





The functional role of oscillatory dynamics in neocortical circuits: A computational perspective

Felix Effenberger^{a,1} , Pedro Carvalho^{a,2}, Igor Dubinin^{a,b,2}, and Wolf Singer^{a,b,c,1} 

Affiliations are included on p. 11.

Contributed by Wolf Singer; received June 26, 2024; accepted December 23, 2024; reviewed by Michael Breakspear and Claudio Mirasso

The dynamics of neuronal systems are characterized by hallmark features such as oscillations and synchrony. However, it has remained unclear whether these characteristics are epiphenomena or are exploited for computation. Due to the challenge of selectively interfering with oscillatory network dynamics in neuronal systems, we simulated recurrent networks of damped harmonic oscillators in which oscillatory activity is enforced in each node, a choice well supported by experimental findings. When trained on standard pattern recognition tasks, these harmonic oscillator recurrent networks (HORNs) outperformed nonoscillatory architectures with respect to learning speed, noise tolerance, and parameter efficiency. HORNs also reproduced a many characteristic features of neuronal systems, such as the cerebral cortex and the hippocampus. In trained HORNs, stimulus-induced interference patterns holistically represent the result of comparing sensory evidence with priors stored in recurrent connection weights, and learning-induced weight changes are compatible with Hebbian principles. Implementing additional features characteristic of natural networks, such as heterogeneous oscillation frequencies, inhomogeneous conduction delays, and network modularity, further enhanced HORN performance without requiring additional parameters. Taken together, our model allows us to give plausible a posteriori explanations for features of natural networks whose computational role has remained elusive. We conclude that neuronal systems are likely to exploit the unique dynamics of recurrent oscillator networks whose computational superiority critically depends on the oscillatory patterning of their nodal dynamics. Implementing the proposed computational principles in analog hardware is expected to enable the design of highly energy-efficient and self-adapting devices that could ideally complement existing digital technologies.

neural oscillations | cerebral cortex | recurrent networks | pattern recognition | traveling waves

Neuronal networks of the cerebral cortex and the hippocampus, and likely also those in homologous structures of nonmammalian species are characterized by several canonical anatomical and physiological features. Among those are recurrent connections between network nodes within and between processing layers (1, 2), the propensity of nodes and subnetworks to oscillate in different preferred frequency ranges (3, 4), heterogeneous but tuned conduction delays (5, 6), and activity-dependent adjustment of the gain of connections (7, 8) by Hebbian synaptic plasticity (9). Certain canonical circuit motifs, such as recurrent inhibition and excitation, play a well-defined role in signal processing. Examples are contrast enhancement, gain control, dynamic range expansion, and competitive interactions. However, such microcircuits also tend to oscillate, and when several such microcircuits interact, additional complex dynamical phenomena emerge. In simultaneous recordings from multiple network nodes, these dynamics manifest themselves as frequency-varying oscillations (10, 11), transient synchronization or desynchronization of discharges (12), resonance (13), entrainment (14), phase shifts (15), and traveling waves (16–18). Yet, the functional significance of many of these dynamical phenomena has remained unclear. Although the role of oscillating neurons and microcircuits in generating motor patterns is well established (19), it remains a matter of discussion whether oscillations support computations in the context of cognitive processes in the cerebral cortex and, if so, how (20–25). Moreover, determining whether these dynamical phenomena serve a functional role for computations is notoriously difficult in physiological experiments because strategies for identifying causal relationships based on loss or gain of function interventions fall short in such complex and highly integrated systems (26). Therefore, virtually all evidence for the functional role of oscillatory dynamics has remained correlative in nature.

Significance

Neocortical circuits are characterized by complex oscillatory dynamics. Whether these oscillations serve computations or are an epiphenomenon is still debated. To answer this question, we designed a computational model of a recurrent network that allows control of oscillatory dynamics (harmonic oscillator recurrent network, HORN). When operating in an oscillatory regime, HORNs outperform nonoscillatory recurrent networks in terms of learning speed, noise tolerance, and parameter efficiency. Moreover, they closely replicate the dynamics of neuronal systems, suggesting that biological neural networks are likely to also exploit the unique properties offered by oscillatory dynamics for computing. The interference patterns provided by wave-based responses allow for a holistic representation and highly parallel encoding of both spatial and temporal relations among stimulus features.

Competing interest statement: F.E. and W.S. have filed a patent application for the hardware implementation of the described network models at the European Patent Office, Application No. EP22183567.1.

Copyright © 2025 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: felix.effenberger@esi-frankfurt.de or wolf.singer@brain.mpg.de.

²P.C. and I.D. contributed equally to this work.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2412830122/-/DCSupplemental>.

Published January 23, 2025.

To overcome this epistemic hurdle and isolate computational principles, we performed *in silico* simulations of recurrent networks (RNNs) and trained them on standard pattern recognition tasks (sequential and permuted MNIST digit recognition, spoken digit recognition, and a Mackey-Glass time-series prediction task), which allowed us to use task performance as a measure of functional relevance. Inspired by physiological evidence (27), we configured network nodes as damped harmonic oscillators (DHOs), a quintessential model of oscillatory dynamics. In the resulting harmonic oscillator recurrent networks (HORNs), oscillatory nodal dynamics act as an inductive bias and permit the comparison of networks in which nodal dynamics is nonoscillatory with various regimes of oscillatory dynamics by solely adjusting the control parameters of the DHOs, leaving other network parameters unchanged.

The enforcement of oscillatory activity in every node is quintessential for our approach. Networks of nonoscillating nodes such as leaky integrators tend to express oscillatory dynamics on the network level. However, the exploitation of the extended dynamical primitives enabled by oscillations such as coding in phase space, (de)synchronization, and resonance, among others, by gradient-based learning schemes is limited, because learning-induced changes in the weight configuration easily change or disrupt the emergent oscillations. In contrast, configuring nodes as oscillators makes the extended dynamical primitives stably available to learning algorithms such as back-propagation through time (BPTT) and permits their exploitation for computational purposes. Controlling the oscillatory properties of the nodes allowed us to quantify the effect of oscillations on network performance.

We found that HORNs outperformed, sometimes by a large margin, their nonoscillating counterparts composed of leaky integrators, with respect to parameter efficiency, task performance, learning speed, and noise tolerance. This superiority was also manifest in comparison with other nonoscillating RNNs that rely on gated architectures [LSTM, GRU (28)], and is particularly pronounced for small system sizes. The finding of increased task performance is supported by previous works that also observed an increase in the performance of recurrent oscillator networks (29–32).

In-depth analyses of the dynamics of HORNs uncovered a powerful computational principle that results from the unique properties of coupled oscillator networks and that uses the superposition and interference patterns of waves for stimulus representation and processing. Without requiring fine-tuning of the parameters for the different experiments, the dynamics of HORNs shared numerous features with those observed in the cerebral cortex, suggesting that natural networks also exploit the uncovered computational principle. To further examine this possibility, we implemented other characteristic features of natural networks in HORNs and found that the inclusion of these biologically inspired features typically resulted in further improved task performance without increasing the number of trainable variables. Because the simulations allowed us to study the functional consequences of implementing known properties of natural neuronal networks—in particular of the cerebral cortex—our synthetic approach in addition provides plausible *a posteriori* explanations for a number of phenomena whose function has so far remained elusive or has given rise to controversial discussions.

In contrast to earlier research in neuroscience and machine learning, which focuses mainly on network dynamics and biological realism (33–36) on the one hand, or task performance (30, 32, 37, 38) on the other, this study combines

elements from both fields by simultaneously addressing the mechanistic modeling of biological network characteristics and their functional validation through benchmark tests.

Results

Oscillating Network Nodes. Choosing DHOs as the model of nodal activity was motivated by several reasons. First, DHOs represent the quintessential implementation of an oscillatory process that allows easy control of its relaxation dynamics through few, interpretable control parameters. Second, damped harmonic oscillations generically result from excitatory–inhibitory interactions or negative feedback subject to a damping force in neuronal microcircuits (39, 40). Third, recent experimental evidence showed that population activity in the visual cortex of macaques is well modeled by driven, damped harmonic oscillators (27). By manipulating the control parameters of the DHOs, we could convert the nodes from oscillators to integrators without interfering with other network properties, which allowed us to compare oscillatory with nonoscillatory regimes. Importantly, a DHO node in our networks should not necessarily be considered as a single biological neuron, but rather as a unit representing an abstract aggregate quantity of the activity of a microcircuit composed of recurrently coupled populations of excitatory and inhibitory neurons such as a (P)ING circuit or alternatively, also a single neuron endowed with pacemaker currents (41–43). In this sense, HORNs can be understood as a mesoscale model of a cortical network (Fig. 1A).

In our model, each DHO node has one state variable x , the oscillator's time-varying amplitude, and three control parameters that jointly determine its relaxation dynamics: the natural frequency ω , the damping coefficient γ , and an excitability coefficient α (*SI Appendix*, Fig. S6 and *Materials and Methods*). These parameters can be thought of as the innate or long-term adaptation of the node to the characteristics of the signals to be processed (*Receptive Fields*). Despite their simplicity, DHO nodes capture the essential dynamical characteristics of various neural mass models in the oscillatory regime (39, 44, 45) (*SI Appendix*). Thus, the oscillatory properties of the nodes implemented in our simulations are both physiologically plausible and experimentally confirmed (27).

Already, a single DHO node possesses computational capabilities that are not accessible to nonoscillating nodes, such as leaky integrators. DHOs convert any input signal into an oscillation, even if the signal is nonoscillatory. Using an oscillatory representation enables a node to encode information not only in amplitude but also in phase, allowing the encoding of stimulus sequence order in phase shifts, for example (*SI Appendix*, Fig. S6). If inputs are oscillatory, which is the case for recurrent inputs from other network nodes but also for some sensory inputs, additional computational capacities can be leveraged. The parameter ω defines a frequency band in which the DHO node selects and amplifies temporally modulated signals through resonance. When integrated in a recurrent network, DHOs convert inputs into wave patterns. These patterns manifest both as standing waves at each network node (a representation not available to nonoscillating architectures) and also propagate throughout the whole network, allowing for interference with waves generated both by internal and reverberating network dynamics, as well as spatially and temporally segregated stimuli. Crucially, when input signals exhibit distinct temporal patterns, the oscillatory behavior of DHO nodes enables them to identify features via resonance, rather than merely through integration over converging input connections, a strategy commonly used in nonoscillating RNNs

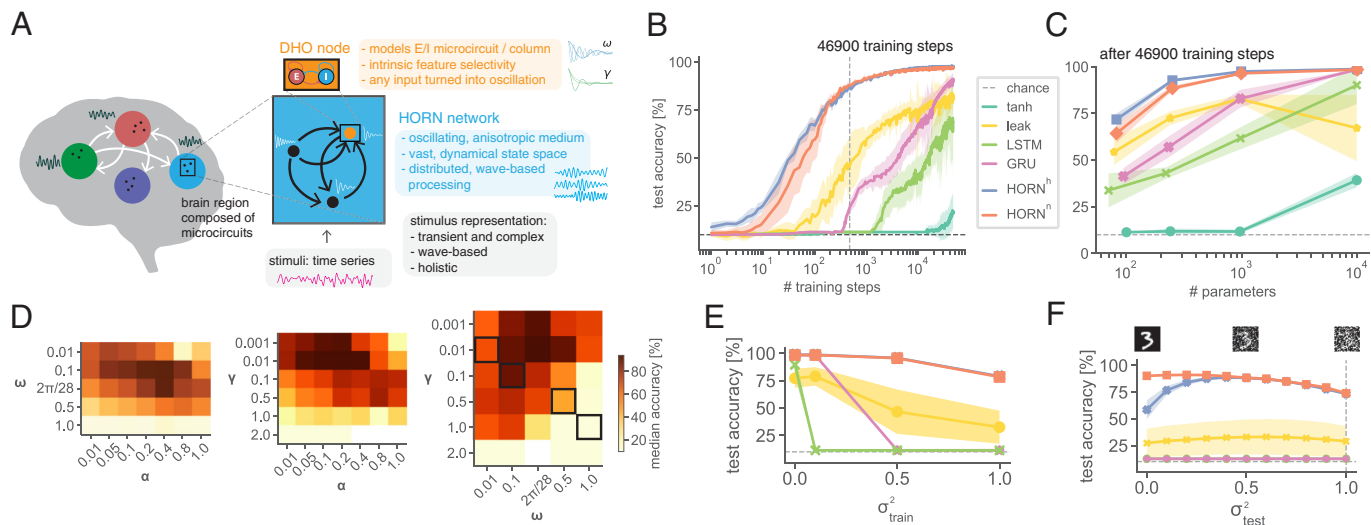


Fig. 1. The architecture of HORN models and their performance in the sMNIST pattern recognition task. (A) A HORN composed of DHO nodes as a mesoscale model of brain activity. Each DHO node models the compound activity of a recurrently connected E-I population as damped oscillations controlled by the parameters ω, γ, α . (B) Test accuracy on sMNIST for different RNN architectures with approximately 2,500 trainable parameters (tanh, leak, HORN 43 nodes, GRU 25 nodes, LSTM 22 nodes) as a function of training steps over 100 training epochs. Lines show mean accuracy over 10 network instances with random weight initialization, shaded areas SD. Legend at the *Top Left*. Note the logarithmic scale in the number of training steps. (C) Maximum test accuracy after 100 training epochs on sMNIST for different RNN architectures as a function of system size (number of nodes for approximately 10^4 training parameters: tanh, leak, HORN 94 nodes, GRU 54 nodes, LSTM 47 nodes). Networks and legend as in B. Note the high task performance of HORNs at very low node and parameter counts. (D) Median test accuracies of HORN^h networks over parameter configurations ω, γ, α from a grid search. Each matrix entry shows the median test accuracy (color scale on the *Right*) for the marginalized remaining parameter. Parameter value pairs (ω, γ) resulting in critical damping ($\omega = \gamma$) marked with a black frame. Note the performance drop for damping values above the critical point where nodes become nonoscillating. (E) Maximum test accuracy of different RNN architectures (10^4 trainable parameters) after training on noisy sMNIST for 100 epochs as a function of training noise level σ_{train}^2 . Lines show mean accuracy over 10 random networks, shaded areas show SD. Colors as in B and C. (F) Test accuracy of different network architectures trained on noisy sMNIST with $\sigma_{\text{train}}^2 = 1$ (dashed vertical line) as a function of test stimulus noise level σ_{test}^2 , calculated over $n = 1,000$ test stimuli. Colors as in E. Note the pronounced noise resilience in HORNs.

and all feed-forward architectures (46). In HORNs, single nodes or assemblies of nodes can further encode information in distinct frequency bands and gate the flow of information as a function of relations among frequencies (Fig. 2E). These features of DHOs endow HORNs with the properties of an anisotropically coupled, nonlinear, analog medium for information processing that can exploit for its computations wave-based representations of stimuli and their interference patterns. This results in a rich and high-dimensional coding space.

The interaction of the input nonlinearity with biologically inspired feedback connections added to the DHO nodes gives them a dynamical repertoire beyond that of a classical damped harmonic oscillator (*SI Appendix*). For example, feedback allows DHO nodes to express self-sustained oscillations and to resonate at fractional harmonics of their natural frequency (*SI Appendix*, Figs. S7–S9).

While modeling neural population activity on a mesoscale with DHO nodes allows us to uncover the generic principles described in this work, it also limits the model's ability to capture other dynamical phenomena observed in neuronal systems, such as dynamical variations in oscillation frequency, spikes and their bursting behavior, and, more generally, any dynamics that are nonoscillatory or cannot be captured at the population level (44).

Configuration and Training of the Networks. We first consider homogeneous HORN^h networks with identical parameter values ω, γ, α for all nodes. As cortical circuits operate in a balanced state (47), emphasizing the significance of fluctuations, we couple DHO nodes on their velocity term \dot{x} (*Materials and Methods*). Stimuli were presented to the networks in the form of time series. Due to its standard nature and widespread use,

we chose the MNIST handwritten digit classification task as our default benchmark, turning MNIST digits into sMNIST time series (*SI Appendix*, Fig. S1A). This transformation converts the geometrical properties of the stimuli into spectral patterns that HORNs can exploit to perform stimulus classification. The approach taken serves as a proxy for other spectral regularities typical of natural signals. To increase task difficulty, we also trained networks on permuted sequential MNIST (psMNIST) stimuli in which a random but fixed permutation is applied to the time series. This operation destroys local luminance correlations in MNIST digits and results in time series stimuli with a flat spectrum, making frequency-based information processing harder (*SI Appendix*, Fig. S10).

We also study the case of geometrically organized inputs (*Geometric Input Drives Self-Organization*) and have trained HORN networks on more challenging datasets, such as a spoken-digit classification task and a Mackey-Glass time series prediction task (*SI Appendix*, Figs. S14 and S15). The results confirm the computational superiority of the oscillatory dynamics implemented in HORNs over nonoscillating architectures also in those cases, particularly for the low-parameter regime.

For the sMNIST classification task, a grid search was performed on the triplet of parameter values ω, γ, α to find a set of DHO control parameter values that resulted in high classification performance (Fig. 1D and *SI Appendix*, Fig. S1B and Table S1). We found that networks with low values of the damping factor γ tended to perform best. Low damping factors put the system in a highly oscillatory state and enhance the memory spans of individual DHOs and the whole network (48). The excitability parameter α had no strong effect on performance, with intermediate values between 0.2 and 0.4 resulting in the best performance.

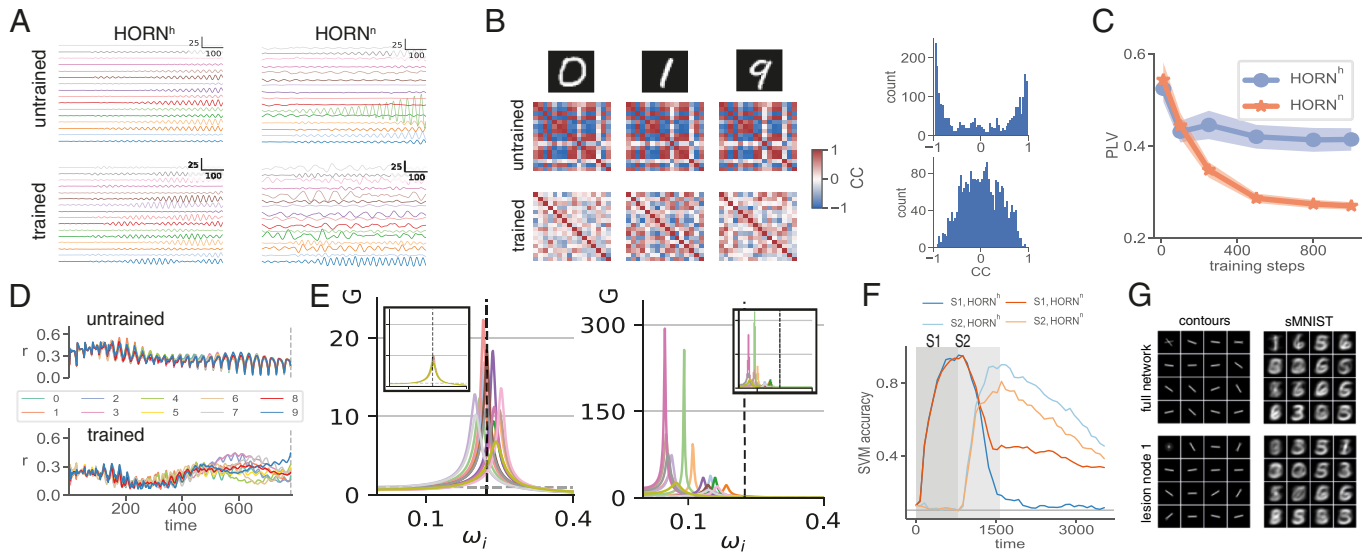


Fig. 2. Dynamics and computational principles in HORNs. (A) Network activity elicited by an sMNIST stimulus (digit 0) of 16-node HORNs trained on sMNIST, before (top row) and after (bottom row) training. Left column HORN_{16}^h , right column HORN_{16}^n . Note the highly irregular temporal dynamics of HORN_{16}^h in the trained state. (B) Pairwise cross-correlation coefficients of node activity for different input stimuli in HORN_{16}^h before (Left, Top row) and after training (Left, Bottom row) on sMNIST for 10 epochs. Distribution of cross-correlation coefficients before (Right, Top) and after (Right, Bottom) training. (C) Pairwise phase locking value (PLV) calculated on $n = 100$ test stimuli as a function of training steps for a HORN_{16}^h (blue) and HORN_{16}^n (orange) network trained on sMNIST. Lines show mean; the shaded region indicates SD. (D) Dynamics of mean Kuramoto order parameter r of HORN_{16}^h activity trained on sMNIST in the untrained network (Top) and after training for 10 epochs (Bottom). Lines show mean values per digit class. The HORN learns to generate dynamic, stimulus-specific patterns of higher-order synchrony. (E) Intrinsic RFs (gain functions G) of nodes as a function of input frequency ω_i of 16-node HORNs before (Insets) and after training on sMNIST. Homogeneous (Left) and heterogeneous (Right) HORN variants. Gain values $G > 1$ indicate amplification. Horizontal dashed lines indicate $G = 1$, vertical dashed lines indicate the frequency $\omega_v = 2\pi/28$ that corresponds to a straight vertical bar in pixel space. Colors indicate node identity. (F) Accuracy of stimulus decoding by linear SVMs trained on the activities of different HORN networks (93 nodes) that were stimulated with a sequence of two temporally segregated stimuli S1 and S2 (shaded areas indicate stimulus timing). Light and dark colored curves indicate SVM decoder performance for S1 and S2 for a homogeneous (HORN^h) and nonhomogeneous (HORN^n) network, respectively. Note the longer persistence and superposition of stimulus-specific information in the HORN^n . (G) Effective RFs of a HORN_{16}^h when mapped with contour stimuli from LSDSa (Top two panels) and sMNIST digits (Bottom two panels). Both cases show the ERFs of the full network and the ERFs after the deletion (lesion) of the first node. Note the reconfiguration of most of the ERFs when a node is inactivated.

The optimal values of the natural frequency parameter ω were found to be around $\omega_v = 2\pi/28$, which corresponds to the fundamental frequency of an sMNIST time series resulting from a straight vertical line in the 28×28 MNIST pixel space. As DHOs resonate with input frequencies around ω_v , this choice of ω enabled the nodes to extract and retain the information in time series corresponding to continuous lines of different slopes in pixel space. The value ω_v also coincides with a peak in the variance of power spectral densities (PSDs) calculated on a representative set of sMNIST samples (SI Appendix, Fig. S10). Thus, by adjusting ω , the network can be adapted to the statistical regularities of the sMNIST stimuli and this setting of priors enhances performance by allowing feature extraction through resonance rather than solely through selective recombination of convergent input connections. This strategy implemented in HORNs is crucially dependent on the presence of nodes or microcircuits with a propensity to oscillate.

Network Performance. Endowing network nodes with oscillatory dynamics improved network performance compared to leaky integrator or gated unit networks with respect to learning speed, parameter efficiency, and noise tolerance (Fig. 1 and SI Appendix, Figs. S11–S15). This superiority was particularly pronounced in the region of low parameter counts (Fig. 1C) and was critically dependent on the configuration of the network nodes to be in a highly oscillatory state ($\omega \gg \gamma$, Fig. 1D). Performance was found to drop for networks that were given low oscillation frequencies, which act more like (leaky) integrators, and for damping values above the critical value $\gamma > \omega$ that abolish nodal oscillations.

This shows the functional benefit of oscillating nodal dynamics in HORNs. When making the task more difficult by decreasing signal-to-noise ratios or increasing the number and similarity of stimulus patterns, the performance differences between HORNs and the other networks tended to increase further (SI Appendix, Figs. S11–S15).

In some cases, even a leaky integrator network performed well when hyperparameter values were optimized (Fig. 1B and C). This finding can be explained by the fact that both HORNs and leaky integrator networks benefit from the residual connections introduced by the discretization scheme of the underlying ODE that introduces stable Lyapunov exponents and stabilizes gradients (SI Appendix). Furthermore, the oscillating dynamics of each DHO node temporally modulate gradients, and this can result in an increase in the practical expressivity of the networks [SI Appendix, Fig. S17 and (48)].

In contrast to the other architectures tested, HORNs were highly noise-resistant and showed only a gradual decline in task performance with increasing noise levels (Fig. 1E). Furthermore, the stimulus representations in HORNs were robust to a mismatch of noise characteristics between training and inference runs (Fig. 1F). The strong attenuation of high-frequency signals can explain this robustness due to the nonlinear, frequency-dependent gain modulation of input signals at each DHO node (Receptive Fields). In particular, this noise tolerance persists even when HORNs are trained on shuffled sequential psMNIST stimuli, in which the dominance of the low-frequency information characteristic for the sMNIST stimuli is removed (SI Appendix, Figs. S10 and S16). As both biological and artificial systems

must learn from noisy stimuli and robustly detect, classify, and process stimuli even under changing noise characteristics, this noise resilience constitutes another attractive property of HORNs.

To assess the influence of the presence of feedback parameters v (amplitude feedback) and w (velocity feedback, diagonal terms of the recurrent weight matrix) on network performance, we trained HORNs with and without DHO feedback connections on sMNIST and psMNIST and measured their task performance (*SI Appendix*, Figs. S18 and S19). We found that even in the absence of feedback in both amplitude and velocity ($v = w = 0$), performance only decreased slightly, but both feedback terms were needed to obtain the fastest learning speed and the highest overall performance.

Before learning, the dynamics of homogeneous HORN^h networks is dominated by large-scale synchronization among the nodes (Fig. 2 *A* and *B*). As learning progresses, global synchronization decreases (Fig. 2 *C*), increasing the dimensionality of network dynamics (*SI Appendix*, Fig. S2*A*). This reduction in global synchronization is accompanied by the emergence of complex, spatiotemporally structured correlations and higher-order synchronization patterns that are stimulus-specific (as assessed via the Kuramoto order parameter that is able to capture higher-order synchronization phenomena; see Fig. 2*D*), well segregated in the high-dimensional activity landscape of the network, and classifiable by a linear read-out (Fig. 2 *A*, *B*, and *D* and *SI Appendix*, Fig. S2 *B* and *C*).

Heterogeneous Networks. The structural and functional organization of mature cortical networks is characterized by heterogeneity (49, 50). Although several recent studies found that this variability can benefit learning and computations (51–53), it is still debated to what extent natural heterogeneity is functionally relevant. To test whether increasing network heterogeneity facilitates learning in HORNs, we simulated nonhomogeneous HORNⁿ networks in which each node had a different natural frequency, damping coefficient, and excitability (*Materials and Methods*). As expected, heterogeneous HORNs responded already in the untrained state with patterns that were more complex and less globally synchronized (Fig. 2*A*). As in the homogeneous case, global synchrony decreased further as learning progressed (Fig. 2*C*). Here again, the decrease in global synchrony led to an increase in the dimensionality of the dynamics (*SI Appendix*, Fig. S2*A*), although the dimensionality of the dynamics was higher compared to their homogeneous counterparts already before training.

A comparison between homogeneous and heterogeneous HORNs revealed superior performance of the latter with respect to learning speed in datasets with more complex spectral structures such as psMNIST (*SI Appendix*, Fig. S12). For such datasets, heterogeneous HORNs performed better than their parameter-optimized homogeneous counterparts. In addition, noise tolerance was enhanced (Fig. 1*F* and *SI Appendix*, Fig. S12). Note that in the case of sMNIST digits that have a reduced complexity of their signal statistics, the final performance of heterogeneous HORNs was found to be on par with that obtained by homogeneous HORNs whose parameters had been optimized for specific stimuli (Fig. 1 *B* and *C* and *SI Appendix* Fig. S11). Thus, heterogeneity allows one to obtain networks with high task performance without the need to find optimal parameter configurations, saving computationally expensive resources. The higher task performance of heterogeneous over homogeneous networks results from the fact that heterogeneity introduces a

multitude of timescales into network dynamics that can be used to process and represent stimuli (48).

Furthermore, heterogeneity brings network dynamics closer to criticality (48, 51, 53, 54). Network dynamics close to the critical point are characterized by scale invariance and divergence of spatial and temporal correlation lengths, leading to an increased dynamic repertoire and longer memory timescales, providing computational advantages for reservoirs that encode information in transient states (54). As HORNs also encode information in transients, the long-lived transient states enabled by dynamics close to criticality increase their computational power and allow them to best meet the trade-off between efficiency and stability of learning (48).

The advantages of heterogeneity were found to be especially prominent for larger networks, more difficult classification problems, and higher levels of stimulus noise (*SI Appendix*, Figs. S11–S15). Because heterogeneous HORNs produced highly structured response landscapes already in the untrained state, we hypothesized that they might also serve as efficient reservoirs and confirmed that this is indeed the case (*SI Appendix*). This finding has most recently also been confirmed in ref. 55.

Conduction Delays. Another source of variability in biological neuronal networks is the scattered nature of conduction delays between nodes (56). To test the influence of introducing coupling delays on task performance, we started with a HORN^h and endowed all recurrent connections with uniformly distributed variable coupling delays [$1, d_{\max}$] (*Materials and Methods*). This manipulation increased HORN performance in both maximal classification accuracy and learning speed on psMNIST (Fig. 3*A*) due to the phase shifts introduced by the delays. Increasing d_{\max} , which results in greater heterogeneity, was found to increase task performance, and this gain of function increased with increasing values of d_{\max} . Thus, like for the preferred oscillation frequencies, heterogeneity in conduction delays enables the generation of more diverse spatiotemporally structured activity landscapes in HORNs, thus increasing the dimensionality of the networks' state space and their performance, particularly for datasets with complex spectral properties.

Multilayer Networks. In the mammalian cerebral cortex, sensory signals are processed in hierarchically organized cortical areas that are reciprocally coupled (57). To investigate the potential benefits of distributed multistage processing, we generated two-layer HORN^h_{32,32} networks consisting of a lower layer (L1) and an upper layer (L2), with each layer consisting of a HORN^h₃₂, and introduced sparse reciprocal connections between the layers (Fig. 3*B* and *Materials and Methods*). The input signals were presented to L1 as before, and the result was read out at the nodes of L2. We performed a grid search on the interlayer feed-forward and feed-back connection probabilities f_F, f_B , as well as the scaling factors f_ω, f_γ that controlled the scaling of the parameters ω, γ of L2 with respect to L1. For each parameter configuration, we trained a HORN^h_{32,32} and assessed the best task performance on psMNIST during 10 training epochs (we chose psMNIST over sMNIST due to the much richer spectral stimulus properties). Note that for each fixed value of each parameter, a parameter configuration usually exists that results in a highly performing two-layer network (*SI Appendix*, Fig. S3).

Importantly, we found that many two-layer HORNs outperformed a single-layer HORN^h at a comparable number of trainable parameters (*SI Appendix*, Fig. S3*C*). In particular, higher task performance resulted when both the preferred

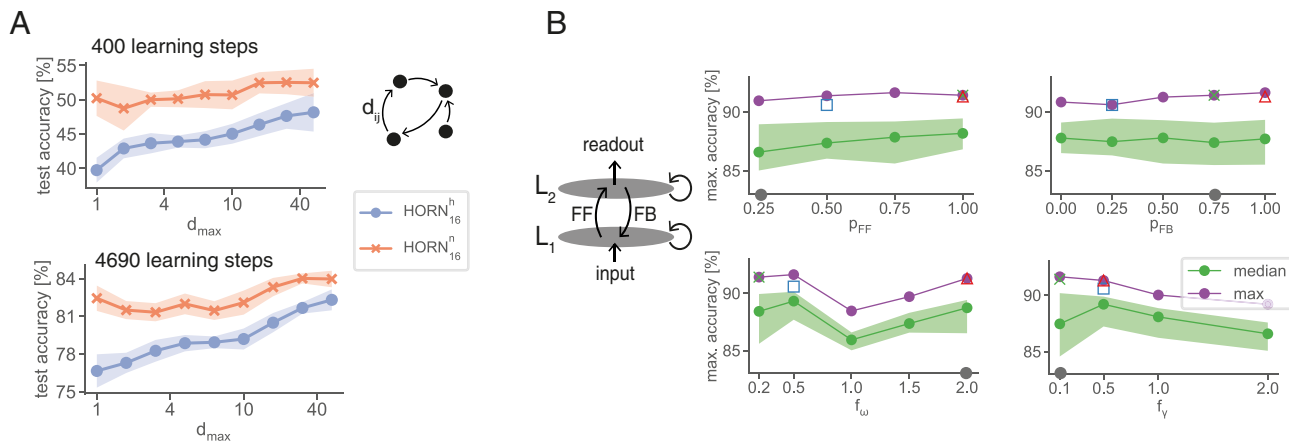


Fig. 3. Biologically inspired extensions to HORN networks. (A) Performance of HORNs with connection delays. Maximum test accuracy on psMNIST after 400 training steps (Left) and after 4,690 training steps (Right), respectively, as a function of maximal synaptic delay d_{max} . For each network, connection delays were sampled from a uniform distribution $U([1, d_{max}])$. A network with $d_{max} = 1$ corresponds to a regular HORN. Lines show mean performance over 10 randomly initialized networks, shaded areas show SD. (B) Two-layer $HORN_{32,32}^h$ performance as a function of the four parameters p_{FF} , p_{FB} (fraction of feed-forward and feed-back connections present, Upper panels), f_w , f_v (scaling factor of ω and γ between the first layer and the second layer, Lower panels). Curves represent median and maximal values of marginal distributions of the maximal accuracy on psMNIST attained during training for each value of the four parameters. Colored symbols refer to five exemplary networks whose parameter constellations and learning curves are shown in SI Appendix Fig. S3. Note the drop in performance when L1 and L2 have the same preferred oscillation frequency and that increasing connection sparsity does not result in a strong performance drop (gray circle marked on axes lies outside of the accuracy range displayed).

oscillation frequency and the damping coefficient were lower in the upper layer (Fig. 3 B, Bottom row). In this setting, the dynamics of the faster lower layer L1 is partly opaque to the slower upper layer L2 due to the stronger attenuation of high frequencies in L2. However, activity in L2 is capable of entraining the nodes in L1, potentially supporting processes such as feature binding by contributing more global binding criteria available to the upper layer, but not to the lower layer. Because the upper layer nodes receive convergent input from the lower layer and operate at a longer time scale, the upper layers can bind longer segments of the stimulus time series. Interestingly, we found that a two-layer network in which the separation of frequency bands across layers breaks down fails to learn (SI Appendix, Fig. S3C, gray curve). In this case, the computations in the different layers are not sufficiently separated into frequency bands, and the cross-talk between layers in the same frequency band hinders the networks from solving the stimulus classification problem successfully.

Fading Memory and Evidence Accumulation. Electrophysiological recordings of neurons in the visual cortex have shown that population activity exhibits fading memory and is capable of simultaneously representing classifiable information on the identity of successively presented stimuli, including their sequence order (58). HORNs share this ability. Following the sequential presentation of two different stimuli, a linear classifier can decode the two stimuli from the same segment of reverberating activity (Fig. 2F). As heterogeneous HORNs possess more diverse memory timescales than homogeneous networks, their ability to represent simultaneously information about temporally segregated stimuli is superior, yet another functional advantage of network heterogeneity.

As in any RNN, the response patterns of HORNs evolve over time due to network dynamics. To determine the time at which the network converges to states of maximal stimulus specificity, a linear SVM was trained on stimulus classification using activity data from a trained homogeneous HORN at different time points throughout the stimulus presentation period. We found that the networks accumulate evidence, with their dynamics allowing

for progressively better decoding of stimulus identity as the network approaches the read-out time on which it was trained (SI Appendix, Figs. S2 B and C and S20).

Receptive Fields. To better understand the principles of computation in HORNs and how stimulus-specific activity patterns emerge during training, we investigated how learning changed the response properties of both individual nodes and the entire network. Each DHO node in a HORN has a gain curve $G(\omega_i)$ that describes how the node modulates the amplitude of a temporally modulated input signal as a function of the input signal frequency ω_i (Fig. 2E and SI Appendix, Fig. S8). Note that this holds both for external, stimulus-dependent signals if they have a temporal structure and always for the oscillatory activity conveyed by the intrinsic recurrent connections. The shape of G is determined by the values of ω , γ , and the self-connection terms v (amplitude feedback) and w (velocity feedback, diagonal terms of the recurrent weight matrix). We call G the intrinsic receptive field (IRF) of the node because it defines the frequency band in which the node shows feature selectivity. During learning, the adjustment of the self-connection weights v and w alters G and therefore drives changes in the IRF of each node, allowing the node to improve its selectivity for stimulus features useful for performing a given task (Fig. 2E). In homogeneous HORNs, the IRFs of all nodes only differ due to learning-induced changes in the values of the feedback parameters v and w that vary between nodes (Fig. 2 E, Left). For heterogeneous networks, the IRFs of the untrained network already cover a larger portion of the frequency space, with nodes tuned to a wider variety of features in the frequency space (Fig. 2 E, Right). Note that the value of ω not only influences the IRF but also sets the frequency band in which the node codes. This increases the diversity of frequency bands available for processing in heterogeneous HORNs.

For comparison with neuronal systems, we mapped the receptive fields (RFs) of the nodes in the same way as is common practice in electrophysiological studies (59). We examined which stimuli activate a node most strongly and address the so-defined RF as effective RF (ERF). Like in neuronal systems (60, 61),

the structure of ERFs varied greatly depending on the nature of the test stimuli. The reason is that the responses depend not only on the external input but also on the network's recurrent dynamics. We determined the ERFs of nodes in our models with two canonical choices of mapping stimuli: i) simple line segments of different orientations, as commonly used for RF mapping in visual experiments (59) (*SI Appendix, Fig. S22*), and ii) the sMNIST digits on which the networks were trained, as an example of complex natural stimuli (61). To measure ERFs, we simulated a heterogeneous HORN_{16}^n trained on sMNIST and calculated for each node the mean stimulus that resulted in maximal node activation (*Materials and Methods*). When stimulated with simple line stimuli, we find simple orientation-selective ERFs that closely resemble the RFs of neurons in the primary visual cortex (Fig. 2 *G, Left* column). When mapped with complex stimuli, some nodes exhibited specificity for particular features of the MNIST digits (Fig. 2 *G, Right* column). Other nodes had ERFs that appeared to be unrelated to the stimuli in the training set, and yet others were inhibited. To determine the influence of the collective network dynamics on individual ERFs, we silenced one node of the network. This resulted in an immediate reconfiguration of the ERFs of the other nodes, which was quite dramatic in some cases (Fig. 2 *G*).

This indicates that after learning, ERFs are, to a large extent, the result of dynamic interactions in the network. Thus, the ERFs of HORN nodes undergo context-dependent dynamic modifications similar to those observed in natural systems (60, 62). To quantitatively describe the ERFs of nodes in frequency space, networks were stimulated with harmonic sine-wave inputs, determining their frequency responses (*SI Appendix, Fig. S23*). Although the nodes in untrained networks resonated best with inputs near their natural frequencies, after training the nodes developed complex and hard-to-predict resonance patterns, confirming the strong influence of network interactions on ERFs.

Learning Priors. In mammalian primary sensory cortices, information about the statistical regularities of the natural environment is stored in the architecture and distribution of synaptic

weights, incorporating Gestalt criteria for feature binding and perceptual grouping (2).

To test how the installation of priors in HORNs influences their learning and task performance, we first installed a set of canonical priors by training a heterogeneous HORN^n (93 nodes) to discriminate simple elongated contours of different orientations placed at random locations in the MNIST pixel matrix (*SI Appendix, Fig. S22*). Such canonical priors enabled pretrained HORNs to create distinct MNIST representations, classifiable with high specificity by just training the readout layer (*SI Appendix, Fig. S24*). Allowing recurrent connections to continue learning after installing priors further increased learning speed (*SI Appendix, Fig. S24*). Prestructured HORNs needed fewer training steps than those trained on sMNIST directly to achieve similar performance, showing that known statistical feature contingencies enhance learning efficiency. Larger pretrained heterogeneous networks also needed fewer steps for high performance (*SI Appendix, Fig. S24*). We predict that large and heterogeneous pretrained HORNs could achieve few-shot learning, akin to natural neuronal systems. Taken together, these findings emphasize the beneficial effects that priors have for the orthogonalization of object representations in artificial and most likely also natural neuronal networks.

Hebbian Learning. We examined whether BPTT-based training of HORNs aligns with Hebbian learning principles by analyzing changes in weight distributions and response activity correlations before and after training. Interestingly, the changes in synaptic weights resulting from BPTT training matched Hebbian predictions (Fig. 4*A* and *SI Appendix, Fig. S4A*). BPTT apparently capitalizes on stimulus-specific correlation structures of network activity to enhance those connections that induce correlation patterns characteristic of particular stimuli, fostering the development of stimulus-specific patterns. Heterogeneous HORNs, with their more diverse dynamics, naturally exploit these structures from the beginning, while homogeneous HORNs must first desynchronize to expand their state space and develop stimulus-specific correlation structures (Fig. 2 *A–C*).

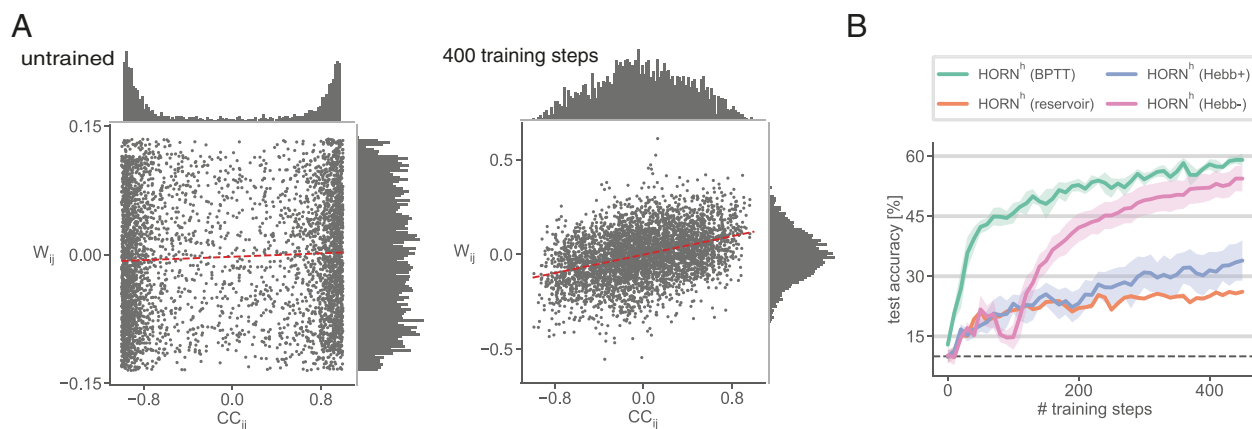


Fig. 4. Backpropagation learning in HORN networks results in Hebbian-like weight changes. (A) Scatter plots of connection weights W_{ij}^{hh} and mean cross-correlation coefficients CC_{ij} of node activities of a homogeneous HORN^h (64 nodes) before training (*Left*) and after training on psMNIST for 500 training steps (*Right*). CC_{ij} computed over 100 samples. Linear regression lines in red. Marginal distributions of W_{ij}^{hh} and CC_{ij} are shown on the *Right* and *Top*, respectively. Note the bimodal distribution of correlation coefficients with modes around $-1, 1$ in the untrained state and the more decorrelated network activity as a result of learning. (B) Performance of a homogeneous HORN^h (64 nodes) network as a function of training steps when instances of the same network are trained with correlation-based Hebbian (suffix “Hebb+”) or anti-Hebbian (suffix “Hebb-”) learning rules compared to instances trained with BPTT (suffix “BPTT”) and when W^{hh} was fixed (suffix “reservoir”). The input and readout parameters are trained with BPTT for all instances. Curves show mean performance over 10 network instances with random weight initialization, shaded areas SD. Note the strong performance of the anti-Hebbian rule for this initially highly synchronized network.

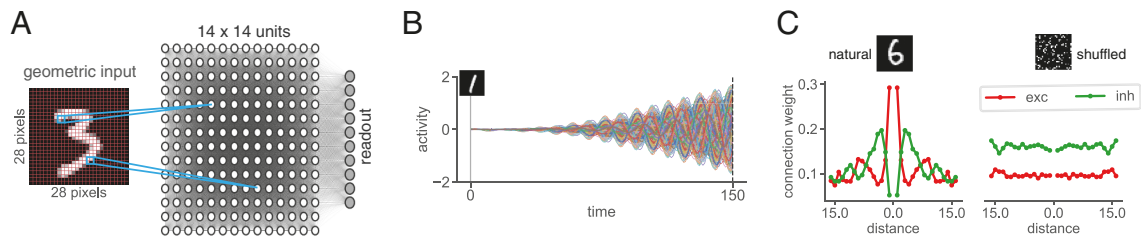


Fig. 5. Self-organization in HORNs driven by geometric input. (A) Schematic of geometric MNIST input provided to an all-to-all connected homogeneous HORN (196 nodes). Each DHO node receives input from its 2×2 pixel receptive field. Input weights are constant and fixed, recurrent and readout weights are trainable. Readout nodes are connected to the entire HORN. (B) Example network dynamics of a HORN^h (196 nodes) stimulated as in A after training with BPTT for 2,000 steps (achieved classification accuracy around 90% for both natural and shuffled stimuli). Input is flashed to the network during one time step at $t = 1$. Readout is performed after 150 time steps of network activity during which no input is given. Curves represent the activity of all nodes in the network (node identity is color-coded). Note the increasing amplitudes of reverberating activity and the drifts in phase of the oscillatory responses. (C) Distance-dependent connection weights of excitatory (red) and inhibitory (green) connections in HORN^h networks after training with BPTT for 2,000 steps on geometric stimuli. Stimuli were either natural (Left) or shuffled (Right) MNIST digits. Note the Mexican-hat-like connectivity structure among nodes in networks trained on natural stimuli and the lack of structure in the connectivity in networks trained with shuffled stimuli.

We examined replacing BPTT with correlation-based learning rules, using unsupervised Hebbian and anti-Hebbian rules to modify recurrent connection weights. Here, a HORN with fixed recurrent weights and plastic input and readout connections served as a baseline (the “reservoir” HORN, Fig. 4B). When recurrent connections were modified by correlation-based learning rules but not BPTT, anti-Hebbian rules in homogeneous networks surprisingly resulted in performance similar to BPTT (Fig. 4B), promoting selective desynchronization and enabling stimulus-specific correlations. In heterogeneous networks, conventional Hebbian rules improved performance above baseline by enhancing existing correlations (SI Appendix, Fig. S4).

In nonoscillating networks, BPTT also led to Hebbian-compatible weights, although less effectively than in HORNs. HORNs quickly acquired stimulus-specific weights due to the dynamics of coupled oscillators, which amplify and sustain synchronous activity patterns suitable for Hebbian mechanisms. BPTT benefits similarly, using resonance to orthogonalize representations in high-dimensional state spaces.

Taken together, these results provide a proof of principle that unsupervised Hebbian learning at the level of recurrent connections in HORNs supports the segregation of stimulus-specific dynamic states and thus facilitates their classification (Fig. 4B and SI Appendix, Fig. S4B).

Geometric Input Drives Self-Organization and Traveling Waves.

To test how HORNs process spatiotemporally structured input rather than just scalar time series, we trained networks that receive spatially organized input (Fig. 5A). To this end, we activated each node for one time step at an intensity corresponding to the sum of the intensity values of the MNIST pixels within its RF. We then gave the network 150 time steps to process the stimulus before performing a linear read-out (Fig. 5B). For training, we used the BPTT algorithm as before, started with an all-to-all random connectivity, and kept the weights of the input layer fixed, while the recurrent and readout weights were plastic. After training for 10 epochs, the best classification accuracy on the test set was found to be 90.21%.

When stimulated, each of the simultaneously activated nodes responded with a damped oscillation (Fig. 5B). These activities spread throughout the network and led to traveling waves and complex interference patterns (Movies S1 and S2). The direction and shape of the waves differed for responses evoked by different stimuli. HORNs are able to sustain oscillatory activity in the form of a standing wave in each DHO node. These waves,

once initiated, are sustained and give rise to global interference patterns. In contrast, traveling waves in RNNs without oscillating nodes interfere only when wave fronts collide.

In our model, nodes coactivated by a flashed stimulus exhibit stimulus-locked synchronized oscillations. As the BPTT algorithm mimics Hebbian plasticity in HORNs, this predicts that during learning these synchronously active nodes should increase their mutual coupling. Consequently, we observed a shift from unspecific all-to-all connections to spatially restricted connections with a distant-dependent decay of coupling strength, following a Mexican hat-like shape (Fig. 5C, Left).

This connectivity captures the essential structural feature of the MNIST stimuli, the continuity of their contours, and the spatial vicinity of activated nodes. During testing (recall) of the trained model, a particular stimulus again induces synchronized oscillations in a respective constellation of nodes, and, due to the enhanced coupling of those nodes in the trained network, these now engage in resonance which leads to an increase in response amplitude.

As expected, such distant dependent connectivity patterns did not emerge when training HORNs on shuffled MNIST stimuli (Fig. 5C, Right), although the networks achieved a comparable classification accuracy of 90.48% on the test set. In the latter case, other priors than spatial continuity and vicinity were installed in the architecture of the coupling connections, namely those representing the statistical contingencies of the shuffled MNIST digits without local spatial structure.

Spontaneous and Evoked Activity. Cortical networks are spontaneously active, and stimulation typically leads to a reduction in variability and the emergence of stimulus-specific substates (63, 64). In spontaneously active HORNs (SI Appendix), spontaneous activity spans large, confined state spaces that envelope stimulation-induced response spaces. In such networks, dynamics rapidly and transiently converge to stimulus-specific substates upon stimulation (SI Appendix, Fig. S5), reproducing experimental phenomena (65).

Discussion

Controlled Oscillations. Implementing characteristics of the mammalian cerebral cortex in RNNs revealed a powerful computational principle based on oscillatory activity. Although RNNs without oscillating nodes naturally produce oscillations, such emergent oscillations are often transient and difficult to

control, hindering their exploitation by gradient-based learning. Enforcing oscillations at each node in HORNs allowed us to study the functional relevance of oscillatory dynamics, identify the computational principle responsible for the increased performance of HORNs, and establish close relationships with the dynamics of natural networks such as the cerebral cortex.

In HORNs, individual nodes turn any input into an oscillation, acquiring the ability to extract features through resonance and, more generally, to modulate gain in a frequency-dependent manner. Networks, in turn, generate holistic transient stimulus representations characterized by wave interference patterns (37, 38, 66). Performance testing on standard pattern recognition benchmarks revealed that a gradient-based learning scheme can capitalize on this extended dynamical repertoire, leading to substantially enhanced performance relative to RNNs without oscillatory nodes. These findings of improved task performance are in line with previous studies in the field of machine learning that investigated RNNs with oscillating nodes (30, 37, 38) (while HORNs have 50% less trainable parameters for the same number of nodes).

Enforcing oscillatory activity in network nodes serves as an inductive bias in RNNs, enhancing model expressivity (48). This oscillatory bias has also been shown to improve task performance in spiking networks by allowing subthreshold membrane potentials to oscillate (67, 68). These findings suggest a universal computational principle based on coupled oscillators, enabling wave-based representations applicable to both neuronal populations, as well as single neurons (69).

When trained with geometrically organized stimuli, HORNs developed local connectivity patterns. Such locally connected oscillator RNNs (37, 38) can be interpreted as a discretization of a neural field model with a specific connection kernel that implements a damped wave equation (70, 71). In this sense, local oscillations and global waves are two sides of the same coin, and the dynamics of HORNs and field models are equivalent under certain conditions, a topic left for future study.

Reasons for Increased Performance. The good performance of HORNs is due to several reasons, and these are closely related to the propensity of network nodes to engage in oscillations.

First, innate preferences for stimulus features (controlled by the parameters ω , γ , α) allow individual nodes to efficiently extract and encode stimulus features already in an untrained network and contribute to the noise resilience of HORNs. If input signals lack temporal structure, the oscillatory properties of the nodes are still beneficial because they transform sustained inputs into oscillatory responses. This transformation allows computations in the common format of temporally modulated signals which prevail in the communication among nodes; see also the fourth point below.

Second, the discretization scheme used for the oscillator differential equations introduces temporal residual connections that stabilize gradients in BPTT learning, and the oscillating dynamics can increase the practical expressivity of the networks by modulating gradients [*SI Appendix* and (30, 48)].

Third, DHOs in a HORN network collectively process stimuli in a fully distributed manner by converting sensory input into waves. Initially, these are standing waves in each oscillator, but then they spread and cause complex interference patterns at the network level (16, 17), findings that are compatible with physiological evidence (66). This representation provides a coding space of massive dimensionality, and, most importantly, permits the superposition of information about multiple spatially

and temporally segregated events. This allows HORNs to analyze and encode simultaneously not only spatial but also temporal relations between a large number of stimulus features and to generate holistic representations of the correlation structure of complex input constellations.

The Virtues of Heterogeneity. Heterogeneity improves the performance of RNNs because it increases the dimensionality of the state spaces of the networks.

Having oscillatory nodes allowed us to increase heterogeneity by varying preferred oscillation frequencies, which increased task performance. In addition to enhancing heterogeneity by varying the preferred oscillation frequencies of the nodes, we induced heterogeneous conduction delays to deliberately induce phase shifts. This further increases the dimensionality of the networks' coding space, which can be exploited for computation.

The advantages of heterogeneity are also documented by simulations of two-layer networks. These networks showed enhanced performance at the same number of parameters, in particular when the higher layer operated at lower preferred frequencies than the lower layer. This allows the multilayer network to operate in different frequency bands and to perform parallel analyses of input patterns at different temporal scales in each layer. This finding was the result of a grid search for optimal parameter settings in two-layer networks and shares similarities with the organization of the cerebral cortex. Here, too, oscillation frequencies decrease as one progresses from lower to higher processing levels (11, 72). Slower oscillations at higher levels can establish relations among temporally segregated stimuli over longer time intervals, which could support chunking. Interestingly, according to basic physics, waves with slower frequencies tend to travel over longer spatial distances, in our case over a larger number of network nodes. In the cerebral cortex, higher areas integrate information from increasingly diverse and spatially remote processing streams, as reflected by their large, often polymodal, and multisensitive receptive fields. Assuming a wave-based representation (66), operating at lower oscillation frequencies would allow these higher areas to integrate information over larger temporal and spatial scales, favoring holistic processing of information and multimodal binding.

Another advantage of heterogeneity is that it brings network dynamics closer to criticality (48, 53, 73). Dynamics close to criticality are a hallmark of cortical networks and provide computational benefits summarized in the "critical brain" hypothesis (74). These benefits are due to the emergence of long-lived transient and metastable states (54). HORNs also encode information in transients and therefore are capable of coding with sequences of metastable states and ghost attractors (75) when in a regime close to criticality. This distinguishes their dynamics from that of attractor networks (76) for which critical slowing down limits computational power near criticality (77). More studies are needed to better understand these transient dynamics in HORNs and to identify related activity in biological networks (75, 78).

In summary, we found that the implementation of physiologically plausible heterogeneity typically increases performance without increasing the number of trainable parameters. Heterogeneity i) gives even untrained networks sensitivity to diverse correlation structures, thereby accelerating learning; ii) enhances the processing of novel or noisy stimuli with varying spectra; iii) expands coding dimensions; iv) allows networks to utilize computational benefits resulting from dynamics closer to criticality, while v) at the same time reducing the need for costly parameter tuning. The gain of function by heterogeneity

was particularly pronounced in larger networks, suggesting that one-shot learning, a hallmark of biological systems, is facilitated in large, heterogeneous networks such as the cerebral cortex. This lets us conclude that the apparent heterogeneity in natural neuronal systems is likely not a reflection of nature's imprecision but rather an efficient solution to computational challenges.

Relations to Neurobiological Systems. HORNs reproduced several characteristic features of the dynamics and organization of natural neuronal systems, particularly the cerebral cortex and probably also the hippocampus. In addition, the simulations allowed us to assign concrete functions to features of natural networks whose role in information processing is still a matter of discussion.

Our simulations show that learning-dependent complex, transient, and stimulus-specific synchronization patterns benefit information processing and identify the oscillatory properties of network nodes as an underlying mechanism. This supports the hypothesis that oscillations and synchrony, also observable in neuronal systems (79), are functionally relevant and not epiphenomena.

Simulations with geometrically organized input patterns processed by the visual system yielded results similar to those obtained with time-series data that do not contain geometric information. Thus, the identified computational principles can handle spatial and temporal relations among input signals similarly and represent computations in the same format. This benefits computations in sensory cortices receiving both temporally and spatially structured input and aiding cross-modal and interareal communication. For spatially structured stimuli, learning led to synaptic weight configurations that decay with distance and capture the Gestalt criteria of continuity and vicinity, a property also known from natural systems (2, 65). In the visual cortex, the basic layout of recurrent connections is genetically determined, but experience-dependent pruning of these connections further enhances their selectivity through a Hebbian mechanism (80).

Stimulation of locally connected HORNs led to traveling waves that closely resemble those observed in natural neuronal networks (16, 66). Traveling waves are also a hallmark of oscillatory RNNs in which local connectivity was enforced by design (37, 38). Wave-based representations allow for very high-dimensional representations and manifold coding strategies. Consequently, numerous hypotheses have been proposed regarding the functional role of traveling waves (17, 69). In a wave-based model of the motor cortex, the direction and wavelength of traveling waves are used to structure commands in a way that is easily decodable by the dendritic arbors of neurons in the descending motor system (69). However, the exact function of traveling waves in the sensory cortices is still not fully understood.

Another similarity between the dynamics of HORNs and the cerebral cortex is the temporal evolution of responses in simulations with geometrically structured stimuli. The initial transient responses were amplified by reverberation, increasing the decodability of the dynamic state due to better segregation of stimulus-specific principal components of the population vector (65). This state can be seen as a highly parallelized search for the best match between sensory evidence and learned priors (2). Thus, one of the core functions of predictive coding, the matching of sensory evidence with stored priors, can be realized through self-organizing dynamic interactions in oscillatory recurrent networks.

During learning, nodes activated by semantically related features increase their mutual coupling, and during recall, these nodes self-organize into a stimulus-specific assembly with synchronized and jointly enhanced responses. This dynamic association of nodes is also observed in the visual cortex for neurons tuned to perceptually bound features (10, 65) and is at the core of the binding by synchrony hypothesis (BBS) (20). HORNs, by exploiting the resonance properties of coupled oscillators, reproduce this important feature of natural cortical networks. The dynamics of spontaneously active HORNs resemble those of natural cortical networks in which stimulation decreases variance (63) and temporarily aligns dynamics to stimulus-specific substates. These substates exist within the subspace of spontaneous activity and arise from comparisons of sensory evidence and stored priors (2). Therefore, spontaneous activity can be seen as a blend of fragments of learned stimulus-specific representations.

In addition to reproducing many physiological phenomena, additional physiological experiments can now be designed to examine specific predictions derived from the present study. These experiments will require massive parallel recordings of neuronal activity both within and across cortical areas with high spatial and temporal resolution to capture the spatiotemporal dynamics of traveling waves and their resulting interference patterns.

Concluding Remarks. Taken together, the present results not only unveil the computational principles accessible to HORNs and other oscillator networks but also allow for a functional interpretation of numerous experimentally verified physiological phenomena whose roles in information processing have been elusive or have caused controversial discussions. Plausible functional roles can now be assigned to i) the propensity of nodes to oscillate and the resulting dynamical phenomena such as synchronization, desynchronization, resonance, entrainment, and traveling waves (12–15, 17), ii) the diversity of preferred oscillation frequencies, their nonstationarity and context dependence (3, 4), iii) the heterogeneity of the conduction velocities of the recurrent connections (5, 6), iv) the decrease of oscillation frequencies in higher areas of the cortical processing hierarchy (11, 72), v) the Hebbian adaptivity of recurrent connections (7, 8), vi) the emergence of context-dependent dynamic receptive fields by network interactions (60, 61), and vii) the reduction of variance in network dynamics during stimulus presentation (63).

The simulations also suggest a physiologically plausible scenario for the rapid and parallel matching of sensory evidence with stored priors through self-organized convergence of network dynamics to classifiable, stimulus-specific, dynamic substates. These substates consist of highly structured, high-dimensional dynamical landscapes that unfold due to interference of wave patterns in amplitude, frequency, and phase space. In essence, the described networks perform highly parallelized analog computations in high-dimensional state spaces that simultaneously relate a large number of spatially and temporally structured input variables, a capacity ideally suited to accomplish context-dependent feature binding and scene segmentation. Consequently, attempts are made to exploit the principle described in this study in machine learning architectures designed to perform scene segmentation (81). Moreover, the computational strategy implemented by HORNs is also well suited to overcome challenges requiring the simultaneous evaluation of multiple nested relations as occurring, for example, in language comprehension. Interestingly, biological systems are at ease with

solving the binding problem, with the segmentation of cluttered scenes and with the analysis of complex time series (e.g., spoken language), while these tasks are notoriously difficult for digital computer architectures that typically rely on serial feedforward processing.

We believe that nature solves such hard problems through analog computations of the kind described in this study. We predict that it will be possible to implement the computational principle presented here in analog hardware that runs at room temperature, is miniaturizable, and is highly energy efficient. Combined with electrical elements mimicking Hebbian synapses, such as memristors, this principle will likely enable the design of self-adapting devices for machine learning applications that can ideally complement existing digital technologies.

Materials and Methods

Network Models. The update equations for a HORN network of n units in discrete time t result from the discretization of a second-order ODE describing a driven damped harmonic oscillator (SI Appendix) and are given by

$$\mathbf{y}_{t+1} = \mathbf{y}_t + \boldsymbol{\alpha} \cdot \tanh\left(\frac{1}{\sqrt{n}}\mathbf{I}_{t+1}^{\text{rec}} + \mathbf{I}_{t+1}^{\text{ext}}\right) - 2\boldsymbol{\gamma} \cdot \mathbf{y}_t - \boldsymbol{\omega}^2 \cdot \mathbf{x}_t,$$

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{y}_{t+1},$$

where vectors and matrices are indicated by boldface symbols, and $\boldsymbol{\omega}$, $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$ are the natural frequencies, damping factors and excitability factors of the network nodes, respectively. Initial conditions are $\mathbf{x}_0 = \mathbf{y}_0 = \mathbf{0}$ unless stated otherwise. $\mathbf{I}_{t+1}^{\text{rec}} = \mathbf{W}^{hh}\mathbf{y}_t + \mathbf{b}^{hh} + \mathbf{v} \cdot \mathbf{x}_t$ and $\mathbf{I}_{t+1}^{\text{ext}} = \mathbf{W}^{ih}\mathbf{s}_{t+1} + \mathbf{b}^{ih}$ denote the recurrent and external input to each node, respectively. Here, the diagonal entries of \mathbf{W}^{hh} and \mathbf{v} denote feedback parameters; see SI Appendix. \mathbf{W}^{ih} , \mathbf{b}^{ih} , and \mathbf{W}^{hh} , \mathbf{b}^{hh} denote the input and hidden weights and biases, respectively, and $\mathbf{S}_{\text{ext}} = (\mathbf{s}_1, \dots, \mathbf{s}_T)$ the external input. HORN networks with conduction delays were simulated by assigning to each connection a delay sampled from a uniform distribution $1 \leq d_{ij} \leq d_{\text{max}}$ of time steps and defining the recurrent input to each node as $\mathbf{I}_{t+1}^{\text{rec}} = \sum_{d=1}^{d_{\text{max}}} \mathbf{M}^d \mathbf{W}^{hh} \mathbf{y}_{t+1-d} + \mathbf{b}^{hh} + \mathbf{v} \cdot \mathbf{x}_t$, where $\mathbf{M}^d \in \{0, 1\}^{n \times n}$, $1 \leq d \leq d_{\text{max}}$ are binary masks and the value of $\mathbf{M}_{ij}^d = 1$ iff the connection from node j to node i has a delay of d , and $\mathbf{M}_{ij}^d = 0$ otherwise, with $\mathbf{y}_i = 0$ for $i < 0$. Similarly, two-layer HORNs were simulated as a HORN of $n = n_1 + n_2$ nodes, where the first n_1 nodes represent layer 1, and the rest of the nodes layer 2. Different connectivity patterns between layers were modeled by modifying the recurrent input using a binary mask as for the networks with delays (SI Appendix). An Elman RNN (label “tanh”) was implemented by the update equation $\mathbf{x}_{t+1} = \tanh(\mathbf{W}^{hh}\mathbf{x}_t + \mathbf{b}^{hh} + \mathbf{I}_{t+1}^{\text{ext}})$. Similarly, a “leak” RNN consisting of leaky integrators was implemented by the update equation $\mathbf{x}_{t+1} = \mathbf{x}_t + h \left[-a\mathbf{x}_t + \tanh(\mathbf{W}^{hh}\mathbf{x}_t + \mathbf{b}^{hh} + \mathbf{I}_{t+1}^{\text{ext}}) \right]$, where h is a microscopic time constant and a a leak parameter that controls memory decay (SI Appendix, Table S1). For GRU and LSTM networks, the default implementations in PyTorch 1.9 were used.

Datasets. The MNIST dataset of handwritten digits comprises 28×28 matrices of intensity values scaled to $[0, 1]$. sMNIST samples are given by 784-length time series representing MNIST digits in scanline order. For psMNIST, a fixed, random permutation is applied to shuffle the pixel positions (SI Appendix, Figs. S10 and S16). For the noisy case, i.i.d., additive white Gaussian noise was added at the pixel level, sampled from $N(0, \sigma^2)$, with stimulus values clamped to $[0, 1]$ after application. The Line segments dataset (LSDS) comprises 28×28 pixel samples with line segments at specific angles, defined by four parameters: the number of angles n_a , maximum segments per sample n_s , the minimum and maximum segment lengths l_{min} and l_{max} , and whether line segment locations are random (r) or centered (c); see SI Appendix. We use LSDS(32,5,11,c) as LSDSa and LSDS(10,3,8,24,r) as LSDSb (SI Appendix, Fig. S22). Further datasets are described in SI Appendix.

Simulations. All networks were simulated using PyTorch (version 1.9) and utilized an affine readout layer at the last time step for training all model parameters with BPTT using a binary cross entropy loss for all classification problems and an MSE loss for regression tasks. AdamW was used as an optimizer and gradient clipping was applied to non-HORN networks to ensure convergence. Optimal hyperparameter configurations were determined by a grid search for each architecture and dataset (SI Appendix).

Synchronization Measures. The PLV of two time series of length T is $\text{PLV}(x_1, x_2) = \frac{1}{T} \left| \sum_{t=1}^T e^{i\Delta\phi(t)} \right| \in [0, 1]$, where $\Delta\phi(t) = \phi_1(t) - \phi_2(t)$ and $\phi_i(t)$ denotes the instantaneous phase of x_i . The time-varying Kuramoto order parameter $r(t)$ is $r(t) = \frac{1}{n} \left| \sum_{k=1}^n \exp(i\phi_k(t)) \right| \in [0, 1]$, where $\phi_k(t)$ is the instantaneous phase of node k .

Receptive Fields. For each DHO node, the intrinsic receptive field (IRF) was calculated as the quotient between the DHO node amplitude, i.e., the stationary forced oscillation, and the input amplitude, as a function of the input frequency ω_i (gain curve) of the node using sinusoidal inputs using an amplitude of 1 over a time of 50,000 time steps to allow transient dynamics to settle. The effective receptive field (ERF) of DHO nodes was calculated as the mean most strongly driving stimulus. For this, network activity was first recorded for 10,000 samples in a test set, and then the average of the 500 stimuli that resulted in the highest mean absolute amplitude over entire stimulus presentation period of a given DHO node was calculated as the ERF.

Hebbian Learning. An additive rule was implemented that modifies the weight of a synaptic connection \mathbf{W}_{ij} from node j to node i by an additive term $\Delta\mathbf{W}_{ij} = \sigma_{\Delta} \lambda_h a_{ij} r(x_i(t), x_j(t))$, where $\sigma_{\Delta} \in \{+1, -1\}$ determines the type of the learning rule ($\sigma_{\Delta} = +1$: conventional Hebb+, $\sigma_{\Delta} = -1$: anti-Hebb-), λ_h denotes the learning rate, $a_{ij} \in \{0, 1\}$ an activity modulator, and $r(x_i(t), x_j(t))$ denotes the Pearson correlation between the activity vectors of the nodes i and j computed over the time steps $t = c, \dots, T$, respectively, where T denotes the stimulus length; see SI Appendix.

Geometric Input. A homogeneous HORN was endowed with geometrically organized nonoverlapping receptive fields (network units on a 14×14 grid with each unit receiving input from the corresponding 2×2 pixel region inside a 28×28 pixel MNIST digit). To simulate a flashed stimulus, the network received input only during the first time step and input connection strengths were kept constant (SI Appendix). The network was trained with BPTT on the MNIST classification task for readout at time $t = 150$. For the shuffled MNIST case, a random, but fixed permutation was applied to all MNIST samples.

Data, Materials, and Software Availability. GitHub repository has been deposited in <https://github.com/exilef/horn> (82).

ACKNOWLEDGMENTS. We thank Martin Vinck, Aandreas Bahmer, and Andreea Lazar for their helpful discussions and comments on an earlier version of this manuscript. We thank Claudia Kernberger for her help with some of the illustrations. We acknowledge computing support from the Ernst Strüngmann Institute and the Max Planck Computing and Data Facility. This study was supported by a Koselleck Grant to W.S. from the German Research Foundation and the Ernst Strüngmann Institute for Neuroscience in Cooperation with Max Planck Society.

Author affiliations: ^aErnst Strüngmann Institute, Frankfurt am Main 60528, Germany; ^bFrankfurt Institute for Advanced Studies, Frankfurt am Main 60438, Germany; and ^cMax Planck Institute for Brain Research, Frankfurt am Main 60438, Germany

Author contributions: F.E. and W.S. designed research; F.E., P.C., I.D., and W.S. performed research; F.E., P.C., I.D., and W.S. contributed new reagents/analytic tools; F.E., P.C., I.D., and W.S. analyzed data; and F.E. and W.S. wrote the paper.

Reviewers: M.B., The University of Newcastle-Newcastle City Campus; and C.M., Universitat de les Illes Balears Institute for Cross-Disciplinary Physics and Complex Systems (IFISC), Spanish National Research Council–University of the Balearic Islands (CSIC-UIB).

1. N. T. Markov *et al.*, The role of long-range connections on the specificity of the macaque interareal cortical network. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 5187–5192 (2013).
2. W. Singer, Recurrent dynamics in the cerebral cortex: Integration of sensory evidence with stored knowledge. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101043118 (2021).
3. R. R. Llinás, The intrinsic electrophysiological properties of mammalian neurons: Insights into central nervous system function. *Science* **242**, 1654–1664 (1988).
4. A. K. Engel, P. König, A. K. Kreiter, T. B. Schillen, W. Singer, Temporal coding in the visual cortex: New vistas on integration in the nervous system. *Trends Neurosci.* **15**, 218–226 (1992).
5. S. G. Waxman, Determinants of conduction velocity in myelinated nerve fibers. *Muscle Nerve* **3**, 141–150 (1980).
6. H. A. Swadlow, Physiological properties of individual cerebral axons studied in vivo for as long as one year. *J. Neurophysiol.* **54**, 1346–1362 (1985).
7. J. C. Magee, D. Johnston, A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science* **275**, 209–213 (1997).
8. H. Markram, J. Lübke, M. Frotscher, B. Sakmann, Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275**, 213–215 (1997).
9. D. Hebb, *The Organization of Behavior* (Psychology Press, 1949).
10. C. M. Gray, P. König, A. K. Engel, W. Singer, Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* **338**, 334–337 (1989).
11. G. Buzsáki, X. J. Wang, Mechanisms of gamma oscillations. *Annu. Rev. Neurosci.* **35**, 203–225 (2012).
12. W. Singer, Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* **55**, 349–374 (1993).
13. J. A. Cardin *et al.*, Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* **459**, 663–667 (2009).
14. K. B. Doelling, D. Poeppel, Cortical entrainment to music and its modulation by expertise. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E6233–E6242 (2015).
15. J. P. Lachaux, E. Rodriguez, J. Martinerie, F. J. Varela, Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* **8**, 194–208 (1999).
16. J. R. Hughes, The phenomenon of travelling waves: A review. *Clin. Electroencephal.* **26**, 1–6 (1995).
17. L. Muller, F. Chavane, J. Reynolds, T. J. Sejnowski, Cortical travelling waves: Mechanisms and computational principles. *Nat. Rev. Neurosci.* **19**, 255–268 (2018).
18. Z. W. Davis, L. Muller, J. Martinez-Trujillo, T. Sejnowski, J. H. Reynolds, Spontaneous travelling cortical waves gate perception in behaving primates. *Nature* **587**, 432–436 (2020).
19. E. Marder, D. Bucher, Central pattern generators and the control of rhythmic movements. *Curr. Biol.* **11**, R986–R996 (2001).
20. W. Singer, Neuronal synchrony: A versatile code for the definition of relations? *Neuron* **24**, 49–65 (1999).
21. M. N. Shadlen, J. Movshon, Synchrony unbound: A critical evaluation of the temporal binding hypothesis. *Neuron* **24**, 67–77 (1999).
22. S. Ray, J. H. R. Maunsell, Differences in gamma frequencies across visual cortex restrict their possible use in computation. *Neuron* **67**, 885–896 (2010).
23. A. L. Giraud, D. Poeppel, Cortical oscillations and speech processing: Emerging computational principles and operations. *Nat. Neurosci.* **15**, 511–517 (2012).
24. G. Thut, C. Miniussi, J. Gross, The functional importance of rhythmic activity in the brain. *Curr. Biol.* **22**, R658–R663 (2012).
25. S. Ray, J. H. Maunsell, Do gamma oscillations play a role in cerebral cortex? *Trends. Cogn. Sci.* **19**, 78–85 (2015).
26. O. Sporns, D. R. Chialvo, M. Kaiser, C. C. Hilgetag, Organization, development and function of complex brain networks. *Trends. Cogn. Sci.* **8**, 418–425 (2004).
27. G. Spyropoulos *et al.*, Spontaneous variability in gamma dynamics described by a damped harmonic oscillator driven by noise. *Nat. Commun.* **13**, 2019 (2022).
28. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, 2016).
29. P. V. Aceituno, G. Yan, Y. Y. Liu, Tailoring echo state networks for optimal learning. *iScience* **23**, 101440 (2020).
30. T. K. Rusch, S. Mishra, “Coupled oscillatory recurrent neural network (coRNN)” in *Proceedings of the ICLR 2021* (2021).
31. M. Goldmann, I. Fischer, C. R. Mirasso, M. C. Soriano, Exploiting oscillatory dynamics of delay systems for reservoir computing. *Chaos* **33**, 093139 (2023).
32. T. Kapoor *et al.*, “Neural oscillators for generalization of physics-informed machine learning” in *Proceedings of the AAAI 2024* (2024), vol. 38, pp. 13059–13067.
33. V. K. Jirsa, J. S. Kelso, Spatiotemporal pattern formation in neural systems with heterogeneous connection topologies. *Phys. Rev. E* **62**, 8462 (2000).
34. M. Breakspear, S. Heitmann, A. Daffertshofer, Generative models of cortical oscillations: Neurobiological implications of the Kuramoto model. *Front. Hum. Neurosci.* **4**, 190 (2010).
35. P. Ashwin, S. Coombes, R. Nicks, Mathematical frameworks for oscillatory network dynamics in neuroscience. *J. Math. Neurosci.* **6**, 1–92 (2016).
36. C. Bick, M. Goodfellow, C. R. Laing, E. A. Martens, Understanding the dynamics of biological and neural oscillator networks through exact mean-field reductions. *J. Math. Neurosci.* **10**, 9 (2020).
37. T. A. Keller, M. Welling, “Neural wave machines” in *Proceedings of the ICML 2023* (PMLR, 2023), vol. 202, pp. 16168–16189.
38. T. A. Keller, L. Muller, T. J. Sejnowski, M. Welling, “Traveling waves encode the recent past and enhance sequence learning” in *Proceedings of the ICLR, 2024* (2024).
39. G. Buzsáki, A. Draguhn, Neuronal oscillations in cortical networks. *Science* **304**, 1926–1929 (2004).
40. B. H. Jansen, V. G. Rit, Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.* **73**, 357–366 (1995).
41. E. Marder, R. L. Calabrese, Principles of rhythmic motor pattern generation. *Physiol. Rev.* **76**, 687–717 (1996).
42. C. M. Gray, D. A. McCormick, Chattering cells. *Science* **274**, 109–113 (1996).
43. I. Onorato *et al.*, A distinct class of bursting neurons with strong gamma synchronization and stimulus selectivity in Monkey V1. *Neuron* **105**, 180–197 (2020).
44. J. D. Cowan, J. Neuman, W. van Drongelen, Wilson–Cowan equations for neocortical dynamics. *J. Math. Neurosci.* **6**, 1 (2016).
45. R. J. Douglas, K. A. Martin, A functional microcircuit for cat visual cortex. *J. Physiol.* **440**, 735–769 (1991).
46. Y. LeCun, Y. Bengio, G. E. Hinton, Deep learning. *Nature* **521**, 436–444 (2015).
47. C. van Vreeswijk, H. Sompolinsky, Chaotic balanced state in a model of cortical circuits. *Neural Comput.* **10**, 1321–1371 (1998).
48. I. Dubinin, F. Effenberger, Fading memory as inductive bias in residual recurrent networks. *Neural Networks* **173**, 106179 (2024).
49. Y. Wang *et al.*, Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat. Neurosci.* **9**, 534–542 (2006).
50. J. D. Murray *et al.*, A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
51. N. Perez-Nieves, V. C. H. Leung, P. L. Dragotti, D. F. M. Goodman, Neural heterogeneity promotes robust learning. *Nat. Commun.* **12**, 5791 (2021).
52. F. Zeldenrust, B. S. Gutkin, S. Denève, Efficient and robust coding in heterogeneous recurrent networks. *PLoS Comput. Biol.* **17**, e1008673 (2021).
53. F. Sánchez-Puig, O. Zapata, O. K. Pineda, G. Iñiguez, C. Gershenson, Heterogeneity extends criticality. *Front. Compl. Syst.* **1**, 1111486 (2023).
54. R. Legenstein, W. Maass, Edge of chaos and prediction of computational performance for neural circuit models. *Neural Networks* **20**, 323–334 (2007).
55. A. Ceni *et al.*, Random oscillators network for time series processing. *PMLR* **238**, 4807–4815 (2024).
56. D. Ferster, S. Lindström, An intracellular analysis of geniculo-cortical connectivity in area 17 of the cat. *J. Physiol.* **342**, 181–215 (1983).
57. N. T. Markov *et al.*, Cortical high-density counterstream architectures. *Science* **342**, 1238406 (2013).
58. D. Nikolić, S. Häusser, W. Singer, W. Maass, Distributed fading memory for stimulus properties in the primary visual cortex. *PLoS Biol.* **7**, e1000260 (2009).
59. D. H. Hubel, T. N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *J. Physiol.* **160**, 106–154 (1962).
60. C. Blakemore, J. Nachmias, P. Sutton, The perceived spatial frequency shift: Evidence for frequency-selective neurones in the human brain. *J. Physiol.* **210**, 727–750 (1970).
61. S. V. David, Natural stimulus statistics alter the receptive field structure of V1 neurons. *J. Neurosci.* **24**, 6991–7006 (2004).
62. A. Das, C. D. Gilbert, Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature* **399**, 655–661 (1999).
63. M. M. Churchland *et al.*, Stimulus onset quenches neural variability: A widespread cortical phenomenon. *Nat. Neurosci.* **13**, 369–378 (2010).
64. P. Berkes, G. Orbán, M. Lengyel, J. Fiser, Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* **331**, 83–87 (2011).
65. A. Lazar, C. Lewis, P. Fries, W. Singer, D. Nikolic, Visual exposure enhances stimulus encoding and persistence in primary cortex. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2105276118 (2021).
66. J. R. King, V. Wyart, The human brain encodes a chronicle of visual events at each instant of time through the multiplexing of traveling waves. *J. Neurosci.* **41**, 7224–7233 (2021).
67. S. Higuchi, S. Kairat, S. M. Bohtë, S. Otte, “Balanced resonate-and-fire neurons” in *Proceedings of the ICML 2024* (2024).
68. M. Baronig, R. Ferrand, S. Sabathiel, R. Legenstein, Advancing spatio-temporal processing in spiking neural networks through adaptation. *arXiv [Preprint]* (2024). <https://arxiv.org/abs/2408.07517> (Accessed 20 August 2024).
69. S. Heitmann, T. W. Boonstra, M. Breakspear, A dendritic mechanism for decoding traveling waves: Principles and applications to motor cortex. *PLoS Comput. Biol.* **9**, e1003260 (2013).
70. S. Coombes, Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**, 91–108 (2005).
71. J. C. Pang *et al.*, Geometric constraints on human brain function. *Nature* **618**, 566–574 (2023).
72. K. Mahjoory, J. M. Schoffelen, A. Keitel, J. Gross, The frequency gradient of human resting-state brain oscillations follows cortical hierarchies. *eLife* **9**, e53715 (2020).
73. F. Zeldenrust, B. S. Gutkin, S. Denève, Efficient and robust coding in heterogeneous recurrent networks. *PLoS Comput. Biol.* **17**, e1008673 (2021).
74. J. M. Beggs, D. Plenz, Neuronal avalanches in neocortical circuits. *J. Neurosci.* **23**, 11167–11177 (2003).
75. D. Koch, A. Nandan, G. Ramesan, A. Koseska, Biological computations: Limitations of attractor-based formalisms and the need for transients. *Biochem. Biophys. Res. Commun.* **720**, 150069 (2024).
76. M. Khona, I. R. Fiete, Attractor and integrator networks in the brain. *Nat. Rev. Neurosci.* **23**, 744–766 (2022).
77. J. R. Tredicce *et al.*, Critical slowing down at a bifurcation. *Am. J. Phys.* **72**, 799–809 (2004).
78. G. Deco, V. K. Jirsa, Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors. *J. Neurosci.* **32**, 3366–3375 (2012).
79. A. Palmigiano, T. Geisel, F. Wolf, D. Battaglia, Flexible information routing by transient synchrony. *Nat. Neurosci.* **20**, 1014–1022 (2017).
80. H. Luhmann, W. Singer, L. Martínez-Millán, Horizontal interactions in cat striate cortex: I. Anatomical substrate and postnatal development. *Eur. J. Neurosci.* **2**, 344–357 (1990).
81. L. H. Liboni *et al.*, Image segmentation with traveling waves in an exactly solvable recurrent neural network. *arXiv [Preprint]* (2023). <https://arxiv.org/abs/2311.16943> (Accessed 1 December 2023).
82. F. Effenberger, HORN model. GitHub. <https://github.com/exilef/horn>. Accessed 7 January 2025.