-state solution of Eq 4 for constant input, which provides the network's desired response if a solution exists. It can be obtained by setting $\dot{x} = 0$ and solving for x by multiplying with the inverse of $\mathbb{1} - W^{\text{rec}}$. The solution reads

$$x = (\mathbb{1} - W^{\text{rec}})^{-1} W^{\text{ff}} r. \quad (6)$$

For the considered excitatory weight matrix with a spectral radius smaller than 1, the inverse exists and can be expanded into a Neumann series, yielding

$$x = (\mathbb{1} + W^{\text{rec}} + (W^{\text{rec}})^2 + \dots) W^{\text{ff}} r. \quad (7)$$

The network response is thus determined by W^{rec} and its higher powers, which reflect the redistribution of feedforward input via poly-synaptic recurrent pathways. As the higher-order terms correspond to longer pathways, they will shape the response at later times. This can be well seen from the approximate, discretized dynamics [47,48]

$$x_i((n+1)\tau) \approx \sum_{j=1}^N W_{ij}^{\text{rec}} x_j(n\tau) + \sum_{j=1}^N W_{ij}^{\text{ff}} r_j, \quad (8)$$

which lead to the same steady state as the time-continuous dynamics. The response to a constant input r after n time constants is

$$x(n\tau) = (\mathbb{1} + W^{\text{rec}} + \dots + (W^{\text{rec}})^{n-1}) W^{\text{ff}} r; \quad (9)$$

higher-power, poly-synaptic terms add to it at successively later times. Viewed differently, a recurrent neural network can be equivalently described by a deep feedforward network that is “unrolled in time” [49], with higher layers generating the results of later computations, through a higher stack of copies of the recurrent weight matrix.

The coding scheme can also be understood as feedforward inputs providing a correction to the response that is mainly constructed from the sparse recurrent input. To clarify this we focus on elementary stationary responses, namely those that are driven by a single unit input from neuron j ; the input activity is $r_k = \delta_{kj}$. Responses to more complicated input patterns are weighted linear sums of such elementary responses. Consider feature neuron i and assume that all other neurons already respond correctly. The desired stationary activity of neuron i in response to a single unit input from neuron j is then RF_{ij} , while the responses of the other network neurons k are RF_{kj} . Eq 4 with $\dot{x}_i = 0$ implies that RF_{ij} is the sum of the RFs of its presynaptic feature neurons and its feedforward connectivity,

$$\text{RF}_{ij} = \sum_k W_{ik}^{\text{rec}} \text{RF}_{kj} + W_{ij}^{\text{ff}}. \quad (10)$$

For the specific network Eq 5 we have

$$\text{RF}_{ij} = w^{\text{rec}} (\text{RF}_{i-1,j} + \text{RF}_{i+1,j}) + w^{\text{ff}} \delta_{ij}, \quad (11)$$

illustrated in Fig 2C. This implies that the desired response of feature neuron i is fully generated by recurrent inputs, unless the unit input comes from input neuron i and there is also a feedforward contribution. The weighted and summed responses of neuron i 's nearest neighbors are thus already very close to neuron i 's target response. This is enabled by the specific exponential shape of the RFs. Only for the preferred input of a feature neuron, the response is too low. The neuron corrects for this by recruiting the missing input through a feedforward connection. Such an “explanatory gap” that is left by the recurrent inputs and can be filled by external input is important for cooperative coding, because the output must depend on external input.

Spatial demand and metabolic cost. To compare the efficiency of the introduced implementations, we focus on two cost dimensions: the space needed to implement the network and the metabolic cost of generating the stationary dynamics. As measure for the space needed for the network we take the number of synaptic connections, or, in other words, the L0 norm of the synaptic weight matrix. In the feedforward network Eq 2, it increases linearly with the width d of the RF if small responses can be neglected. This holds in particular when using our convention that the number of relevant synapses equals $2d + 1$ (see section “Feedforward implementation”). In the recurrent network Eq 5 three synapses per neuron suffice to generate the desired stationary response regardless of the RF size. We show after Eq S18 in S1 Appendix that the recurrent network Eq 5 therefore minimizes the L0 norm.

For the metabolic cost of generating the stationary dynamics, we may focus on the cost of generating the postsynaptic currents, which is proportional to their L1 norm (see S1 Appendix). This is because all other contributors, such as the neuronal activity, are identical between both network architectures. Since in both implementations all modeled synaptic currents are excitatory, the L1 norm of synaptic currents equals the total synaptic current. In the stationary state this current is the same in both implementations, because neurons have the

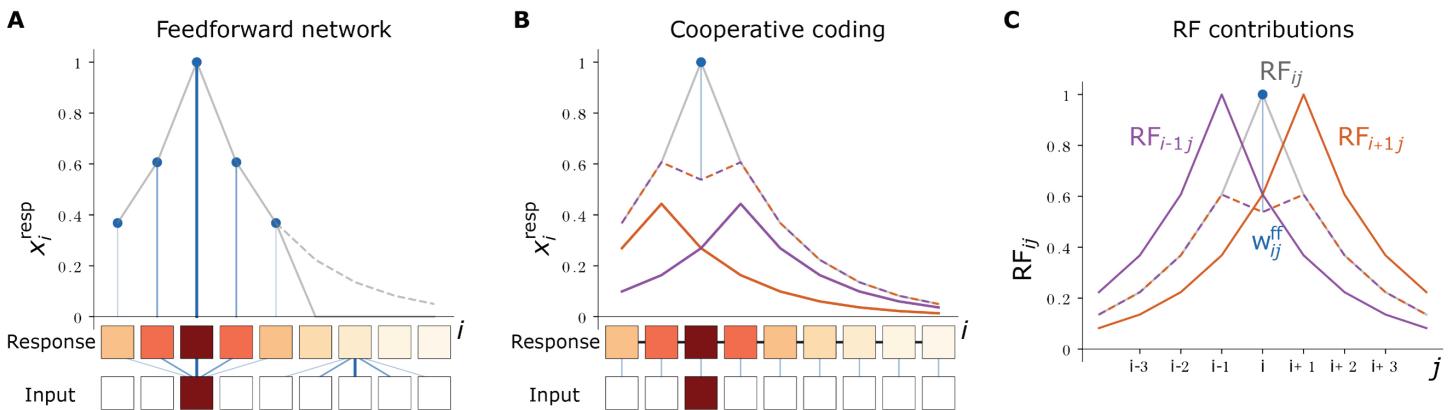


Fig 2. Schematics of feedforward and cooperatively coding networks. (A) Top: In the feedforward network, the response x_i^{resp} (gray solid curve) to an isolated input is fully generated by the neurons' feedforward inputs (blue lines and dots, line thickness represents input strength). For the displayed RF width $d = 2$, five neurons receive feedforward input, so that the network response (gray solid curve) represents $\approx 63\%$ of the summed target response (gray dashed curve). Bottom: Feature and input neuron activities as in Fig 1. Outgoing feedforward synapses from the active input neuron $j = 2$ and incoming feedforward synapses to feature neuron $i = 6$ are shown in blue. (B) Top: In the cooperatively coding network model, the network response (gray solid curve) is the sum of feedforward input (blue line and dot) and recurrent input (brown-purple dashed curve). For the displayed case of an isolated input, only one neuron receives feedforward input, which induces a part of the stationary response of the most active feature neuron. The rest of the response and all other responses are induced by recurrent input from neighboring neurons. The total recurrent input that each feature neuron receives is the sum of recurrent input from the right (brown solid curve) and left neighbor (purple solid curve). Bottom: Each feature neuron receives one feedforward synapse (blue lines) and two recurrent synapses (black lines, all recurrent connections are bidirectional). (C) The RF of feature neuron i (RF_{ij} for varying j , gray solid curve) is the weighted sum (brown-purple dashed curve) of the RFs of its left ($RF_{i-1,j}$, purple) and right neighbors ($RF_{i+1,j}$, brown) plus a contribution from feedforward input (W_{ij}^{ff} , blue line and dot). All shown RFs have width $d = 2$.

<https://doi.org/10.1371/journal.pcbi.1012156.g002>

same stationary activity; therefore also the cost is the same. We conclude that the metabolic cost for maintaining the stationary network state is the same in both the feedforward and the recurrent network implementation.

Response speed. In the feedforward network, activity converges with the intrinsic time constant τ , which we define as its response time. In the cooperatively coding network, the excitatory recurrent connectivity increases the response time: Fig 3 shows the dynamics of the response formation.

The network activity splits into independently evolving, orthogonal eigenmodes that approach their steady state values at different speeds, see Eq S8 in [S1 Appendix](#). We use the L1-norm of the deviation of the response from the steady state,

$$L(t) = |x(t) - x^{\text{steady}}|_1, \quad (12)$$

as a loss measure. The linearized loss ($\sum_i x_i^{\text{steady}} - x_i(t)$) equals a constant offset minus the projection of the activity $x(t)$ onto the vector of ones $(1, \dots, 1)$, which coincides with the slowest-decaying eigenmode. The linearized loss, as well as the full loss for the constant-zero activity initialization used here, thus decays exponentially. We define the response time τ_{resp} of the network as the time constant of this decay,

$$\tau_{\text{resp}} \equiv \frac{\tau}{1 - w_{\text{sum}}^{\text{rec}}}, \quad (13)$$

(see Fig 4), where $w_{\text{sum}}^{\text{rec}} = \sum_j W_{ij}^{\text{rec}}$ is the sum of recurrent weights arriving at (or, equivalently, originating from) a neuron. For generic initial conditions, this provides the time constant of the slowest-converging activity mode and thus the time constant that dominates the long-term convergence of the full loss; the linearized loss decays with time constant τ_{resp} during the entire time evolution. τ_{resp} scales inversely with the difference of the largest eigenvalue of the

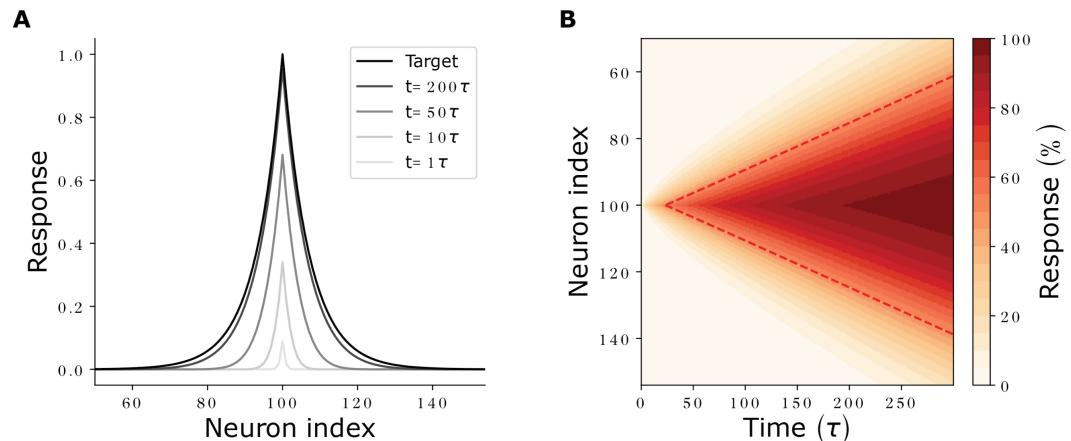


Fig 3. Response formation and activity propagation. (A) Network activity at different times (shaded curves) after r_{100} has been set from 0 to 1. For long times, network activity approaches the target response (black curve). (B) Development of the activity of neurons (y axis) with time (x axis), measured relative to their target activities. The diagonal fronts of equal relative activities indicate propagation of activity with constant propagation speed. The points where neurons reach 50% of their final activity are connected by a red dashed line. Parameters: $w^{\text{rec}} = 0.5 + 1/(1 - 1/100)$, such that $\tau_{\text{resp}} = 100\tau$ (see Eq 13), $N = 200$ neurons.

<https://doi.org/10.1371/journal.pcbi.1012156.g003>

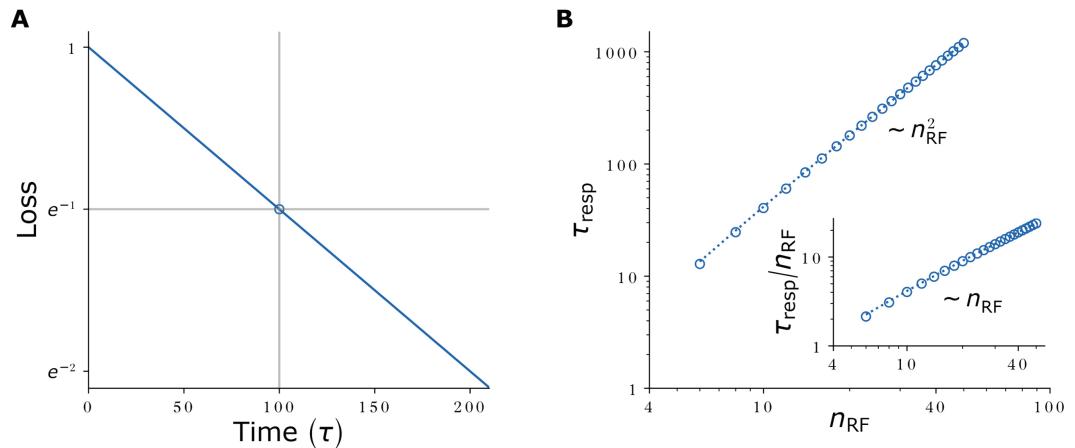


Fig 4. Loss evolution and response speed - synapse number trade-off. (A) Exemplary loss evolution for a network with $w_{\text{rec}} = 0.5 \cdot (1 - 1/100)$ so that $\tau_{\text{resp}} = 100\tau$. Experimentally, τ_{resp} is determined as the time (gray vertical line) at which the loss drops to e^{-1} (gray horizontal line, blue open circle). (B) Response times τ_{resp} (circles: simulation results; dotted line: analytical solution Eq 14) for target RFs of different widths n_{RF} . Data was created by scanning n_{RF} , setting $w_{\text{sum}}^{\text{rec}}$ to yield an RF of size n_{RF} and determining τ_{resp} from the loss dynamics.

<https://doi.org/10.1371/journal.pcbi.1012156.g004>

network, which is equal to $w_{\text{sum}}^{\text{rec}}$ (cf. Eq S12 in S1 Appendix), from 1. In particular, it depends only on the summed recurrent weights.

Eq 13 holds generally, for networks of the type Eq 4 with purely excitatory circulant recurrent weight matrix and convergent dynamics. We now specialize the result to networks with nearest-neighbor coupling Eq 5 that generate the RFs Eq 1. In these networks, $w_{\text{sum}}^{\text{rec}} = 2w^{\text{rec}}$. Inserting $w^{\text{rec}} = 1/(\gamma + \gamma^{-1})$ and $\gamma = \exp(-1/d)$ relates the response time to the RF width. By approximating $\exp(\pm 1/d) \approx 1 \pm 1/d + 1/(2d^2)$ for large d , we obtain $w^{\text{rec}} \approx 1/(2 + 1/d^2)$ and, inserting this into Eq 13,

$$\tau_{\text{resp}} \approx \frac{\tau}{1 - \frac{2}{2 + \frac{1}{d^2}}} = (1 + 2d^2) \tau \approx 2d^2 \tau \approx \frac{1}{2} n_{\text{RF}}^2 \tau. \quad (14)$$

In the last part of the equation we used that $n_{\text{RF}} = 1 + 2d \approx 2d$ for large d . Eq 14 shows that wide RFs require long equilibration time. This is because they need strong recurrent weights with a largest eigenvalue close to 1. Further the equation reveals the trade-off between response time and number of employed synapses: The feedforward implementation Eq 2 needs n_{RF} synapses and has a response time τ . The recurrent implementation thus saves $n_{\text{RF}} - 3 \approx n_{\text{RF}}$ synapses per feature neuron. Eq 14 shows that the response time increases quadratically in the number of saved synapses, see also Fig 4B.

The quadratic dependence of the response time on d reflects that, as the RF becomes wider, not only does activity have to spread further, it also spreads more slowly: this is consistent with the idea that the settling of a neuron depends (indirectly) more on activity propagating back from more distant neurons that settle after it.

Faster response with spike frequency adaptation. For activity to rapidly spread through the network, neurons need to be able to cause a large activity change in their neighbors within a short period of time. To achieve this, they need strong recurrent weights. However, recurrent weights are restricted to $w_{\text{sum}}^{\text{rec}} < 1$ to not cause runaway activity. We now show how spike frequency adaptation (SFA) can help ease this conflict and speed up network dynamics. SFA is

typical for excitatory principal neurons and induces a reduction of their response to constant inputs in the long run [42,50,51].

We model SFA through a negative-feedback adaptation current $u(t)$, which is triggered by neuronal activity $x(t)$ and characterized by its scale a_{SFA} and time constant τ_{SFA} ,

$$\tau \dot{x}_i(t) = -x_i(t) + \sum_{j=1}^N W_{ij}^{\text{rec}} x_j(t) + \sum_{j=1}^N W_{ij}^{\text{ff}} r_j(t) - a_{\text{SFA}} u_i(t), \quad (15)$$

$$\tau_{\text{SFA}} \dot{u}_i(t) = -u_i(t) + x_i(t). \quad (16)$$

This model is a slightly simplified version of that in [52] and the same as in [53,54]. Setting $\dot{x}_i(t) = 0$ and $\dot{u}_i(t) = 0$ yields the steady state. We see immediately that it implies $u_i = x_i$. Inserting this into Eq 15 shows that in the stationary state the spike frequency adaptation results in a stronger leak current, $-(1 + a_{\text{SFA}})x_i$. Dividing by $1 + a_{\text{SFA}}$ yields

$$0 = -x_i + \sum_{j=1}^N \frac{W_{ij}^{\text{rec}}}{1 + a_{\text{SFA}}} x_j + \sum_{j=1}^N \frac{W_{ij}^{\text{ff}}}{1 + a_{\text{SFA}}} r_j. \quad (17)$$

Consequently, in order to implement the same response as a network without SFA ($a_{\text{SFA}} = 0$, cf. Eq 4), the recurrent and feedforward weights have to be scaled up by a factor of $1 + a_{\text{SFA}}$. The additional excitatory synaptic input compensates in the steady state the added inhibitory adaptation current.

To understand the network dynamics, it is instructive to consider the limit $\tau_{\text{SFA}} \rightarrow 0$ where $u_i(t) \rightarrow x_i(t)$ as in the steady state. Inserting this into Eq 15 and again dividing by $1 + a_{\text{SFA}}$ yields an equation equivalent to Eq 4 with smaller neuronal time constant and smaller weights,

$$\frac{\tau}{1 + a_{\text{SFA}}} \dot{x}_i(t) \stackrel{\tau_{\text{SFA}} \rightarrow 0}{=} -x_i(t) + \sum_{j=1}^N \frac{W_{ij}^{\text{rec}}}{1 + a_{\text{SFA}}} x_j(t) + \sum_{j=1}^N \frac{W_{ij}^{\text{ff}}}{1 + a_{\text{SFA}}} r_j(t). \quad (18)$$

We see that a network with arbitrarily fast SFA and appropriately upscaled weights has the same dynamics as a network without SFA, but with its time constant reduced by $1 + a_{\text{SFA}}$. This factor only depends on a_{SFA} and is independent of the RF width that the network implements. We might thus expect that introducing SFA with a given a_{SFA} and small τ_{SFA} causes a constant speedup, but still results in a quadratic dependence of $\tau_{\text{resp}}^{\text{SFA}}$ on n_{RF} (see Eq 14).

There is, however, an additional possibility: SFA might yield faster dynamics for finite, nonzero τ_{SFA} . This is because then $u_i(t)$ lags behind $x_i(t)$, which creates a temporal “window of opportunity”. Within this window, the up-scaled weights can mediate strong interactions that are not yet cancelled by the retarded adaptation currents of the receiving neurons. In our networks, this leads to the following concept to exploit SFA: During the initial response phase, strong weights should cause a fast response while SFA keeps the steady state before and after an input change at the desired activity values as well as dynamically stable. In particular, the modified recurrent synaptic weights may then be (and to optimally exploit SFA: should be) so strong that without the SFA current the network dynamics are unstable.

To incorporate SFA in a cooperatively coding network, we modify the weights in Eq 5 as described above and add the SFA current. For the neuron activities, this yields the dynamical

equation

$$\begin{aligned} \tau \dot{x}_i(t) = & -x_i(t) + (1 + a_{\text{SFA}}) w^{\text{rec}}(x_{i+1}(t) + x_{i-1}(t)) \\ & + (1 + a_{\text{SFA}}) w^{\text{ff}} r_i(t) - a_{\text{SFA}} u_i(t), \end{aligned} \quad (19)$$

where the adaptation current obeys Eq 16. We find that the second of the above-described possibilities applies to such networks: Measuring the response time as a function of τ_{SFA} , we observe that it first decreases when increasing τ_{SFA} from zero and reaches a minimum at a nonzero, optimal value of the SFA time scale (Fig A in S1 Appendix). Increasing τ_{SFA} further eventually causes diverging activity, because the retarded adaptation current $u(t)$ becomes so slow that it never compensates the stronger input due to the upscaled weights. We note that also when keeping τ_{SFA} at a fixed value, there is an optimal nonzero value of the inhibitory feedback strength a_{SFA} , which we define to minimize the integrated loss (Fig A in S1 Appendix). Importantly, we find that introducing SFA with finite τ_{SFA} and optimal a_{SFA} improves the scaling of $\tau_{\text{resp}}^{\text{SFA}}$ with n_{RF} from quadratic as without or with arbitrarily fast SFA to linear. The accelerating effect is a form of “balanced amplification” [21], where the matrix governing a dynamical system is non-normal, featuring a hidden feedforward structure from difference modes (here: high neuronal activity but still low adaptation currents) to sum modes (here: high neuronal activity and adaptation currents).

We now further analyze the balanced amplification [21] in cooperative coding networks with SFA, by reducing the dynamics Eq 19 to effective single neuron dynamics with feedback. For this we assume that all neuronal activities receive the same inputs $r_i(t) = r(t)$ and follow the same time course $x_i(t) = x(t)$, so that we can replace $x_{i-1}(t) + x_{i+1}(t) = 2x(t)$. We note that we thereby study the eigenmode related to the L1 loss, see Eq S13 in Appendix S7. The resulting effective single neuron dynamics read

$$\begin{pmatrix} \dot{x}(t) \\ \dot{u}(t) \end{pmatrix} = \begin{pmatrix} \frac{2(1+a_{\text{SFA}})w^{\text{rec}}-1}{\tau_{\text{SFA}}} & -\frac{a_{\text{SFA}}}{\tau_{\text{SFA}}} \\ \frac{1}{\tau_{\text{SFA}}} & -\frac{1}{\tau_{\text{SFA}}} \end{pmatrix} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} + (1 + a_{\text{SFA}}) w^{\text{ff}} \begin{pmatrix} r(t) \\ 0 \end{pmatrix}. \quad (20)$$

A complex Schur decomposition of the matrix defining the effective neuron’s intrinsic 2D linear dynamics (homogeneous part of Eq 20) reveals a strong feedforward coupling from a difference mode (real parts of the eigenvector components have opposite sign) to a sum mode (real parts of the eigenvector components have the same sign). In our numerical evaluations we consider networks where the inhibitory feedback strength a_{SFA} is optimized such that the integrated loss is minimal. For these networks $2(1 + a_{\text{SFA}})w^{\text{rec}} - 1$ is positive (as the network is unstable without adaptation); furthermore, both the sum and difference mode are oscillatory (as oscillations help to reduce the integrated error), i.e., the matrix in Eq 20 has complex eigenvalues. This is similar to the two-population networks in Ref [21], which, however, have mostly real, non-positive eigenvalues and thus non-oscillatory modes without Hebbian amplification.

Concerning the use of resources, SFA does not require additional synaptic connections, so the spatial demand of the cooperatively coding network is the same as in the original model Eq 5. The increased weights, however, lead to stronger synaptic currents. Together with the added adaptation currents, this increases the energetic cost of maintaining the stationary state.

Balanced networks

The networks we studied so far had only excitatory synapses, while biological neural networks also have recurrent inhibition, which balances the excitation [32,55]. These are likely required for a range of reasons, such as ensuring network stability and maintaining irregular spiking activity [56–60]. Given their existence, we here show how inhibition can be used to speed up the network response, in an architecture that still relies on few synapses.

Experiments show that individual excitatory and inhibitory currents can be much larger than their sum and precisely temporally balanced with a lag smaller than the neuronal time constant [61,62]. Further, many inhibitory neurons are rather sharply tuned [62–64], sometimes similarly sharply as excitatory ones. To incorporate inhibition consistent with these findings, we add inhibitory neurons to the existing network of excitatory feature neurons.

Specifically, we assume that there are as many inhibitory neurons as feature neurons and that each inhibitory neuron follows the activity of one feature neuron with a small delay, τ_{lag} , such that we do not need to introduce a separate dynamical equation for it. Eq 4 thus becomes

$$\tau \dot{x}_i(t) = -x_i(t) + \sum_{j=1}^N W_{ij}^{\text{rec,E}} x_j(t) + \sum_{j=1}^N W_{ij}^{\text{rec,I}} x_j^I(t) + \sum_{j=1}^N W_{ij}^{\text{ff}} r_j(t) \quad (21)$$

$$= -x_i(t) + \sum_{j=1}^N W_{ij}^{\text{rec,E}} x_j(t) + \sum_{j=1}^N W_{ij}^{\text{rec,I}} x_j(t - \tau_{\text{lag}}) + \sum_{j=1}^N W_{ij}^{\text{ff}} r_j(t), \quad (22)$$

where $x_i^I(t)$ is the inhibitory activity, which equals the delayed excitatory feature neuron activity $x_i(t - \tau_{\text{lag}})$. $W_{ij}^{\text{rec,I}} \leq 0$ is the coupling from inhibitory neuron j to feature neuron i . We note that an alternative, more common choice is to model inhibitory activity as low-pass filtered version of excitatory activity, $\tau_I \dot{x}_j^I(t) = -x_j^I(t) + x_j^E(t)$ with $\tau_I = \tau_{\text{lag}}$. We exemplarily checked that this leads to qualitatively similar results (Fig D in [S1 Appendix](#)).

We now introduce the state change of feature neuron i between $t - \tau_{\text{lag}}$ and t ,

$$\Delta x_i(t) = x_i(t) - x_i(t - \tau_{\text{lag}}). \quad (23)$$

To rewrite the network dynamics in terms of a net interaction and a balanced interaction, we define the new weights

$$W_{ij}^{\text{rec,net}} = W_{ij}^{\text{rec,E}} + W_{ij}^{\text{rec,I}}, \quad (24)$$

$$W_{ij}^{\text{rec,bal}} = -W_{ij}^{\text{rec,I}}, \quad (25)$$

which we call net and balanced weights, respectively. The balanced weights $W_{ij}^{\text{rec,bal}} \geq 0$ describe the part of recurrent excitation that is in the stationary regime balanced (canceled) by inhibition; the net weights describe the unbalanced remainder of the recurrent interaction, which could in principle also be inhibitory. These definitions allow to rewrite Eq 22 as

$$\tau \dot{x}_i(t) = -x_i(t) + \sum_{j=1}^N W_{ij}^{\text{rec,net}} x_j(t) + \sum_{j=1}^N W_{ij}^{\text{rec,bal}} \Delta x_j(t) + \sum_{j=1}^N W_{ij}^{\text{ff}} r_j(t). \quad (26)$$

We now insert the values of the cooperatively coding network Eq 5 and further assume that the inhibitory neurons inhibit and balance the same sets of neurons that their driving feature neurons excited, i.e., $W_{ij}^{\text{rec,E}} = w^{\text{rec,E}} (\delta_{i+1,j} + \delta_{i-1,j})$ and $W_{ij}^{\text{rec,I}} = w^{\text{rec,I}} (\delta_{i+1,j} + \delta_{i-1,j})$. Eq 22 then

becomes

$$\begin{aligned}\tau \dot{x}_i(t) = & -x_i(t) + w^{\text{rec},\text{E}}(x_{i+1}(t) + x_{i-1}(t)) \\ & + w^{\text{rec},\text{I}}(x_{i+1}(t - \tau_{\text{lag}}) + x_{i-1}(t - \tau_{\text{lag}})) + w^{\text{ff}} r_i(t).\end{aligned}\quad (27)$$

Further, defining the net weights $w^{\text{rec},\text{net}} = w^{\text{rec},\text{E}} + w^{\text{rec},\text{I}}$ and the balanced weights $w^{\text{rec},\text{bal}} = -w^{\text{rec},\text{I}}$, Eq 27 becomes

$$\begin{aligned}\tau \dot{x}_i(t) = & -x_i(t) + w^{\text{rec},\text{net}}(x_{i+1}(t) + x_{i-1}(t)) \\ & + w^{\text{rec},\text{bal}}(\Delta x_{i+1}(t) + \Delta x_{i-1}(t)) + w^{\text{ff}} r_i(t).\end{aligned}\quad (28)$$

The balanced interaction term (proportional to $w^{\text{rec},\text{bal}}$) describes the combined effect of delayed inhibition and immediate excitation of the same strength. It depends on $\Delta x_{i\pm 1}(t)$ and therefore only acts when there are activity changes during the preceding brief E-I lag. Any change δx_j in the activity of neuron j causes postsynaptic activity changes in neuron i that integrate to $\frac{\tau_{\text{lag}}}{\tau} W_{ij}^{\text{rec},\text{bal}} \delta x_j$ (see Eq S21 in S1 Appendix). Once $\frac{\tau_{\text{lag}}}{\tau} w_{\text{sum}}^{\text{rec},\text{bal}} > 1$, where $w_{\text{sum}}^{\text{rec},\text{bal}} = \sum_i W_{ij}^{\text{rec},\text{bal}}$, an activity change δx in one neuron causes directly further activity changes that are, integrated over time and neurons, larger than δx . Correspondingly, and taking into account the stabilizing contracting dynamics due to the net interactions, the network dynamics become unstable once $\frac{\tau_{\text{lag}}}{\tau} w_{\text{sum}}^{\text{rec},\text{bal}}$ becomes slightly larger than 1, Fig 5.

Henceforth we consider $w^{\text{rec},\text{net}}$ and $w^{\text{rec},\text{bal}}$ to be the independent variables. This means, in particular, that increasing the balanced weights, or equivalently the strength of the inhibitory weights, implies a concurrent increase of the excitatory weights to keep the net weights invariant. To connect Eq 28 to our previous, unbalanced network Eq 5, we set

$$w^{\text{rec},\text{net}} = w^{\text{rec}}. \quad (29)$$

The dynamical equations then agree if $w^{\text{rec},\text{bal}} = 0$ or $\Delta x_i(t) = 0$. The latter is satisfied in the steady state. The steady state is thus independent of the strength of the EI-balance, which is given by $w^{\text{rec},\text{bal}}$. In particular, the steady state is the same as in Eq 5 (which is the special case of $w^{\text{rec},\text{bal}} = 0$). The stability of the steady state, however, depends on $w^{\text{rec},\text{bal}}$. We may thus think of $w^{\text{rec},\text{net}}$ as defining the RF, and of $w^{\text{rec},\text{bal}}$ as affecting the dynamics by modulating the EI-balance. During the build-up of the response strong excitation ramps up slightly before the balancing inhibition. For networks with large $w^{\text{rec},\text{net}}$, the analytical solution of Eq 28 and our stability analysis (see Sec. ‘Response Speed in the balanced network’ below and Fig B in S1 Appendix) show that throughout this window of opportunity excitation may be up to approximately $\tau/\tau_{\text{lag}} + 1$ times larger than the net interaction without destabilizing the network. This strong interaction allows a much quicker propagation of activity, convergence to the steady state, and decay of the loss function.

Spatial demand and metabolic cost in the balanced network. Compared to the unbalanced network, the balanced network requires three additional synapses per principal feature neuron, one E-to-I and two I-to-E synapses, i.e., a total of six synapses. This is again independent of the RF width, such that for large RFs, the balanced, cooperatively coding network still saves synapses compared to the feedforward network. It requires additional space for the inhibitory neurons, which may, however, be needed for other purposes anyways.

Also the metabolic maintenance cost increases, since there are more neurons. Further, there is an increased metabolic cost to sustain the synaptic currents in the stationary state: In

this state, large parts of the excitatory and inhibitory currents cancel to give rise to a net current that equals the one in the purely excitatory recurrent network, see Eq 24 and Eq 29. In the L1 norm of the synaptic currents, the excitatory currents and the absolute inhibitory currents, however, add up. The metabolic cost thus increases by the amount of excitatory and inhibitory currents that cancel each other.

Response speed in the balanced network. Under some additional assumptions, the evolution of the L1-loss Eq 12 can be analytically approximated as the solution of a linear delay differential equation, see Eq S28 in S1 Appendix for details. The resulting dynamics are those of a damped oscillator, see Fig 5A: For weak EI-balance, characterized by small $w_{\text{rec},\text{bal}}$, they are “overdamped” in the sense that they are well described by the sum of two exponentials with different decay rates. At a specific intermediate balance, the two decay rates agree and we have “critical damping”. For stronger balance the dynamics are “underdamped” in the sense that the loss behaves as the absolute value of an oscillation with exponentially decaying amplitude. Overly strong balance, and hence for fixed net interactions overly strong excitation, causes divergence of the dynamics (see Fig 5B and Fig Ba in S1 Appendix).

In the overdamped regime, the smaller decay rate is the relevant one, as it dominates the speed of the decay for longer times. The larger decay rate rather describes how quickly faster dynamics, that may be present due to the initial conditions, are suppressed and the dynamics converge to the slower mode. The smaller decay rate increases when the balance approaches its critical strength. The same holds for the single decay rate in the oscillatory regime. At the critical balance the overall decay of the loss is thus fastest, see Fig 5B.

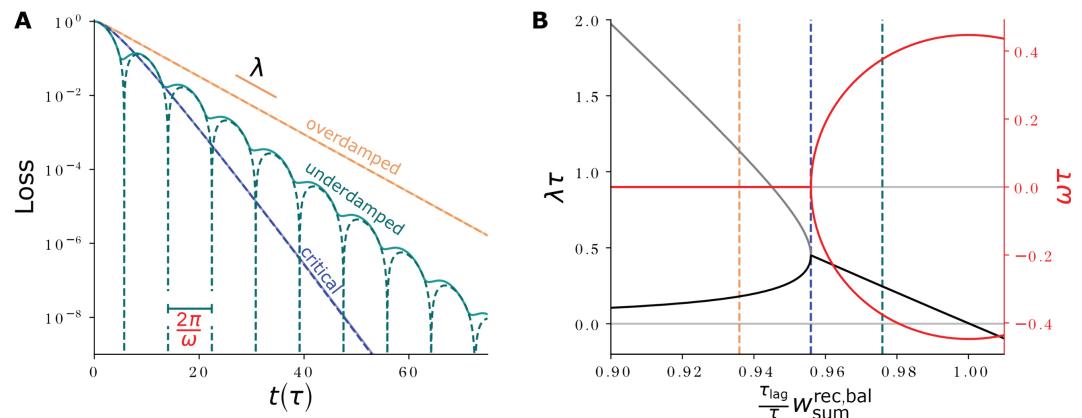


Fig 5. Loss evolution for different strengths of EI-balance. (A) Loss evolution (dashed: analytical approximation (cf. Eq S30 and S41 in S1 Appendix, partly occluded; solid: network simulation) for balance strengths that are slightly weaker (orange), equal (blue) or slightly stronger (teal) than the critical balance, on a logarithmic scale. The slope of the decay is given by λ (see (B)), explicitly highlighted for the overdamped dynamics. The oscillation period of the underdamped dynamics is $T_{\text{osci}} = 2\pi/\omega$. In case of oscillations, the analytic approximation briefly reaches zero loss once in a period (sharp dips in dashed curve). In the network simulation there is also a pronounced oscillation, but there always remains a finite error. (B) Real part (decay rate λ , black/gray) and imaginary part (oscillation frequency ω times ± 1 , red) of the complex frequency of the exponential loss evolution, scaled by τ . For weak EI-balance, measured by $w_{\text{sum}}^{\text{rec},\text{bal}}$, there are two exponentially decaying modes (λ , black and gray curve). At the critical balance $w_{\text{sum},c}^{\text{rec},\text{bal}}$ (blue dashed vertical line), there is only a single decay rate and no oscillation; the decay rate (in the overdamped case: of the relevant slower-decaying mode) is maximized. For stronger balance, network activity begins to oscillate (nonzero ω , red), and diverges once λ becomes negative. This happens approximately at $\frac{\tau_{\text{lag}}}{\tau} w_{\text{sum}}^{\text{rec},\text{bal}} \approx 1 + \frac{\tau_{\text{lag}}}{3\tau_{\text{resp}}}$, which is slightly larger than 1 because of the stabilizing effect of the contracting dynamics of the unbalanced network. Dashed vertical lines show the balance strengths scaled by τ_{lag}/τ for the curves in (A) ($(\tau_{\text{lag}}/\tau) w_{\text{sum},c}^{\text{rec},\text{bal}} - 0.02$, $(\tau_{\text{lag}}/\tau) w_{\text{sum},c}^{\text{rec},\text{bal}}$, $(\tau_{\text{lag}}/\tau) w_{\text{sum},c}^{\text{rec},\text{bal}} + 0.02$). Parameters: $w_{\text{sum}}^{\text{rec},\text{net}} = 0.99$, $\tau = 1$, $\tau_{\text{lag}} = 0.1$, and $N = 200$ for the network simulation.

<https://doi.org/10.1371/journal.pcbi.1012156.g005>

We find analytically that its decay time constant is approximately proportional to the geometric mean of the response time in absence of inhibition and of the inhibitory delay,

$$\tau_{\text{resp}}^{\text{bal},c} \approx \sqrt{\frac{\tau_{\text{resp}} \tau_{\text{lag}}}{2}}, \quad (30)$$

(cf. Eq S55 in S1 Appendix); the superscript “*c*” indicates that the result holds for critical balance.

Importantly, this implies that the scaling of the response time with the RF width and size improves compared to the purely excitatory network. This is because $\tau_{\text{resp}}^{\text{bal},c} \sim \sqrt{\tau_{\text{resp}}}$. Inserting Eq 14 into Eq 30 for the 1D network, we obtain

$$\tau_{\text{resp}}^{\text{bal},c} \approx \sqrt{\frac{(1 + 2d^2) \tau \tau_{\text{lag}}}{2}} \approx d \sqrt{\tau \tau_{\text{lag}}} \approx \frac{1}{2} n_{\text{RF}} \sqrt{\tau \tau_{\text{lag}}}, \quad (31)$$

which is only linear in the RF width d and size n_{RF} , instead of quadratic as in the case of $w_{\text{sum}}^{\text{rec,bal}} = 0$, compare Eq 31 with Eq 14 and in Fig 6A the red and blue dotted curves. As a consequence also the speedup gained through the balance, $\tau_{\text{resp}}/\tau_{\text{resp}}^{\text{bal},c}$, increases for wider RFs.

We finally note that the balanced interactions mediated by $W^{\text{rec,bal}}$ can also be thought of as implementing an excitatory transmission of activity changes: an activity change in neuron

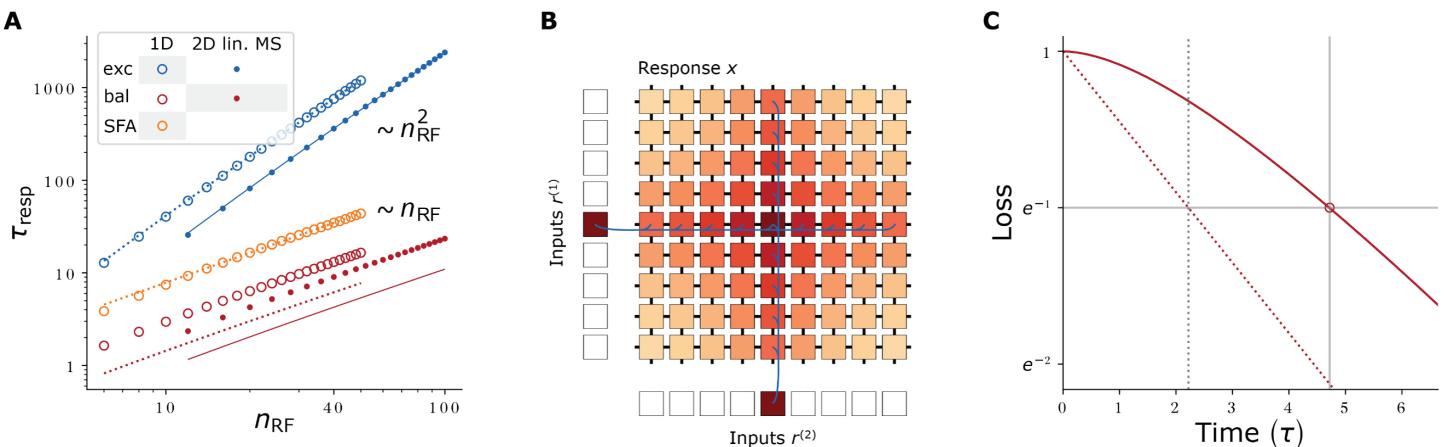


Fig 6. Response speed of networks with inhibition and linear MS. (A) Response times for the 1D network and for the 2D linear MS network. The quadratic scaling of τ_{resp} with n_{RF} for the excitatory networks (blue) can be improved to a linear dependence by introducing balancing, delayed inhibition (red) or SFA (orange). Open (1D network) and filled (MS network) circles display numerical results. Alike-colored dotted (1D network) or continuous (MS network) curves show theoretical estimates (Eqs 14, 31, 37, 38) or, for the SFA network, fit results (monomial fit: $\tau_{\text{resp}}^{\text{SFA}}(n_{\text{RF}}) = 0.66724(n_{\text{RF}})^{1.07063}$). We use the slowest-decaying eigenmode to theoretically estimate the response times (see (C) and Eq S46 in S1 Appendix). Since the balanced networks are not initialized in this eigenmode (in contrast to the purely excitatory networks), the numerically measured response times (red markers) lie above the theoretical values (red lines). (B) Schematic of a 2D network with linear MS. Feature neurons are arranged on a two-dimensional grid (labeled “Response x ”). Each receives feedforward input from two arrays of input neurons (labeled “Inputs $r^{(1|2)}$ ”) and four recurrent inputs. Feedforward and recurrent synapses are shown in blue (exemplarily) and black, respectively. Input and feature neuron activities are color-coded. The (linear) network response is the sum of the responses to input one and input two. (C) Exemplary loss evolution of a 1D network with lagged inhibition. Due to the temporally constant initialization ($x_i(0) = 0$, $\Delta x_i(0) = 0$), the network activity (solid red curve) converges initially more slowly than the network’s slowest eigenmode (dotted red line). The experimentally measured response time (continuous vertical gray line) is defined as the time when the loss has decayed by $1/e$ (red open circle, horizontal gray line), see also Fig 4A. It is larger than that of the network’s eigenmode (dotted gray line), which we use as analytical estimate of the response time. We created the data in (A) by scanning n_{RF} , setting $w_{\text{sum}}^{\text{rec,net}}$ to yield an RF of size n_{RF} , setting $w_{\text{sum}}^{\text{rec,bal}}$ to 0 or its critical value, and determining τ_{resp} or $\tau_{\text{resp}}^{\text{bal}}$ from the loss dynamics. For the SFA network we set $\tau_{\text{SFA}} = \tau$, scanned a_{SFA} and used the value that minimized the temporally integrated loss.

<https://doi.org/10.1371/journal.pcbi.1012156.g006>

j adds an activity change with the same sign to neuron i , because $W^{\text{rec,bal}} = -W^{\text{rec,I}}$ in Eq 26 is positive. Thus activity changes in different neurons in the network amplify each other. In the limit of small τ_{lag} , the temporal derivative is transmitted, see below Eq S21 in S1 Appendix.

Linear mixed selectivity

Neurons often respond selectively to more than one stimulus or input feature [65]. This phenomenon is called mixed selectivity (MS). Since our networks are linear, we consider linear mixed selectivity, where neuronal responses are linear functions of multiple stimuli. This is a simplification compared to the nonlinear mixed selectivity that is ubiquitous in the brain [65, 66]. Concretely, neurons with activities x_{ij} , $i, j = 1, \dots, N$, are arranged on a two-dimensional grid and respond with equal selectivity to two input features, represented by input neurons $r_k^{(1)}$ and $r_l^{(2)}$ with $k, l = 1, \dots, N$,

$$x_{ij}^{\text{resp}} = \sum_{k=1}^N \text{RF}_{ijk}^{(1)} r_k^{(1)} + \sum_{l=1}^N \text{RF}_{ijl}^{(2)} r_l^{(2)}. \quad (32)$$

Due to the linearity in the input representation and in the network, the total response is the sum of the responses to the single input features. We take the grid axes to be aligned with the stimulus dimensions, so that the first index in x_{ij} determines its response to $r^{(1)}$ and the second that to $r^{(2)}$. We model this dependence as the same localized, exponentially decaying shape as for the 1D network (cf. Fig 6B),

$$\text{RF}_{ijk}^{(1)} = \gamma^{|i-k|} \quad \text{RF}_{ijk}^{(2)} = \gamma^{|j-l|}. \quad (33)$$

The desired network response can be generated as the steady state of a recurrent network that is equivalent to the 1D network Eq 5 in each dimension of the 2D grid (see next section),

$$\begin{aligned} \tau \dot{x}_{ij} = & -x_{ij} + w^{\text{rec,MS}} (x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1}) \\ & + w^{\text{ff,MS}} (r_i^{(1)} + r_j^{(2)}), \end{aligned} \quad (34)$$

with the modified constants $w^{\text{rec,MS}} = \frac{w^{\text{rec}}}{1+2w^{\text{rec}}} = \frac{\gamma}{(1+\gamma)^2}$ and $w^{\text{ff,MS}} = \frac{w^{\text{ff}}}{1+2w^{\text{rec}}} = \frac{1-\gamma}{1+\gamma}$. Each neuron receives two external inputs and is connected to its nearest neighbors along each stimulus axis. The network has thus only six synapses per neuron, regardless of the RF width. Also a feedforward network where each dimension of the 2D grid is equal to the 1D network Eq 2 generates the desired response. This implementation requires $n_{\text{RF}}^{\text{MS}} = 2n_{\text{RF}} = 2(2d + 1)$ synapses per neuron, a number that increases linearly with the RF width.

Mapping to a 1D system. In the following, we trace the network dynamics Eq 34 back to those of the 1D system Eq 5. Due to the linearity of Eq 34, network responses again superpose. It thus suffices to study the network in the case where only one input neuron is active: we choose $r_i^{(1)}$, which specifies a property of the first stimulus, to be nonzero. Since the input is independent of j , the dynamics Eq 34 are (for initial conditions homogeneous in j such as $x_{ij}(0) = 0$) independent of j , $x_{ij}(t) = x_i(t)$. The recurrent inputs $w^{\text{rec,MS}}(x_{i,j+1}(t) + x_{i,j-1}(t)) = 2w^{\text{rec,MS}}x_i(t)$ then simply amount to a modification of the leak current to $-(1 - 2w^{\text{rec,MS}})x_i$,

$$\tau \dot{x}_i = -(1 - 2w^{\text{rec,MS}})x_i + w^{\text{ff,MS}}r_i^{(1)} + w^{\text{rec,MS}}(x_{i+1} + x_{i-1}). \quad (35)$$

After dividing by $(1 - 2w^{\text{rec,MS}})$, the differential equation for x_i becomes

$$a\tau\dot{x}_i = -x_i + aw^{\text{rec,MS}}(x_{i+1} + x_{i-1}) + aw^{\text{ff,MS}}r_i^{(1)}, \quad (36)$$

where we introduced $a = (1 - 2w^{\text{rec,MS}})^{-1}$ for brevity. With the values of the constants $w^{\text{rec,MS}}$ and $w^{\text{ff,MS}}$ highlighted after Eq 34 this is equivalent to the one-dimensional network dynamics Eq 5 up to a different neuronal time constant $a\tau$ instead of τ . (We note that we obtained the modified constants such that this holds. For example, equating the prefactors of the recurrent term in Eq 36 and Eq 5 gives $w^{\text{rec}} = aw^{\text{rec,MS}} = (1 - 2w^{\text{rec,MS}})^{-1}w^{\text{rec,MS}}$, which then can be solved for $w^{\text{rec,MS}}$.) As a direct consequence, while the 1D network must have recurrent coupling strength of $w^{\text{rec}} < 0.5$ for being stable, the 2D MS network must have $w^{\text{rec,MS}} < 0.25$. This is because in the MS network, a neuron receives direct recurrent input from four nearest neighbors instead of two as in the 1D case.

Eq 36 means that the RF of the MS network, along one axis, has the same shape and width d as the equivalent one-dimensional network. In particular, d is related to the recurrent weight strength $aw^{\text{rec,MS}}$ via $aw^{\text{rec,MS}} = w^{\text{rec}} = \frac{\gamma}{(1+\gamma)^2}$ and $\gamma = \exp(-1/d)$; the two RF components in Eq 33 are the same as the RFs in Eq 1, for example $\text{RF}_{ijk}^{(1)} = \text{RF}_{ik}$.

Response speed. From the mapping of the MS to the 1D system, Eq 36, we see that the MS dynamics behave in response to a single input like the 1D dynamics with the neuronal time constant τ enlarged by a factor of a . The response time is thus given by Eq 14, but with enlarged neuronal time constant, $\tau \rightarrow a\tau$. For sufficiently large d , we have $\gamma \approx 1$ (reflecting the spatially slow RF decay), $w^{\text{rec}} \approx 1/2$, $w^{\text{rec,MS}} \approx 1/4$ and thus $a \approx 2$. The scaling of the response time with the RF width d and size $n_{\text{RF}}^{\text{MS}}$ is thus again quadratic,

$$\tau_{\text{resp}}^{\text{MS}} \approx (1 + 2d^2)2\tau \approx 4d^2\tau \approx \frac{1}{4} (n_{\text{RF}}^{\text{MS}})^2 \tau. \quad (37)$$

In the last equation we used $n_{\text{RF}}^{\text{MS}} = 2(2d + 1) \approx 4d$. Compared to the 1D case (Eq 14), the response time as a function of d is therefore larger by a factor $a \approx 2$. In contrast, it is smaller by a factor $1/2$ as a function of the RF size, compare Eq 37 with Eq 14 and the blue continuous and dotted curves in Fig 6A. In other words: the trade-off between response time and number of needed synapses improves for sufficiently large RFs by a constant factor of about $1/2$ compared to the 1D network. This is because the MS network effectively implements two 1D RFs (Eq 32).

Balanced network. We now incorporate the effect of inhibitory neurons into the MS network. As in the 1D case, we assume that the generated inhibition precisely tracks excitation with a short time delay. We thus add to each recurrent excitatory connection an inhibitory one that is slightly delayed. This results in a delayed differential equation like Eq 28 for the balanced MS network dynamics. The parameters are given by those of the 1D balanced system up to a factor $a = 1 + w^{\text{rec,net}}$, like in Eq 36. Further, it is again sufficient to study the response dynamics to a single input, which can be reduced to those of the 1D balanced network Eq 28 with adapted parameters. As in the purely excitatory case, for a fair comparison of response times, we consider MS and 1D networks with the same neuronal time constant τ . The effective time constant of the MS dynamics is then $a\tau$. Therefore the response time of the MS network is given by that of the 1D network Eq 31 with neuronal time constant τ replaced by $a\tau \approx 2\tau$,

$$\tau_{\text{resp}}^{\text{bal,MS}} \approx \sqrt{(1 + 2d^2) \tau \tau_{\text{lag}}} \approx \sqrt{2}d\sqrt{\tau \tau_{\text{lag}}} \approx \frac{1}{2\sqrt{2}}n_{\text{RF}}^{\text{MS}}\sqrt{\tau \tau_{\text{lag}}}. \quad (38)$$

At the critical balance, the response time thus scales again with the square root of the response time of the unbalanced network (Eq 30). Therefore it scales linearly with the RF width d and size $n_{\text{RF}}^{\text{MS}}$. The response time is as a function of the RF size by a factor of about $1/\sqrt{2}$ smaller than that in the 1D case, Eq 31, see Fig 6A, red continuous and dotted curves. In other words, the trade-off between response time and number of required synapses improves by a factor of $1/\sqrt{2}$. This is again because the MS network effectively implements two RFs in the MS case; the RF size doubles for the same width compared to the 1D network.

As for the 1D stimulus, the balanced networks have twice as many neurons and additional synapses: Each (excitatory) feature neuron drives one inhibitory neuron, which mirrors its activity. This inhibitory neuron in turn forms inhibitory synapses to the four nearest neighbors that its presynaptic feature neuron excites. In total, the balanced, cooperatively coding network thus requires eleven synapses per feature neuron, instead of six for the unbalanced network. This is independent of the RF width, such that the balanced, cooperatively coding network saves synapses for sufficiently wide RFs.

Higher-dimensional linear MS. We can straightforwardly extend the introduced scheme to networks that have MS with $P > 2$ stimuli. Neurons are then arranged on a hyper-grid with one grid axis per stimulus dimension, so that N^P feature neurons respond to PN input neurons. In the cooperatively coding network, each neuron receives P feedforward and $2P$ recurrent inputs, requiring a total of $3P$ synapses per neuron. The feedforward network, in contrast, needs for each stimulus dimension $2d + 1$ synapses, in total $P(2d + 1)$ synapses per neuron. The number of saved synapses thus grows linearly with the number of encoded stimulus dimensions and the RF width.

Encoding a 2D stimulus

We finally consider the encoding of a two-dimensional stimulus, with both input and feature neurons arranged on a two-dimensional grid, see Fig 7A. Two-dimensional input appears for example in vision [36] or planar navigation tasks [67]. Each feature neuron responds to inputs that are close to its preferred input in both stimulus dimensions. The RFs of neighboring neurons thus overlap and neuronal responses tile the represented stimulus space. A purely excitatory cooperative coding network generating such activity as stationary state is given by

$$\tau \dot{x}_{ij} = -x_{ij} + w^{\text{rec},2\text{D}}(x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1}) + w^{\text{ff},2\text{D}}r_{ij}. \quad (39)$$

It has the same recurrent connectivity as the network with linear MS Eq 34, but the feed-forward input is arranged on a grid. The activity of feature neuron ij in the stationary state is

$$x_{ij}^{\text{steady}} = \sum_{kl=1}^N \text{RF}_{ijkl} r_{kl}. \quad (40)$$

The neuron thus responds to a combination of two input features represented by input neurons r_{ij} with $i, j = 1, \dots, N$. The RF is explicitly given by (see Eq S62 in S1 Appendix)

$$\text{RF}_{ijkl} \approx c \cdot K_0(\gamma^{2\text{D}} \rho_{ijkl}), \quad (41)$$

with $c = \frac{1}{2\pi} \frac{w^{\text{ff},2\text{D}}}{w^{\text{rec},2\text{D}}}$, $\gamma^{2\text{D}} = \sqrt{\frac{1-4w^{\text{rec},2\text{D}}}{w^{\text{rec},2\text{D}}}}$ and $\rho_{ijkl} = \sqrt{|i-k|^2 + |j-l|^2}$. K_0 is the zeroth modified Bessel function of the second kind, which decays with distance ρ approximately as $K_0(\rho) \approx \frac{\pi}{2} e^{-\rho} / \sqrt{\rho + 1/8}$ [68]. The RF is thus approximately radially symmetric. The RF size depends

on $w^{\text{rec},2D}$ and the response amplitude also on $w^{\text{ff},2D}$. The network requires 5 synapses per neuron.

We also construct a balanced network by introducing for each excitatory recurrent input a delayed inhibitory one, like in the balanced 1D and MS networks. A balanced implementation with explicit inhibitory neurons requires twice as many neurons and synapses than the purely excitatory network: it requires additionally one inhibitory neuron per principal neuron, one inhibitory synapse for each excitatory recurrent synapse and one synapse from each principal neuron to its corresponding inhibitory neuron.

Response speed. To estimate the dependence of the response speed on the RF size, we first need to appropriately adapt the definition of the RF size, which we introduced after Eq 3. For the one-dimensional network, this definition can be reformulated as follows: we count the number of synapses that are necessary to generate the largest (around the center) RF responses such that these responses summed together amount to a fraction of about $1 - e^{-1} \approx 63\%$ of the summed non-truncated RF. Accordingly, for the 2D network at hand we define the RF size as the number of feedforward synapses that are necessary to implement the largest RF entries, such that together they account for a fraction of approximately 63% of the summed nontruncated RF. We denote the so-defined RF sizes by \hat{n}_{RF} .

As for the 1D and linear MS networks, the response time with or without lagged inhibition depends only on the summed excitatory weights or on the summed net and inhibitory weights. It is thus given by Eq 13 or by Eq 30 in terms of $w_{\text{sum}}^{\text{rec}} = 4w^{\text{rec},2D}$ or in terms of the alike obtained $w_{\text{sum}}^{\text{rec},\text{net}}$ and $w_{\text{sum}}^{\text{rec},\text{bal}}$. Fig 7B shows that the scaling of the response time with the RF size is linear for unbalanced and square-root-like for balanced networks. We give a geometric argument for this general scaling in the next paragraph. The scaling is more economical than for the 1D and 2D linear MS networks, cf. Fig 6.

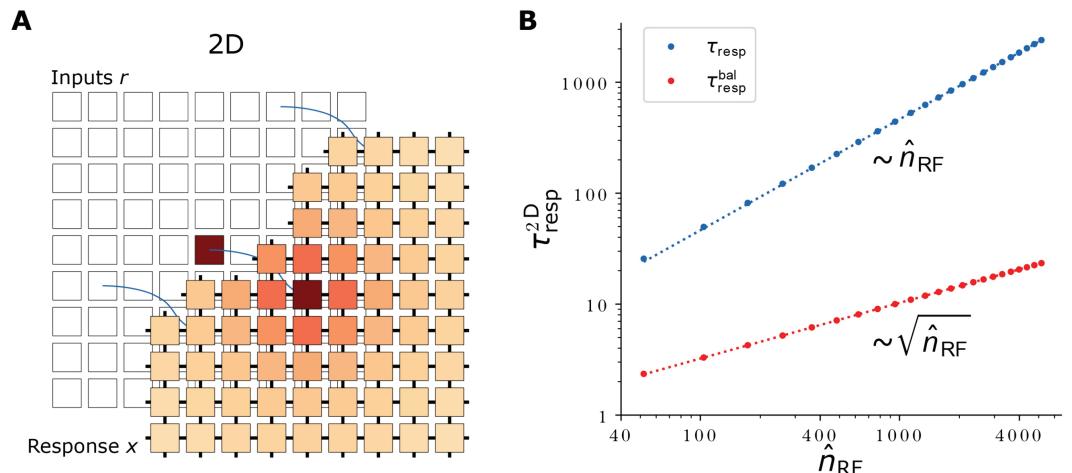


Fig 7. 2D network schematic and response times versus RF size. (A) Schematic of a two-dimensional network responding to a two-dimensional stimulus. Feature neurons (labeled “Response x ”) and input neurons (labeled “Inputs r ”) are arranged on two-dimensional grids. In the cooperatively coding network each feature neuron receives one feedforward and four recurrent inputs; activities and shown connections are color-coded as in Fig 2. (B) Response times in the cooperatively coding 2D network increase linearly (without inhibition, blue. Monomial fit: $\tau_{\text{resp}}^{2D} = 0.46115 (\hat{n}_{\text{RF}})^{1.00085}$) or square-root-like (with inhibition, red. Monomial fit: $\tau_{\text{resp}}^{\text{bal},2D} = 0.31995 (\hat{n}_{\text{RF}})^{0.50202}$) with the RF size \hat{n}_{RF} . Dotted lines represent the monomial fits. Data was created by scanning $w_{\text{sum}}^{\text{rec},\text{net}}$, setting $w_{\text{sum}}^{\text{rec},\text{bal}}$ to 0 or its critical value, and determining \hat{n}_{RF} and τ_{resp} or $\tau_{\text{resp}}^{\text{bal}}$, respectively, from the response curves after network activity converged.

<https://doi.org/10.1371/journal.pcbi.1012156.g007>

Multi-dimensional stimuli. For $P \geq 2$ -dimensional stimuli, the fact that activity simultaneously propagates along all dimensions suggests that the scaling of the response time with d , the characteristic RF width along one of the dimensions, does not change with the number of dimensions. However, the RF size n_{RF} , the number of synapses needed in a purely feed-forward implementation, can be assumed to scale as $n_{RF} \sim (2d + 1)^P$ (number of neurons in a cube with $2d + 1$ neurons at each edge), with a prefactor that depends on geometry. This reasoning suggests that the response time of networks encoding higher-dimensional stimuli scales like $\tau_{\text{resp}} \sim n_{RF}^{2/P}$ (spreading time of activity along one of the dimensions, since spreads in all dimensions happen simultaneously) and $\tau_{\text{resp}}^{\text{bal}} \sim n_{RF}^{1/P}$ (spreading time of activity along one dimension for the balanced network), which we verified for $P = 1, 2$. For higher-dimensional stimuli, the trade-off between response time and saved synapses would thus become highly beneficial: the response time τ_{resp} or $\tau_{\text{resp}}^{\text{bal}}$ would grow only slowly with the number of saved synapses due to the strongly sublinear relationship with n_{RF} for larger P .

Cooperative coding in spiking neural networks

Excitatory networks. So far, we implemented and analyzed cooperative coding in linear rate networks. We now show that the main results concerning synaptic savings as well as scaling of response speed with and without balanced amplification also apply to cooperatively coding spiking neural networks. We demonstrate this for a one-dimensional feature layer with periodic boundary conditions (Fig 8A, cf. Fig 1). Each feature (rate) neuron becomes a feature population composed of multiple spiking neurons. Specifically, we use leaky integrate-and-fire (LIF) neurons (see Methods, Eq 47). Synaptic coupling is sparse and random; couplings are restricted to neurons of the same and neighboring feature populations (Fig 8A). The feedforward input is modeled as a constant mean drive, which is the same for all neurons of a given feature population. Every neuron also receives independent Gaussian white noise mimicking balanced background activity. For a moderate amount of noise and when omitting the absolute refractory period, the single neuron transfer function becomes approximately threshold-linear (Fig 8B). Since we require threshold-linearity only for the network's operating regime between 0 Hz and a peak rate x_{max} , small absolute refractory periods can be incorporated without deviating too much from threshold-linearity.

We use an analytical approximation to construct appropriate networks: We first approximate synaptic inputs with a diffusion approximation. Using the threshold-linear approximation of the single neuron transfer function, we then derive a self-consistency condition for the stationary feature population rates (Eq S64 in [S1 Appendix](#)). This yields for any target RF with size n_{RF} and peak rate x_{max} the feedforward stimulation $I_{\text{ext}}^{\text{on}}$ and the recurrent synaptic weight J_{EE} for which the network is expected to settle into the target rates (Eq S74 – S77 in [S1 Appendix](#)).

For small RF sizes, the analytically obtained stimulation and weight sizes directly yield networks that generate the desired RFs (Fig 8C, gray crosses). Larger RFs depend more sensitively on the recurrent synaptic coupling strength, so we further optimize J_{EE} numerically to obtain the desired RFs (Fig 8C, black circles; see also Discussion and Fig G in [S1 Appendix](#)). We decided not to co-tune the feedforward strength $I_{\text{ext}}^{\text{on}}$ since the simpler, one-dimensional optimization of only J_{EE} already yields satisfactory matches to the target fields.

Analogously to the rate model, the spiking network can exhibit cooperative coding, i.e., the neurons can have wide RFs despite narrow feedforward and limited recurrent connectivity. The stationary spiking activity in our numerically simulated networks is asynchronous (Fig 8D, 8E). For an isolated, constant input to one feature population (black arrow in Fig 8A), the stationary firing rates, averaged within feature populations and over time (Fig 8D,

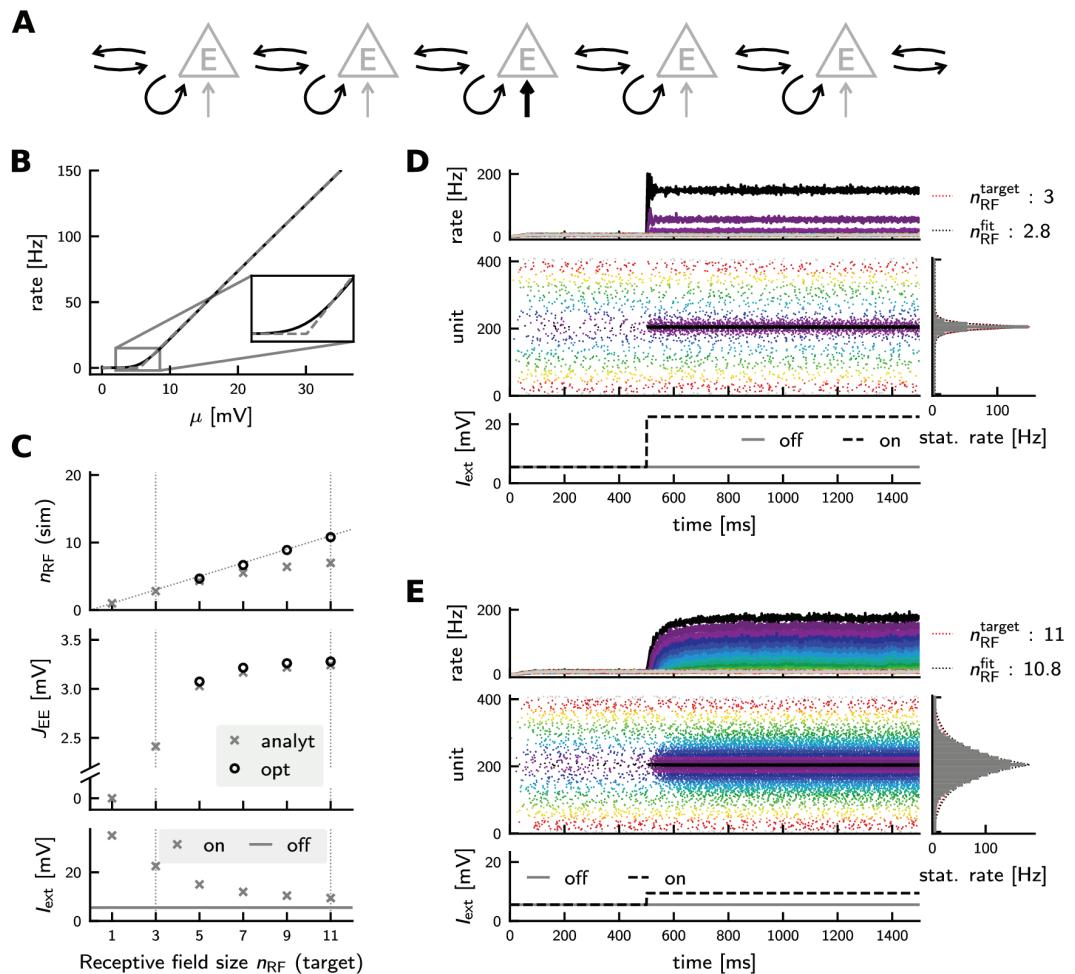


Fig 8. Cooperative coding in an excitatory spiking network of leaky integrate-and-fire neurons. (A) Network wiring diagram. Connections are marked in black if their strength depends on RF size, and in gray otherwise. (B) Single neuron transfer function (black solid) and threshold-linear fit (gray dashed). Inset highlights onset nonlinearity. (C) Top: Simulated vs. target RF size. Light gray crosses: networks with analytically computed weights; black circles: networks with numerically optimized weights. Middle: excitatory synaptic weights (normalized for arbitrary indegree: the voltage increase in response to one spike is given by $W_{EE} = J_{EE}/K_{EE}$). Bottom: Feedforward input to stimulated population (on) and to others (off), used both for networks with analytically or numerically computed weights. Left and right dashed vertical lines mark example simulation shown in (D) and (E). The target RF peak rate was set here to $x_{\max} = 150$ Hz. (D) Example simulation for target RF size $n_{RF} = 3$. Top: Binned average firing rates of 41 feature populations. Middle: Raster plot showing spikes of 10 exemplary neurons from each feature population. Color codes distance from the stimulated population (population index 21, the population includes neurons 200–209 shown in black). Bottom: Feedforward stimulation. Right: Stationary population rates (gray), exponential fit (black dashed), and target rate profile (red). (E) Example simulation for target RF size $n_{RF} = 11$. All panels are as in (D).

<https://doi.org/10.1371/journal.pcbi.1012156.g008>

8E, right panel, gray bars), decay with distance from the stimulated population, as desired. The slight deviations from the exponentially decaying target profile (Fig 8D, 8E, right panel, red dotted line) likely arise from the onset nonlinearity of the transfer function (see inset in Fig 8B).

For the simulations shown in Fig 8, we fixed the indegree of the recurrent synaptic coupling and used relatively large feature populations to stabilize the dynamics and facilitate the estimation of stationary rates and response speeds. We confirmed in exemplary simulations

that cooperative coding is still possible with smaller (e.g. $N_E = 500$) feature populations with random Erdős-Rényi connectivity and under Poisson spiking input (Fig H in [S1 Appendix](#)).

Number of synaptic connections. As in the rate model, cooperative coding can save synapses compared to a feedforward network: With pure feedforward coding, every feature neuron needs to receive input from as many input populations as the RF is large ($\sim n_{RF}$, cf. [Fig 2A](#)) — as well as recurrent input from its peers, if we include within-population coupling for comparability with the cooperatively coding network. On average this yields

$$N_{FF} = n_{RF}K_{FF} + K_{EE}^{in} \quad (42)$$

synaptic inputs per feature neuron, where K_{FF} denotes the average feedforward indegree from an input to a feature population and K_{EE}^{in} denotes the average indegree of recurrent coupling within a feature population. In the cooperatively coding network, the number of synaptic inputs is independent of RF size: Every feature population receives input from only one input population, from itself, and from its two neighboring feature populations ([Fig 2B](#), [Fig 8A](#)):

$$N_{CC} = K_{FF} + K_{EE}^{in} + 2K_{EE}^{cross}, \quad (43)$$

where K_{EE}^{cross} denotes the average indegree for connections across two neighboring feature populations. We note that synaptic coupling within feature populations ($K_{EE}^{in} > 0$) is not required for either feedforward or cooperative coding, but was included here for biological plausibility. In biological neural networks, it might be used for cooperative coding between identically tuned neurons to save further feedforward synapses. In our simplest model we assume that all indegrees are equal, $K_{FF} = K_{EE}^{in/cross} =: K$. In this case, the number of synapses per feature neuron is smaller in the cooperative coding network for all RF sizes larger than three:

$$N_{CC} = 4K < (1 + n_{RF})K = N_{FF} \quad \text{for } n_{RF} > 3. \quad (44)$$

The total number of synapses in the network depends also on the size N_E of the N_F feature populations. We have $N_{FF}^{\text{total}} = N_F N_E N_{FF}$ and $N_{CC}^{\text{total}} = N_F N_E N_{CC}$ synapses for feedforward and cooperative coding architectures, respectively. Furthermore, if we assume a fixed connection probability p_{EE} between neurons of connected feature populations, $K_{EE}^{in/cross}$ increases with N_E . The total number of synapses in the network then becomes

$$N_{FF}^{\text{total}} = N_F N_E N_{FF} = n_{RF} N_F K_{FF} N_E + p_{EE} N_F N_E^2, \quad (45)$$

$$N_{CC}^{\text{total}} = N_F N_E N_{CC} = N_F K_{FF} N_E + 3p_{EE} N_F N_E^2. \quad (46)$$

(If, instead, the indegree of each neuron stayed constant, i.e., independent of N_E , the scalings would only be linear in N_E .) We observe that for larger feature populations synaptic savings happen from larger field sizes on, [Fig 9](#). The cooperative scheme relies on averaging recurrent synaptic input and therefore requires a certain minimal population size N_E . Our estimate [Fig 9](#) suggests that a cooperative spiking network with 1000 neurons per feature population may save synapses for RFs larger than ~ 3 .

Response speed and balanced networks. In the purely excitatory spiking network, the response time increases approximately quadratically with RF size for larger RFs ([Fig 10C](#), blue). In rate networks, we observed a speedup and even an improved scaling of the response time for balanced networks. To investigate whether this also occurs in spiking networks, we implement a spiking version of the balanced rate network described in [Eq 21](#) (see Methods,

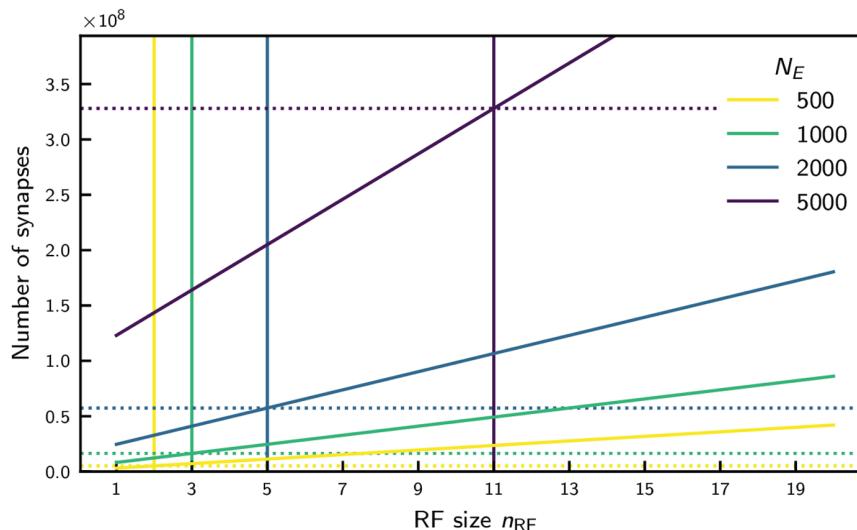


Fig 9. Number of synaptic connections for the feedforward and cooperative coding architecture depending on RF size, for different sizes of feature populations. Total number of synapses in a network of $N_F = 41$ feature populations as a function of RF size n_{RF} and number of neurons N_E per feature population (color-coded). In the feedforward network the number of synapses N_{FF}^{total} (solid lines) depends linearly on RF size with a slope proportional to N_E . In the cooperative network the number of synapses N_{CC}^{total} (dotted lines) is independent of RF size but increases quadratically with N_E . Vertical lines mark the minimal RF size for which the cooperative network saves synapses compared to the feedforward network ($N_{CC}^{\text{total}} \leq N_{FF}^{\text{total}}$). Synapse numbers are shown here for $K_{FF} = 100$ and $p_{EE} = 0.1$.

<https://doi.org/10.1371/journal.pcbi.1012156.g009>

Eq 48 and Eq 49, and Fig 10A). Taking advantage of the approximately threshold-linear transfer functions of single neurons, we tune the synaptic strength from excitatory to inhibitory feature populations such that the inhibitory populations fire at approximately the same stationary rate as their excitatory counterparts, analogous to our rate models (Eq S78 in S1 Appendix). To construct a balanced network, we increase the excitatory recurrent coupling J_{EE} by a factor $s > 1$, and tune the inhibitory-to-excitatory projections such that the balanced network exhibits an RF of approximately the same size as the excitatory reference network (Fig 10B, Eq S81 in S1 Appendix). For each target RF size n_{RF} we broadly grid-search for the scaling factor s that yields the shortest rate response time (Fig 10C). For such optimal amplification s_{opt} , the scaling of the response speed with respect to RF size improves from quadratic to linear as in the rate model (cf. Fig 6). As expected from the theory (Eq S43 in S1 Appendix), the optimal scaling factor increases with RF size (Fig 10C, bottom). The speedup of the response is likely due to balanced amplification as in the rate network.

We note that, depending on the nature and strength of the noise, spiking neurons can exhibit fast rate responses with transient onmodulated oscillations to step current input [69–71]. We observe such oscillations in our purely excitatory spiking networks for small RF sizes (where we have strong feedforward input), e.g. Fig 8D. This is in contrast to our purely excitatory rate networks. We also find transient oscillatory rate modulations in our spiking EI networks, e.g. Fig 10E.

Given their often rather sharp tuning [62–64], it is conceivable that not only excitatory, but also inhibitory feature populations receive tuned feedforward input. We thus also construct a network where the feedforward input to both subpopulations is the same (see Eq S101 in S1 Appendix). Inspired by “Model A” of [72] we furthermore use the same connectivity and biophysical parameters for both excitatory and inhibitory subpopulations, such that they have

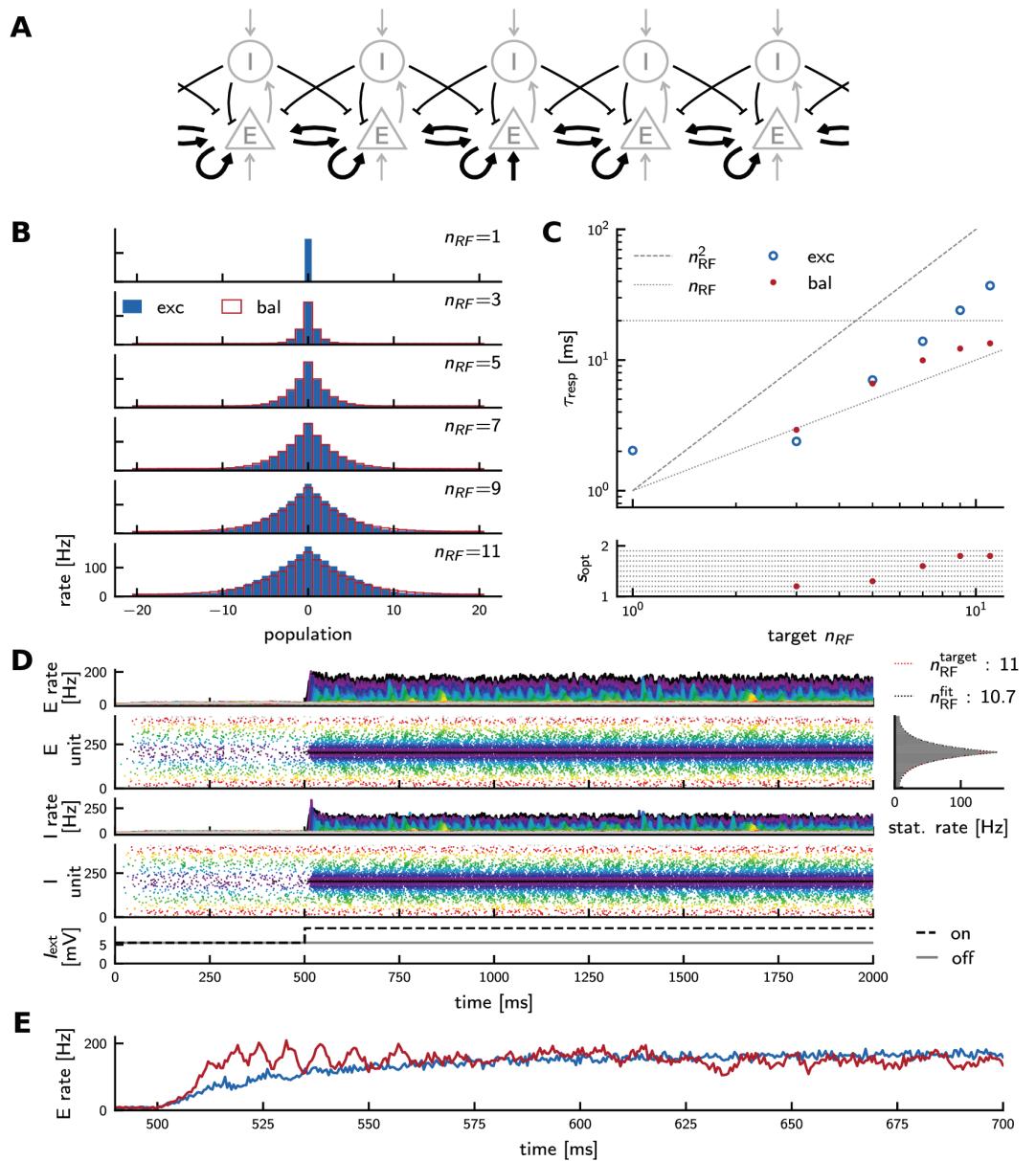


Fig 10. Response times of cooperatively coding spiking neural networks (A) Network wiring diagram of the balanced excitatory-inhibitory spiking network. Connections are marked in black if their strength depends on RF size, and in gray otherwise. (B) Stationary rate profiles of excitatory (blue) and balanced networks (red) in response to isolated inputs match. There is no balanced network simulation shown for $n_{RF} = 1$, since there is no recurrent coupling between or within populations in that case. (C) Top: Response times. Blue: purely excitatory network (cf. Fig 8). Red: balanced excitatory-inhibitory network. Dashed horizontal line: membrane time constant. Quadratic and linear scaling is indicated by the dashed and dotted lines, respectively. Bottom: scaling factor s used to scale up the excitatory synaptic coupling strength J_{EE} in the balanced network. (D) Activity of a balanced network with RF size 11. Panels as in Fig 8D, 8E; top: excitatory, bottom: inhibitory populations. (E) Direct comparison of the rate of the stimulated population (index 21) around stimulus onset in the purely excitatory network (blue, cf. black trace in Fig 8E), and in the balanced network (red, cf. black trace in panel D, top). Rates are shown for a finer binsize of 0.5 ms.

<https://doi.org/10.1371/journal.pcbi.1012156.g010>

the same stationary rates. We fix the coupling strengths within EI feature populations and only tune the across-population coupling and feedforward input to achieve the desired RF size

(Eq S108 and S110) in [S1 Appendix](#)). Interestingly, in this network it seems that the balanced amplification mechanism cannot come into effect: In the terminology of [21], both feedforward and recurrent inputs recruit only sum, not difference modes. Nevertheless, we find that the response speed of this network is enhanced compared to an excitatory-only network with the same excitatory within-population coupling strength (Eq S85 and Fig I in [S1 Appendix](#)). This may be a consequence of the fact that embedding neurons in balanced EI nodes slightly flattens their transfer function (Fig Ib in [S1 Appendix](#)). Thus, the across-population coupling strength required for a certain RF size to emerge is larger than in the purely excitatory network (Fig Ic in [S1 Appendix](#)). This increase in cross-population coupling may induce the speed-up of signal propagation.

Discussion

Cooperative coding. In this work, we have studied networks that encode continuous variables with neurons that have overlapping response properties. We developed a *cooperative coding* scheme, which enables similarly tuned neurons to share and distribute computations, crucially using (net) excitatory connections. In general, the signature of cooperative coding is that the network trades feedforward and less specific recurrent synapses for fewer specific recurrent ones.

Saving synapses. How can a network save synapses by constructing a given response from different sets of recurrent and feedforward connections? The key observation is that the outputs of few similarly tuned neurons already provide a “large part” of a neuron’s input-output transformation, as well as indirect access to many input neurons. For the simplest considered networks this sharing of computations minimizes the number of required synapses while the total amount of synaptic current remains the same as in a purely feedforward implementation. For networks of neurons that represent higher-dimensional stimuli [65,66], the number of saved synapses is especially large. Our results thus suggest the number of synapses and space constraints as a possible normative reason underlying the cortical like-to-like excitation, which currently lacks such explanation.

Response time. The saving of synapses comes at the cost of longer response times. In our most simple, purely excitatory cooperatively coding networks the response is slowed down compared to that of single neurons due to recurrent excitation, which implements a positive feedback loop. This type of amplification has been termed “Hebbian amplification” in [21]. Using rate neuron models, we find, however, that neurons with SFA and neurons in networks in which excitation is largely balanced by delayed inhibition can use the window of opportunity between the arrival of excitation and inhibition to significantly speed up their convergence to the steady state response. This balanced amplification [21] decreases response times by orders of magnitude and improves their scaling with RF size. Implementing cooperative coding in spiking neural networks, we find that these show the same scaling and improvement through balancing inhibition.

Our balanced spiking networks exhibit response times of ~2–20 ms for RFs of sizes 1–11. This is roughly consistent with experimental measurements of response times, which vary between 5–10 ms measured in V1 single cells from response onset [73] to ≥50 ms measured in populations from stimulus onset [74,75]. However, a direct comparison of our model’s response time with experimental data is challenging, since (1) the onset of feedforward input currents is often unknown in experiments, and (2) our biophysical model parameters have not been matched to any cortical region in particular. Response times are generally thought to increase along the cortical hierarchy [76–79]. This is consistent with an accumulation

of response slowdowns and with increased response times for larger RFs, which are both expected in our model.

Model choices. Linear rate networks are sufficient for the main part of our study, as we do not model nonlinear phenomena such as feature competition [18]. They allow to capture in an intuitively well-accessible manner the core insight that a desired response can be constructed using few, specific excitatory inputs from similarly tuned neurons (and few feed-forward inputs), instead of many feedforward inputs. In addition, linear rate networks can be mathematically well analyzed. This enables a detailed understanding of their response dynamics. In particular, we analytically obtain the network response times and their scaling behavior.

We test the results in spiking neural networks of LIF neurons for a one-dimensional stimulus. Encouragingly, we observe the same qualitative rate dynamics and RFs as in the linear rate networks.

Networks with SFA. To speed up the response, we first introduce SFA, a typical feature of excitatory principal neurons [42,50,51]. In the networks with SFA, excitation still dominates, i.e., we have on the one hand Hebbian amplification. On the other hand, we have balanced amplification [21], which speeds up responses. The delayed inhibition thereby originates from private adaptation currents, instead of inhibitory neurons as for balanced amplification in EI-networks.

Balanced networks. We model the inhibitory activity in our rate networks as mirroring excitatory activity with an explicit lag. This enables an analysis of the convergence toward equilibrium with techniques from the theory of delay differential equations. Depending on the strength of the balance and the lag of inhibition, this revealed qualitatively different types of dynamics, which are familiar from the harmonic oscillator, namely overdamped, critical, and underdamped dynamics. In our networks with SFA and in balanced amplification networks of others [21], the effective lag of inhibitory feedback originates instead from the fact that inhibitory currents are evoked by a low-pass filtered version of the excitatory activity. In contrast to such models, the inhibitory activity in our balanced networks contains the undamped high-frequency components of the excitatory activity and shifts them by the same delay as the low frequency ones. Because we evaluate our balanced networks at the critical balance, where the time scale of the network dynamics is much larger than the lag, high-frequency components are likely unimportant. Therefore, we do not expect qualitative differences and only small quantitative differences between both implementations in our linear rate networks. Exemplary simulations confirm this expectation, Fig D in [S1 Appendix](#). The reasoning also explains why we find the same qualitative dynamics in our spiking networks, where the inhibitory neurons are fully modeled as LIF neurons with biologically plausible membrane time constants and synaptic delays that govern the delayed feedback to excitatory populations. The basic mechanism of shortening the impulse response and speeding up the reaction to inputs is the same in all these cases.

RFs in model and experiment. In the brain the responses of neurons from lower areas are combined to determine the responses of hierarchically higher ones [80,81]. This offers the general opportunity to harness the concept of cooperative coding. RF size often increases along the processing hierarchy, for example by a factor of about 3 to 10 along the ventral visual stream of humans and macaques [82,83]. Ref. [40] suggests a ~10-fold increase in place field size from dentate gyrus [84] to CA3 [85]. For such changes in RF size, our cooperative coding scheme predicts prominent savings in the number of required synapses compared to purely feedforward networks.

The exponentially decaying RFs that we consider act as tractable models for experimentally encountered localized, overlapping and broadening RFs [2,38,40]. They allow to elegantly illustrate how neurons can use recurrent interactions to cooperatively share feedforward information and shape the network response. However, experimentally measured RFs have different and often more complex shapes [7–9,36]. We are optimistic that these can be approximated by the steady state of a cooperatively coding recurrent network with sparse connectivity, although more synapses will be required. Determining the necessary network parameters might involve minimizing a loss with L0 regularization, which is challenging.

Related work. Conceptually, our 1D model is a ring model, operating in the input-driven regime with a single stable ground state [42]. Ring models have been proposed to model orientation selectivity in the visual cortex [23,53] (see also [86]), head direction cells [87] and spatial memory [88]. Similarly, our 2D and 2D mixed selectivity models have a toroidal or, when removing the periodicity, a planar structure. Such networks may be important for spatial navigation [67]. Previous models have broad coupling fields or ranges of coupling probabilities, equivalent to many recurrent synaptic connections that extend over neurons with quite different preferred stimuli [23,42,53,67,87,88]. In contrast, in cooperatively coding networks, we have very sparse synaptic connections between neurons with highly similar tuning.

Our work considers the encoding of continuous variables in a scheme with minimal numbers of required synapses. Ref. [89] investigated a different but related problem: the binary and multinomial classification of random patterns in large networks of neurons with limited and fixed indegrees. The study finds that if an intermediate layer (which is the analogue to our feature layer) is equipped with sparse, excitatory like-to-like recurrent connections, then sparse feedforward connectivity and a sparse readout are sufficient for classification regardless of network size. This connectivity and the resulting saving of synapses imply that the networks realize cooperative coding in our sense. We note that in their Hopfield network intermediate layers could implement cooperative coding by realizing stronger weights with higher (instead of equal) probability.

In a 1D-ring model, ref. [18] had to incorporate strong nearest-neighbor-like excitatory interactions to match experimentally found responses, as they stabilize network responses in the presence of input noise. In our model such connections even determine the RF.

Ref [90] found that local recurrent connectivity in Hebbian assemblies of spiking neurons can reduce the number of feedforward connections between assemblies required for memory replay. The total number of synapses in their model is, however, minimized by a purely feedforward architecture.

Consistent with our results for shallow networks, intermediate-depth ML networks featuring recurrent and feedback connections can match the performance of much deeper feedforward networks while requiring less units and parameters [91]. It would be interesting to investigate whether the recurrent connectivity in such networks is also like-to-like. If so, this would indicate that cooperative coding naturally appears also in ML networks. It may be helpful in particular in convolutional networks, like our models, to save feedforward connections and rely on very sparse recurrent connectivity instead.

Properties of connectivity. Our cooperative coding scheme relies on the presence of few strong recurrent excitatory connections between similarly tuned cells; inhibition needs to leave the functional connectivity excitatory. This fits data in visual cortex, which shows that pyramidal neurons with similar RFs connect at higher rates and with stronger synapses [14–16]. Recurrent connections are generally sparse in the cortex [92–95]. Furthermore, net functional connectivity is excitatory between (spatially close) neurons with similar tuning [19] and most correlated responses [18]. Refs. [18,19] also show an inhibitory effect on largely differently tuned neurons. Ref [18] found net inhibition between rather similarly tuned neurons

as well. This is assumed to implement feature competition, which we did not include in our model.

A particular benefit of our cooperative coding scheme is that it allows feedforward connections to be sparse. This fits for example experimental observations in V1, where the vast majority of inputs are local recurrent ones, while only a few percent are feedforward inputs [96,97]. Ref. [98] estimated based on experimental studies [99] that a single hypercolumn in primate V1 receives only 10-30 feedforward inputs from the magnocellular layer of dorsal LGN mediating retinal input, with single cells in L4 α receiving as little as 0 - 6 inputs. Pyramidal cells in the hippocampal region CA3 may receive input from only about 50 dentate gyrus neurons but from 6000 other CA3 pyramidal cells [100,101]. This is in line with cooperative coding of spatial inputs from dentate gyrus in CA3 and could explain the enlargement of place fields along this pathway [40]. Note, however, that CA3 also receives inputs from entorhinal cortex [102–104].

Experiments that aim to disentangle feedforward from recurrent contributions to orientation selectivity resulted in mixed findings. Ref. [105] showed that excitatory postsynaptic potentials in simple cells in L4 of cat V1 exhibit orientation tuning to drifting gratings, even when recurrent inputs are suppressed by cortical cooling. In line with this, ref. [17] found that thalamic and cortical contributions to the first harmonic (F1) of the response curve to drifting gratings are co-tuned. However, the temporally averaged response (F0) is tuned only in cortical but not in thalamic inputs. A recent study, ref. [106], suggests that the total input current from L4 of mouse primary visual cortex to L2/3 may lack orientation tuning and that orientation selectivity is determined by recurrent inputs from within L2/3.

Optimality. We assessed optimality in terms of energy, synapse numbers, and response speed. We find that the simplest, purely excitatory cooperatively coding network minimizes the number of required synapses and has a similar metabolic cost as a feedforward implementation. Adding SFA or balancing inhibition reduces response times but increases metabolic cost. Balancing inhibition also requires additional synapses and neurons. We conclude that the brain might use cooperative coding to save synapses and space compared to a purely feedforward or more wasteful recurrent implementation, but might invest some synapses, neurons, space and energy in balancing inhibition to retain a reasonable response speed.

Previous studies often minimized the number of spikes or, more generally, the neuronal activity needed to represent encoded features. Refs. [30,32,34] follow this approach and suggest that tight EI-balance may be a signature of a highly coordinated and competitive code that, despite the irregular firing, is orders of magnitude more precise than a Poisson rate code. This spike-code depends on an extremely structured, dense connectivity, through which similarly coding neurons quickly inhibit each other to prevent redundant spiking. From this standpoint the findings of excitatory functional connectivity between very similarly tuned neurons [15,16,18,19] seem counter-intuitive.

Spiking networks. We verified that the central insights from linear rate networks hold also for biologically more detailed models, by creating cooperatively coding spiking neural networks. In particular, cooperative coding can still save synapses for reasonably sized RFs and introducing balancing inhibition improves the scaling of the response time with the RF size from quadratic to linear. The spiking networks consist of a set of discrete feature populations, with the same connectivity statistics for all neurons within one population. In the stationary state, each feature population encodes as latent variable the stationary activity of a feature neuron of our linear rate networks. The stationary spiking activity thus lies in a low-dimensional manifold of the space spanned by all neurons [107]; the dimensionality is given by the number of feature populations. However, different circuit structures can give rise to

the same low-dimensional activity [33,108,109]. We expect that similar dominant latent variables as in our current spiking networks can be reached with spiking neurons arranged quasi-continuously along a ring (or in 2D space) and with tuning-dependent sparse like-to-like connectivity that realizes cooperative coding.

Sensitivity to recurrent weight strength. To produce larger RFs, the cooperative network requires larger recurrent weights, which converge toward a critical value. Thus, for larger RFs, the cooperative network is increasingly sensitive toward variation in the recurrent weight and an accurate numerical integration of the spiking networks requires increasingly fine discretization time steps (cf. Fig E in [S1 Appendix](#)). Stability can be increased in a hybrid coding model, which uses a small fraction of the inputs from the feedforward model, and complements them with cooperatively coding recurrent inputs. We expect that this may again save synapses compared to the feedforward model.

Experimental predictions. In our cooperatively coding models, feedforward input is sparse [98,99,110] and, compared to recurrent input, weak [17]. It directly contributes to only a small part of a neuron's receptive field or response. In our models, only the center neuron/population receives direct feedforward input. When taken literally, our model predicts strong excitatory connections only between very similarly tuned cells, consistent with some recent experimental findings [18]. A hybrid coding model would yield co-tuning of feedforward and recurrent inputs [17,81,105]. In either case, feedforward input is amplified through recurrent connections, which redistribute it and thereby establish the full response. Our models thus predict that removal of recurrence should lead to responses and receptive fields that are much smaller in amplitude, and deficient in the sense that they lack responses to many inputs within the full RF.

Conclusion. To conclude, net excitatory connectivity between similarly tuned neurons is compatible with a novel cooperative coding scheme that generates network responses with a minimal number of synapses. This suggests space constraints as an important factor in shaping neural networks, providing a possible normative explanation for excitatory like-to-like connectivity. The window of opportunity between excitation and balancing, delayed adaptation or inhibition may be harnessed to rapidly propagate activity changes through the network, speeding up equilibration times by orders of magnitude.

Methods

Rate models

All simulations have periodic boundary conditions. Fixed network parameters are the number of neurons N for 1D and N^2 for 2D networks, the neuronal time-constant τ and, in networks with inhibition, the EI-lag τ_{lag} . We set $N = 200$, $\tau = 1$ and $\tau_{\text{lag}} = 0.1$. In the networks with SFA we use a fixed value of $\tau_{\text{SFA}} = \tau = 1$ and, for each RF size, obtain the value of a_{SFA} that minimizes the temporal mean of the normalized L1-loss, $(1/T) \int_0^T dt |x(t) - x^*(t)|_1 / |x^*|_1$, through a linear grid search. Here $T = 500\tau$ is the length of a trial as described in Fig Aa in [S1 Appendix](#) and $x^*(t)$ the target corresponding to the present input. Fig Ae,c show the scans over a_{SFA} and the individual loss curves for the optimal a_{SFA} values. In all networks with inhibition, we set $w_{\text{sum}}^{\text{rec,bal}}$ to its critical value given by Eq S43 in [S1 Appendix](#).

We simulate our networks with SFA using the Euler method and all other networks using the midpoint method with stepsize $dt = 0.01$. To simulate the networks with delayed inhibition, we also need midpoint values of the delayed activity. We obtain them by copying the midpoint values of the non-delayed activity τ_{lag} (τ_{lag}/dt simulation steps) before.

For the data in [Figs 4](#) and [6](#), we obtain RFs with different sizes by setting $w_{\text{sum}}^{\text{rec}}$ or $w_{\text{sum}}^{\text{rec,net}}$ to appropriate values: In the case of 1D networks with and without SFA and in the case of 2D

linear MS networks, we have analytical expressions for the RF sizes as a function of $w_{\text{sum}}^{\text{rec}}$ or $w_{\text{sum}}^{\text{rec,net}}$. We thus chose $w_{\text{sum}}^{\text{rec}}$ or $w_{\text{sum}}^{\text{rec,net}}$ such that the RF sizes are sampled linearly from $n_{\text{RF}}^{\text{1D}} = 6$ to $n_{\text{RF}}^{\text{1D}} = 50$ in steps of two. For the 2D network, we simulate networks with 20 different values of $w_{\text{sum}}^{\text{rec}}$ or $w_{\text{sum}}^{\text{rec,net}}$ and measure the RF sizes that the networks generate after convergence. We obtain $w_{\text{sum}}^{\text{rec}}$ or $w_{\text{sum}}^{\text{rec,net}}$ as $w_{\text{sum}}^{\text{rec}}$ (or $w_{\text{sum}}^{\text{rec,net}}$) = $1 - \tau/\tau_{\text{resp}}$ by varying τ_{resp} from 10 to 1,000 with equal spacing on a logarithmic scale.

To numerically determine a network's response time, we first simulate the network for a long time, clearly longer than the convergence time, and define the resulting state as the final, target state x^* . The loss is the L1 norm of the difference between x^* and the current state. We then simulate the network for a second time. We obtain τ_{resp} or $\tau_{\text{resp}}^{\text{bal}}$ as the earliest time at which the loss drops and stays below e^{-1} times the initial loss.

Spiking models

Excitatory network. We model N_F feature populations arranged on a one-dimensional ring. Every feature population contains N_E excitatory neurons, which are modeled as leaky integrate-and-fire (LIF) neurons. Feature populations are connected randomly and sparsely to their neighbors, as well as recurrently within themselves.

The membrane potential of neuron k in feature population i is described by the stochastic differential equation (SDE)

$$\tau_m \dot{v}_{ik} = -v_{ik} + V_{\text{rest}} + I_{\text{ext}}^i + w_{\text{EE}} \tau_m \sum_{\Gamma_{ik}^{\text{E}}} \sum_{\hat{t}} \delta(t - \hat{t} - \Delta_{\text{EE}}) + \sqrt{2\tau_m} \sigma \xi_{ik}(t). \quad (47)$$

We set $V_{\text{rest}} = 0$ mV and leave it out in the following to simplify the notation. When the membrane potential reaches the threshold V_{thr} , a spike is generated and the membrane potential is reset to V_{reset} . There is no absolute refractory period. All neurons of a feature population i receive the same feedforward input I_{ext}^i : For the stimulated population (denoted as $i = 0$ in the following) this feedforward input is stronger ($I_{\text{ext}}^0 = I_{\text{ext}}^{\text{on}}$), while all other populations receive a lower background input ($I_{\text{ext}}^i = I_{\text{ext}}^{\text{off}}$, $i \neq 0$). Each neuron receives independent Gaussian white noise ξ with $\langle \xi_{ik}(t) \rangle = 0$, and $\langle \xi_{ik}(t) \xi_{jl}(t') \rangle = \delta_{ij} \delta_{kl} \delta(t - t')$; σ specifies its strength. Further, a neuron receives excitatory input from a pool $\Gamma_{ik}^{\text{E}} = \Gamma_{ik}^{\text{E,in}} \cup \Gamma_{ik}^{\text{E,out}}$ of neurons from its own population i ($\Gamma_{ik}^{\text{E,in}}$) as well as from the two neighboring populations $i \pm 1$ ($\Gamma_{ik}^{\text{E,out}}$). The pool of presynaptic neurons is constructed by drawing from each admissible population $K_{\text{EE}} = p_{\text{EE}} N_E$ presynaptic neurons (excluding autapses) — i.e., the indegree is fixed. The inner sum in Eq 47 is taken across all spike times \hat{t} of all neurons in these presynaptic pools. Each presynaptic spike induces a jump of the postsynaptic membrane potential of size w_{EE} after a delay Δ_{EE} . Synaptic delays are drawn randomly from a uniform distribution between 0 and 2 ms. Table 1 summarizes all parameter values. The tuning of the recurrent synaptic weight strength w_{EE} and the feedforward input $I_{\text{ext}}^{\text{on}}, I_{\text{ext}}^{\text{off}}$ is described in Eq S74 – S77 in S1 Appendix.

Balanced network. We add N_F inhibitory feature populations, each containing N_I inhibitory neurons. The membrane potential of a neuron k in excitatory feature population i is given by the SDE

$$\begin{aligned} \tau_m \dot{v}_{ik}^E = -v_{ik}^E + I_{\text{ext}}^i + \bar{w}_{\text{EE}} \tau_m \sum_{\Gamma_{ik}^{\text{E,in}} \cup \Gamma_{ik}^{\text{E,out}}} \sum_{\hat{t}} \delta(t - \hat{t} - \Delta_{\text{EE}}) \\ + w_{\text{EI}} \tau_m \sum_{\Gamma_{ik}^{\text{I,in}} \cup \Gamma_{ik}^{\text{I,out}}} \sum_{\hat{t}} \delta(t - \hat{t} - \Delta_{\text{EI}}) + \sqrt{2\tau_m} \sigma \xi_{ik}^E(t). \end{aligned} \quad (48)$$

Table 1. Parameters of spiking network models. Top: excitatory-only network (Fig 8). Bottom: additional parameters for balanced network (Fig 10). Equation numbers refer to S1 Appendix.

Parameter	Value	Definition
N_F	41	Number of feature populations
N_E	4000	Excitatory neurons per feature population
p_{EE}	0.1	Connection probability from E to E
w_{EE}	Eq S75	Synaptic weight (E to E)
V_{thr}	10 mV	Spike threshold
V_{rest}	0 mV	Resting voltage
V_{reset}	0 mV	Reset voltage
τ_m	20 ms	Membrane time constant
τ_{ref}	0 ms	Refractory period
Δ_{EE}	$\sim \mathcal{U}(0, 2)$ ms	Synaptic delay from E to E
dt	0.01 ms	Simulation time step
N_I	1000	Inhibitory neurons per feature population
p_{EI}	0.1	Connection probability from I to E
p_{IE}	0.1	Connection probability from E to I
\bar{w}_{EE}	Eq S79	(increased) Synaptic weight (E to E)
w_{IE}	Eq S78	Synaptic weight (E to I)
w_{EI}	Eq S81	Synaptic weight (I to E)
Δ_{IE}	$\sim \mathcal{U}(0, 2)$ ms	Synaptic delay from E to I
Δ_{EI}	$\sim \mathcal{U}(0, 2)$ ms	Synaptic delay from I to E

<https://doi.org/10.1371/journal.pcbi.1012156.t001>

The membrane potential of a neuron k in inhibitory feature population i similarly obeys

$$\tau_m \dot{v}_{ik}^I = -v_{ik}^I + I_{\text{ext}}^{\text{off}} + w_{IE} \tau_m \sum_{\Gamma_{ik}^{\text{E,in}}} \sum_{\hat{t}} \delta(t - \hat{t} - \Delta_{IE}) + \sqrt{2\tau_m} \sigma \xi_{ik}^I(t). \quad (49)$$

Excitatory and inhibitory populations are connected as described for the balanced rate network and illustrated in Fig 10A: An inhibitory neuron k in population i receives excitatory synaptic input from the excitatory partner population ($\Gamma_{ik}^{\text{E,in}}$). An excitatory neuron k in population i receives synaptic inputs from the inhibitory partner population ($\Gamma_{ik}^{\text{I,in}}$), as well as its two neighbors ($\Gamma_{ik}^{\text{I,out}}$). Excitatory-to-excitatory connections are as in the purely excitatory network described above. For all pathways, synapses are again drawn randomly with a connection probability of 10%, while imposing a fixed indegree. Inhibitory neurons have the same biophysical parameters as excitatory cells. All synaptic delays are drawn randomly from a uniform distribution between 0 and 2 ms. Table 1 summarizes all parameter values. The feedforward input to inhibitory neurons is untuned. The tuning of the weights between excitatory and inhibitory populations is described in Eq S78 and S81 in S1 Appendix.

Numerical simulations. All spiking network simulations are performed using the spiking network simulator Brian2 [111]. For RF sizes of $n_{\text{RF}} \leq 13$ we simulate $N_F = 41$ feature populations. For larger fields ($n_{\text{RF}} \geq 15$) we use $N_F = 61$ to reduce boundary effects. The feedforward stimulation is chosen to target the central population (i.e., population 21 for $N_F = 41$ or population 31 for $N_F = 61$, respectively). First, the network is simulated for 500 ms with only background-level feedforward input. Then the feedforward input to the stimulated population is increased instantaneously and the network is simulated for three seconds to ensure that a stable state has been reached.

We estimate the stationary firing rates by averaging over the last second of the simulation (second 2–3 after stimulus onset). Response times are estimated as described for the rate model, based on the L1 loss of the instantaneous population rates after $t = 500$ ms, using an

exponential fit. For response time estimation, we use unfiltered population rates, computed as the average number of spikes per simulation time bin. For plotting, population rates are displayed in larger bins of 1 ms.

We numerically optimized the excitatory synaptic weights in Fig 8C by repeatedly simulating the network with fine variations in the weight parameter, until a rate profile within the range of $[n_{RF}^{\text{target}} - 0.5, n_{RF}^{\text{target}}]$ was found. Note that we did not enforce a strict match of the peak rate x_{\max} in this optimization process, which would require a covariation of the stimulation strength $I_{\text{ext}}^{\text{on}}$.

The “optimal” scaling factors in Fig 10C were found by simulating balanced excitatory-inhibitory networks for each field size with scaling factors varied in the range of $s \in [1.1, 1.9]$ in steps of 0.1 (see horizontal dashed lines in Fig 10C, bottom). For each scaling factor the (analytically predicted) balancing inhibitory synaptic strength (Eq S81 in S1 Appendix) had to be optimized numerically in order to assure a close match of the EI-rate profile to the target.

Numerical integration was performed in Brian2 using a simple Euler-Maruyama scheme with a discrete time step of 0.01 ms. As the RF size increases, the recurrent weights of the cooperative network approach the point where the network becomes unstable. Close to this instability point, an accurate “clock-driven” numerical integration of the network dynamics requires increasingly small time steps. With spike times restricted to the “grid” set by $\{0, dt, 2dt, 3dt, \dots\}$, spikes can be missed or recorded late [112]. This affects not only the slope of the rate response at the onset of a step stimulus, but also the steady state into which the rates settle. This effect can already be observed in a single population with recurrent coupling approaching the instability point, Fig F in S1 Appendix. In our cooperative network with multiple coupled populations, the time step sensitivity increases for larger fields. For the simulations shown here we used a timestep of 0.01 ms. At this resolution we find satisfactory convergence of the dynamics of networks tuned to RF sizes $n_{RF} \leq 11$. Larger RF sizes would require even smaller time steps, see Fig E in S1 Appendix.

Supporting information

S1 Appendix. Details, derivations, proofs and supporting simulations. In Sec A we derive the stationary state and the evolution of eigenmodes for the 1D cooperatively coding networks without SFA and balancing inhibition. Section B derives the evolution of their L1 loss. In Sec C we proof that their architecture indeed minimizes the number of synapses. Section D argues that we can compare the metabolic cost of generating the stationary state in the different implementations by the L1 norms of their synaptic currents. Sec E and Fig A in S1 Appendix study the dynamics of a 1D cooperatively coding network with SFA. In Sec F we define and calculate the effective strength of the balanced interaction. Section G derives the evolution of the L1 loss for cooperatively coding networks with delayed, balancing inhibition, and expressions for the critical balance strength and network response time. Fig B in S1 Appendix characterizes the loss evolution for different balance strengths. In Sec H and Fig C in S1 Appendix we explain the difference between the loss evolution after constant initialization and in the slowest-decaying eigenmode. Section I and Fig D in S1 Appendix demonstrate that dynamics are similar whether inhibition is modeled with an explicit lag or leakily integrates excitation. In Sec J we derive the stationary response of networks that cooperatively encode 2D stimuli. In Sec K we derive an analytical tuning of the excitatory and balanced spiking networks as a function of RF size and peak rate. We comment on the additional numerical optimization of recurrent synaptic weights. In Sec L we illustrate the need for small discretization time steps for the numerical integration of spiking networks with large RFs (Fig E in S1 Appendix), and analyze the sensitivity of RF formation with respect to

the recurrent weight strength (Fig G in S1 Appendix), and finite size fluctuations (Fig H in S1 Appendix). In Sec M we present an alternative, homogeneous network architecture, for which we observe a similar response speedup when balancing inhibition is added (Fig I in S1 Appendix).

(PDF)

Acknowledgments

We thank Felipe Kalle Kossio for fruitful discussions.

Author contributions

Conceptualization: Paul Züge, Raoul-Martin Memmesheimer.

Formal analysis: Paul Züge, Natalie Schieferstein, Raoul-Martin Memmesheimer.

Funding acquisition: Raoul-Martin Memmesheimer.

Investigation: Paul Züge, Natalie Schieferstein, Raoul-Martin Memmesheimer.

Methodology: Paul Züge, Natalie Schieferstein, Raoul-Martin Memmesheimer.

Software: Paul Züge, Natalie Schieferstein.

Supervision: Raoul-Martin Memmesheimer.

Validation: Paul Züge, Natalie Schieferstein, Raoul-Martin Memmesheimer.

Visualization: Paul Züge, Natalie Schieferstein, Raoul-Martin Memmesheimer.

Writing – original draft: Paul Züge, Natalie Schieferstein, Raoul-Martin Memmesheimer.

Writing – review & editing: Paul Züge, Natalie Schieferstein, Raoul-Martin Memmesheimer.

References

1. Dayan P, Abbott LF. Theoretical neuroscience: computational and mathematical modeling of neural systems. Massachusetts Institute of Technology Press; 2001.
2. Trappenberg T. Fundamentals of Computational Neuroscience. OUP Oxford; 2010. <https://books.google.de/books?id=4PDsA1EVCx0C>
3. HUBEL DH, WIESEL TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*. 1962;160(1):106–54. <https://doi.org/10.1113/jphysiol.1962.sp006837> PMID: 14449617
4. Nauhaus I, Nielsen KJ, Callaway EM. Efficient receptive field tiling in primate V1. *Neuron*. 2016;91(4):893–904. <https://doi.org/10.1016/j.neuron.2016.07.015> PMID: 27499086
5. Taube S. Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *J Neurosci*. 1995.
6. DiCarlo JJ, Johnson KO, Hsiao SS. Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey. *J Neurosci*. 1998;18(7):2626–45. <https://doi.org/10.1523/JNEUROSCI.18-07-02626.1998> PMID: 9502821
7. O'Keefe J, Dostrovsky J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res*. 1971;34(1):171–5. [https://doi.org/10.1016/0006-8993\(71\)90358-1](https://doi.org/10.1016/0006-8993(71)90358-1) PMID: 5124915
8. Hafting T, Fyhn M, Molden S, Moser M-B, Moser EI. Microstructure of a spatial map in the entorhinal cortex. *Nature*. 2005;436(7052):801–6. <https://doi.org/10.1038/nature03721> PMID: 15965463
9. Stensola H, Stensola T, Solstad T, Frøland K, Moser M-B, Moser EI. The entorhinal grid map is discretized. *Nature*. 2012;492(7427):72–8. <https://doi.org/10.1038/nature11649> PMID: 23222610

10. Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybernetics*. 1980;36(4):193–202. <https://doi.org/10.1007/bf00344251>
11. Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*. 1996;381(6583):607–9. <https://doi.org/10.1038/381607a0> PMID: 8637596
12. Banino A, Barry C, Uria B, Blundell C, Lillicrap T, Mirowski P, et al. Vector-based navigation using grid-like representations in artificial agents. *Nature*. 2018;557(7705):429–33. <https://doi.org/10.1038/s41586-018-0102-6> PMID: 29743670
13. Cueva CJ, Wei XX. Emergence of grid-like representations by training recurrent neural networks to perform spatial localization. In: International Conference on Learning Representations; 2018. <https://openreview.net/forum?id=B17JTOe0>
14. Ding Z, Fahey PG, Papadopoulos S, Wang EY, Celii B, Papadopoulos C, et al. Functional connectomics reveals general wiring rule in mouse visual cortex. *bioRxiv*. 2024;2023.03.13.531369. <https://doi.org/10.1101/2023.03.13.531369> PMID: 36993398
15. Ko H, Hofer SB, Pichler B, Buchanan KA, Sjöström PJ, Mrsic-Flogel TD. Functional specificity of local synaptic connections in neocortical networks. *Nature*. 2011;473(7345):87–91. <https://doi.org/10.1038/nature09880> PMID: 21478872
16. 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. <https://doi.org/10.1038/nature14182> PMID: 25652823
17. Lien AD, Scanziani M. Tuned thalamic excitation is amplified by visual cortical circuits. *Nat Neurosci*. 2013;16(9):1315–23. <https://doi.org/10.1038/nn.3488> PMID: 23933748
18. Chettih SN, Harvey CD. Single-neuron perturbations reveal feature-specific competition in V1. *Nature*. 2019;567(7748):334–40. <https://doi.org/10.1038/s41586-019-0997-6> PMID: 30842660
19. Oldenburg IA, Hendricks WD, Handy G, Shamardani K, Bounds HA, Doiron B, et al. The logic of recurrent circuits in the primary visual cortex. *Nat Neurosci*. 2024;27(1):137–47. <https://doi.org/10.1038/s41593-023-01510-5> PMID: 38172437
20. Valente M, Pica G, Bondanelli G, Moroni M, Runyan CA, Morcos AS, et al. Correlations enhance the behavioral readout of neural population activity in association cortex. *Nat Neurosci*. 2021;24(7):975–86. <https://doi.org/10.1038/s41593-021-00845-1> PMID: 33986549
21. Murphy BK, Miller KD. Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron*. 2009;61(4):635–48. <https://doi.org/10.1016/j.neuron.2009.02.005> PMID: 19249282
22. Lim S, Goldman MS. Balanced cortical microcircuitry for maintaining information in working memory. *Nat Neurosci*. 2013;16(9):1306–14. <https://doi.org/10.1038/nn.3492> PMID: 23955560
23. Ben-Yishai R, Bar-Or RL, Sompolinsky H. Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A*. 1995;92(9):3844–8. <https://doi.org/10.1073/pnas.92.9.3844> PMID: 7731993
24. Miller P, Brody CD, Romo R, Wang X-J. A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cereb Cortex*. 2003;13(11):1208–18. <https://doi.org/10.1093/cercor/bhg101> PMID: 14576212
25. Zhang W-H, Wu S, Josić K, Doiron B. Sampling-based Bayesian inference in recurrent circuits of stochastic spiking neurons. *Nat Commun*. 2023;14(1):7074. <https://doi.org/10.1038/s41467-023-41743-3> PMID: 37925497
26. Attwell D, Laughlin SB. An energy budget for signaling in the grey matter of the brain. *J Cereb Blood Flow Metab*. 2001;21(10):1133–45. <https://doi.org/10.1097/00004647-200110000-00001> PMID: 11598490
27. Zalc B. The acquisition of myelin: a success story. *Novartis Found Symp*. 2006;276:15–21; discussion 21–5, 54–7, 275–81. <https://doi.org/10.1002/9780470032244.ch3> PMID: 16805421
28. Howarth C, Gleeson P, Attwell D. Updated energy budgets for neural computation in the neocortex and cerebellum. *J Cereb Blood Flow Metab*. 2012;32(7):1222–32. <https://doi.org/10.1038/jcbfm.2012.35> PMID: 22434069
29. Garcia KE, Kroenke CD, Bayly PV. Mechanics of cortical folding: stress, growth and stability. *Philos Trans R Soc Lond B Biol Sci*. 2018;373(1759):20170321. <https://doi.org/10.1098/rstb.2017.0321> PMID: 30249772
30. Boerlin M, Machens CK, Denève S. Predictive coding of dynamical variables in balanced spiking networks. *PLoS Comput Biol*. 2013;9(11):e1003258. <https://doi.org/10.1371/journal.pcbi.1003258> PMID: 24244113
31. Thalmeier D, Uhlmann M, Kappen HJ, Memmesheimer R-M. Learning universal computations with spikes. *PLoS Comput Biol*. 2016;12(6):e1004895. <https://doi.org/10.1371/journal.pcbi.1004895> PMID: 27309381

32. Denève S, Machens CK. Efficient codes and balanced networks. *Nat Neurosci*. 2016;19(3):375–82. <https://doi.org/10.1038/nn.4243> PMID: 26906504
33. Abbott LF, DePasquale B, Memmesheimer R-M. Building functional networks of spiking model neurons. *Nat Neurosci*. 2016;19(3):350–5. <https://doi.org/10.1038/nn.4241> PMID: 26906501
34. Kadmon J, Timcheck J, Ganguli S. Predictive coding in balanced neural networks with noise, chaos and delays. In: NeurIPS Proceedings. 2020. p. 12.
35. Pulido C, Ryan TA. Synaptic vesicle pools are a major hidden resting metabolic burden of nerve terminals. *Sci Adv*. 2021.
36. Jones JP, Palmer LA. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J Neurophysiol*. 1987;58(6):1233–58. <https://doi.org/10.1152/jn.1987.58.6.1233> PMID: 3437332
37. Dumoulin SO, Wandell BA. Population receptive field estimates in human visual cortex. *NeuroImage*. 2008;39(2):647–60. <https://doi.org/10.1016/j.neuroimage.2007.09.034>
38. Wandell BA, Winawer J. Computational neuroimaging and population receptive fields. *Trends Cogn Sci*. 2015;19(6):349–57. <https://doi.org/10.1016/j.tics.2015.03.009> PMID: 25850730
39. Serre T. Hierarchical models of the visual system. *Encyclopedia of computational neuroscience*. New York: Springer; 2014. p. 1–12. https://doi.org/10.1007/978-1-4614-7320-6_345-1
40. Neher T, Azizi AH, Cheng S. From grid cells to place cells with realistic field sizes. *PLoS One*. 2017;12(7):e0181618. <https://doi.org/10.1371/journal.pone.0181618> PMID: 28750005
41. Dayan P, Abbott L. Theoretical neuroscience: computational and mathematical modeling of neural systems. Cambridge: MIT Press. 2001.
42. Gerstner W, Kistler WM, Naud R, Paninski L. Neuronal dynamics - from single neurons to networks and models of cognition. Cambridge: Cambridge University Press. 2014.
43. Jetschke G. Mathematik der Selbstorganisation. Frankfurt am Main: Harri Deutsch; 2009.
44. Hirsch MW, Smale S. Differential equations, dynamical systems, and linear algebra. San Diego [U.A.]: Academic Press; 1974.
45. Pernice V, Staude B, Cardanobile S, Rotter S. How structure determines correlations in neuronal networks. *PLoS Comput Biol*. 2011;7(5):e1002059. <https://doi.org/10.1371/journal.pcbi.1002059> PMID: 21625580
46. Kossio FYK, Goedeke S, van den Akker B, Ibarz B, Memmesheimer R-M. Growing critical: self-organized criticality in a developing neural system. *Phys Rev Lett*. 2018;121(5):058301. <https://doi.org/10.1103/PhysRevLett.121.058301> PMID: 30118252
47. White OL, Lee DD, Sompolinsky H. Short-term memory in orthogonal neural networks. *Phys Rev Lett*. 2004;92(14):148102. <https://doi.org/10.1103/PhysRevLett.92.148102> PMID: 15089576
48. Ganguli S, Huh D, Sompolinsky H. Memory traces in dynamical systems. *Proc Natl Acad Sci U S A*. 2008;105(48):18970–5. <https://doi.org/10.1073/pnas.0804451105> PMID: 19020074
49. Lillicrap TP, Santoro A, Marris L, Akerman CJ, Hinton G. Backpropagation and the brain. *Nat Rev Neurosci*. 2020;21(6):335–46. <https://doi.org/10.1038/s41583-020-0277-3> PMID: 32303713
50. Izhikevich EM. Dynamical systems in neuroscience: the geometry of excitability and bursting. Cambridge: MIT Press. 2007.
51. Gutkin B, Zeldnerust F. Spike frequency adaptation. *Scholarpedia*. 2014;9(2):30643. <https://doi.org/10.4249/scholarpedia.30643>
52. Treves A. Mean-field analysis of neuronal spike dynamics. *Netw: Comput Neural Syst*. 1993;4(3):259–84. https://doi.org/10.1088/0954-898x_4_3_002
53. Hansel D, Sompolinsky H. Modeling feature selectivity in local cortical circuits. In: Koch C, Segev I, editors. *Methods in neuronal modeling*. 2nd ed. Cambridge, MA: MIT Press; 1998. p. 499–566.
54. Fuhrmann G, Markram H, Tsodyks M. Spike frequency adaptation and neocortical rhythms. *J Neurophysiol*. 2002;88(2):761–70. <https://doi.org/10.1152/jn.2002.88.2.761> PMID: 12163528
55. Dehghani N, Peyrache A, Telenczuk B, Le Van Quyen M, Halgren E, Cash SS, et al. Dynamic balance of excitation and inhibition in human and monkey neocortex. *Sci Rep*. 2016;6:23176. <https://doi.org/10.1038/srep23176> PMID: 26980663
56. Gerstein GL, Mandelbrot B. Random walk models for the spike activity of a single neuron. *Biophys J*. 1964;4(1 Pt 1):41–68. [https://doi.org/10.1016/s0006-3495\(64\)86768-0](https://doi.org/10.1016/s0006-3495(64)86768-0) PMID: 14104072
57. Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci*. 1998;18(10):3870–96. <https://doi.org/10.1523/JNEUROSCI.18-10-03870.1998> PMID: 9570816
58. van Vreeswijk C, Sompolinsky H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*. 1996;274(5293):1724–6. <https://doi.org/10.1126/science.274.5293.1724> PMID: 8939866

59. Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A. The asynchronous state in cortical circuits. *Science*. 2010;327:587–90.
60. Herstel LJ, Wierenga CJ. Network control through coordinated inhibition. *Curr Opin Neurobiol*. 2021;67:34–41. <https://doi.org/10.1016/j.conb.2020.08.001> PMID: 32853970
61. Okun M, Lampl I. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat Neurosci*. 2008;11(5):535–7. <https://doi.org/10.1038/nn.2105> PMID: 18376400
62. Znamenskiy P, Kim M-H, Muir DR, Iacaruso MF, Hofer SB, Mrsic-Flogel TD. Functional specificity of recurrent inhibition in visual cortex. *Neuron*. 2024;112(6):991–1000.e8. <https://doi.org/10.1016/j.neuron.2023.12.013> PMID: 38244539
63. Runyan CA, Schummers J, Van Wart A, Kuhlman SJ, Wilson NR, Huang ZJ, et al. Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron*. 2010;67(5):847–57. <https://doi.org/10.1016/j.neuron.2010.08.006> PMID: 20826315
64. Najafi F, Elsayed GF, Cao R, Pnevmatikakis E, Latham PE, Cunningham JP, et al. Excitatory and inhibitory subnetworks are equally selective during decision-making and emerge simultaneously during learning. *Neuron*. 2020;105(1):165–179.e8. <https://doi.org/10.1016/j.neuron.2019.09.045> PMID: 31753580
65. Rigotti M, Barak O, Warden MR, Wang X-J, Daw ND, Miller EK, et al. The importance of mixed selectivity in complex cognitive tasks. *Nature*. 2013;497(7451):585–90. <https://doi.org/10.1038/nature12160> PMID: 23685452
66. Fusi S, Miller EK, Rigotti M. Why neurons mix: high dimensionality for higher cognition. *Curr Opin Neurobiol*. 2016;37:66–74. <https://doi.org/10.1016/j.conb.2016.01.010> PMID: 26851755
67. McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser M-B. Path integration and the neural basis of the “cognitive map”. *Nat Rev Neurosci*. 2006;7(8):663–78. <https://doi.org/10.1038/nrn1932> PMID: 16858394
68. Yang Z-H, Chu Y-M. On approximating the modified Bessel function of the second kind. *J Inequal Appl*. 2017;2017(1):41. <https://doi.org/10.1186/s13660-017-1317-z> PMID: 28250694
69. Knight BW. Dynamics of encoding in a population of neurons. *J Gen Physiol*. 1972;59(6):734–66. <https://doi.org/10.1085/jgp.59.6.734> PMID: 5025748
70. Gerstner W. Population dynamics of spiking neurons: fast transients, asynchronous states, and locking. *Neural Comput*. 2000;12(1):43–89. <https://doi.org/10.1162/089976600300015899> PMID: 10636933
71. Brunel N, Chance FS, Fourcaud N, Abbott LF. Effects of synaptic noise and filtering on the frequency response of spiking neurons. *Phys Rev Lett*. 2001;86(10):2186–9. <https://doi.org/10.1103/PhysRevLett.86.2186> PMID: 11289886
72. Sanzeni A, Histed MH, Brunel N. Response nonlinearities in networks of spiking neurons. *PLoS Comput Biol*. 2020;16(9):e1008165. <https://doi.org/10.1371/journal.pcbi.1008165> PMID: 32941457
73. Marsálek P, Koch C, Maunsell J. On the relationship between synaptic input and spike output jitter in individual neurons. *Proc Natl Acad Sci U S A*. 1997;94(2):735–40. <https://doi.org/10.1073/pnas.94.2.735> PMID: 9012854
74. Zugaro MB, Arleo A, Berthoz A, Wiener SI. Rapid spatial reorientation and head direction cells. *J Neurosci*. 2003;23(8):3478–82. <https://doi.org/10.1523/JNEUROSCI.23-08-03478.2003> PMID: 12716956
75. Ito J, Joana C, Yamane Y, Fujita I, Tamura H, Maldonado PE, et al. Latency shortening with enhanced sparseness and responsiveness in V1 during active visual sensing. *Sci Rep*. 2022;12(1):6021. <https://doi.org/10.1038/s41598-022-09405-4> PMID: 35410997
76. Murray JD, Bernacchia A, Freedman DJ, Romo R, Wallis JD, Cai X, et al. A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci*. 2014;17(12):1661–3. <https://doi.org/10.1038/nn.3862> PMID: 25383900
77. Chaudhuri R, Knoblauch K, Gariel M-A, Kennedy H, Wang X-J. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron*. 2015;88(2):419–31. <https://doi.org/10.1016/j.neuron.2015.09.008> PMID: 26439530
78. Li S, Wang X-J. Hierarchical timescales in the neocortex: Mathematical mechanism and biological insights. *Proc Natl Acad Sci U S A*. 2022;119(6):e2110274119. <https://doi.org/10.1073/pnas.2110274119> PMID: 35110401
79. Siegle JH, Jia X, Durand S, Gale S, Bennett C, Graddis N, et al. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature*. 2021;592(7852):86–92. <https://doi.org/10.1038/s41586-020-03171-x>

80. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol.* 1962;160(1):106–54. <https://doi.org/10.1113/jphysiol.1962.sp006837> PMID: 14449617
81. Reid RC, Alonso JM. Specificity of monosynaptic connections from thalamus to visual cortex. *Nature.* 1995;378(6554):281–4. <https://doi.org/10.1038/378281a0> PMID: 7477347
82. Smith AT, Singh KD, Williams AL, Greenlee MW. Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cereb Cortex.* 2001;11(12):1182–90. <https://doi.org/10.1093/cercor/11.12.1182> PMID: 11709489
83. Freeman J, Simoncelli EP. Metamers of the ventral stream. *Nat Neurosci.* 2011;14(9):1195–201. <https://doi.org/10.1038/nn.2889> PMID: 21841776
84. Neunuebel JP, Knierim JJ. Spatial firing correlates of physiologically distinct cell types of the rat dentate gyrus. *J Neurosci.* 2012;32(11):3848–58. <https://doi.org/10.1523/JNEUROSCI.6038-11.2012> PMID: 22423105
85. Mizuseki K, Royer S, Diba K, Buzsáki G. Activity dynamics and behavioral correlates of CA3 and CA1 hippocampal pyramidal neurons. *Hippocampus.* 2012;22(8):1659–80. <https://doi.org/10.1002/hipo.22002> PMID: 22367959
86. von der Malsburg C. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik.* 1973;14(2):85–100. <https://doi.org/10.1007/BF00288907> PMID: 4786750
87. Redish AD, Elga AN, Touretzky DS. A coupled attractor model of the rodent head direction system. *Network.* 1996;7(4):671–85.
88. Renart A, Song P, Wang X-J. Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. *Neuron.* 2003;38(3):473–85. [https://doi.org/10.1016/s0896-6273\(03\)00255-1](https://doi.org/10.1016/s0896-6273(03)00255-1) PMID: 12741993
89. Kushnir L, Fusi S. Neural classifiers with limited connectivity and recurrent readouts. *J Neurosci.* 2018;38(46):9900–24. <https://doi.org/10.1523/JNEUROSCI.3506-17.2018> PMID: 30249794
90. Chenkov N, Sprekeler H, Kempter R. Memory replay in balanced recurrent networks. *PLoS Comput Biol.* 2017;13(1):e1005359. <https://doi.org/10.1371/journal.pcbi.1005359> PMID: 28135266
91. Nayebi A, Sagastuy-Brena J, Bear DM, Kar K, Kubilius J, Ganguli S, et al. Recurrent connections in the primate ventral visual stream mediate a trade-off between task performance and network size during core object recognition. *Neural Comput.* 2022;34(8):1652–75. https://doi.org/10.1162/neco_a_01506 PMID: 35798321
92. Braintenberg V, Schüz A. *Cortex: statistics and geometry of neuronal connectivity.* Berlin: Springer; 1998.
93. Holmgren C, Harkany T, Svennenfors B, Zilberman Y. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *J Physiol.* 2003;551(Pt 1):139–53. <https://doi.org/10.1113/jphysiol.2003.044784> PMID: 12813147
94. Cuturidis V, Graham BP, Cobb S, Vida I. *Hippocampal microcircuits.* 2nd ed. Cham, Switzerland: Springer; 2018.
95. Seeman SC, Campagnola L, Davoudian PA, Hoggarth A, Hage TA, Bosma-Moody A, et al. Sparse recurrent excitatory connectivity in the microcircuit of the adult mouse and human cortex. *Elife.* 2018;7:e37349. <https://doi.org/10.7554/elife.37349> PMID: 30256194
96. da Costa NM, Martin KAC. The proportion of synapses formed by the axons of the lateral geniculate nucleus in layer 4 of area 17 of the cat. *J Comp Neurol.* 2009;516(4):264–76. <https://doi.org/10.1002/cne.22133> PMID: 19634180
97. Markov NT, Misery P, Falchier A, Lamy C, Vezoli J, Quilodran R, et al. Weight consistency specifies regularities of macaque cortical networks. *Cereb Cortex.* 2011;21(6):1254–72. <https://doi.org/10.1093/cercor/bhq201> PMID: 21045004
98. Chariker L, Shapley R, Young L-S. Orientation selectivity from very sparse LGN inputs in a comprehensive model of macaque V1 cortex. *J Neurosci.* 2016;36(49):12368–84. <https://doi.org/10.1523/JNEUROSCI.2603-16.2016> PMID: 27927956
99. Angelucci A, Sainsbury K. Contribution of feedforward thalamic afferents and corticogeniculate feedback to the spatial summation area of macaque V1 and LGN. *J Comp Neurol.* 2006;498(3):330–51. <https://doi.org/10.1002/cne.21060> PMID: 16871526
100. Amaral DG, Ishizuka N, Claiborne B. Neurons, numbers and the hippocampal network. *Prog Brain Res.* 1990;83:1–11. [https://doi.org/10.1016/s0079-6123\(08\)61237-6](https://doi.org/10.1016/s0079-6123(08)61237-6) PMID: 2203093
101. Sammons RP, Vezir M, Moreno-Velasquez L, Cano G, Orlando M, Sievers M, et al. Structure and function of the hippocampal CA3 module. *Proc Natl Acad Sci U S A.* 2024;121(6):e2312281120. <https://doi.org/10.1073/pnas.2312281120> PMID: 38289953

102. Steward O. Topographic organization of the projections from the entorhinal area to the hippocampal formation of the rat. *J Comp Neurol.* 1976;167(3):285–314. <https://doi.org/10.1002/cne.901670303> PMID: 1270625
103. Amaral DG, Witter MP. The three-dimensional organization of the hippocampal formation: a review of anatomical data. *Neuroscience.* 1989;31(3):571–91. [https://doi.org/10.1016/0306-4522\(89\)90424-7](https://doi.org/10.1016/0306-4522(89)90424-7) PMID: 2687721
104. Witter MP. Intrinsic and extrinsic wiring of CA3: indications for connectional heterogeneity. *Learn Mem.* 2007;14(11):705–13. <https://doi.org/10.1101/lm.725207> PMID: 18007015
105. Ferster D, Chung S, Wheat H. Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature.* 1996;380(6571):249–52. <https://doi.org/10.1038/380249a0> PMID: 8637573
106. Biggiogera J, Sanzeni A. Feature tuning and network dynamics in mouse visual cortex: insights from connectomics; 2023. <http://doi.org/10.12751/nncn.bc2023.217>
107. DePasquale B, Sussillo D, Abbott LF, Churchland MM. The centrality of population-level factors to network computation is demonstrated by a versatile approach for training spiking networks. *Neuron.* 2023;111(5):631–649.e10. <https://doi.org/10.1016/j.neuron.2022.12.007> PMID: 36630961
108. Maheswaranathan N, Williams AH, Golub MD, Ganguli S, Sussillo D. Universality and individuality in neural dynamics across large populations of recurrent networks. *Adv Neural Inf Process Syst.* 2019;2019:15629–41. PMID: 32782422
109. Pezon L, Schmutz V, Gerstner W. Linking neural manifolds to circuit structure in recurrent networks; 2024.
110. Ringach DL. Sparse thalamocortical convergence. *Curr Biol.* 2021;31(10):2199–2202.e2. <https://doi.org/10.1016/j.cub.2021.02.032> PMID: 33705713
111. Goodman DFM, Brette R. The brian simulator. *Front Neurosci.* 2009;3(2):192–7. <https://doi.org/10.3389/neuro.01.026.2009> PMID: 20011141
112. Hansel D, Mato G, Meunier C, Neltner L. On numerical simulations of integrate-and-fire neural networks. *Neural Comput.* 1998;10(2):467–83. <https://doi.org/10.1162/089976698300017845> PMID: 9472491