scientific reports

Check for updates

OPEN

Critically synchronized brain waves form an effective, robust and flexible basis for human memory and learning

Vitaly L. Galinsky[1]✉ & Lawrence R. Frank[1,2]

The effectiveness, robustness, and flexibility of memory and learning constitute the very essence of human natural intelligence, cognition, and consciousness. However, currently accepted views on these subjects have, to date, been put forth without any basis on a true physical theory of how the brain communicates internally via its electrical signals. This lack of a solid theoretical framework has implications not only for our understanding of how the brain works, but also for wide range of computational models developed from the standard orthodox view of brain neuronal organization and brain network derived functioning based on the Hodgkin–Huxley ad-hoc circuit analogies that have produced a multitude of Artificial, Recurrent, Convolution, Spiking, etc., Neural Networks (ARCSe NNs) that have in turn led to the standard algorithms that form the basis of artificial intelligence (AI) and machine learning (ML) methods. Our hypothesis, based upon our recently developed physical model of weakly evanescent brain wave propagation (WETCOW) is that, contrary to the current orthodox model that brain neurons just integrate and fire under accompaniment of slow leaking, they can instead perform much more sophisticated tasks of efficient coherent synchronization/desynchronization guided by the collective influence of propagating nonlinear near critical brain waves, the waves that currently assumed to be nothing but inconsequential subthreshold noise. In this paper we highlight the learning and memory capabilities of our WETCOW framework and then apply it to the specific application of AI/ML and Neural Networks. We demonstrate that the learning inspired by these critically synchronized brain waves is shallow, yet its timing and accuracy outperforms deep ARCSe counterparts on standard test datasets. These results have implications for both our understanding of brain function and for the wide range of AI/ML applications.

The mechanisms of human memory remains one of the great unsolved mysteries in modern science. As a critical component of human learning, the lack of a coherent theory of memory has far-reaching implications for our understanding of cognition as well. Recent advances in experimental neuroscience and neuroimaging have highlighted the importance of considering the interactions of the wide-range of spatial and temporal scales at play in brain function, from the microscales of subcellular dendrites, synapses, axons, somata, to the mesoscales of the interacting networks of neural circuitry, the macroscales of brain-wide circuits. Current theories derived from these experimental data suggest that ability of humans to learn and adapt to ever-changing external stimuli is predicated on the development of complex, adaptable, efficient, and robust circuits, networks, and architectures derived from a flexible arrangements among the variety of neuronal and non-neuronal cell types in the brain. A viable theory of memory and learning must therefore be predicated on a physical model capable of producing multiscale spatiotemporal phenomena consistent with observed data.

At the heart of all current models for brain electrical activity is the neuron spiking model formulated by Hodgkin and Huxley (HH)[1] that has provided quantitative descriptions of Na^+/K^+ fluxes, voltage- and time-dependent conductance changes, the waveforms of action potentials, and the conduction of action potentials along nerve fibers[2]. Unfortunately, although the HH model has been useful in fitting multiparametric set of equations to local membrane measurements, the model has been of limited utility in deciphering complex functions arising in interconnected networks of brain neurons[3]. From a practical standpoint, the original HH

[1]Center for Scientific Computation in Imaging, University of California at San Diego, La Jolla, CA 92037-0854, USA. [2]Center for Functional MRI, University of California at San Diego, La Jolla, CA 92037-0677, USA. ✉email: vit@ucsd.edu

nature portfolio

1

model is too complicated to to describe even relatively small networks[4–6]. This has resulted in the development of optimization techniques[7–10] based on a much reduced model of a leaky integrate-and-fire (LIF) neuron that is simple enough for use in neural networks, as it replaces all these multiple gates, currents, channels and thresholds with just a single threshold and time constant. A majority of spiking neural network (SNN) models use this simplistic LIF neuron for the so called "deep learning"[11–14] claiming that this is inspired by brain functioning. While multiple LIF models are used for image classification on large datasets[15–19], most applications of SNNs are still limited to less complex datasets, due to the complex dynamics of even the oversimplified LIF model and non-differentiable operations of LIF spiking neurons. Some remarkable studies have applied SNNs for object detection tasks[20–22]. Spike based methods were also used for object tracking[23–26]. A research is booming in using LIF spiking networks for online learning[27], braille letter reading[28], different neuromorphic synaptic devices[29] for detection and classification of biological problems[30–36]. Significant research is focused on making human-level control[37], optimizing back-propagation algorithms for spiking networks[38–40], as well as penetrating much deeper into ARCSes core[41–44] with smaller number of time steps[41], using an event-driven paradigm[36, 40, 45, 46], applying batch normalization[47], scatter-and-gather optimizations[48], supervised plasticity[49], time-step binary maps[50], and using transfer learning algorithms[51]. In concert with this broad range of software applications, there is a huge amount of research directed at developing and using these LIF SNN in embedded applications with the help of the *neuromorphic hardware*[52–57], the generic name given to hardware that is nominally based on, or inspired by, the structure and function of the human brain. However, while the LIF model is widely accepted and ubiquitous in neuroscience, it is nevertheless problematic in that it does not generate any spikes per se.

A single LIF neuron can formally be described in differential form as

$$\tau_m \frac{\partial U}{\partial t} = -(U - U_{rest}) + RI, \tag{1}$$

where $U(t)$ is the membrane potential, U_{rest} is the resting potential, τ_m is the membrane time constant, R is the input resistance, and $I(t)$ is the input current[58]. It is important to note that Eq. (1) does *not* describe actual spiking. Rather, it integrates the input current $I(t)$ in the presence of an input membrane voltage $U(t)$. In the absence of the current $I(t)$, the membrane voltage rapidly (exponentially) decays with time constant τ_m to its resting potential U_{rest}. In this sense the integration is "leaky". There is no structure in this equation that even approximates a system resonance that might be described as "spiking". Moreover, both the decay constant τ_m and the resting potential U_{rest} are not only unknowns, but assumed constant, and therefore significant oversimplifications of the actual complex tissue environment.

It is a curious development in the history of neuroscience that the mismatch between the observed spiking behavior of neurons and a model of the system that is incapable of producing spiking was met not with a reformulation to a more physically realistic model, but instead with what can only be described as an ad-hoc patchwork fix: the introduction of a "firing threshold" Θ that defines when a neuron finally stops integrating the input, resulting in a large action potential almost magically shared with its neighboring neurons, after which the membrane voltage U is reset by hand back to the resting potential U_{rest}. Adding these conditions results in (1) being only capable of describing the dynamics that happen when the membrane potential U is below this spiking ruler threshold Θ. It is important to recognize that this description of the "sub-threshold" dynamics of the membrane potential until it has reached its firing threshold describes a situation where neighboring neurons are *not effected* by what is essentially a description of sub-threshold noise.

In short, the physical situation described by (1) is contradictory to many careful neuroscience experiments that show, for example, that (1) the neuron is anisotropically activated following the origin of the arriving signals to the membrane; (2) a single neuron's spike waveform typically varies as a function of the stimulation location; (3) spatial summation is absent for extracellular stimulations from different directions; (4) spatial summation and subtraction are not achieved when combining intra- and extra-cellular stimulations, as well as for nonlocal time interference[59]. Such observation have lead to calls "to re-examine neuronal functionalities beyond the traditional framework"[59].

Such a re-examination has been underway in our lab, where we have developed a physics based model of brain electrical activity. We have demonstrated that in the inhomogeneous anisotropic brain tissue system, the underlying dynamics is not necessarily restricted by reaction–diffusion type only. The recently developed theory of weakly evanescent brain waves (WETCOW)[60–62] shows from a physical point of view that propagation of electromagnetic fields through the highly complex geometry of inhomogeneous and anisotropic domain of real brain tissues can also happen in a wave-like form . This wave-like propagation agrees well with the results of the above neuronal experiments[59] as well as in general explains the broad range of observed seemingly disparate brain spatiotemporal characteristics. The theory produces a set of nonlinear equations for both the temporal and spatial evolution of brain wave modes that include all possible nonlinear interaction between propagating modes at multiple spatial and temporal scales and degrees of nonlinearity. The theory bridges the gap between the two seemingly unrelated spiking and wave 'camps' as the generated wave dynamics includes the complete spectra of brain activity ranging from incoherent asynchronous spatial or temporal spiking events, to coherent wave-like propagating modes in either temporal or spatial domains, to collectively synchronized spiking of multiple temporal or spatial modes.

In this paper we highlight some particular aspects of the WETCOW theory directly related to biological learning through wave dynamics, and demonstrate how these principles can not only augment our understanding of cognition, but provide the basis for a novel class of engineering analogs for both software and hardware learning systems that can operate with the extreme energy and data efficiency characteristics of biological systems that facilitate adaptive resilience in dynamic environments.

We would like to emphasize that a major motivation for our work is the recognition that there has been a rapidly growing focus in the research community in recent years on theories of memory, learning, and consciousness rely on networks of HH (LIF) neurons as biological and/or physical basis[63]. Every single neuron in this case is assumed to be an element (or a node) with fixed properties that isotropically collects input and fires when enough has been collected. The learning algorithms are then discussed as processes that update network properties, e.g., connection strength between those fixed nodes through plasticity[64], or number of participating fixed neuron nodes in the network through birth and recruitment of new neuron nodes[65], etc. In our paper, we focus on different aspect of network functioning—we assume that network is formed not by fixed nodes (neurons) but by a flexible pathways encompassing propagating waves, or wave packets, or wave modes. Formally those wave modes play in any network of wave modes the same role as single HH (LIF) node in network of neurons, therefore, we often interchangeably use and substitute 'wave mode' for 'network node'. But, as any single neuron may encounter multiple wave modes arriving from any other neuron, and synchronization with or without spiking will manifest as something that looks like anisotropic activation depending on the origin of the arriving signals[59], our wave network paradigm is capable of characterizing much more complex and subtle coherent brain activity and thus shows more feature-rich possibilities for "learning" and memory formation.

The test examples based on our WETCOW inspired algorithms show excellent performance and accuracy and can be expected to be resilient to catastrophic forgetting, will demonstrate real-time sensing, learning, decision making, and prediction. Due to very efficient, fast, robust and very precise spike synchronization, the WETCOW based algorithms are able to respond to novel, uncertain, and rapidly changing conditions in real-time, and will enable appropriate decisions based on small amounts of data over short time horizons. These algorithms can include uncertainty quantification for data of high sparsity, large size, mixed modalities, and diverse distributions, and will be pushing the bounds on out-of-distribution generalization.

The WETCOW model is supposed to capture several different memory phenomena. In a non-mathematical way it can be described as

- Critical encoding–the WETCOW model shows how independent oscillators in a heterogeneous network with different parameters form a coherent state (memory) as a result of critical synchronization.
- Synchronization speed—the WETCOW model shows that due to coordinated work of amplitude-phase coupling this synchronization process is significantly more fast than memory formation in the spiking network of integrate-and-fire neurons.
- Storage capacity—the WETCOW model shows that a coherent memory state with predicted encoding parameters can be formed with as low as two nodes, thus potentially allows for significant increase of memory capacity comparing to the traditional spiking paradigm.
- Learning efficiency—the WETCOW model shows that processing of a new information by a mean of synchronization of network parameters in a near critical range allows a natural development of continuous learner-type memory representative of human knowledge processing.
- Memory robustness—the WETCOW model shows that memory state formed in non-planar critically synchronized network potentially more stable, continuous learning prone, and resilient to catastrophic forgetting.

The experiments presented below are clearly beyond human training capacity, but nevertheless represent a very good set of preliminary stress tests to provide support for all our claims, from Critical Encoding to Memory Robustness. The demonstration that back propagation step is not generally required for very good performance of continues learning as well as small footprint of nodes and input data involved in memory formation are representative of human few-shot learning[66–69] as well as relevant to the larger issues introduced in our paper.

Weakly evanescent brain waves

A set of derivations that lead to the WETCOW description was presented in details in[60–62] and is based on considerations that follow from the most general form of brain electromagnetic activity expressed by Maxwell equations in inhomogeneous and anisotropic medium[70–72]

$$\nabla \cdot \boldsymbol{D} = \rho, \quad \nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \quad \Rightarrow \quad \frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{J} = 0.$$

Using the electrostatic potential $\boldsymbol{E} = -\nabla \Psi$, Ohm's law $\boldsymbol{J} = \boldsymbol{\sigma} \cdot \boldsymbol{E}$ (where $\boldsymbol{\sigma} \equiv \{\sigma_{ij}\}$ is an anisotropic conductivity tensor), a linear electrostatic property for brain tissue $\boldsymbol{D} = \varepsilon \boldsymbol{E}$, assuming that the scalar permittivity ε is a "good" function (i.e. it does not go to zero or infinity everywhere) and taking the change of variables $\partial x \rightarrow \varepsilon \partial x'$, the charge continuity equation for the spatial–temporal evolution of the potential Ψ can be written in terms of a permittivity scaled conductivity tensor $\boldsymbol{\Sigma} = \{\sigma_{ij}/\varepsilon\}$ as

$$\frac{\partial}{\partial t}\left(\nabla^2 \Psi\right) = -\boldsymbol{\nabla} \cdot \boldsymbol{\Sigma} \cdot \nabla \Psi + \mathcal{F}, \tag{2}$$

where we have included a possible external source (or forcing) term $\mathcal{F}$. For brain fiber tissues the conductivity tensor $\boldsymbol{\Sigma}$ might have significantly larger values along the fiber direction than across them. The charge continuity without forcing i.e., ($\mathcal{F} = 0$) can be written in tensor notation as

$$\partial_t \partial_i^2 \Psi + \Sigma_{ij} \partial_i \partial_j \Psi + \left(\partial_i \Sigma_{ij}\right)\left(\partial_j \Psi\right) = 0, \tag{3}$$

where repeating indices denote summation. Simple linear wave analysis, i.e. substitution of $\Psi \sim \exp{(-i(\boldsymbol{k} \cdot \boldsymbol{r} - \Omega t))}$, where $\boldsymbol{k}$ is the wavenumber, $\boldsymbol{r}$ is the coordinate, Ω is the frequency and t is the time, gives the following complex dispersion relation:

$$D(\Omega, \boldsymbol{k}) = -i\Omega k_i^2 - \Sigma_{ij} k_i k_j - i\partial_i \Sigma_{ij} k_j = 0, \tag{4}$$

which is composed of the real and imaginary components:

$$\gamma \equiv \Im\Omega = \Sigma_{ij} \frac{k_i k_j}{k^2} \qquad \omega \equiv \Re\Omega = -\frac{\partial_i \Sigma_{ij} k_j}{k^2} \tag{5}$$

Although in this general form the electrostatic potential Ψ, as well as the dispersion relation $D(\Omega, \boldsymbol{k})$, describe three dimensional wave propagation, we have shown[60, 61] that in anisotropic and inhomogeneous media some directions of wave propagation are more equal than others with preferred directions determined by the complex interplay of the anisotropy tensor and the inhomogeneity gradient. While this is of significant practical importance, in particular because the anisotropy and inhomogeneity can be directly estimated from non-invasive methods, for the sake of clarity we focus here on the one dimensional scalar expressions for spatial variables x and k that can be easily generalized for the multi dimensional wave propagation as well.

Based on our nonlinear Hamiltonian formulation of the WETCOW theory[62], there exists an anharmonic wave mode

$$H^s(a, a^\dagger) = \Gamma aa^\dagger + aa^\dagger \left[\beta_a a + \beta_{a^\dagger} a^\dagger - 2\alpha \left(aa^\dagger \right)^{1/2} \right] \tag{6}$$

where a is a complex wave amplitude and $a^\dagger$ is its conjugate. The amplitude a denotes either temporal $a_k(t)$ or spatial $a_\omega(x)$ wave mode amplitudes that are related to the spatiotemporal wave field $\Psi(x, t)$ through a Fourier integral expansions

$$a_k(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi(x, t) e^{-i(kx + \omega_k t)} dx, \tag{7}$$

$$a_\omega(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi(x, t) e^{-i(k_\omega x + \omega t)} dt, \tag{8}$$

where for the sake of clarity we use one dimensional scalar expressions for spatial variables x and k, but it can be easily generalized for the multi dimensional wave propagation as well. The frequency ω and the wave number k of the wave modes satisfy the dispersion relation $D(\omega, k) = 0$, and ω_k and k_ω denote the frequency and the wave number roots of the dispersion relation (the structure of the dispersion relation and its connection to the brain tissue properties has been discussed in[60]).

The first term $\Gamma aa^\dagger$ in (6) denotes the harmonic (quadratic) part of the Hamiltonian with either the complex valued frequency $\Gamma = i\omega + \gamma$ or wave number $\Gamma = ik + \lambda$ that both include a pure oscillatory parts (ω or k) and possible weakly excitation or damping rates, either temporal γ or spatial λ. The second anharmonic term is cubic in the lowest order of nonlinearity and describes the interactions between various propagating and nonpropagating wave modes, where α, β_a and $\beta_{a^\dagger}$ are the complex valued strengths of those different nonlinear processes. This theory can be extended to a network of interacting wave modes of the form (6) which can be described by a network Hamiltonian form that describes discrete spectrum of those multiple wave modes as[62]

$$H(\boldsymbol{a}, \boldsymbol{a}^\dagger) = \sum_n \left[H^s(a_n, a_n^\dagger) + \sum_{m \neq n} \left(a_n r_{nm} a_m^\dagger + a_n^\dagger r_{nm}^* a_m \right) \right] \tag{9}$$

where the single mode amplitude a_n again denotes either a_k or a_ω, $\boldsymbol{a} \equiv \{a_n\}$ and $r_{nm} = w_{nm} e^{i\delta_{nm}}$ is the complex network adjacency matrix with w_{nm} providing the coupling power and δ_{nm} taking into account any possible differences in phase between network nodes. This description includes both amplitude $\Re(a)$ and phase $\Im(a)$ mode coupling and as shown in[62] allows for significantly unique synchronization behavior different from both phase coupled Kuramoto oscillator networks[73–75] and from networks of amplitude coupled integrate-and-fire neuronal units[58, 76, 77].

An equation for the nonlinear oscillatory amplitude a then can be expressed as a derivative of the Hamiltonian form

$$\frac{da}{dt} = \frac{\partial H^s}{\partial a^\dagger} \equiv \Gamma a + \beta_{a^\dagger} aa^\dagger + \beta_a a^2 - \alpha a(aa^\dagger)^{1/2}, \tag{10}$$

after removing the constants with a substitution of $\beta_{a^\dagger} = 1/2\tilde{\beta}_{a^\dagger}$ and $\alpha = 1/3\tilde{\alpha}$ and dropping the tilde. We note that although (10) is an equation for the temporal evolution, the spatial evolution of the mode amplitudes $a_\omega(x)$ can be described by a similar equation substituting temporal variables by their spatial counterparts, i.e., $(t, \omega, \gamma) \rightarrow (x, k, \lambda)$.

Splitting (10) into an amplitude/phase pair of equations using $a = Ae^{i\phi}$ and making some rearrangements these equations can be rewritten as[70–72]

$$\frac{dA}{d\tau} = \gamma A + A^2 \left[w^a \cos(\phi - \psi) - \alpha \right], \tag{11}$$

$$\frac{d\phi}{d\tau} = \omega + A w^\phi \cos\phi, \tag{12}$$

where ψ, w^a and w^ϕ some model constants.

Single mode firing rate

The effective period of spiking T_s (or its inverse—either the firing rate $1/T_s$ or the effective firing frequency $\omega_s = 2\pi/T_s$) was estimated in[70–72] as

$$T_s = \frac{2\pi}{\omega\sqrt{1 - \gamma^2/\gamma_c^2}} = \frac{2\pi}{\omega\sqrt{1 - \omega_c^2/\omega^2}}, \tag{13}$$

$$\omega_s = \omega\sqrt{1 - \gamma^2/\gamma_c^2} = \sqrt{\omega^2 - \omega_c^2}, \tag{14}$$

where the critical frequency ω_c or the critical growth rate γ_c can be expressed as

$$\omega_c = \gamma w, \qquad \gamma_c = \frac{\omega}{w}, \tag{15}$$

where

$$w = \frac{w^\phi \cos\phi^c}{\alpha + w^a \cos(\phi^c + \psi)}, \tag{16}$$

and

$$\phi^c = \arctan\left[\frac{w^a \sin\psi}{\sqrt{\alpha^2 - (w^a \sin\psi)^2}} \right]. \tag{17}$$

Figure 1 compares the single node results (13) to (15) with peak-to-peak period/frequency estimates from direct simulations of the system (11) to (12). Several inserts show shapes of numerical solution generated at the correspondent level of criticality c_r

$$c_r = \frac{\gamma}{\gamma_c} = w\frac{\gamma}{\omega}. \tag{18}$$
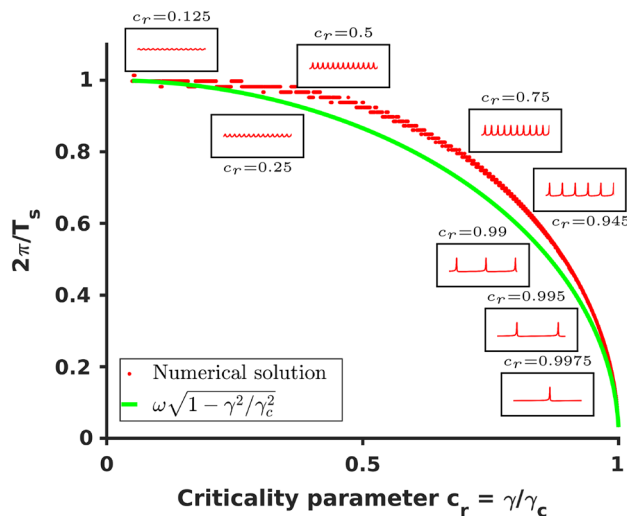
Figure 1. Plot of the analytical expression (14) for the effective spiking frequency $\omega_s = 2\pi/T_s$ (green) and the frequency estimated from numerical solutions of (11) and (12) (red) with several inserts showing the numerical solution with indicated value of the criticality parameter $c_r = \gamma/\gamma_c$ (detailed plots of numerical solutions used for generating inserts are included in "Appendix"). In the numerical solution only γ was varied and the remaining parameters were the same as parameters reported in[62].

The above analytically derived single node results (13) to (15) can be directly used to estimate firing of interconnected networks as they express the rate of spiking as a function of a distance from criticality, and the criticality value can be in turn expressed through other system parameters.

A set of coupled equations for a network of multiple modes can be derived similarly to single mode set (11) and (12) by taking a derivative of the network Hamiltonian form (9) and appropriately changing variables. That gives for the amplitude A_i and the phase ϕ_i a set of coupled equations

$$\frac{dA_i}{dt} = \gamma_i A_i + A_i^2 \big(w_i^a \cos(\phi_i - \psi_i) - \alpha_i \big) + \sum_{j \neq i} w_{ij} A_j \cos(\phi_j - \phi_i - \delta_{ij}), \tag{19}$$

$$A_i \frac{d\phi_i}{dt} = \omega_i A_i + A_i^2 w_i^\phi \cos \phi_i + \sum_{j \neq i} w_{ij} A_j \sin(\phi_j - \phi_i - \delta_{ij}). \tag{20}$$

In the small (and constant) amplitude limit ($A_i = $ const) this set of equations turns into a set of phase coupled harmonic oscillators with a familiar $\sin(\phi_j - \phi_i \cdots)$ form of phase coupling. But in its general form (19) and (20) include also phase dependent coupling of amplitudes ($\cos(\phi_j - \phi_i \cdots)$) that dynamically defines if the input from j to i will either play excitatory ($|\phi_j - \phi_i + \cdots| < \pi/2$) or inhibitory ($|\phi_j - \phi_i + \cdots| > \pi/2$) roles (this is in addition to any phase shift introduced by the static network attributed phase delay factors δ_{ij}).

Synchronized network memory of a single node sensory response

Let us start with a single unconnected mode that is excited by a sensory input. Based on the strength of excitation the mode can be in any of the states shown in Fig. 1, with activity ranging from small amplitude oscillations in linear regime, to nonlinear anharmonic oscillations, to spiking with different rates (or effective frequencies) in sub-critical regime, to a single spike-like transition following by silence in supercritical range of excitation. The type of activity is determined by the criticality parameter $c_r = (\gamma_0 + \gamma_i)/\gamma_c$ where γ_c depends on the parameters of the system (15) and γ_0 determines the level of sensory input and γ_i is the level of background activation (either excitation or inhibition). Hence, for any arbitrary ith mode

$$\frac{dA_i}{dt} = (\gamma_0 + \gamma_i)A_i + A_i^2 \big(w_i^a \cos(\phi_i - \psi_i) - \alpha_i \big) \tag{21}$$

$$\frac{d\phi_i}{dt} = \omega_i + A_i w_i^\phi \cos \phi_i. \tag{22}$$

As a result, the mode i will show nonlinear oscillation with an effective frequency ω_s

$$\omega_s = \sqrt{\omega_i^2 - (\gamma_0 + \gamma_i)^2 w_i^2}, \tag{23}$$

$$w_i = \frac{w_i^\phi \cos \phi_i^c}{\alpha_i + w_i^a \cos(\phi_i^c + \psi_i)}, \tag{24}$$

$$\phi_i^c = \arctan \left[\frac{w_i^a \sin \psi_i}{\sqrt{\alpha_i^2 - (w_i^a \sin \psi_i)^2}} \right]. \tag{25}$$

Next we assume that instead of a single mode we have some network of modes described by (11) and (12) where the sensory excitation is absent ($\gamma_0 = 0$) and for simplicity we first assume that all the parameters (γ_i, ω_i, α_i, ψ_i, w_i^a, and w_i^ϕ) are the same for all modes and only the coupling parameters w_{ij} and δ_{ij} can vary. The mean excitation level for the network $\gamma_1 \equiv \gamma_i$ ($i = 1 \ldots N$) determines the type of activity the unconnected modes would be operating and it may be in any of the liner, nonlinear, sub-critical or supercritical range. Of course, the activity of individual nodes in network (11) and (12) depends on the details of coupling (parameters w_{ij} and δ_{ij}) and can be very complex. Nevertheless, at it was shown in[62], one of the features of the phase–amplitude coupled system (11) and (12), that distinguishes it both from networks of phase coupled Kuramoto oscillators and from networks of amplitude coupled integrate and fire neurons (or actually from any networks that are based on spike summation generated by neurons of Hodgkin–Huxley type or it's derivations), is that even for relatively weak coupling the synchronization of some modes in network (11) and (12) may happen in a very efficient manner. The conditions for coupling coefficients when this synchronized state is realized and every mode i of the network produce the same activity pattern as sensory excited single mode, but without any external excitation, can be expressed for every mode i as

$$\sum_j w_{ij} \cos(\delta_{ij}) = \gamma_0, \tag{26}$$

$$\sum_j w_{ij} \sin(\delta_{ij}) = 0. \tag{27}$$

This is necessary (but not sufficient) condition that shows that every recurrent path through the network, that is every brain wave loop that do not introduce nonzero phase delays, should generate the same level of amplitude excitation.

Even for this already oversimplified case of identical parameters, the currently agreed lines of research proceed with even more simplifications and either employ constant (small) amplitude phase synchronization approach (Kuramoto oscillators) assuming that all δ_{ij} equal to $-\pi/2$ or $\pi/2$, or use amplitude coupling (Hodgkin–Huxley neuron and the like) with δ_{ij} equal to 0 (excitatory) or π (inhibitory). Both of these cases are extremely limited and do not provide a framework for the effectiveness, flexibility, adaptability, and robustness characteristic of human brain functioning. The phase coupling is only capable of generating very slow and inefficient synchronization. The amplitude coupling is even less efficient as it completely ignores the details of the phase of the incoming signal, thus is only able to produce sporadic and inconsistent population level synchronization.

Of course, (26) and (27) are used as an idealized illustrative picture of critically synchronized memory state formation in phase–amplitude coupled network (11) and (12). In practice, in the brain the parameters of network (11) and (12), including frequencies, excitations, and other parameters of a single mode Hamiltonian (6), may be different between modes. But even in this case the formation of critically synchronized state follows the same outlined above procedure, and requires that for all modes total inputs to the phase and the amplitude parts ($\bar{\omega}_i$ and $\bar{\gamma}_i$) generate together the same effective frequency ω_s satisfying the relation

$$\omega_s = \sqrt{\bar{\omega}_i^2 - \bar{\gamma}_i^2 w_i^2}, \tag{28}$$

where

$$\bar{\gamma}_i = \gamma_i + \sum_j w_{ij} \cos(\delta_{ij}), \tag{29}$$

$$\bar{\omega}_i = \omega_i + \sum_j w_{ij} \sin(\delta_{ij}). \tag{30}$$

Overall, the critically synchronized memory can be formed by making a loop from as few as two modes. Of course, this may require too large an amount of amplitude coupling and will not produce the flexibility and robustness of multimode coupling with smaller steps of adjustment of amplitude–phase coupling parameters. Figures 2 and 3 show two examples of network synchronization with effective frequencies that replicate the original single mode effective frequency without sensory input. Ten modes were shown with the same parameters of $w_i^a = w_i^\phi = \sqrt{5}, \psi_i = 2\arctan(1/3), \phi_i^c = \arctan(1/2), w_i = 1/2$ but with a set of uniformly distributed frequencies ω_i, (with a mean of 1 and a standard deviation of 0.58–0.59). The network coupling w_{ij} and δ_{ij} were also selected from a range of values (from 0 to 0.2 for w_{ij} and $-\pi/2$ to $\pi/2$ for δ_{ij}).

Again, for phase only coupling (δ_{ij} equal to $-\pi/2$ or $\pi/2$) the synchronization is very inefficient and only happening as a result of an emergence of forced oscillations at common frequency in some parts of the network or in the whole network dependent on the details of the coupling parameters. The amplitude coupling of Hodgkin–Huxley and the like neurons is even less effective than phase-only coupling as it does not even consider the oscillatory and wave-like propagation nature of the subthreshold signals that contribute to the input and collective generation of spiking output. Therefore, expressions (28) to (30) are not applicable for HH and LIF models as phase information, as well as frequency dependence, is lost by those models and replaced by ad-hoc sets of thresholds and time constants.

Contrary to the lack of efficiency, flexibility, and robustness demonstrated by those state-of-the-art curtailed phase-only and amplitude-only approaches, the presented model of memory shows that when both phase and amplitude are operating together, a critical behavior emerging in the nonlinear system (9) gives birth to an efficient, flexible, and robust synchronization characteristic of human memory, appropriate for any type of coding, being it either rate or time.

Application to neural networks and machine learning

The presented critically synchronized memory model based on the theory of weakly evanescent brain waves—WETCOW[60–62]—has several very important properties. First of all, the presence of both amplitude w_{ij} and phase δ_{ij} coupling makes if possible to construct an effective and accurate recurrent networks that do not require extensive and time consuming training. The standard back propagation approach can be very expensive in terms of both computations, memory requirements, and large amount of communications involved, therefore may be poorly suited to the hardware constraints in computers and neuromorphic devices[78]. However, with the WETCOW based model it is easy to construct a small shallow network that will replicate the spiking produced by any input condition using the interplay of the amplitude–phase coupling (19) to (20) and the explicit analytical conditions for spiking rate (13) and (15) as a function of criticality. The shallow neural networks constructed using those analytical conditions give very accurate results with very little amount of training and very little memory requirements.

Comparison of a schematic diagram for the typical workflow of traditional multi-layer ARCSe neural network with a diagram for the critically synchronized WETCOW inspired shallow neural network is shown in Fig. 4. For the traditional multi-layer ARCSe neural network the diagram (Left panel) includes the connection weights
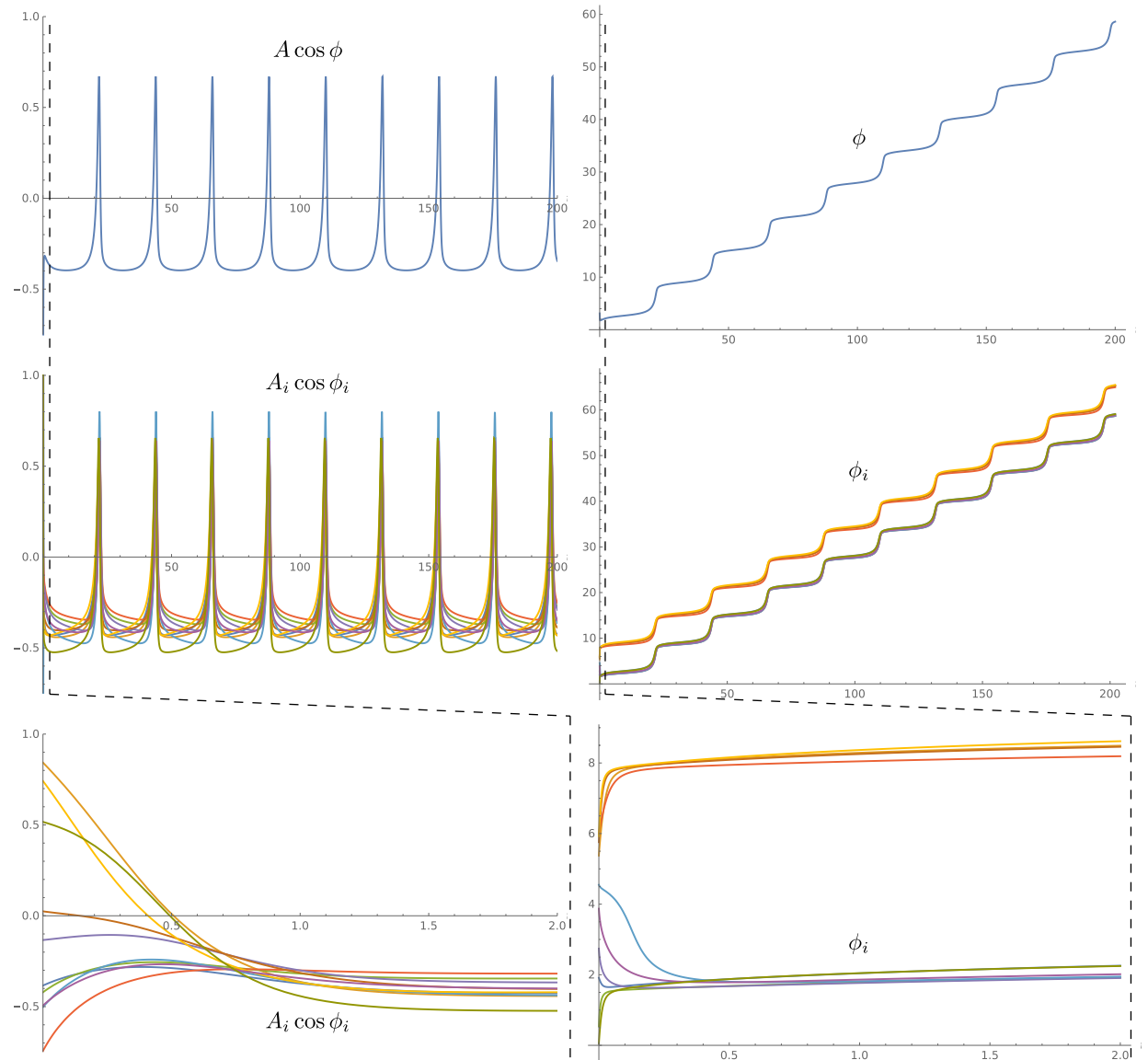
Figure 2. (Top) The amplitude and phase of a single mode subcritical spiking. (Middle) The spiking of multiple modes with different linear frequencies ω_i critically synchronized at the same effective spiking frequency (the units are arbitrary). The details of wavefront shapes for each mode are different, but the spiking synchronization between modes is very strong and precise. (Bottom) Expanded view of the initial part of the amplitude and phase of the mode shows the efficiency of synchronization—synchronization happens even faster than the single period of linear oscillations.

that are approximated by varying path width. The Schematic diagram for the critically synchronized WETCOW inspired shallow neural network in addition to amplitude weighting factors (that again shown by varying recurrent path widths) has an additional phase parameter that is shown by the non-planarity of the connection paths. The presence of non-planarity in a single layer (shallow) amplitude-phase synchronized neural network allows more efficient computations, memory requirements, and learning capabilities than multi-layer deep ARCSes of traditional AI/ML neural networks.

The non-planarity of the critically synchronized WETCOW inspired shallow neural network shown in Fig. 4 illustrates and emphasizes another important advantage in comparison to the traditional multi-layer ARCSe neural networks. It is well known that the traditional multi-layer deep learning ARCSe neural network models suffer from the phenomenon of catastrophic forgetting—a deep ARCSe neural network, carefully trained and back and forth massaged to perform some important task can unexpectedly lose its generalization ability on this task after additional training on a new task has been performed[79–81]. This typically happens because a new task overrides the previous weights that have been learned in the past. This means that continues learning degrades or even destroys the model performance for the previous tasks. This is a tremendous problem meaning that the traditional deep ARCSe neural network represents a very bad choice to function as a continuous learner, as it
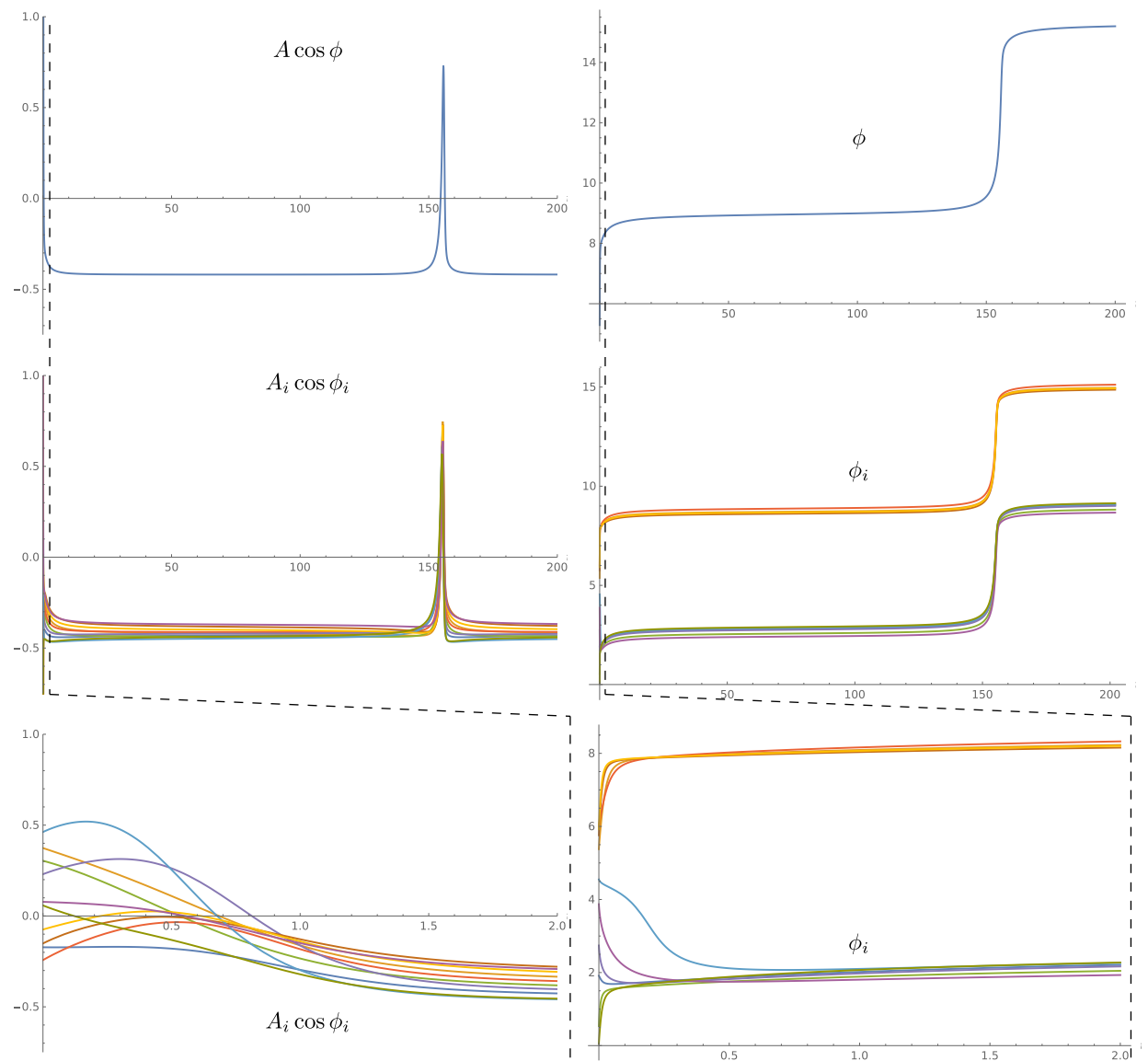
Figure 3. (Top) The amplitude and phase of a single mode spiking in a close to critical regime. (Middle) The spiking of multiple modes with different linear frequencies ω_i critically synchronized at the same effective spiking frequency that is close to critical frequency (the units are arbitrary). Similar to subcritical spiking in Fig. 2, the details of wavefront shapes for each mode are different, but the spiking synchronization between modes is very strong and precise. (Bottom) Expanded view of the initial part of the amplitude and phase of the mode shows the efficiency of synchronization–synchronization happens.

constantly forgets the previously learned knowledge being exposed to a bombardment of a new information. As any new information added to the traditional multi-layer deep learning ARCSe neural network inevitably modifies the network weights confined to the same plane and shared with all previously accumulated knowledge produced by a hard training work, this catastrophic forgetting phenomena is generally not a surprise. The non-planarity of the critically synchronized WETCOW inspired shallow neural network provides an additional way to encode new knowledge with a different out-of-plane phase-amplitude choice, thus preserving previous accumulated knowledge. This makes the critically synchronized WETCOW inspired shallow neural network model more suitable for the use in a continuous learning scenario.

Another important advantage of the WETCOW algorithms is their numerical stability, which makes them robust even in the face of extensive training. Because the system (19) and (20) describes the full range of dynamics, from linear oscillations to spiking in the perfectly differentiable form, it is perfectly differentiable. They thus are not subject to one of the major limitations of current standard models—the non-differentiability of the spiking nonlinearity for LIF (and similar) models, whose derivative is zero everywhere except at $U = \Theta$, and even at $U = \Theta$ the derivatives is not just large, but strictly speaking they are not defined.
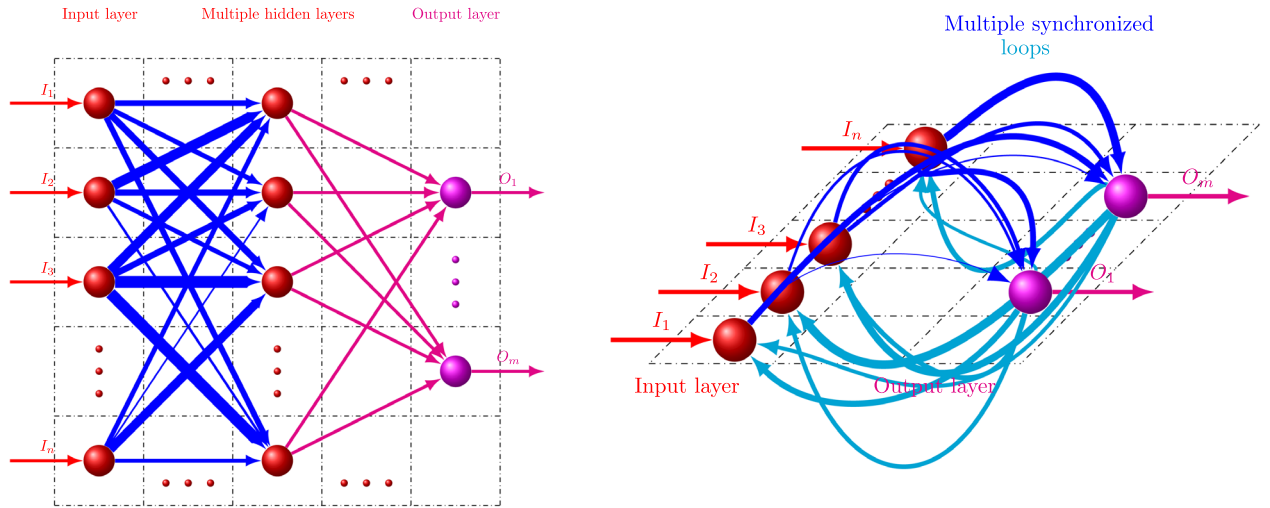
Figure 4. (Left) Schematic graph of typical multi-layer neural network where connection weights are shown by varying path width. (Right) Schematic graph of critically synchronized WETCOW inspired shallow neural network, where amplitude weighting factors are again shown by varying path widths, but an additional phase parameter controls the non-planar recurrent paths behavior. The interplay of amplitude-phase synchronization, shown by a non-planarity of a shallow—comprised of a single layer of synchronized loops—neural network, allows more efficient computations, memory requirements, and learning capabilities than multi-layer deep ARCSes of traditional AI/ML neural networks.

MNIST digits and MNIST fashion tests

The performance and accuracy of WETCOW based learning approaches is easily demonstrated on two commonly used databases: MNIST[82] and Fashion-MNIST[83]. Both the original handwritten digits MNIST database Fig. 5 (top) and an MNIST-like fashion product database—dataset of Zalando's article images designed as a direct drop-in replacement for the original MNIST dataset—Fig. 5 (bottom) contain 60,000 training images and 10,000 testing images. Each individual image is a 28×28 pixels grayscale image, associated with a single label from 10 different label classes.

The results for our WETCOW based model for a shallow recurrent neural network applied to the MNIST handwritten digits Fig. 5 (top) and MNIST fashion images Fig. 5 (bottom) are summarized in Table 1. In both cases the networks were generated for 7×7 downsampled images moved and rescaled to the common reference

Figure 5. (Top) Several example images from MNIST database of handwritten digits. (Bottom) Several example images from MNIST-like fashion product database of Zalando's article images designed as a direct drop-in replacement for the original MNIST dataset.

	Accuracy	Time	Others
MNISTDigits (without training)	0.9858–0.9883 (117–142 errors per 10,000 samples)	Several seconds	
MNISTDigits (with training)	0.9986 (14 errors per 10,000 samples)	Several minutes	0.88–0.998 14 h for 0.9977 accuracy
MNISTFashion (without training)	0.9385 (615 errors per 10,000 samples)	Several seconds	
MNISTFashion (with training)	0.9742 (258 errors per 10,000 samples)	Several minutes	0.444–0.897 From 1 to 50 h

Table 1. Summary of accuracy and timing results obtained by shallow learning processing of the original handwritten digits MNIST dataset[82] and the Fashion-MNIST dataset[83] using WETCOW inspired algorithm and test implementation based on the ideas of critically synchronized learning.

10

system. For each of the datasets, Table 1 shows two entries, the first corresponds to an initial construction of a recurrent network that involves just a single iteration, without any back propagation and retraining steps. In both cases this initial step produces very good initial accuracy, on par or even exceeding the final results of some of the deep ARCSes[84, 85]. The second entry for each dataset shows highest accuracy achieved and the corresponding training times. Both entries confirm that to achieve the accuracies that are higher than the accuracies obtained by any of the deep ARCSes orders of magnitude smaller training times are required.

Conclusion

This paper presents arguments and test results showing that recently developed physics based theory of wave propagation in the cortex—the theory of weakly evanescent brain waves—WETCOW[60–62]—provides both a theoretical and computational framework with which to better understand the adaptivity, flexibility, robustness, and effectiveness of human memory, and, hence, can be instrumental in development of novel learning algorithms. Those novel algorithms potentially allow the achievement of extreme data efficiency and adaptive resilience in dynamic environments, characteristic of biological organisms. The test examples based on our WETCOW inspired algorithms show excellent performance (orders of magnitude faster than current state-of-the-art deep ARCSe methods) and accuracy (exceeding the accuracy of current state-of-the-art deep ARCSe methods) and can be expected to be resilient to catastrophic forgetting, and will demonstrate real-time sensing, learning, decision making, and prediction. Due to very efficient, fast, robust and very precise spike synchronization, the WETCOW based algorithms are able to respond to novel, uncertain, and rapidly changing conditions in real-time, and will enable well-informed decisions based on small amounts of data over short time horizons. The WETCOW based algorithms can include uncertainty quantification for data of high sparsity, large size, mixed modalities, and diverse distributions, and will push the bounds on out-of-distribution generalization.

The paper presents ideas of how to extract principles, not available from current neural network approaches, by which biological learning occurs through wave dynamic processes arising in neuroanatomical structures, and in turn provides a new framework for the design and implementation of highly efficient and accurate engineering analogs of those processes and structures that could be instrumental in the design of novel learning circuits.

Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Appendix: Several examples of single mode solutions
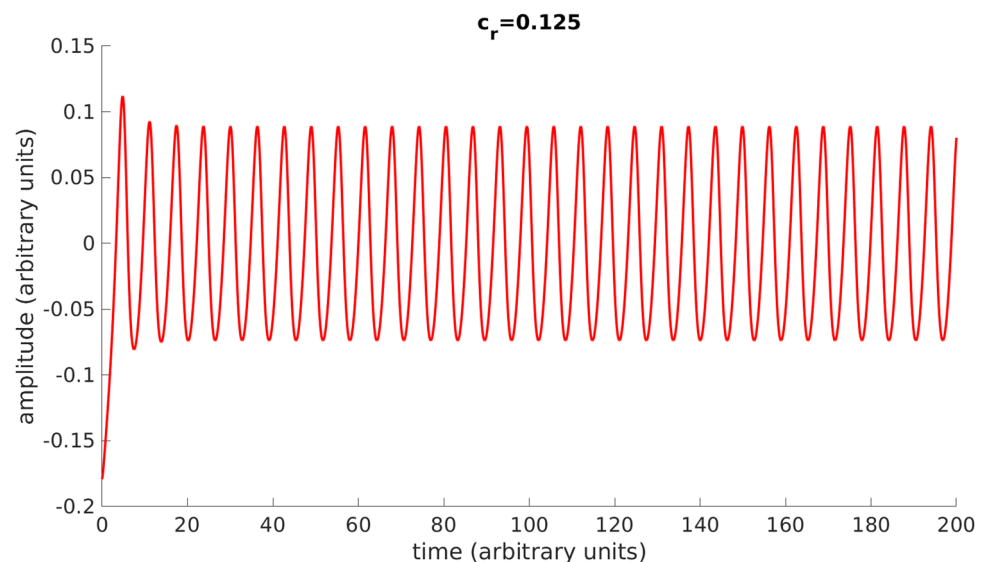
See Figs. 6, 7, 8, 9, 10, 11, 12 and 13.

Figure 6. The amplitude and phase of a single mode for $c_r = 0.125$.
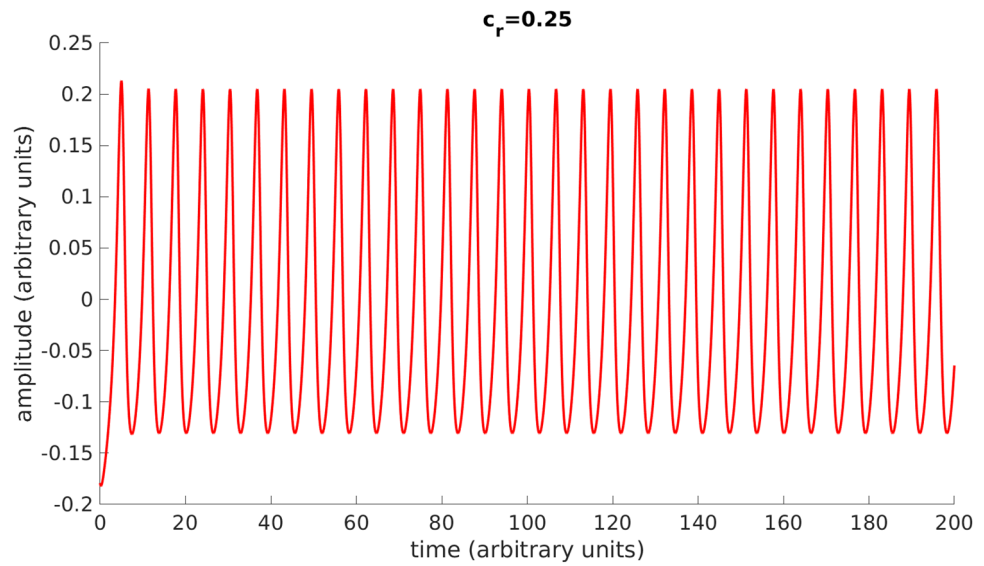
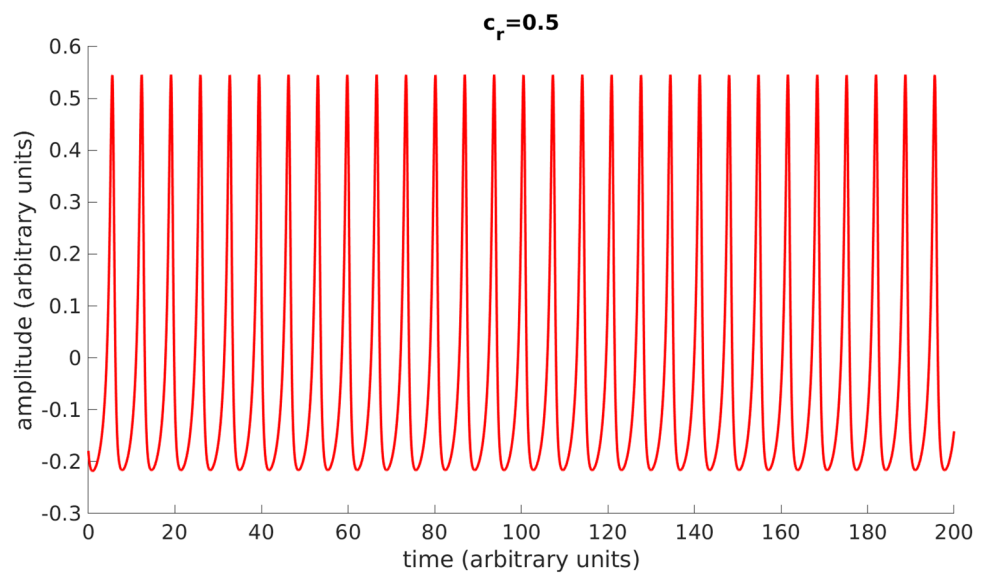**Figure 7.** The amplitude and phase of a single mode for $c_r = 0.25$.

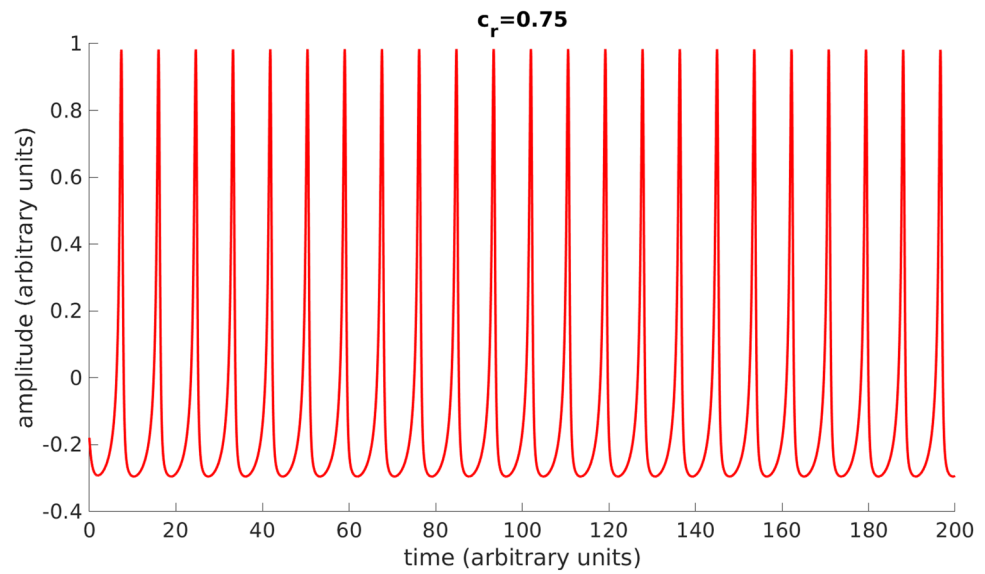**Figure 8.** The amplitude and phase of a single mode for $c_r = 0.5$.

c_r=0.75

Figure 9. The amplitude and phase of a single mode for $c_r = 0.75$.
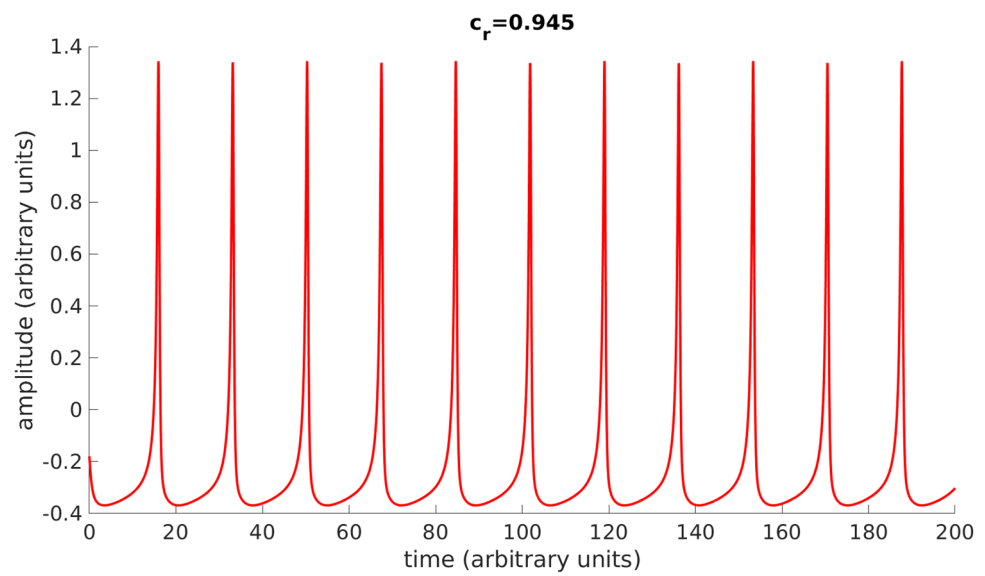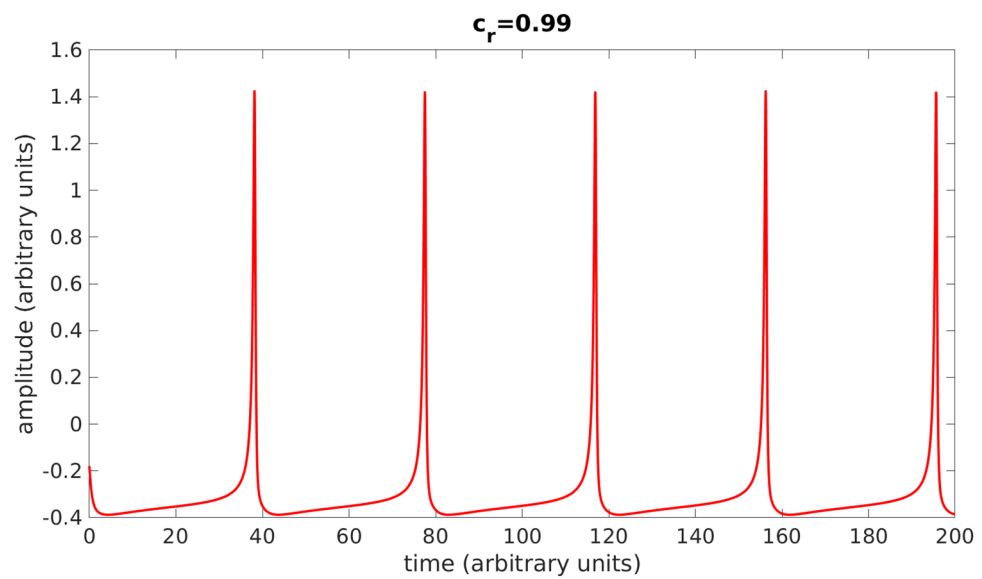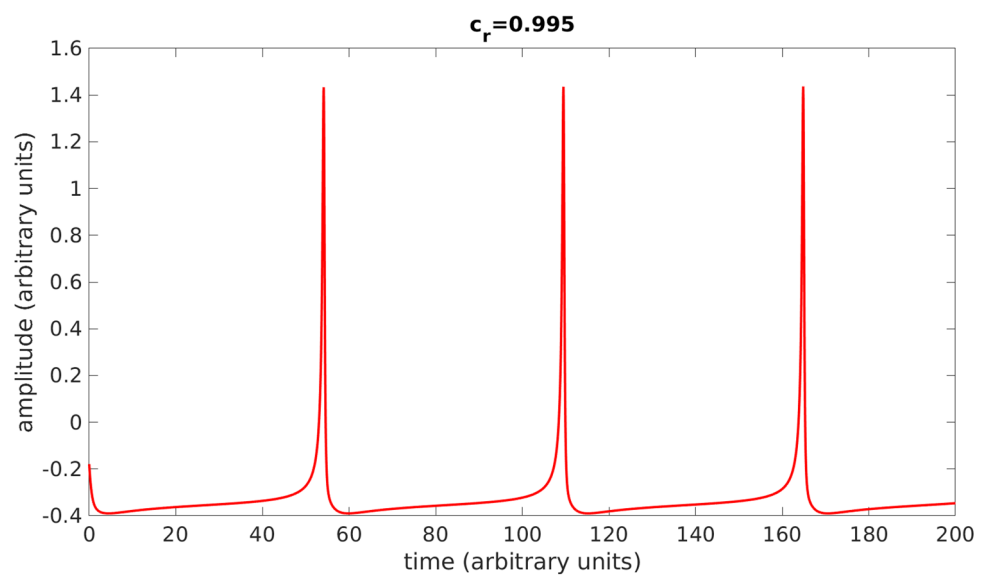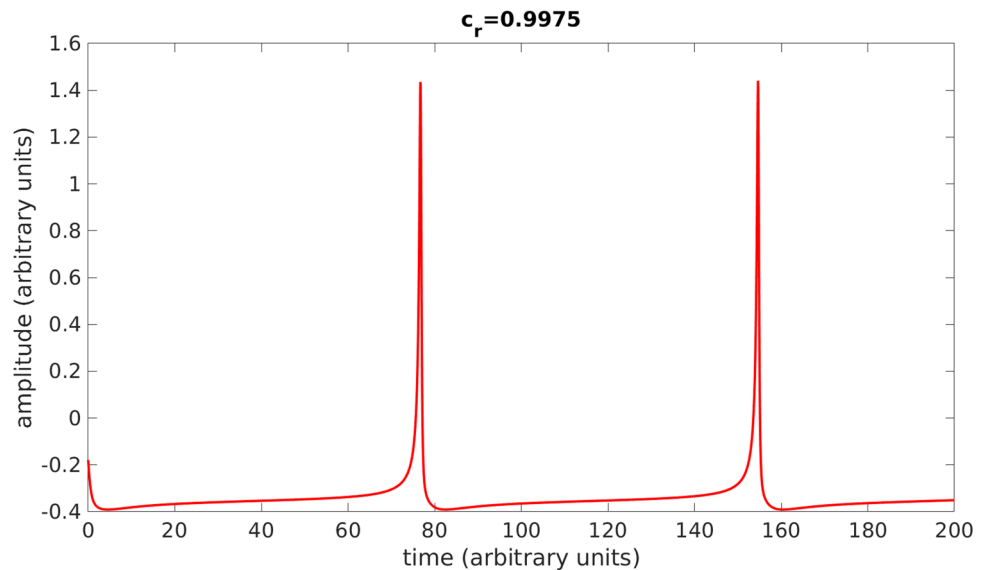
c_r=0.945

Figure 10. The amplitude and phase of a single mode for $c_r = 0.945$.

Figure 11. The amplitude and phase of a single mode for $c_r = 0.99$.

Figure 12. The amplitude and phase of a single mode for $c_r = 0.995$.

Figure 13. The amplitude and phase of a single mode for $c_r = 0.9975$.

References

1. Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (Lond.)* **117**, 500–544 (1952).
2. Johnston, D. & Wu, S.M.-S. *Foundations of Cellular Neurophysiology (A Bradford Book)* hardcover. (A Bradford Book, 1994).
3. Giannari, A. & Astolfi, A. Model design for networks of heterogeneous Hodgkin–Huxley neurons. *Neurocomputing* **496**, 147–157. https://doi.org/10.1016/j.neucom.2022.04.115 (2022).
4. Strassberg, A. F. & DeFelice, L. J. Limitations of the Hodgkin–Huxley formalism: Effects of single channel kinetics on transmembrane voltage dynamics. *Neural Comput.* **5**, 843–855 (1993).
5. Meunier, C. & Segev, I. Playing the devil's advocate: Is the Hodgkin–Huxley model useful?. *Trends Neurosci.* **25**, 558–563 (2002).
6. Yamazaki, K., Vo-Ho, V.-K., Bulsara, D. & Le, N. Spiking neural networks and their applications: A review. *Brain Sci.* **12**, 863 (2022).
7. Fitzhugh, R. Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.* **1**, 445–466 (1961).
8. Nagumo, J., Arimoto, S. & Yoshizawa, S. An active pulse transmission line simulating nerve axon. *Proc. IRE* **50**, 2061–2070. https://doi.org/10.1109/jrproc.1962.288235 (1962).
9. Morris, C. & Lecar, H. Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* **35**, 193–213 (1981).
10. Izhikevich, E. M. Simple model of spiking neurons. *IEEE Trans. Neural Netw.* **14**, 1569–1572 (2003).
11. Zenke, F. & Ganguli, S. SuperSpike: Supervised learning in multilayer spiking neural networks. *Neural Comput.* **30**, 1514–1541 (2018).
12. Kaiser, J., Mostafa, H. & Neftci, E. Synaptic plasticity dynamics for deep continuous local learning (DECOLLE). *Front. Neurosci.* **14**, 424 (2020).
13. Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T. & Maida, A. Deep learning in spiking neural networks. *Neural Netw.* **111**, 47–63 (2019).
14. Bellec, G. *et al.* A solution to the learning dilemma for recurrent networks of spiking neurons. *Nat. Commun.* **11**, 3625 (2020).
15. Hunsberger, E. & Eliasmith, C. Training spiking deep networks for neuromorphic hardware. arXiv:1611.05141. https://doi.org/10.13140/RG.2.2.10967.06566 (2016).
16. Sengupta, A., Ye, Y., Wang, R., Liu, C. & Roy, K. Going deeper in spiking neural networks: VGG and residual architectures. *Front. Neurosci.* **13**, 95 (2019).
17. Rueckauer, B., Lungu, I.-A., Hu, Y., Pfeiffer, M. & Liu, S.-C. Conversion of continuous-valued deep networks to efficient event-driven networks for image classification. *Front. Neurosci.* **11**, 682 (2017).
18. Xu, Q. *et al.* Hierarchical spiking-based model for efficient image classification with enhanced feature extraction and encoding. *IEEE Trans. Neural Netw. Learn. Syst.* **20**, 1–9 (2022).
19. Shen, J., Zhao, Y., Liu, J. K. & Wang, Y. HybridSNN: Combining bio-machine strengths by boosting adaptive spiking neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **20**, 1–15 (2021).
20. Kheradpisheh, S. R., Ganjtabesh, M., Thorpe, S. J. & Masquelier, T. Stdp-based spiking deep convolutional neural networks for object recognition. *Neural Netw.* **99**, 56–67 (2018).
21. Kim, S., Park, S., Na, B. & Yoon, S. Spiking-yolo: Spiking neural network for energy-efficient object detection. *Proc. AAAI Conf. Artif. Intell.* **34**, 11270–11277 (2020).
22. Zhou, S., Chen, Y., Li, X. & Sanyal, A. Deep scnn-based real-time object detection for self-driving vehicles using lidar temporal data. *IEEE Access* **8**, 76903–76912 (2020).
23. Luo, Y. *et al.* Siamsnn: Spike-based siamese network for energy-efficient and real-time object tracking. arXiv:2003.07584 (arXiv preprint) (2020).
24. Bertinetto, L., Valmadre, J., Henriques, J. F., Vedaldi, A. & Torr, P. H. Fully-convolutional siamese networks for object tracking. In *Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8–10 and 15–16, 2016, Proceedings, Part II 14*, 850–865 (Springer, 2016).
25. Patel, K., Hunsberger, E., Batir, S. & Eliasmith, C. A spiking neural network for image segmentation. arXiv:2106.08921 (arXiv preprint) (2021).
26. Rasmussen, D. Nengodl: Combining deep learning and neuromorphic modelling methods. *Neuroinformatics* **17**, 611–628 (2019).

27. Rostami, A., Vogginger, B., Yan, Y. & Mayr, C. G. E-prop on SpiNNaker 2: Exploring online learning in spiking RNNs on neuromorphic hardware. *Front. Neurosci.* **16**, 1018006 (2022).
28. Muller Cleve, S. F. *et al.* Braille letter reading: A benchmark for spatio-temporal pattern recognition on neuromorphic hardware. *Front. Neurosci.* **16**, 951164 (2022).
29. Lee, S. T. & Bae, J. H. Investigation of deep spiking neural networks utilizing gated Schottky diode as synaptic devices. *Micromachines (Basel)* **13**, 25 (2022).
30. Paul, A., Tajin, M. A. S., Das, A., Mongan, W. M. & Dandekar, K. R. Energy-efficient respiratory anomaly detection in premature newborn infants. *Electronics (Basel)* **11**, 25 (2022).
31. Petschenig, H. *et al.* Classification of Whisker deflections from evoked responses in the somatosensory barrel cortex with spiking neural networks. *Front. Neurosci.* **16**, 838054 (2022).
32. Patino-Saucedo, A., Rostro-Gonzalez, H., Serrano-Gotarredona, T. & Linares-Barranco, B. Liquid state machine on SpiNNaker for spatio-temporal classification tasks. *Front. Neurosci.* **16**, 819063 (2022).
33. Li, K. & Ncipe, J. C. Biologically-inspired pulse signal processing for intelligence at the edge. *Front. Artif. Intell.* **4**, 568384 (2021).
34. Syed, T., Kakani, V., Cui, X. & Kim, H. Exploring optimized spiking neural network architectures for classification tasks on embedded platforms. *Sensors (Basel)* **21**, 25 (2021).
35. Fil, J. & Chu, D. Minimal spiking neuron for solving multilabel classification tasks. *Neural Comput.* **32**, 1408–1429 (2020).
36. Saucedo, A., Rostro-Gonzalez, H., Serrano-Gotarredona, T. & Linares-Barranco, B. Event-driven implementation of deep spiking convolutional neural networks for supervised classification using the SpiNNaker neuromorphic platform. *Neural Netw.* **121**, 319–328 (2020).
37. Liu, G., Deng, W., Xie, X., Huang, L. & Tang, H. Human-level control through directly trained deep spiking Q-networks. *IEEE Trans. Cybern.* **20**, 20 (2022).
38. Zhang, M. *et al.* Rectified linear postsynaptic potential function for backpropagation in deep spiking neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **33**, 1947–1958 (2022).
39. Kwon, D. *et al.* On-chip training spiking neural networks using approximated backpropagation with analog synaptic devices. *Front. Neurosci.* **14**, 423 (2020).
40. Lee, J., Zhang, R., Zhang, W., Liu, Y. & Li, P. Spike-train level direct feedback alignment: Sidestepping backpropagation for on-chip training of spiking neural nets. *Front. Neurosci.* **14**, 143 (2020).
41. Meng, Q. *et al.* Training much deeper spiking neural networks with a small number of time-steps. *Neural. Netw.* **153**, 254–268 (2022).
42. Chen, Y., Du, J., Liu, Q., Zhang, L. & Zeng, Y. Robust and energy-efficient expression recognition based on improved deep ResNets. *Biomed. Tech. (Berl.)* **64**, 519–528 (2019).
43. Pfeiffer, M. & Pfeil, T. Deep learning with spiking neurons: Opportunities and challenges. *Front. Neurosci.* **12**, 774 (2018).
44. Stromatias, E. *et al.* Robustness of spiking Deep Belief Networks to noise and reduced bit precision of neuro-inspired hardware platforms. *Front. Neurosci.* **9**, 222 (2015).
45. Thiele, J. C., Bichler, O. & Dupret, A. Event-based, timescale invariant unsupervised online deep learning with STDP. *Front. Comput. Neurosci.* **12**, 46 (2018).
46. Neftci, E., Das, S., Pedroni, B., Kreutz-Delgado, K. & Cauwenberghs, G. Event-driven contrastive divergence for spiking neuromorphic systems. *Front. Neurosci.* **7**, 272 (2013).
47. Kim, Y. & Panda, P. Revisiting batch normalization for training low-latency deep spiking neural networks from scratch. *Front. Neurosci.* **15**, 773954 (2021).
48. Zou, C. *et al.* A scatter-and-gather spiking convolutional neural network on a reconfigurable neuromorphic hardware. *Front. Neurosci.* **15**, 694170 (2021).
49. Liu, F. *et al.* SSTDP: Supervised spike timing dependent plasticity for efficient spiking neural network training. *Front Neurosci* **15**, 756876 (2021).
50. Zhang, L. *et al.* A cost-efficient high-speed VLSI architecture for spiking convolutional neural network inference using time-step binary spike maps. *Sensors (Basel)* **21**, 25 (2021).
51. Zhan, Q., Liu, G., Xie, X., Sun, G. & Tang, H. Effective transfer learning algorithm in spiking neural networks. *IEEE Trans. Cybern.* **52**, 13323–13335 (2022).
52. Detorakis, G. *et al.* Neural and synaptic array transceiver: A brain-inspired computing framework for embedded learning. *Front. Neurosci.* **12**, 583 (2018).
53. Guo, W., Fouda, M. E., Yantir, H. E., Eltawil, A. M. & Salama, K. N. Unsupervised adaptive weight pruning for energy-efficient neuromorphic systems. *Front. Neurosci.* **14**, 598876 (2020).
54. Miranda, E. J. Memristors for neuromorphic circuits and artificial intelligence applications. *Materials (Basel)* **13**, 25 (2020).
55. Gale, E. M. Neuromorphic computation with spiking memristors: Habituation, experimental instantiation of logic gates and a novel sequence-sensitive perceptron model. *Faraday Discuss* **213**, 521–551 (2019).
56. Esser, S. K. *et al.* Convolutional networks for fast, energy-efficient neuromorphic computing. *Proc. Natl. Acad. Sci. USA* **113**, 11441–11446 (2016).
57. Davidson, S. & Furber, S. B. Comparison of artificial and spiking neural networks on digital hardware. *Front. Neurosci.* **15**, 651141 (2021).
58. Gerstner, W., Kistler, W. M., Naud, R. & Paninski, L. *Neuronal dynamics: From single neurons to networks and models of cognition* (Cambridge University Press, 2014).
59. Sardi, S., Vardi, R., Sheinin, A., Goldental, A. & Kanter, I. New types of experiments reveal that a neuron functions as multiple independent threshold units. *Sci. Rep.* **7**, 18036 (2017).
60. Galinsky, V. L. & Frank, L. R. Universal theory of brain waves: From linear loops to nonlinear synchronized spiking and collective brain rhythms. *Phys. Rev. Res.* **2**(023061), 1–23 (2020).
61. Galinsky, V. L. & Frank, L. R. Brain waves: Emergence of localized, persistent, weakly evanescent cortical loops. *J. Cogni. Neurosci.* **32**, 2178–2202 (2020).
62. Galinsky, V. L. & Frank, L. R. Collective synchronous spiking in a brain network of coupled nonlinear oscillators. *Phys. Rev. Lett.* **126**, 158102. https://doi.org/10.1103/PhysRevLett.126.158102 (2021).
63. Seth, A. K. & Bayne, T. Theories of consciousness. *Nat. Rev. Neurosci.* **23**, 439–452 (2022).
64. Humeau, Y. & Choquet, D. The next generation of approaches to investigate the link between synaptic plasticity and learning. *Nat. Neurosci.* **22**, 1536–1543 (2019).
65. Miller, S. M. & Sahay, A. Functions of adult-born neurons in hippocampal memory interference and indexing. *Nat. Neurosci.* **22**, 1565–1575 (2019).
66. Feng, S. & Duarte, M. F. Few-shot learning-based human activity recognition. *Expert Syst. Appl.* **138**, 112782. https://doi.org/10.1016/j.eswa.2019.06.070 (2019).
67. Wang, Y., Yao, Q., Kwok, J. T. & Ni, L. M. Generalizing from a few examples. *ACM Comput. Surv.* **53**, 1–34. https://doi.org/10.1145/3386252 (2020).
68. Drori, I. *et al.* A neural network solves, explains, and generates university math problems by program synthesis and few-shot learning at human level. *Proc. Natl. Acad. Sci.* https://doi.org/10.1073/pnas.2123433119 *(2022).*

69. Walsh, R., Abdelpakey, M. H., Shehata, M. S. & Mohamed, M. M. Automated human cell classification in sparse datasets using few-shot learning. *Sci. Rep.* **12**, 2924 (2022).
70. Galinsky, V. L. & Frank, L. R. Critical brain wave dynamics of neuronal avalanches. *Front. Phys.* https://doi.org/10.3389/fphy.2023.1138643 *(2023)*.
71. Galinsky, V. L. & Frank, L. R. Neuronal avalanches: Sandpiles of self organized criticality or critical dynamics of brain waves?. *Front. Phys.* **18**, 45301. https://doi.org/10.1007/s11467-023-1273-7 (2023).
72. Galinsky, V. L. & Frank, L. R. Neuronal avalanches and critical dynamics of brain waves. arXiv:2111.07479 (eprint) (2021).
73. Kuramoto, Y. Self-entrainment of a population of coupled non-linear oscillators. In *Mathematical Problems in Theoretical Physics, vol 39 of Lecture Notes in Physics* (ed. Araki, H.) 420–422 (Springer, Berlin, 1975). https://doi.org/10.1007/BFb0013365.
74. Kuramoto, Y. & Battogtokh, D. Coexistence of coherence and incoherence in nonlocally coupled phase oscillators. *Nonlinear Phenom. Complex Syst.* **5**, 380–385 (2002).
75. Kuramoto, Y. Reduction methods applied to non-locally coupled oscillator systems. In *Nonlinear Dynamics and Chaos: Where do We Go from Here?* (eds Hogan, J. *et al.*) 209–227 (CRC Press, 2002).
76. Kulkarni, A., Ranft, J. & Hakim, V. Synchronization, stochasticity, and phase waves in neuronal networks with spatially-structured connectivity. *Front. Comput. Neurosci.* **14**, 569644 (2020).
77. Kim, R. & Sejnowski, T. J. Strong inhibitory signaling underlies stable temporal dynamics and working memory in spiking neural networks. *Nat. Neurosci.* **24**, 129–139 (2021).
78. Neftci, E. O., Mostafa, H. & Zenke, F. Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks. *IEEE Signal Process. Mag.* **36**, 51–63 (2019).
79. Kirkpatrick, J. *et al.* Overcoming catastrophic forgetting in neural networks. *Proc. Natl. Acad. Sci. USA* **114**, 3521–3526 (2017).
80. Kemker, R., McClure, M., Abitino, A., Hayes, T. & Kanan, C. Measuring catastrophic forgetting in neural networks (2017). arXiv:1708.02072.
81. Ramasesh, V. V., Dyer, E. & Raghu, M. Anatomy of catastrophic forgetting: Hidden representations and task semantics (2020). arXiv:2007.07400.
82. LeCun, Y., Cortes, C. & Burges, C. Mnist handwritten digit database. *ATT Labs [Online].* http://yann.lecun.com/exdb/mnist**2** (2010).
83. Xiao, H., Rasul, K. & Vollgraf, R. Fashion-mnist: A novel image dataset for benchmarking machine learning algorithms (2017). arXiv:1708.07747 [cs.LG].
84. Xiao, H., Rasul, K. & Vollgraf, R. A mnist-like fashion product database. Benchmark—Github. https://github.com/zalandoresearch/fashion-mnist (2022). Accessed 21 Ma 2022.
85. MNIST database. Mnist database—Wikipedia, the free encyclopedia (2022). Accessed 31 July 2022.

Acknowledgements

Author contributions

V.G. and L.F. wrote the main manuscript text. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to V.L.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.