Biological Cybernetics (2026) 120:16
https://doi.org/10.1007/s00422-026-01039-3

Biological Cybernetics

PERSPECTIVE

Dynamical mechanisms for coordinating long-term working memory based on the precision of spike-timing in cortical neurons

Terrence J. Sejnowski[1,2]

Received: 1 December 2025 / Accepted: 17 February 2026

Abstract

In the last century, most sensorimotor studies of cortical neurons relied on average firing rates. Rate coding is efficient for fast sensorimotor processing that occurs within a few seconds. Much less is known about the neural mechanisms underlying long-term working memory with a time scale of hours (Ericsson and Kintsch in Psychol Rev 102(2):211, 1995). Cognitive states may not have sensory or motor correlates. For example, you can sit in a quiet room making plan without moving or sensory processing. You can also make plans while out walking. This suggests that the neural substrate for cognitive states neither depends on nor interferes with ongoing sensorimotor brain activity. In this perspective, I make the case for a possible second tier of neural activity that coexists with the well-established sensorimotor tier. The prominent physiological feature of the second tier is coordinated spike timing activity. The interplay of data supporting this hypothesis involves three puzzling yet highly intriguing experimental observations, without any obvious indication that they might actually represent different aspects of a single functional organization. First, consider the precision of spiking in individual neurons. The discovery of millisecond-precision spike initiation in cortical neurons was unexpected (Mainen and Sejnowski in Science 268:1503–1506, 1995). Even more striking was the precision of spiking in vivo, in response to rapidly fluctuating sensory inputs. Second, high temporal resolution can also mediate spike timing-dependent plasticity (STDP) by controlling the relative timing of presynaptic and postsynaptic spikes at the millisecond scale. Third, we observe waves across many frequency bands traveling across the cortex. Strikingly, their timing is highly precise. Gamma waves, for example, which are triggered by attention, can plausibly trigger STDP that lasts for hours in cortical neurons. This temporary cortical network, ostensibly a second tier of functionality, rides astride the long-term sensorimotor network and could support cognitive processing and long-term working memory.

Keywords Cerebral Cortex · Long-term Working Memory · Traveling Waves · STDP · Spike-time Codes

1 Introduction

We are at a crossroads in systems neuroscience. In the twentieth century, progress was made in characterizing the response properties of single neurons using sharp microelectrodes during behavioral tasks. The cortical responses of sensory neurons correlated with sensory inputs, those of motor neurons with movements, and those of neurons in association areas with higher cognitive functions. The tasks typically lasted a few seconds and required extensive training. However, in the wild, behavior is self-generated and is coordinated over much longer time spans. Recordings from freely moving rodents, for example, revealed that hippocampal neurons responded selectively to places in the environment, which would have been missed in head-fixed experiments. Self-generated cognition that occurs without any sensory inputs or motor outputs over minutes and hours, as in remembering and planning, is much more difficult to study than tasks that require working memory over seconds.

During the 1980s, I focused on network models of vision (Ballard, Hinton, and Sejnowski, 1983), grounded in rich psychophysical research and neural recordings from the visual cortex. I was inspired by pioneering network models, such as the Marr and Poggio (1976) model of stereopsis.

Communicated by Paul Tiesinga.

✉ Terrence J. Sejnowski
terry@salk.edu

[1] Computational Neurobiology Laboratory, Salk Institute for Biological Sciences, La Jolla, San Diego, CA 92037, USA

[2] Department of Neurobiology, University of California, La Jolla, San Diego, CA 92093, USA

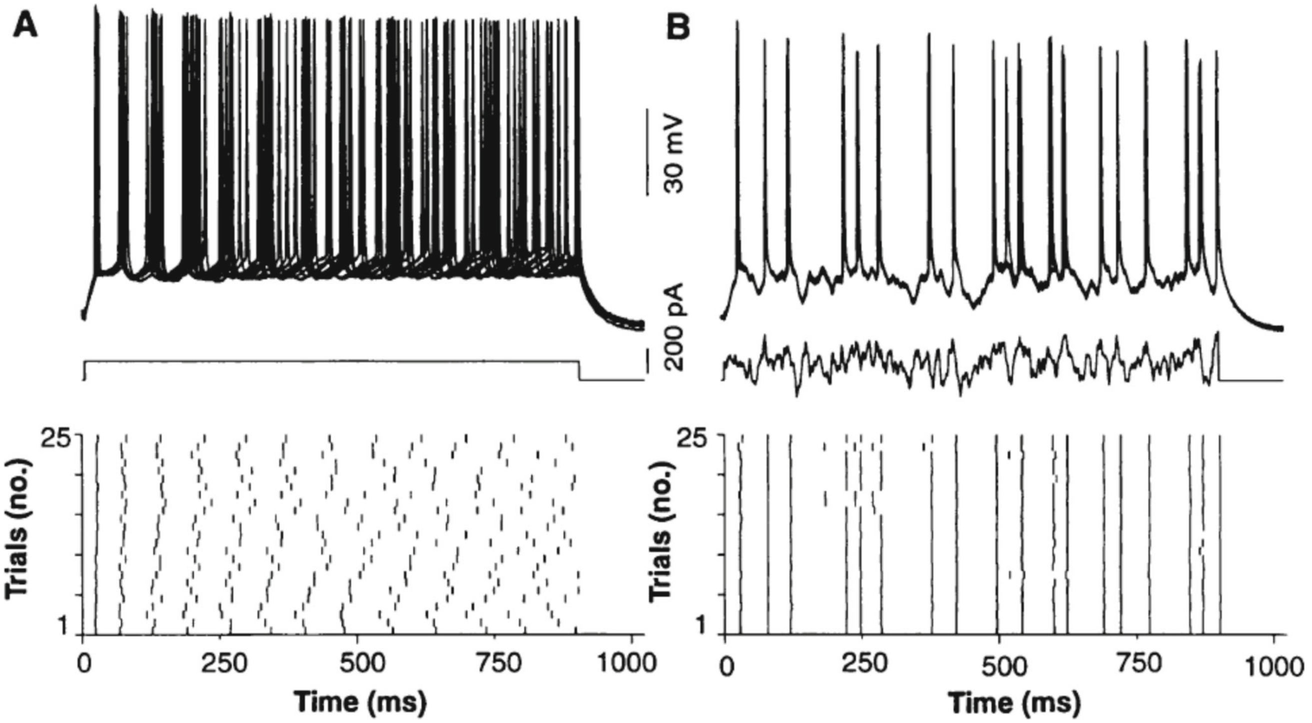

Fig. 1 Precise spike timing in response to frozen noise. **A** Constant current injection into a cortical pyramidal neuron results in a drifting spike raster. **B** Precisely timed spike responses to repeated injection of the same fluctuating current (Adapted from Mainen and Sejnowski 1995)

Vision research in that era focused on images and object recognition. I knew that the visual system integrated information across eye movements and was curious how it was done. For example, as you read this article, your eyes make fast, saccadic movements across the page, taking in small groups of words in your fovea three times per second. Each saccade is a snapshot that must be integrated with previous words to build a conceptual understanding of what is being conveyed.

Psychologists call this long-term working memory (Ericsson and Kintsch 1995). After reading this article, your brain will think about it in the context of experiences and thoughts previously stored in long-term memory. After listening and watching a lecture for an hour, you can retain enough details that you heard and the slides you saw to ask a relevant question. During a concert, recurring themes are expected and variations detected. Long-term working memory has a time scale of hours, much longer than short-term working memory, which lasts seconds and minutes, such as remembering a phone number by rehearsal.

We can now record distributed neural activity from many thousands of neurons simultaneously throughout the cortex. We need a new conceptual framework for how cognition arises from global activity on long time scales. Self-generated long-term working memory could be supported by the dynamical neural mechanisms proposed here.

This paper reviews a wide range of established cortical observations from a personal perspective and synthesizes a forward-looking two-tier hypothesis for cortical processing. Pieces of the puzzle are beginning to fit together. Key assumptions made along the way require testing and experimental confirmation. I recently received an NIH Director's Pioneer Award to investigate the functions of spike timing and traveling waves in cortical processing.

2 The precision of spike timing in cortical neurons

While transitioning from theoretical physics to neuroscience, I was fortunate to work as a postdoctoral fellow with Stephen Kuffler in the Department of Neurobiology at Harvard Medical School. I gained skills in intracellular recordings from neurons and synaptic physiology, which introduced me to a pioneering tradition in electrophysiology going back to Alan Hodgkin, Andrew Huxley, and Bernard Katz in the 1950s. My first job was in the Thomas C. Jenkins Biophysics Department at Johns Hopkins University in the early 1980s, where I set up an experimental lab and started simulating Hodgkin-Huxley models of neurons using the NEURON computer program.

When I moved from Johns Hopkins to San Diego in 1989, I established the Computational Neurobiology Laboratory at the Salk Institute and the Institute for Neural Computation at the University of California at San Diego. My lab attracted some of the best and brightest graduate students and postdoctoral fellows. Zachary Mainen was a new graduate student in my lab, having come from Tom Brown's lab at Yale. We had daily afternoon teas where new ideas and advances were discussed. Tony Zador, who was in Chuck Stevens' lab next door and previously Zach's graduate supervisor at Yale, would often join the discussions. Neural coding was a hot topic, and one day we had a lively debate over spikes. Tony Bell, a postdoctoral fellow, believed every spike mattered. This went against the seemingly random spike times in in vivo cortical recordings, consistent with the consensus view that cortical processing was noisy and that a rate code carried most of the information in cortical networks. If spike times matter, what are they coding and how are they decoded?

Spike timing is determined by the highly fluctuating membrane potentials observed in intracellular recordings in vivo, which are driven by barrages of excitatory and inhibitory synaptic inputs from ongoing cortical activity. How would a cortical neuron respond if a controlled, randomly fluctuating current were to be injected into the cell body? One of the advantages of having a wet lab is that you do not have to wait for someone else to do your experiment. It was not long before Zack had the answer (Mainen and Sejnowski 1995). When he repeatedly injected the same fluctuating current waveform, called frozen noise, the precision of action potential timing was in the millisecond range (Fig. 1). This revealed that the spike initiation mechanism in pyramidal neurons was surprisingly robust and raised an interesting question: Why?

Our precision paper has been cited over 2,500 times. The citation rate on most papers falls exponentially, but 30 years later, this paper continues to receive around 100 citations per year. The difference between constant-current injection and frozen noise in Fig. 1 is what most remember from the paper. We also asked what feature in the current waveform triggered the spike. We expected a fast depolarizing fluctuation, but spike-triggered averaging revealed a consistent hyperpolarizing dip preceding the spike (Fig. 2). A strong inhibitory input controls spike timing by post-inhibitory rebound. In Hodgkin-Huxley models, sodium channels are partially inactivated near threshold, and hyperpolarization deinactivates them, making a neuron more excitable and ensuring precise spike timing.

There are many types of inhibitory neurons in the cortex. Which could provide strong inhibitory input to control spike timing? The soma of a pyramidal neuron receives a calyx of feedback inhibition from parvalbumin-positive (PV) basket cells. Synaptic inputs from basket cells are functionally equivalent to injecting current into the soma. Basket cells also receive strong excitatory inputs from pyramidal neurons. This inhibitory feedback circuit is also responsible for precisely timed bursts of spikes during 30–80 Hz gamma oscillations (Jadi and Sejnowski 2014).

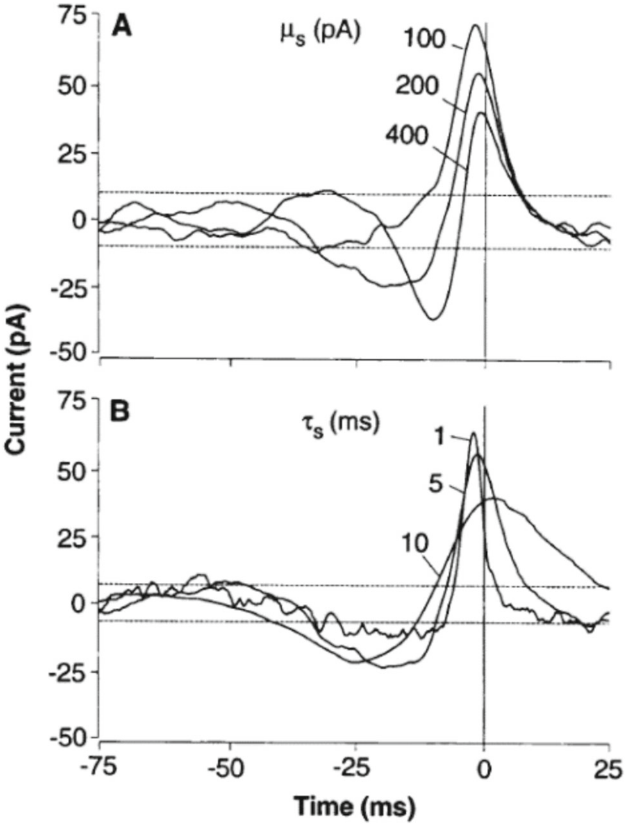

Fig. 2 Spike-triggered average of the frozen noise injected into a cortical pyramidal neuron reveals a postinhibitory rebound mechanism underlying the precision of spike timing in Fig. 1. The mean and time constant of the filtered noise was varied (Adapted from Mainen and Sejnowski 1995)

3 In vivo spike timing precision

Current injected directly into the soma of neurons produces precise spike timing in slices. In vivo conditions may be different because current is injected throughout the dendritic tree. Thalamic neurons preserve the millisecond timing of a flickering visual stimulus (Reinagel and Reid 2002) (Fig. 3). Is sensory timing information preserved downstream? A good place to look is area MT, which receives direct input from area V1 and is several synapses deeper in the visual cortex. First identified by John Allman and John Kaas in the medial temporal area of the owl monkey's visual cortex (Allman and Kaas 1971), MT neurons selectively respond to the direction of motion of visual stimuli. Several researchers have focused on neurons in MT, including Tom Albright, Tony Movshon, and Bill Newsome (Britten et al. 1992).

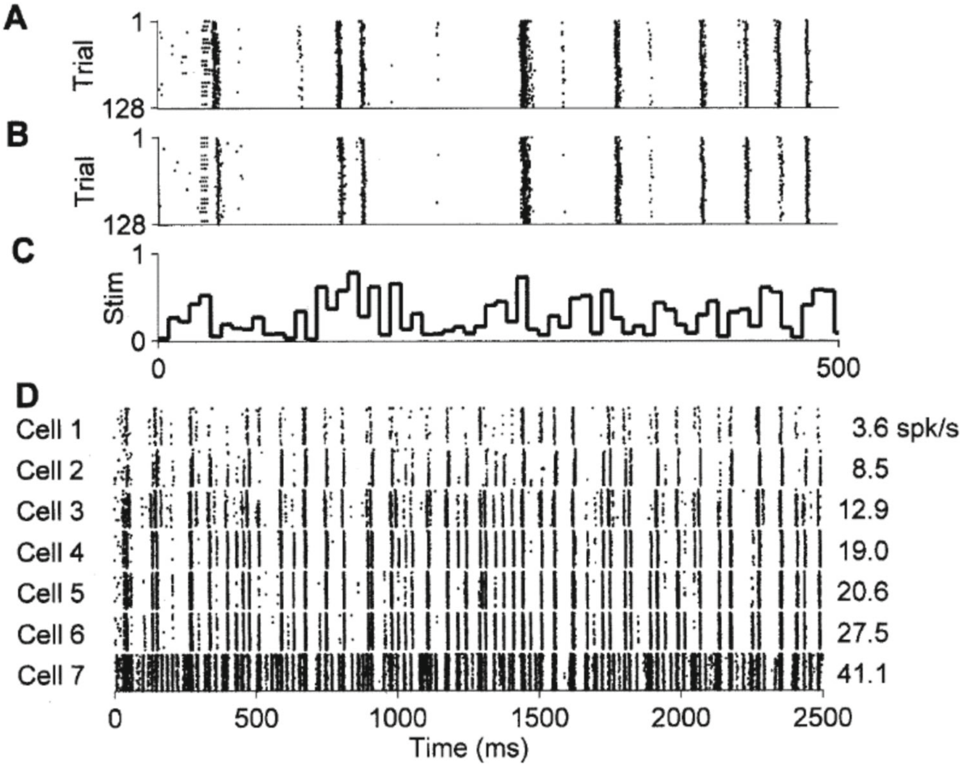

Fig. 3 Response of thalamic relay neurons in the cat lateral geniculate nucleus to a fluctuating visual stimulus. **A, B** Spike rasters. **C** Time course of the intensity of the visual stimulus. **D** Spike-time rasters for seven thalamic relay neurons with averge firing rates varying over an order of magnitude. (Adapted from Reinagel and Reid 2002)

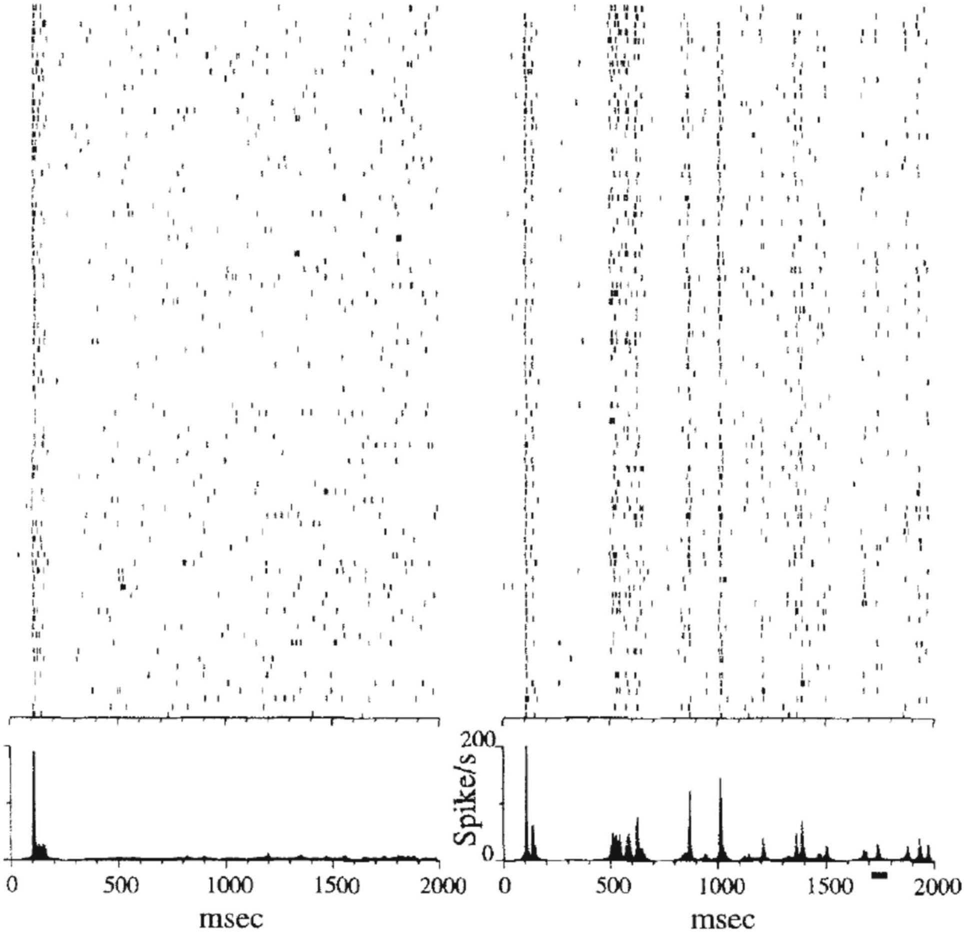

Fig. 4 Responses of neurons in Monkey area MT in response to drifting random dots. Left: Display program initiated with a random seed on each trial. Right: Display program initiated with the same random seed on each trial (Adapted from Bair and Koch 1996)

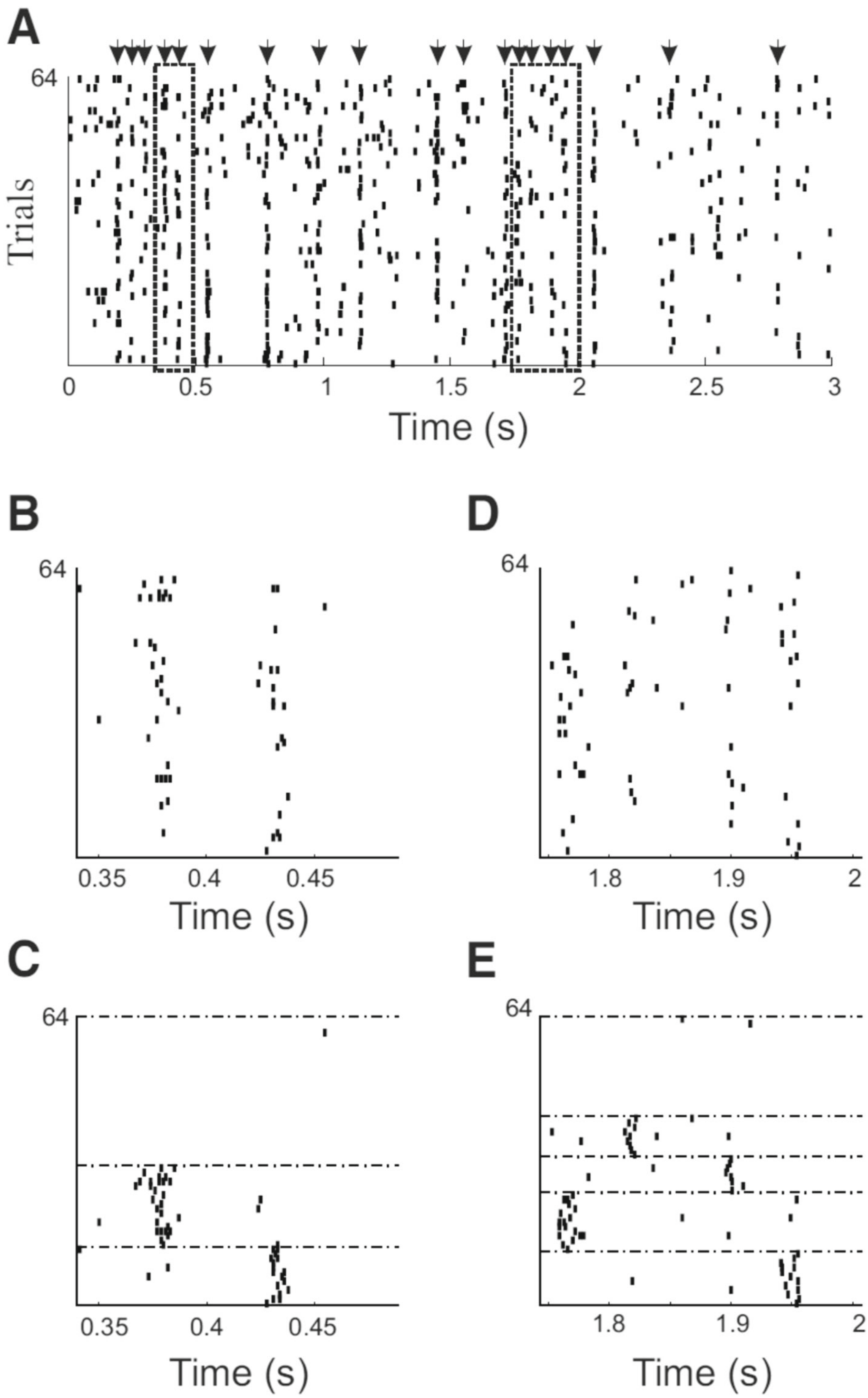

Fig. 5 Multiple spike patterns in response to the same fluctuating visual stimulus in monkey area MT. **A** spike responses to from 64 trials to the same fluctuating visual stimulus. **B**, **D** Blow-up of two segments of the raster in panel A. **C**,**E** Clustering of the trails in panels B and D, revealed two and four spike-time patterns, respectively. (Adapted from Fellous, Tiesinga, Thomas, Sejnowski, 2004)

Newsome trained monkeys to detect the direction of random dots on a screen, with a fraction of them briefly moving in one direction (Bair and Koch 1996). He varied the fraction and found the detection threshold and the rising psychometric function. Wyeth Bair and Christof Koch reanalyzed Newsome's recordings. They found that when trials with the same input motion correlation were displayed in a raster format, spike peaks appeared with 3 ms timing precision, similar to those found in vitro (Fig. 4, right). The random-dot display program used by the Newsome lab, obtained from Movshon, did not reset the random seed, so the same random dots appeared at the same locations and times in every trial for trials with the same motion correlation. This was a fortuitous

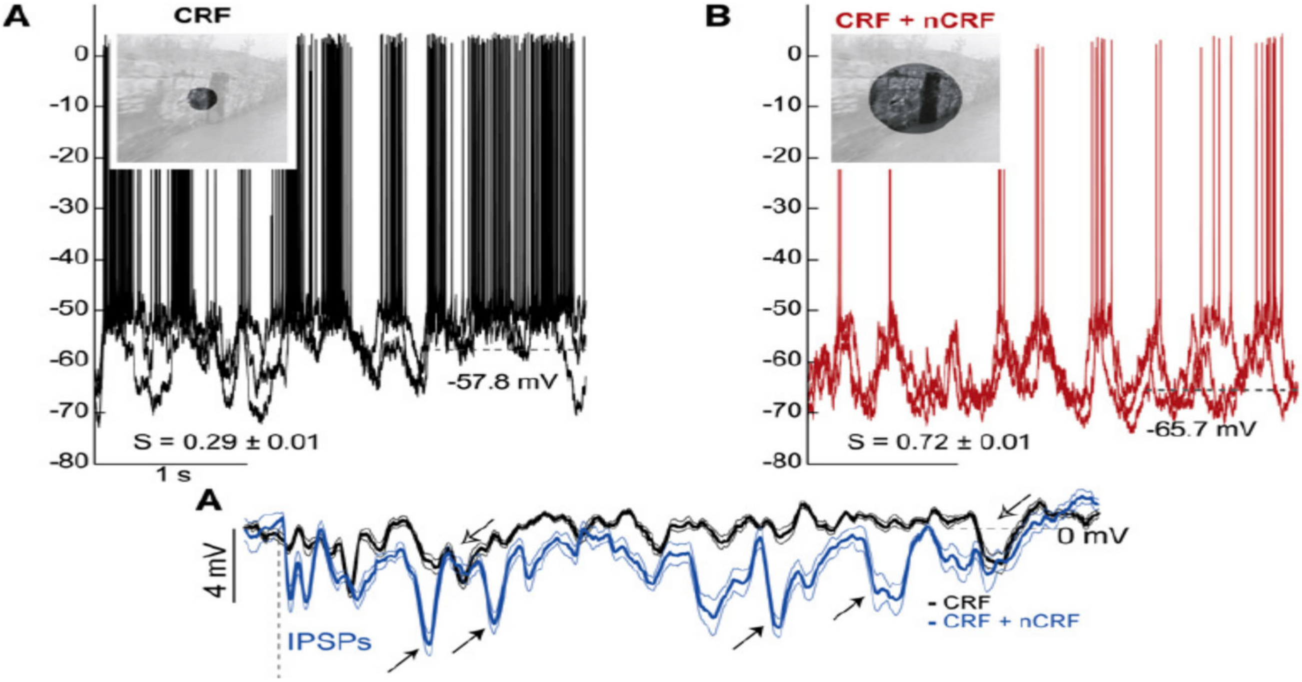

Fig. 6 Intracellular responses of a cortical pyramidal neuron to a repeated video clip: **A** when restricted to the classical receptive field (CRF) and **B** when widened into the nonclassical receptive field (nCRF). Below: Intracellular membrane potential in response to the CRF (black) and CRF + nCRF (blue). The modulatory influence of visual inputs from outside of the classical receptive field was first observed in area MT by John Allman and has subsequently been shown to occur throughout the cortex (Allman et al. 1985). (Adapted from Haider et al. 2010)

oversight since otherwise the precision of spike timing would not have been apparent (Fig. 4 left).

Shortly thereafter, researchers in Albright's lab confirmed that the spike timing of neurons in monkey area MT in response to fluctuations in moving visual stimuli was preserved with 3 ms precision (Buracas, Zador, DeWeese, Albright, 1998). In a reanalysis of these data, clustering algorithms revealed that the same fluctuating stimulus produced not a single precisely timed spike sequence, but several precise spike patterns (Fellous, Tiesinga, Thomas, Sejnowski, 2004) (Fig. 5). A few trials in Fig. 1 show offsets, suggesting something similar may happen in cortical circuits perhaps by a small perturbation, leading to multiple, interleaved dynamical spike patterns to the same fluctuating stimulus. The discovery that the cortex could respond nearly deterministically to sensory inputs with precisely timed spikes led to research into how diverse patterns of precisely timed spikes emerged from Hodgkin-Huxley models (Tiesinga, Fellous, Sejnowski, 2008).

4 Multiple cortical neural codes

The concept of a receptive field was an organizing principle for interpreting the properties of single neurons in the twentieth century (Sejnowski 1976a). The receptive field of a cortical neuron is the specific region of sensory space in which an appropriate stimulus elicits a change in its firing rate. Because of variation in the number and timing of spikes, the response is typically averaged over trials. Evidently, these neurons use firing rate to code sensory information. What was left for the cortex to process from sensory inputs that needed millisecond precision? Could it be that both types of coding are present and used for different purposes (Singer 2021)?

David McCormick recorded intracellular membrane potentials from neurons in cat visual cortex in response to repeated video clips (Haider et al. 2010). They masked the visual stimulus within a neuron's receptive field and compared the spike trains to responses when the mask was widened (Fig. 6). When the stimulus was restricted to the classical receptive field, the neuron responded vigorously, with spike timing patterns that varied trial-to-trial. However, when the mask was widened, the neuron hyperpolarized, lowering the firing rate and revealing precisely timed spikes. Large inhibitory inputs appeared in the nonclassical stimulus condition (Fig. 6, lower panel). Compare these inhibitory events with the hyperpolarizing events preceding spikes in Fig. 2): The hyperpolarization was responsible for the precisely timed spikes. These results show that, in addition to modulating the average firing rate, inputs from the nonclassical receptive field shift neural circuits toward precise spike timing.

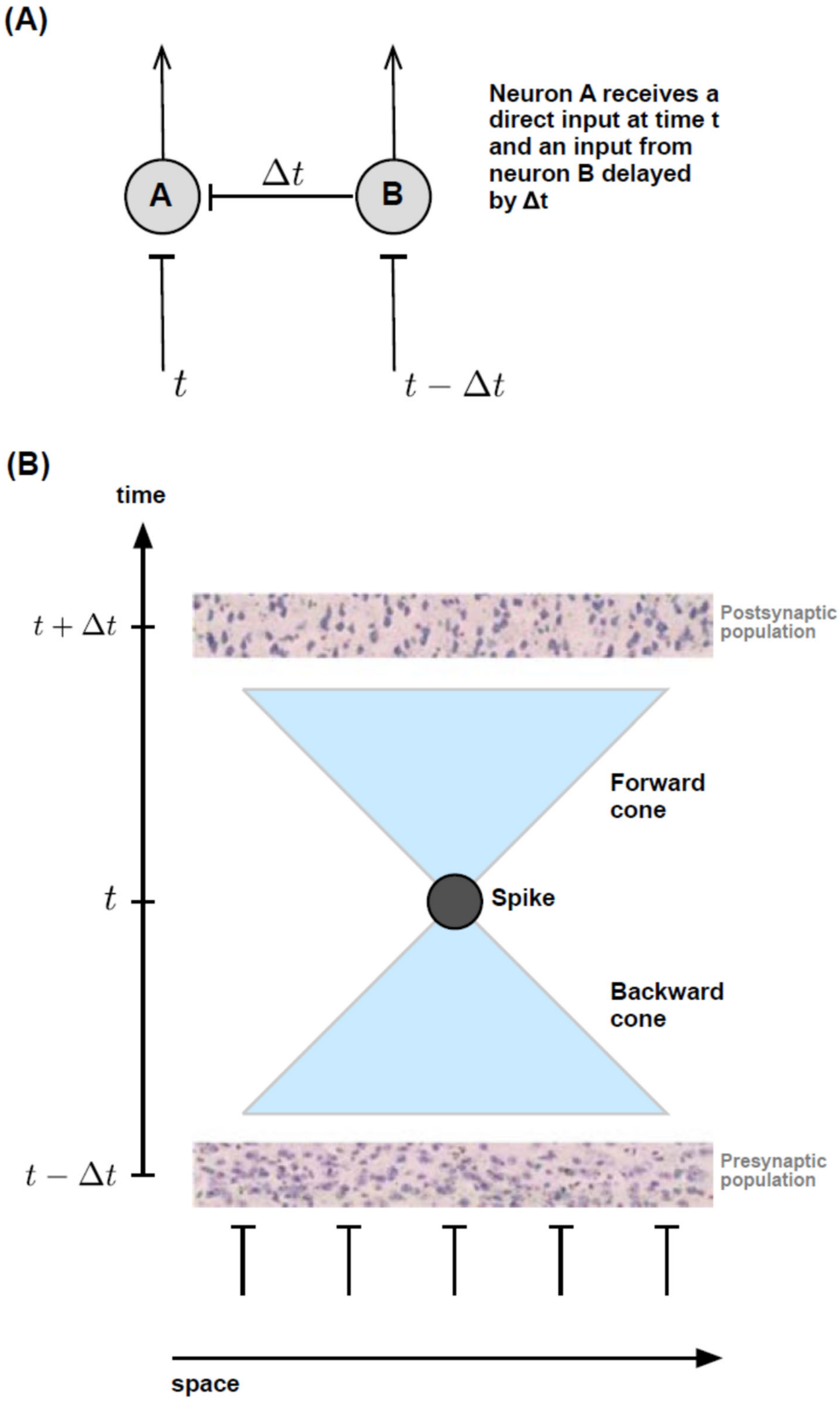

Fig. 7 Spacetime population coding. **A** Hassenstein-Reichardt model for detecting the direction of motion of a visual stimulus (Reichardt 1957). The visual stimulus moves from location B to location A during a time delay Δt. Neuron A responds only when it receives a direct visual input at time t and an indirect visual input with a time delay Δt from neuron B. **B** The simple circuit in panel A can be expanded to include a wide range of converging inputs from neighboring neurons with a range of time delays. The timing of the spike at time t mixes information within the lower blue spacetime cone and, in turn, influences neurons in the upper blue cone. (Adapted from Muller et al. 2024)

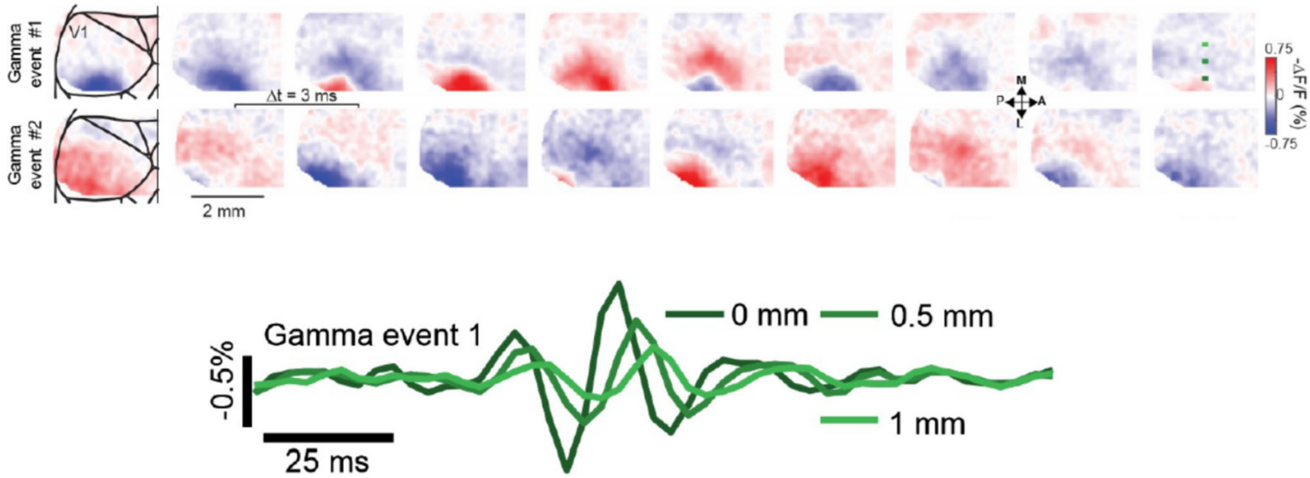

Fig. 8 Membrane potentials of neurons in mouse visual cortex measured with a voltage-sensitive dye. Top: two sequences of 3 ms frames illustrating waves traveling from the bottom to the top of the frame. The frequency is approximately 37 Hz. Bottom: Time course of voltage measurements at three locations in the top image showing the waveform and phase differences. Traveling waves are wave packets. (Adapted from Haziza et al. 2025)

These results have interesting implications. First, many neurons in the visual cortex are in a nonclassical state in response to full-field visual inputs; second, there is a mixture of single-spike responses and more rate-like responses in this state. This blend suggests that both are present and may convey different kinds of information. Rate coding and spike timing codes could coexist within the same population of neurons, serving different purposes at different time scales. A third, less obvious implication is that inputs from the nonclassical surround, in addition to mixing spatial information, also mix temporal information. The long-range horizontal axons in the cortex that link neurons in cortical columns travel many millimeters. The axons are thin, unmyelinated, and their spikes travel slowly at 0.1 m/sec, arriving with time delays of 10–100 ms (Johnson and Frostig 2016).

5 A new spacetime code for temporal context

Nonclassical receptive fields provide both spatial context and temporal context. For example, a moving stimulus can be detected by sequential inputs in neighboring locations. This is the kernel of the Reichardt detector and Adelson-Bergen motion energy model (Fig. 7A) (Reichardt 1957; Adelson and Bergen 1985). Cortical neurons receive delayed signals from many nearby neurons; the farther away, the longer the delay (Fig. 7B).

Thus, a population of neurons could represent not just what is happening at a single moment, but also over a window of space and time (Muller et al. 2024). Moving stimuli produce moving patterns of activity in the cortex, called an equivariant flow, which can be generalized to more abstract dynamic representations (Keller, Muller, Sejnowski, Welling, 2024b).

6 Traveling waves mix temporal and spatial information

Cortical oscillations are ubiquitous in recordings of local field potentials, which average extracellular signals from neighboring neurons. These oscillations have generally been interpreted as evidence for synchronous spiking. Recordings from electrode arrays and optical recordings have instead revealed that local oscillations in all frequency bands extend spatially as traveling waves (Muller, Chavane, Reynolds, Sejnowski, 2018; Keller, Muller, Sejnowski, and Welling, 2024b; Cruddas, Pang and Fornito. 2025). Both stimulus-induced and spontaneous cortical travelling waves have been observed. The stimulus-induced Event-Related Potential (ERP) is a phase reset of ongoing alpha wave activity (Makeig, Westerfield, Jung, Enghoff, Townsend, Courchesne, and Sejnowski, 2002).

Figure 8 shows an optical recording of a traveling wave in the gamma band in mouse visual cortex. The dynamical patterns of spikes in traveling waves observed in the cortex allow information in relative spike timing to be propagated across the cortex by long-range horizontal axons (Gilbert and Wiesel 1989). Spiking is sparse in cortical traveling waves, allowing many different subsets of interconnected neurons to represent many different spatiotemporal features. The circuit mechanisms for the origin of sparse mesoscale waves in

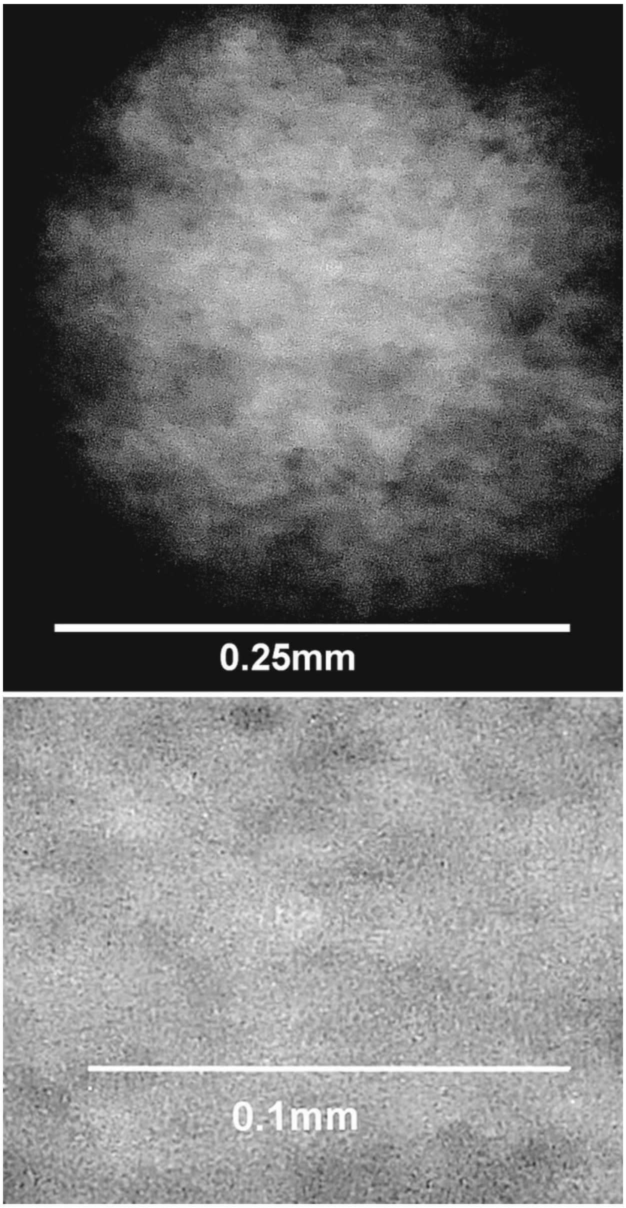

Fig. 9 This is a small part of a transmission hologram photographic plate viewed through a microscope at two levels of magnification. The hologram captured images of a toy van and a car. Holographic information is encrypted in the speckle pattern. It is no more possible to detect the objects in the hologram from this pattern than it is to identify what music has been recorded by looking at the surface of a DVD. The three-dimensional objects can be recovered by shining a laser onto the hologram. (https://en.wikipedia.org/wiki/Holography)

the cortex have been identified and mathematically modeled (Davis et al. 2021).

Spacetime representations are counterintuitive, but recall that a three-dimensional object floating in space can be reconstructed from the speckles in a two-dimensional hologram that our visual system cannot interpret (Fig. 9). The object can be reconstructed by shining a laser on the hologram, or on a portion at reduced resolution. In a holographic spacetime representation, one of the dimensions is time. There is no reason for evolution to make neural encoding easy for us to understand. We face the same challenge interpreting activity patterns in large language models. Karl Pribram (1971) drew an analogy between memory and holograms, proposing a holonomic brain theory.

Dynamical traveling waves convey more information than synchronous oscillations and stabilize the representation and processing of time-varying sensory signals (Keller, Muller, Sejnowski, Welling, 2024b). Because the same process is repeated at every layer in cortical hierarchies, the time window increases from 100 ms in V1 to 10s in the inferotemporal cortex (Hasson et al. 2008). There is also an increasing gradient of NMDA receptors up the cortical hierarchy, with the highest density in the prefrontal cortex (Wang 2020).

Recurrent neural networks. In addition to spatial traveling waves, which mix spatial and temporal information, circulating electrical activity within a cortical recurrent neural network mixes temporal information at a single spatial location. Highly recurrent networks are found in layers 2/3 of the cortex and area CA3 in the hippocampus (Chen, Zhang, Cameron, Sejnowski, 2024). Circulating activity can be stored by learning algorithms such as predictive coding, sequence learning, and forcing in reservoir computing. These sequences can be replayed spontaneously and converted into traveling waves.

Impact on behavior. Traveling waves can affect perception when a visual stimulus is near threshold. As spontaneous waves pass through area MT of a marmoset, their phase modulates spike responses to low-contrast stimuli. The depolarizing phase is highly correlated with behavioral reports that the weak signal was perceived (Davis, Muller, Martinez-Trujillo, Sejnowski, Reynolds, 2020). This modulation of perception reveals a link to behavior, though weakly, consistent with its influence on Tier 2. Waves do not change percepts above threshold or directly cause an action on the sensorimotor Tier 1. The coordination of traveling waves is mediated by shared subthreshold membrane potentials, organized by massive but weak synaptic signaling that modulates spike timing. This is a form of *E pluribus unum* in large populations of neurons. Traveling waves do not consciously change perception under normal conditions. Conscious thinking does not directly affect perception either.

Absence epilepsy, common in children, lasts for 10–20 s, during which there is a staring spell and a pause in motor behaviors (Destexhe and Sejnowski 2023). At the end of the seizure, the conversation and thinking continue as before the pause. Remarkably, absence epilepsy is characterized by massive, low-frequency thalamic bursting that sweeps across the cortex. How is the previous trace maintained? The continuity of thoughts could be sustained by fast synaptic weight changes lasting minutes to hours, preserved across the electrical storm. These temporary weight changes could serve as the substrate for long-term working memory.

7 Fast but temporary synaptic plasticity

A candidate for rapid weight change is spike-timing-dependent plasticity (STDP) (Markram, 1997; Feldman 2012) (Fig. 10). STDP occurs when an excitatory input occurs within 10 ms of a backward-propagating spike in the postsynaptic dendrite. The strength of the synapse is potentiated when the presynaptic input precedes the postsynaptic spike (pre before post) and is depressed when the postsynaptic spike precedes the presynaptic input (post before pre). STDP has been demonstrated in slice preparations but remains elusive in vivo, partly because the mechanisms that regulate cortical spike timing in vivo are unknown.

Could traveling waves trigger STDP? As a traveling wave passes a pyramidal neuron in a neighboring column, excitatory synapses are activated on both the pyramidal neuron and neighboring PV interneurons. Strong transient activation of PV neurons could trigger inhibitory rebound and a backward-propagating spike in the pyramidal dendrites, arriving a few milliseconds after the feedforward excitatory inputs (Fig. 6). This dual pathway to the pyramidal neurons could coordinate the elusive 10 ms time window for pairing presynaptic and postsynaptic spikes during STDP.

Traveling waves are typically packets of spikes at 10–40 Hz, which also match the bursts used to elicit STDP in vitro (Markram et al. 1997). The simultaneous activation of local basket cells by the wave front could also trigger the ING mechanism for synchronizing reciprocally connected inhibitory neurons, thereby amplifying the transient hyperpolarizing pulse in the somas of nearby pyramidal neurons (Fig. 11) (Van Vreeswijk, Abbott, Ermentrout, 1994; Tiesinga and Sejnowski, 2009). Not all the basket cells, however, may synchronize. In a model of short-term working memory, activity in one group of basket cells could switch the network's output by inhibiting another group (Kim and Sejnowski 2021). A chimera of locally synchronized basket cells could control both activity and plasticity.

Only a small fraction of cortical neurons spike as a traveling wave passes. In contrast, the traveling wave during a seizure is dense and is followed by postictal depression, during which the cortex is silent for 5–30 s. However, even if only 1% of cortical neurons in a cortical column of 100,000 neurons spike during a sparse traveling wave, only 1000 would be activated. In a cortical column, around 10% of pyramidal neurons are reciprocally connected, so around 100 neurons could undergo STDP together, strengthening their mutual connectivity and forming a recurrent microcircuit, which Hebb called an assembly (Yoshimura et al. 2005; Perin et al. 2011).

Protocols that induce STDP produced plasticity that lasted for an hour in vitro, which were called long-term potentiation and long-term depression. However, another possibility is that STDP in vivo induces temporary plasticity for only a few hours, the timescale of long-term working memory. If so, then which cortical synapses are most likely to support STDP? Excitatory synapses on small spines are weaker than those with large spines, and entering calcium increases the calcium concentration in their small volume faster, making them more susceptible to STDP. They are also more labile and turn over with learning (Yang et al. 2014; Berry and Nedivi 2017). Because the distribution of spine head sizes in the cortex is log-normal, the small, labile spines constitute the majority of excitatory synapses on pyramidal neurons (Loewenstein et al., 2011). Thus, the majority of cortical excitatory synapses could be recruited for cognitive processing over a moving hours-long time window. The number of small synapses decreases with age, paralleling the lower capacity of long-term working memory in older adults (Holtmaat et al. 2005).

8 Origins of traveling waves

Sleep is a fascinating brain state. During the day, responses to uncontrolled sensory inputs are difficult to separate from self-generated activity. Sensory inputs are suppressed during sleep, and self-generated brain rhythms constantly ring the changes across sleep stages at night. It is possible to study purely internally generated traveling waves during sleep and uncover their underlying neural mechanisms. This was my goal for 20 years, first in collaboration with Mircea Steriade and David McCormick (Steriade et al. 1993), and later with Syd Cash (Muller et al. 2016) and April Benasich. Based on these recordings, we developed detailed biophysical models of thalamic and cortical neurons during slow-wave sleep, especially during sleep spindles.

Sleep spindles are global 10–14 Hz circular traveling waves lasting 1–2s that coordinate the selective consolidation of experiences into long-term cortical memories (Destexhe and Sejnowski 2023). Spindles originate in the thalamus and entrain cortical circuits, which in turn project back to the thalamus. We discovered that sleep spindles in humans are global, circular traveling waves (Muller et al. 2016).

Sleep spindles are triggered at night by the replay of experiences stored in the hippocampus during the day. The replay originates in area CA3 of the hippocampus, a highly recurrent network (Chen, Zhang, Cameron, Sejnowski, 2024). Initializing a recurrent neural network with weights that generate traveling waves makes learning input sequences much more efficient than starting with random values (Keller,

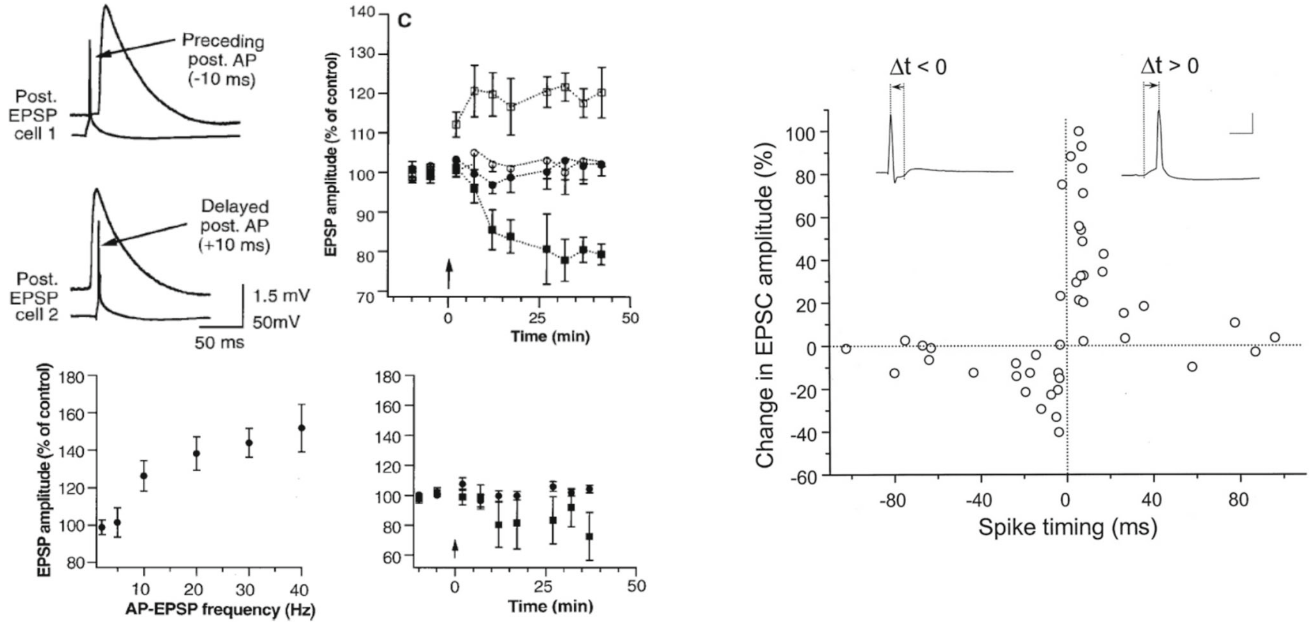

Fig. 10 Spike-timing dependent plasticity. Left: Pairing presynaptic inputs with postsynaptic spikes with a 10 ms delay in rodent cortical slices produces fast weight increases (pre before post) and decreases (post before pre) at synapses. Bursts of 4 pairings were repeated 10 times, spaced by 1s, at frequencies ranging from 4 to 40 Hz. (Adapted from Markram et al. 1997). Right: The interval between the presynaptic input and the postsynaptic spike pairing was varied in cultured hippocampal neurons. (Adapted from Bi and Po, 1998)

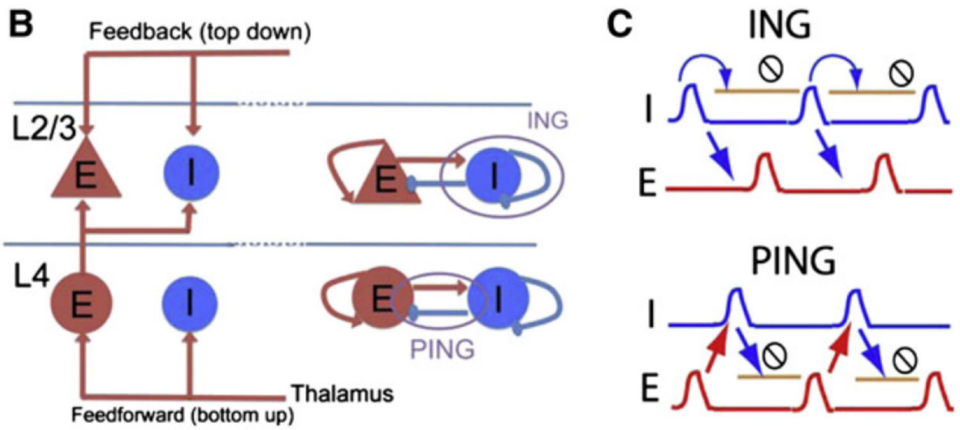

Fig. 11 Neural mechanisms for synchronizing spiking in populations of neurons **B** (Left) Simplified representation of the cortical laminar structure. (Right) Synchrony can be achieved by feedback between excitatory and inhibitory populations (PING) or reciprocal interactions between inhibitory populations (ING). The gray loops in the ING and PING insets encircle the essential interactions that lead to synchrony. Note that the inhibitory neurons are reciprocally connected. **C** For the ING mechanism, I cells are sufficiently excited to spike synchronously without excitatory network activity. The light-brown lines with a stop sign indicate the period during which the network is inactive due to the high inhibitory conductance. (Adapted from Tiesinga and Sejnowski, 2009)

Muller, Sejnowski, Welling, 2024a). Mexican hat connectivity helps maintain stability while learning an antisymmetric component that creates flow of the sequence (Wagner, Chen, Karuvally, Cameron, and Sejnowski, 2026). This network inductive bias is learned by evolution and forms early in development.

Circular traveling waves occur in the monkey prefrontal cortex during specific phases of working memory tasks (Bhattacharya et al. 2022; Batabyal et al. 2025) and in the human cortex (Koller et al. 2024). Spontaneous circular traveling waves in rodent somatosensory cortex are coordinated with subcortical spiking patterns in the thalamus, striatum, and midbrain (Ye et al. 2023). Given their ubiquity and apparent involvement in so many aspects of cortical function, these traveling waves deserve our attention.

9 Phase precession

Little is known about the relative timing of spikes in a traveling wave as it passes through the cortex. Do all the spikes in a wave arrive synchronously, or are they offset? Only a few milliseconds separate potentiation from depression in STDP (Fig. 10). How is this knife-edge controlled in vivo?

Traveling waves in the theta band (4–8 Hz) occur in the hippocampus and entorhinal cortex of rodents as they explore the environment (Lubenov and Siapas 2009; Patel et al. 2012; Hafting et al. 2008). Neurons in the hippocampus have spatial place fields that release bursts of spikes at progressive phases of the theta wave as the rodent traverses the place field, a phenomenon called phase precession. Phase precession also occurs in the grid cells of the entorhinal cortex (Fig. 12). As the wave moves across the cortex, membrane potentials in a local population of neurons are sequentially swept across threshold in order of their initial depolarization levels. Phase precession has also been observed in recordings from the hippocampus, entorhinal cortex, parahippocampal gyrus, anterior cingulate cortex, orbitofrontal cortex, and amygdala of neurosurgical patients undergoing clinical treatment for drug-resistant epilepsy (Qasim, Fried, Jacobs, 2021).

When the time-ordered spikes arrive downstream, the earliest will activate basket cells, eliciting a backpropagating action potential that potentiates the early presynaptic inputs on dendrites (pre before post) but depresses subsequent synaptic inputs (post before pre). Repetitive pairing in the same order will ensure that a sparse subset of synapses is strengthened; however, because of phase precession, earlier spikes drop out on subsequent cycles, allowing synaptic inputs from later spikes to become potentiated. An advantage of this scheme is that it orders spike timing and maintains a balance between potentiated and depressed synapses. Zucker has explored spike-timing computation by small cliques of pyramidal neurons (Miller and Zucker 1999). Computing with time-ordered sequences of spikes, called polychrony, has been explored by Izhikevich (2006, 2025).

10 Decoding spike-time codes

Neural coding in the prefrontal cortex (PFC). The PFC is involved in high-level decision-making and planning, encoding decision variables and time intervals using a combination of rate and precise temporal codes. The specific timing of spikes can carry more information and greater stability for complex stimuli and behaviors than rate codes alone. The PFC sends excitatory glutamatergic projections to the striatum of the basal ganglia (BG).

Timing-to-Rate Conversion. Striatal neurons exhibit rich, variable population dynamics that encode timing information and temporal context for subsequent actions (Jin et al. 2009). The convergence of precisely timed spikes from many PFC neurons onto many fewer striatal neurons summates their excitatory postsynaptic inputs, effectively converting the temporal pattern into a stronger, suprathreshold signal that influences firing rates in downstream targets. The precise timing of PFC spikes, after processing through the BG circuit, potentially involving complex feedback loops and post-inhibitory rebound in the thalamus, influences the average firing rates of thalamic neurons.

The conversion of a decision represented by spike timing in Tier 2 of the PFC into a firing-rate code in the motor cortex is a complex process that involves several subcortical structures. The cortex projects topographically to the BG and loops back to the cortex through the thalamus. This key circuit transforms temporal dynamics into motor commands. The BG effectively acts as a filter and integrator, translating timing-based signals into amplified firing-rate “decision” and “command” signals that the motor system reliably uses to execute a specific sequence of movements.

Thalamocortical Neurons. The cortex projects widely to the thalamic reticular nucleus, a sheath of inhibitory neurons that surrounds the thalamus. Sleep spindles are generated by reciprocal bursting interactions between reticular and thalamocortical neurons (Destexhe and Sejnowski 2023). Thalamocortical bursts also occur in awake monkeys at the onset of visual detection (Ortuño et al. 2014), and in rats at the onset of whisking (Fanselow et al. 2001). This “wake-up” call could be triggered by the widespread corticofugal projections from the prefrontal cortex to the thalamic reticular neurons (Zikopoulos and Barbas, 2006). Thalamocortical neurons have relatively few synapses on cortical pyramidal neurons and depend on synchronous input spikes to control the firing rates of cortical targets (Wang et al. 2010).

Thalamic relay neurons convey information to the motor cortex and premotor cortices. This feedforward input to the cortex is integrated with cortical feedback from layer 6 (Berry 2025). Thalamic inputs directly modulate the firing rates of motor cortex neurons, which then initiate and refine the movements and actions.

11 Spacetime codes in transformers

Transformers in large language models are feedforward deep learning networks trained to predict the next word in a sentence. The temporal context of a word in a sentence helps us understand its meaning. In “Bob, who is hopeful, likes Alice,”

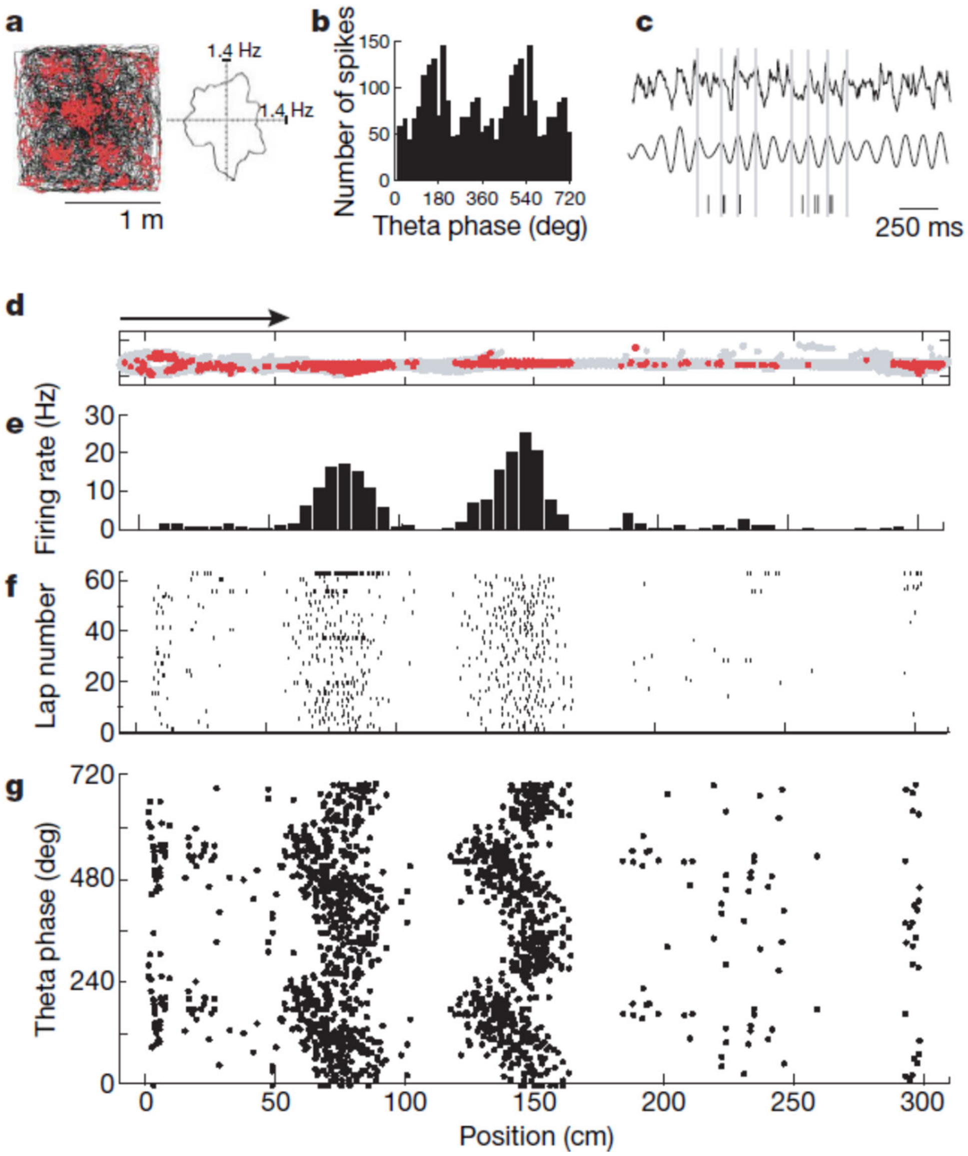

Fig. 12 Phase precession in a layer 2 pyramidal neuron in the entorhinal cortex. **a** Left: firing field of a grid cell during running in an open field (black, trajectory; red, individual spikes). Right: firing rate as a function of running direction (bin size 6°). **b** Distribution of firing rate within the theta cycle for the cell during running on a linear track (bins of 20°, two theta cycles, peak of local theta rhythm is 0°). **c** Local entorhinal EEG with spike times for the layer 2 cell in *a* and *b* during 2.1 s of track running. Vertical ticks are individual spikes. Top: unfiltered EEG trace. Bottom: filtered at 6–11 Hz. Vertical grey lines indicate 0°. **d–g**: Rate distribution and phase relationship for spikes of the cell in *a–c* during running from left to right. **d** Trajectory (gray) with locations of individual spikes (red). The arrow indicates the running direction. **e** Linearized spatial firing rate map (bins of 5 cm). **f**, Raster plot indicating spike positions (vertical ticks) on the track. **g** Theta phase as a function of position (two theta cycles. Note the gradual advance of firing phase as the rat passes through each field. (Adapted from Hafting et al. 2008)

“Bob” is positively associated with “hopeful,” “while” “Alice” is negatively associated. The antecedent of the word “this” in a sentence depends on semantics, which can also be extracted from the temporal context.

These associations are represented in Transformers as “self-attention” (Fig. 13). “Attention” here refers to the strength of association between words, not to the way the term “attention” is used in psychology and neuroscience. Self-attention is a matrix that captures the associations between words in a sentence. These associations help to extract semantic clues from syntactic structure.

Remarkably, self-attention can be replaced by a linear state-space model that generates traveling waves (Karuvaly, et al., 2025). The state-space model has a matrix that generates traveling waves, which, like those in the cortex, mix spatial and temporal information (Keller et al., 2023). Moreover, the banded diagonal matrix has only N parameters, compared with N^2 in self-attention matrices, making it much more efficient.

When watching an axe fall on a wood block at a distance, the visual signal precedes the auditory signal, but they are perceived simultaneously. Spacetime codes in cortical association areas can represent, within a single population, a window of time spanning both signals, like words in a sentence. A lesson from deep learning is that information is not lost when combining information across time from multiple sources and modalities, and that recoding in this way enriches neural representations.

Topographically organized loops between the cortex and the BG, shown in Fig. 13, could provide temporal context for cortical representations. The BG can learn sequences of

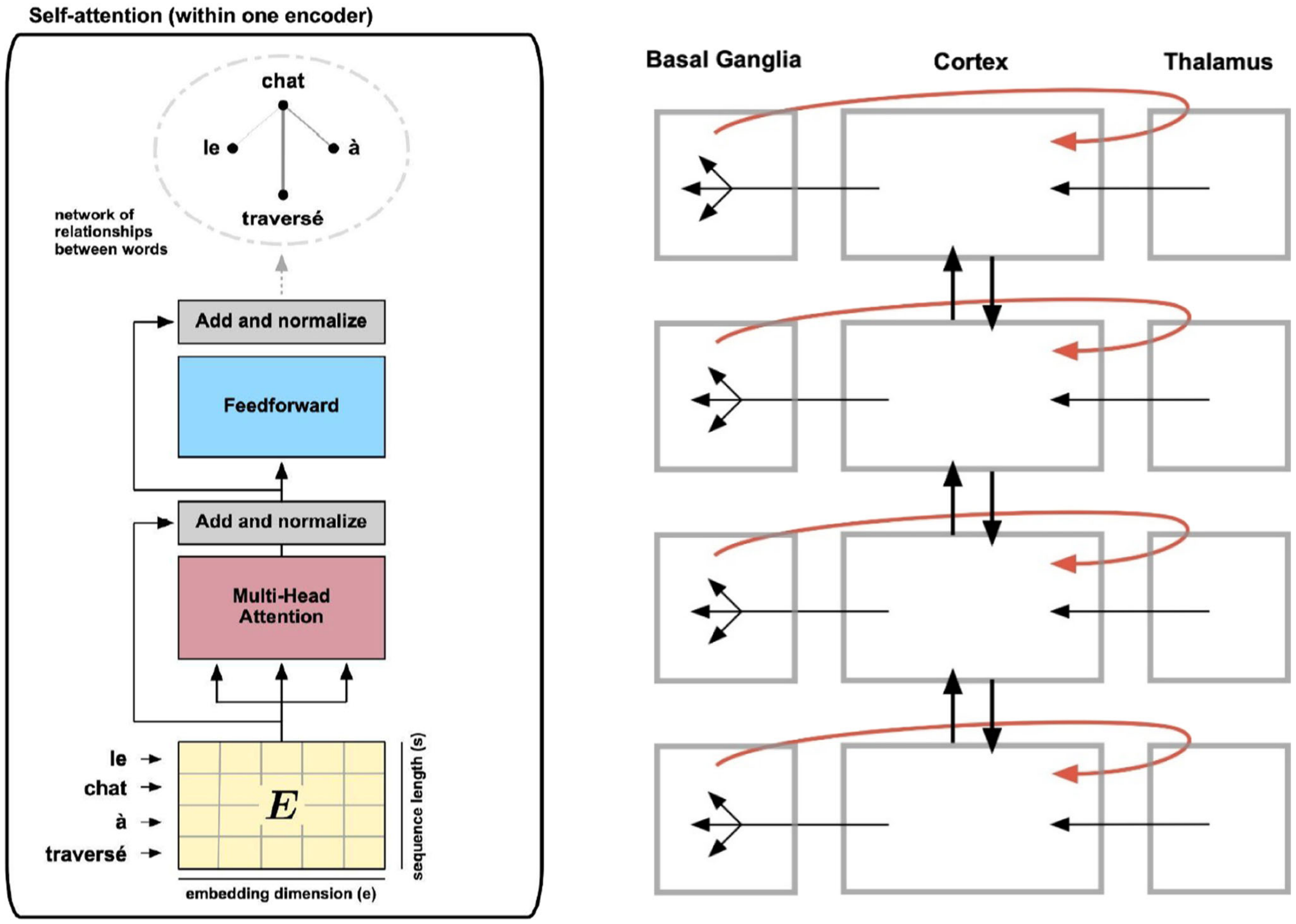

Fig. 13 Left: One module in a feedforward Transformer with input on the bottom and output from the top. Right: Hierarchy of cortical areas receiving inputs from the thalamus and a feedback loop from the BG. (Adapted from Muller et al. 2024)

actions to achieve goals and rewards. Dopamine neurons in the substantia nigra pars compacta (SNc) compute reward prediction error for actions and spoken words, and in the ventral tegmental area (VTA) for thoughts and plans. Synaptic strengths are then modified to reduce prediction error via the temporal difference learning algorithm (TD) (Montague et al. 1996). TD learning constructs a value function for model-free reinforcement learning, enabling fast decision-making based on experience. The value function is distributed across neurons in the ventral striatum and the orbital frontal cortex. These same structures could learn associations within sequences of words in the posterior cortex and thoughts in the prefrontal cortex, which, in Transformers, are implemented with self-attention.

12 BOLD fMRI and Cognition

Thus far, the evidence for 2 tiers of cortical processing has been indirect. A stronger case can be made by following the energy needs for each tier. During routine sensorimotor processing in Tier 1, the average firing rate of a cortical neuron is less than 1 Hz, and there is minimal synaptic plasticity. In contrast, Tier 2 depends on constantly changing synaptic strengths, which is metabolically more expensive than the spiking activity. This is illustrated for a sensory and a motor task.

The Blood Oxygen Level Dependent (BOLD) fMRI response is a measure of overall cortical energy demand. It has been interpreted as an indirect measure of spiking activity, but there is a more parsimonious interpretation. Visual attention elicits a strong BOLD fMRI response in the visual cortex, far exceeding much smaller changes in firing rates (Boynton 2011). This long-standing paradoxical dissociation between fMRI brain imaging responses and spiking responses during visual attention could be explained by the two-tier hypothesis: Attention is a cognitive activity that requires more energy than the sensory signals that support the visual percept. Percepts normally do not change when you pay attention, but the way you interpret those signals does change because of cognitive processing.

The cortical BOLD signal is much stronger during practice when learning sequences of actions. It is greatly reduced when the behavior is automatized. This decrease in the BOLD signal has been interpreted as a shift from cortical control to automatic basal ganglia control. Alternatively, the decrease in the BOLD signal could mainly reflect reduced need for cognitive control rather than a decrease in the energy needed for the learned sensorimotor actions themselves.

The BOLD signal is strongly correlated with the local field potential (LFP), which is generated primarily by synaptic currents and perisynaptic astrocytes rather than spikes (Ekstrom 2010). Cortical oscillations and traveling waves are prominent in the LFP, which is another link between traveling waves and cognition. The cognitive interpretation of the BOLD signal has far-reaching consequences for interpreting previous experimental results and for planning future experiments.

13 Feedback projections

The primary pathway into the cortex passes through the thalamus. Visual inputs from the retina project to thalamocortical neurons in the topographically organized lateral geniculate nucleus. The thalamic projection to the cortex terminates in layer 4 with a collateral projection to layer 6. In addition to this core thalamic sensory and motor system, a second set of intralaminar matrix thalamic nuclei projects to layers 1 and 5. This thalamic pathway receives input from the brainstem reticular formation, the basal ganglia, the cerebellum, and the deep layers of the cerebral cortex, brain areas involved in arousal and cognitive functions, with major inputs conveying information about attention, pain, and, via the superior colliculus, orienting sensory signals.

Feedback from higher cortical areas terminates in layers 1 and 5/6, avoiding layer 4. The neocortex is topographically organized, but layer 1 is an exception (Fig. 14). Deep-layer 5 pyramidal neurons have apical dendrites that extend into layer 1, where they sprout dendritic tufts that span millimeters across many cortical columns. These dendrites in layer 1 receive “top-down” feedback pathways from higher cortical areas, which can trigger strong, regenerative calcium currents in the tufts that, when coinciding with bottom-up input to the basal dendrites, cause the layer 5 neuron to fire bursts of action potentials (Fig. 14) (Larkum 2013). Layer 1 is densely innervated by neuromodulatory inputs that regulate inhibitory interneurons (Ibrahim et al. 2020).

Sleep spindles during slow wave sleep originate in the thalamus and recruit cortical traveling waves in both the core thalamic and matrix pathways (Hagler et al. 2018). Sleep spindles recorded from scalp electrodes primarily capture the matrix pathway in the superficial cortex, which are closest to the skull. Spindles originating in the topographically organized core pathway recruit the more spatially distributed matrix spindles (Bonjean et al. 2012). The intralaminar nuclei have not been as well studied as the core thalamic pathways and are a “crowning mystery” (Ibrahim et al. 2020). However, the evidence points to separate thalamic systems for the two tiers: the core thalamic pathway for sensorimotor control in Tier 1 and the matrix pathway, which dominates cortical feedback, for the control of cognitive processing in Tier 2. Sleep spindles in layer 1 also suggest an involvement with memory consolidation.

Joe Bogan, the neurosurgeon who performed split-brain operations for Roger Sperry, once told me that, among the critical brain areas where lesions have a devastating effect in humans, such as the motor and language areas of the cortex, even minor lesions of the intralaminar nuclei can lead to loss of consciousness.

14 Thinking and thought disorders

You are more likely to remember something the next day if you pay attention to it. Attention releases the neuromodulator acetylcholine, shifting the state of local cortical circuits and triggering bursts of gamma waves (Fig. 8), which are grouped within lower-frequency theta waves (Howe et al. 2017). Cortical states marked by attention in the hippocampus can then be transferred to long-term memory via global traveling waves during sleep spindles (Muller et al. 2016). These traveling waves control what we remember.

Abnormal gamma waves are linked to thought disorders (Uhlhaas, and Singer, 2010). Schizophrenia patients have reduced amplitude gamma waves, which are associated with symptoms that include delusions and disorganized thinking. Similar symptoms appear after serial use of a party drug, Special K. After several days of raving under the influence of Special K, ravers present with symptoms indistinguishable from a schizophrenic break. Fortunately, the symptoms go away after a few days. Special K is ketamine, an NMDA receptor antagonist and an anesthetic at high doses. At low doses, ketamine induces hallucinations, dissociative out-of-body experiences, and reduced gamma waves. There are few memories of events while under the influence of the drug, which may be related to the blockade of NMDA receptors.

This ketamine model of human psychosis has been studied in mice, which has revealed the underlying neural mechanisms (Behrens and Sejnowski 2009). Ketamine downregulates the enzyme that synthesizes the inhibitory neurotransmitter GABA in basket cells. Reduced inhibition unbalances cortical circuits, making them hyperactive and reducing gamma activity, as observed chronically in schizophrenia patients (Jadi, Behrens, Sejnowski, 2016). Ketamine is effective at relieving clinical depression, perhaps by boosting reduced cortical activity to normal levels.

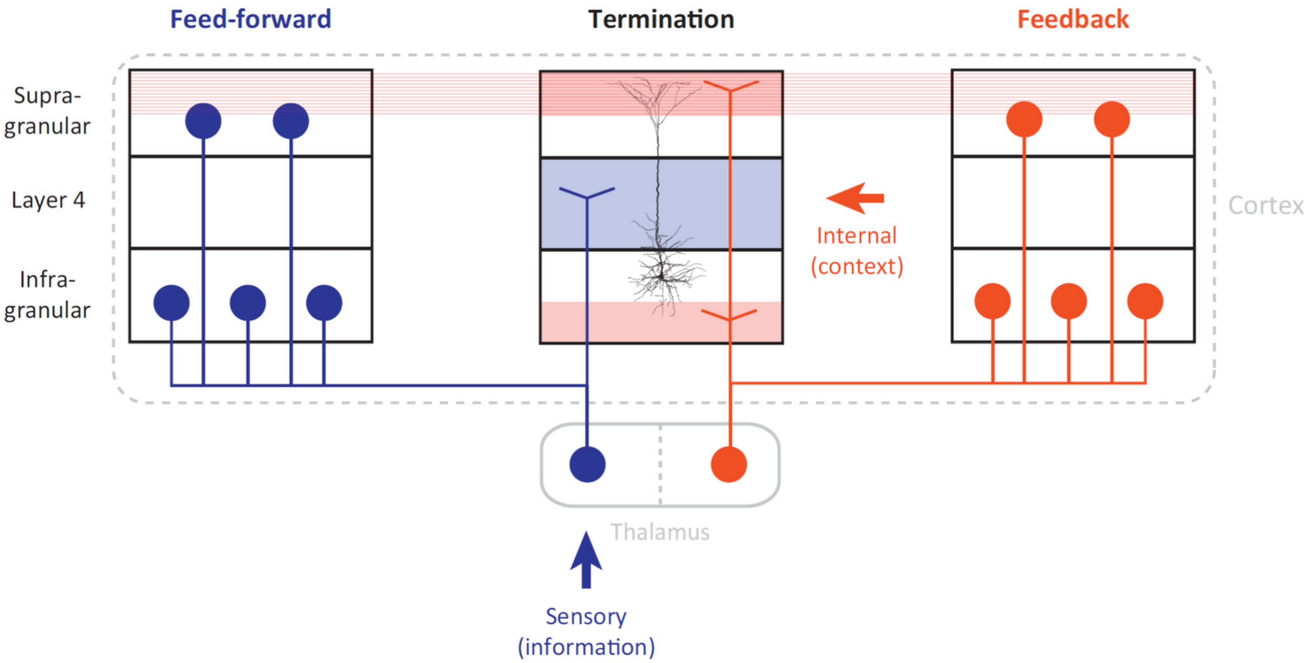

Fig. 14 Long-range architecture of feedforward and feedback connectivity between cortical areas, including thalamo-cortical interconnections. The feedforward stream is driven by external information influencing the sensory apparatus. The feedback stream is driven by an internal representation built from prior experiences. Blue neurons are core feedforward pathways (Tier 1) and red neurons are matrix feedback pathways (Tier 2). (Adapted from Larkum 2013)

15 Discussion

It has been 30 years since Mainen and Sejnowski (1995) established the millisecond precision of spike initiation in cortical pyramidal neurons in vitro. Since then, evidence has accumulated for precise spike timing in vivo, but a function for this level of precision has yet to be established. Cortical traveling waves have been reported in hundreds of papers (Muller, Chavane, Reynolds, Sejnowski, 2018). Without a known function, it has been difficult to generate much interest in cortical waves. The traditional approach to assigning function to spikes is by their impact on behavior. Because spikes on Tier 2 are sparse, they do not directly move muscles or interfere with the sensorimotor rate code. Traveling waves are more ephemeral and are not anchored to the Broadman map of cortical areas. Thinking is also ephemeral and can activate information stored throughout the cortex.

STDP depends on precise spike timing in the millisecond range. Together, waves and STDP could create long-term working memories that last for hours. Interfering with spike timing, traveling waves, or STDP should disrupt the cortical processing that underlies long-term working memory. This can be tested by reversibly blocking STDP with NMDA receptor antagonists such as ketamine, dizocilpine (MK801), and phencyclidine (PCP).

Prefrontal Cortex. The prefrontal cortex has been implicated in working memory. There is an increasing gradient in NMDA receptor density from the posterior cortex to the prefrontal cortex, especially the slow NR2B subunit, which opens for 300–400 ms (Wang 2020; Peng et al. 2025). STDP is triggered by NMDA receptors, which could extend the time scale of working memory to hours. In recordings from monkey dorsolateral prefrontal cortex, low doses of ketamine selectively disrupted sequences of activity and working memory on virtual reality tasks, but not sensorimotor performance (Roussy, et al., 2012; Busch, et al. 2024; Peng, et al. 2025). Working memory recovered when the ketamine was cleared. Blocking NMDA receptors also blocks STDP, supporting the hypothesis that long-term working memory depends on synaptic plasticity.

Reciprocal inhibitory influence between inhibitory neurons could contribute to working memory (Kim and Sejnowski 2021). Inhibitory PV basket cells have a decreasing gradient in the cortical hierarchy, parallel to an increasing gradient of somatostatin inhibitory neurons (Kim et al. 2017). Large-scale spiking cortical models could be used to explore the involvement of inhibitory neurons in controlling long-term working memory.

Noise. Does every spike matter? Reaction times to sensory stimuli have a broad distribution, suggesting noise in the sensorimotor chain. The release and recycling of neurotransmitters at cortical synapses is expensive. The release is probabilistic, with the probability of release proportional to the strength of the synapse (Schikorski and Stevens 1997). The largest synapses are the most reliable. Paradoxically, the majority of synapses are small and have the lowest release probabilities. We can appeal to the Central Limit Theorem, which states that for large sample sizes, the distribution of sample means from any distribution will be approximately Gaussian. This mitigates, but does not eliminate, synaptic noise, making possible the observed spike-timing precision in vivo (Figs. 3–5). The precision of spike timing on the cognitive Tier 2 is compatible with the presence of noise on the sensorimotor Tier 1 that does not depend on spike timing and is consistent with the mixture of precisely timed spikes and rate-coded spikes in Fig. 6. Brain noise remains an open question, but not as much as we once imagined.

Traveling waves during development. Traveling waves are ubiquitous not only in the mature brain but also during brain development. Traveling waves sculpt and refine projections in the visual pathway. Spontaneous traveling waves that crisscross the retina before eye opening and segregate projections to the left and right eye laminae in the lateral geniculate nucleus of the thalamus, and for tuning orientation-selective neurons in the cortex (Katz and Shatz, 1996; Feller et al., 1996).

Cortical neurons receive spatially correlated inputs from the direction of the traveling wave on the retina. Synaptic plasticity could be triggered as described in Section 7 and maintained by repetitive traveling waves. After eye opening, the orientation tuning is sharpened, and long-range horizontal axons selectively connect orientation columns with the same orientation tuning (Gribizis and Fitzpatrick, 2025). Traveling waves in long-range horizontal axons in other cortical areas may also connect neurons with similar properties.

Thinking. LLMs have a feedforward hierarchy of layers. In contrast, cortical hierarchies have massive feedback between layers and recurrent connectivity within layers. These feedback connections support attention, decision making, and learning, which are regulated by powerful neuromodulatory systems.

16 Conclusion

Predictions. Two intertwined working hypotheses are proposed: 1) Traveling waves create spacetime codes that encode temporal context for sequences of cortical inputs; 2) Traveling waves trigger STDP, which maintains traces of sensory inputs over time in long-term working memory on the scale of hours and supports cognitive processing.

Spacetime population codes are a type of holographic representation based on neural dynamics on timescales appropriate for encoding temporal context. If so, then it should be possible to reconstruct videos using a deep learning model trained on spike-time recordings from large-scale neural populations to predict past movie frames over time windows that increase up the cortical hierarchy.

The debate between rate coding and spike timing is a false dichotomy. There is room for multiple codes to serve multiple functions. For sensorimotor tasks, rate coding is well established. The evidence presented here for high-precision spike timing underpinning long-term working memory is provisional and needs direct experimental tests of predictions. Methods that selectively manipulate spike timing are needed. Large-scale spiking models can explore the integration of the neural mechanisms underlying long-term working memory.

Protocols that induce STDP by pairing pre- and postsynaptic bursts of spikes at 10–40 Hz mesh well with the characteristics of traveling waves (Markram et al. 1997). There are many open questions: How do neuromodulators shape temporary STDP and the long-term dynamics in the Tier 2 network? How does STDP interact with synaptic tagging and capture and active currents in dendrites (O'Donnell and Sejnowski 2014)? Do fMRI signals primarily reflect synaptic plasticity? How are working memories selected for integration into long-term memories? Could a Tier 2 temporary network form a global workspace for cognition (Baars 1988)?

Brains are self-generative in a way that does not exist in current large language models (LLMs). When your dialog with an LLM ends, its internal activity stops. We continue to think and plan without any external prompting. The mechanisms underlying self-generative activity in brains could be implemented in transformers, endowing them with what researchers in AI call System 2 capabilities (Bengio, 2017).

Many philosophical questions about the "mind" could have scientific answers if we knew more about the neural mechanisms underlying cognition. Descartes (1637) put it simply: "Cogito, ergo sum." A corollary of this dictum is: "Once we understand thinking, we will know ourselves." We should be open to more surprises as we explore the global dynamics of spiking neurons and their underlying subthreshold dynamics over a wide range of time scales.

Appendix

My Physics Ph.D. thesis at Princeton was on "A Stochastic Model of Nonlinearly Interacting Neurons," which collected several papers that I had previously published based on traditional rate-coded recurrent neural network models with a sigmoid activation function (Sejnowski 1976a). In one of these papers, I proved that the covariance equation is linear when the membrane potentials are Gaussian distributed

(Sejnowski 1976b). The interaction matrix governing the covariance is modulated by the mean activity levels of the network's nonlinearly interacting neurons, and only the sparse network of neurons near threshold propagates correlated activity (Sejnowski 1981).

Theorem *(Sejnowski 1976b): A network of neurons is stochastically interacting.*

$$\tau\frac{d}{dt}\phi_a+\phi_a=\eta_a+\sum_b K_{ab}L_b(\phi_b)$$

where ϕ_a *are the membrane potentials,* K_{ab} *is the interaction matrix,* $L_b(\phi_b)$ *is a sigmoid activation function, and* η_a *are the inputs. If the membrane potentials are Gaussian distributed, then the covariance equation is linear with the interaction matrix:*

$$K'_{ab}(t)=K_{ab}P'_b(\hat{\phi}_b(t))$$

where P'_b *is the derivative of the expectation of the sigmoid* $L_b(\phi_b)$ *with respect to the mean membrane potential* φ_b*. Intuitively,* P'_b *is nearly zero when the membrane potential is well below or well above threshold, so that covariance only depends on the interactions between those neurons near threshold. The covariance interaction matrix rapidly changes with the mean membrane potentials. Traveling waves are sparse, and only a sparse skeleton of connections between them controls the dynamics of their covariance during a wave (Sejnowski 1981). I also worked on a model of synaptic plasticity to store the covariance in long-term memory (Sejnowski 1977a, b). Covariance is a second-order measure of coordinated activity and is all that is needed for Gaussian statistics.*

This model is stochastic but not spike-based. In the Linear-Nonlinear-Poisson (LNP) cascade model, a rate-modulated Poisson generator converts a continuous signal into random spikes. Neither of these models can fully capture the higher-order statistics of spikes in neural circuits. Spike synchrony is a higher-order property, and traveling wave dynamics in a spatial array of neurons could support more complex computational systems (Keller, Muller, Sejnowski, Welling, 2024b).

My research has come full circle, back to where it began many decades ago. The missing link was spike timing, and the turning point was Mainen and Sejnowski (1995).

Acknowledgements This paper is based on a talk given on November 14, 2025, at a gathering of former lab members celebrating the 35th anniversary of the Computational Neurobiology Laboratory (CNL) at the Salk Institute. The focus was on research arising from CNL. An extended, more complete history of spike timing is planned.

Author contribution I am the sole author on this manuscript.

Funding Office of Naval Research,N00014-23-1-2069, National Institutes of Health,DP1NS149613

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no conflict of interests.

References

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. J Opt Soc Am A 2(2):284–299

Allman JM, Kaas JH (1971) A representation of the visual field in the caudal third of the middle temporal gyrus of the owl monkey (*Aotus trivirgatus*). Brain Res 31(1):85–105

Allman J, Miezin F, McGuinness E (1985) Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local–global comparisons in visual neurons. Annu Rev Neurosci. https://doi.org/10.1146/annurev.ne.08.030185.002203

Baars BJ (1988) A Cognitive Theory of Consciousness. Cambridge University Press, Cambridge

Bair W, Koch C (1996) Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. Neural Comput 8(6):1185–1202

Ballard DH, Hinton GE, Sejnowski TJ (1983) Parallel visual computation. Nature 306:21–26

Batabyal, T., Brincat, S. L., Donoghue, J. A., Lundqvist, M., Mahnke, M. K. and Miller, E. K. (2025). State–Space Trajectories and Traveling Waves Following Distraction. Journal of Cognitive Neuroscience, 1–21.

Behrens MM, Sejnowski TJ (2009) Does schizophrenia arise from oxidative dysregulation of parvalbumin-interneurons in the developing cortex? Neuropharmacology 57(3):193–200

Bengio, Y., 2017. The consciousness prior. *arXiv preprint* arXiv:1709.08568

Berry KP, Nedivi E (2017) Spine dynamics: are they all the same? Neuron 96(1):43–55

Berry. M. J., A Theory of Thalamocortical Loops in Evidence Accumulation and Decision-Making (2025). bioRxiv. https://www.biorxiv.org/content/https://doi.org/10.64898/2025.12.23.696263v1

Bhattacharya S, Brincat SL, Lundqvist M, Miller EK (2022) Traveling waves in the prefrontal cortex during working memory. PLoS Comput Biol 18(1):e1009827

Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J Neurosci 18(24):10464–10472

Bonjean M, Baker T, Bazhenov M, Cash S, Halgren E, Sejnowski TJ (2012) Interactions between core and matrix thalamocortical projections in human sleep spindle synchronization. J Neurosci 32(15):5250–5263

Boynton GM (2011) Spikes, BOLD, attention, and awareness: a comparison of electrophysiological and fMRI signals in V1. J vis 11(5):12–12

Britten KH, Shadlen, MN, Newsome, WT, Movshon JA (1992) The analysis of visual motion a comparison of neuronal and psychophysical performance. J Neurosci 12(12):4745-4765

Buracas GT, Zador AM, DeWeese MR, Albright TD (1998) Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. Neuron 20(5):959–969

Busch A, Roussy M, Luna R, Leavitt ML, Mofrad MH, Gulli RA, Corrigan B, Mináč J, Sachs AJ, Palaniyappan L, Muller L (2024) Neuronal activation sequences in lateral prefrontal cortex encode visuospatial working memory during virtual navigation. Nat Commun 15(1):4471

Buzsáki G (2019) The brain from inside out. Oxford University Press, Oxford

Chen Y, Zhang H, Cameron M, Sejnowski T (2024) Predictive sequence learning in the hippocampal formation. Neuron 112(15):2645–2658

Cruddas J, Pang JC, Fornito A (2026) Cortical traveling waves in time and space: Physics, physiology, and psychology. Neuron 114:985–1005

Davis ZW, Muller L, Martinez-Trujillo J, Sejnowski T, Reynolds JH (2020) Spontaneous travelling cortical waves gate perception in behaving primates. Nature 587(7834):432–436

Davis ZW, Benigno GB, Fletterman C, Desbordes T, Steward C, Sejnowski TJ, Reynolds JH, Muller L (2021) Spontaneous traveling waves naturally emerge from horizontal fiber time delays and travel through locally asynchronous-irregular states. Nat Commun 12(1):6057

Descartes, R. (1637). *Discours de la Méthode pour bien conduire sa raison, et chercher la vérité dans les sciences*.

Destexhe A, Sejnowski T (2023) Thalamocortical Assemblies: Sleep spindles, slow waves and epileptic discharges. Oxford University Press, Oxford

Ekstrom A (2010) How and when the fMRI BOLD signal relates to underlying neural activity: the danger in dissociation. Brain Res Rev 62(2):233–244

Ericsson KA, Kintsch W (1995) Long-term working memory. Psychol Rev 102(2):211

Fanselow EE, Sameshima K, Baccala LA, Nicolelis MA (2001) Thalamic bursting in rats during different awake behavioral states. Proc Natl Acad Sci 98(26):15330–15335

Feldman DE (2012) The spike-timing dependence of plasticity. Neuron 75(4):556–571

Fellous JM, Tiesinga PH, Thomas PJ, Sejnowski TJ (2004) Discovering spike patterns in neuronal responses. J Neurosci 24:2989–3001

Feller, M.B., Wellis, D.P., Stellwagen, D., Werblin, F.S. and Shatz, CJ (1996). Requirement for cholinergic synaptic transmission in the propagation of spontaneous retinal waves. Sci 272(5265):1182-1187.

Gilbert CD, Wiesel TN (1989) Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. J Neurosci 9(7):2432–2442

Gribizis, A and Fitzpatrick, D (2025). Retinal waves reveal axial biases in modular patterns of cortical activity that predict future orientation preferences. bioRxiv. 2025.07.09.66373

Hafting T, Fyhn M, Bonnevie T, Moser MB, Moser EI (2008) Hippocampus-independent phase precession in entorhinal grid cells. Nature 453(7199):1248–1252

Hagler DJ, Ulbert I, Wittner L, Erőss L, Madsen JR, Devinsky O, Doyle W, Fabo D, Cash SS, Halgren E (2018) Heterogeneous origins of human sleep spindles in different cortical layers. J Neurosci 38(12):3013–3025

Haider B, Krause MR, Duque A, Yu Y, Touryan J, Mazer JA, McCormick DA (2010) Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. Neuron 65(1):107–121

Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N (2008) A hierarchy of temporal receptive windows in human cortex. J Neurosci 28(10):2539–2550

Haziza S, Chrapkiewicz R, Zhang Y, Kruzhilin V, Li J, Li J, Delamare G, Swanson R, Buzsáki G, Kannan M, Vasan G, Lin MZ, Zeng H, Daigle TL, Schnitzer MJ (2025) Imaging high-frequency voltage dynamics in multiple neuron classes of behaving mammals. Cell 188(16):4401–4423

Holtmaat AJ, Trachtenberg JT, Wilbrecht L, Shepherd GM, Zhang X, Knott GW, Svoboda K (2005) Transient and persistent dendritic spines in the neocortex in vivo. Neuron 45(2):279–291

Howe WM, Gritton HJ, Lusk NA, Roberts EA, Hetrick VL, Berke JD, Sarter M (2017) Acetylcholine release in prefrontal cortex promotes gamma oscillations and theta–gamma coupling during cue detection. J Neurosci 37(12):3215–3230

Ibrahim LA, Schuman B, Bandler R, Rudy B, Fishell G (2020) Mining the jewels of the cortex's crowning mystery. Curr Opin Neurobiol 63:154–161

Izhikevich EM (2006) Polychronization: computation with spikes. Neural Comput 18(2):245–282

Izhikevich, E. M. (2025). Spiking Manifesto. arXiv:2512.11843

Jadi MP, Sejnowski TJ (2014) Cortical oscillations arise from contextual interactions that regulate sparse coding. Proc Natl Acad Sci USA 111:6780–6785

Jadi MP, Behrens MM, Sejnowski TJ (2016) Abnormal gamma oscillations in N-methyl-D-aspartate receptor hypofunction models of schizophrenia. Biol Psychiatry 79(9):716–726

Jin DZ, Fujii N, Graybiel AM (2009) Neural representation of time in cortico-basal ganglia circuits. Proc Natl Acad Sci USA 106(45):19156–19161

Johnson BA, Frostig RD (2016) Long, intrinsic horizontal axons radiating through and beyond rat barrel cortex have spatial distributions similar to horizontal spreads of activity evoked by whisker stimulation. Brain Struct Funct 221:3617–3639

Karuvally, A., Nowak, F., Keller, A.T., Alonso, C.A., Sejnowski, T.J. and Siegelmann, H.T., 2025 (2025). Bridging Expressivity and Scalability with Adaptive Unitary SSMs. *arXiv preprint* arXiv:2507.05238.

Katz, LC. and Shatz, CJ (1996). Synaptic activity and the construction of cortical circuits. Science, 274(5290):1133-1138.

Keller, T. A., Muller, L., Sejnowski, T., Welling, M. (2024a). Traveling Waves Encode the Recent Past and Enhance Sequence Learning, *International Conference on Learning Representations* (*ICLR*).

Keller, T. A., Muller, L., Sejnowski, T., Welling, M. (2024b). A space-time perspective on dynamical computation in neural information processing systems. *arXiv preprint* arXiv:2409.13669.

Kim R, Sejnowski TJ (2021) Strong inhibitory signaling underlies stable temporal dynamics and working memory in spiking neural networks. Nat Neurosci 24(1):129–139

Kim Y, Yang GR, Pradhan K, Venkataraju KU, Bota M, Del Molino LCG, Fitzgerald G, Ram K, He M, Levine JM, Mitra P, Huang AJ, Wang X-J, Osten P (2017) Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. Cell 171(2):456–469

Koller DP, Schirner M, Ritter P (2024) Human connectome topology directs cortical traveling waves and shapes frequency gradients. Nat Commun 15(1):3570

Larkum M (2013) A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. Trends Neurosci 36(3):141–151

Loewenstein Y, Kuras A, Rumpel S (2011) Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. J Neurosci 31(26):9481–9488

Lubenov EV, Siapas AG (2009) Hippocampal theta oscillations are travelling waves. Nature 459(7246):534–539

Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. Science 268:1503–1506

Makeig S, Westerfield M, Jung TP, Enghoff S, Townsend J, Courchesne E, Sejnowski TJ (2002) Dynamic brain sources of visual evoked responses. Science 295(5555):690–694

Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science 275(5297):213–215

Marr D, Poggio T (1976) Cooperative computation of stereo disparity: a cooperative algorithm is derived for extracting disparity information from stereo image pairs. Science 194(4262):283–287

Miller DA, Zucker SW (1999) Computing with self-excitatory cliques: a model and an application to hyperacuity-scale computation in visual cortex. Neural Comput 11(1):21–66

Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16(5):1936–1947

Muller L, Piantoni G, Koller D, Cash SS, Halgren E, Sejnowski TJ (2016) Rotating waves during human sleep spindles organize global patterns of activity that repeat precisely through the night. Elife 5:e17267

Muller L, Chavane F, Reynolds J, Sejnowski TJ (2018) Cortical travelling waves: mechanisms and computational principles. Nat Rev Neurosci 19(5):255–268

Muller L, Churchland PS, Sejnowski TJ (2024) Transformers and cortical waves: encoders for pulling in context across time. Trends Neurosci 47(10):788–802

O'Donnell C, Sejnowski TJ (2014) Selective memory generalization by spatial patterning of protein synthesis. Neuron 82(2):398–412

Ortuño T, Grieve KL, Cao R, Cudeiro J, Rivadulla C (2014) Bursting thalamic responses in awake monkey contribute to visual detection and are modulated by corticofugal feedback. Front Behav Neurosci 8:198

Patel J, Fujisawa S, Berényi A, Royer S, Buzsáki G (2012) Traveling theta waves along the entire septotemporal axis of the hippocampus. Neuron 75(3):410–417

Peng J, Wang G, Han Y, Qu L, Wang M, Arnsten AF, Cai J, Li B (2025) Prefrontal cortical NR2B-containing NMDA receptors are essential for spatial working memory performance. Transl Psychiatry 15(1):357

Perin R, Berger TK, Markram H (2011) A synaptic organizing principle for cortical neuronal groups. Proc Natl Acad Sci USA 108(13):5419–5424

Pribram KH (1971) Languages of the Brain: Experimental Paradoxes and Principles in Neuropsychology. Prentice-Hall, Englewood Cliffs

Qasim SE, Fried I, Jacobs J (2021) Phase precession in the human hippocampus and entorhinal cortex. Cell 184(12):3242–3255

Reichardt W (1957) Autokorrelations-Auswertung als Funktionsprinzip des Zentralnervensystems: (bei der optischen Bewegungswahrnehmung eines Insektes). Z Naturforsch B 12(7):448–457

Reinagel P, Reid RC (2002) Precise firing events are conserved across neurons. J Neurosci 22(16):6837–6841

Roussy M, Luna R, Duong L, Corrigan B, Gulli RA, Nogueira R, Moreno-Bote R, Sachs AJ, Palaniyappan L, Martinez-Trujillo JC (2021) Ketamine disrupts naturalistic coding of working memory in primate lateral prefrontal cortex networks. Mol Psychiatry 26(11):6688–6703

Schikorski T, Stevens CF (1997) Quantitative ultrastructural analysis of hippocampal excitatory synapses. J Neurosci 17:5858–5867

Sejnowski TJ (1976a) On global properties of neuronal interaction. Biol Cybern 22:85–95

Sejnowski TJ (1976b) On the stochastic dynamics of neuronal interaction. Biol Cybern 22:203–211

Sejnowski TJ (1977a) Statistical constraints on synaptic plasticity. J Theor Biol 69:385–389

Sejnowski TJ (1977b) Storing covariance with nonlinearly interacting neurons. J Math Biol 4:203–211

Sejnowski TJ (1981) Skeleton filters in the brain. In: Hinton GE, Anderson JA (eds) Parallel models of associative memory. Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp 189–212

Sejnowski, T. J., (2025) Thinking about thinking: AI offers theoretical insights into human memory, The Transmitter, https://www.thetransmitter.org/human-neurotechnology/thinking-about-thinking-ai-offerstheoretical-insights-into-human-memory/

Singer W (2021) Recurrent dynamics in the cerebral cortex: integration of sensory evidence with stored knowledge. Proc Natl Acad Sci USA 118(33):e2101043118

Steriade M, McCormick DA, Sejnowski TJ (1993) Thalamocortical oscillations in the sleeping and aroused brain. Science 262:679–685

Tiesinga P, Fellous JM, Sejnowski TJ (2008) Regulation of spike timing in visual cortical circuits. Nat Rev Neurosci 9:97–109

Van Vreeswijk C, Abbott LF, Ermentrout BG (1994) When inhibition not excitation synchronizes neural firing. J Comput Neurosci 1(4):313–321

Tiesinga P, Sejnowski TJ (2009) Cortical enlightenment: are attentional gamma oscillations driven by ING or PING? Neuron 63(6):727–732

Uhlhaas PJ, Singer W (2010) Abnormal neural oscillations and synchrony in schizophrenia. Nature Reviews Neuroscience 11(2):100–113

Wagner, M., Chen, Y., Karuvally, A., Cameron, M. and Sejnowski, T. J. 2026. Balancing Stability and Flow in Hippocampal Networks via Inductive Bias and Learned Symmetry Breaking, *Philosophical Transactions*, bioRxiv

Wang XJ (2020) Macroscopic gradients of synaptic excitation and inhibition in the neocortex. Nat Rev Neurosci 21(3):169–178

Wang HP, Spencer D, Fellous JM, Sejnowski TJ (2010) Synchrony of thalamocortical inputs maximizes cortical reliability. Science 328(5974):106–109

Yang G, Lai CSW, Cichon J, Ma L, Li W, Gan WB (2014) Sleep promotes branch-specific formation of dendritic spines after learning. Science 344(6188):1173–1178

Ye, Z., Bull, M. S., Li, A., Birman, D., Daigle, T. L., Tasic, B., Zeng, H., Steinmetz, N. A. (2023). Brain-wide topographic coordination of traveling spiral waves. *bioRxiv*, 2023–12.

Yoshimura Y, Dantzker JL, Callaway EM (2005) Excitatory cortical neurons form fine-scale functional networks. Nature 433(7028):868–873

Zikopoulos B, Barbas H (2006) Prefrontal projections to the thalamic reticular nucleus form a unique circuit for attentional mechanisms. J Neurosci 26(28):7348–7361