Switches from tonic to burst firing enable memory consolidation through late-phase synaptic plasticity

Authors
Kathleen Jacquerie[1,2], Danil Tyulmankov[3,4], Pierre Sacré[1] and Guillaume Drion[1]

Author affiliations
[1]Biology Department, Brandeis University, Waltham, MA, United States
[2]Department of Electrical Engineering and Computer Science, University of Liège, Liège, Belgium
[3]Center for Theoretical Neuroscience, Columbia University, New York, NY, USA
[4]Viterbi School of Engineering, University of Southern California, Los Angeles, CA, USA

Corresponding authors
Kathleen Jacquerie: kathleen.jacquerie@gmail.com

Declaration of interests
The authors declare no competing interests.

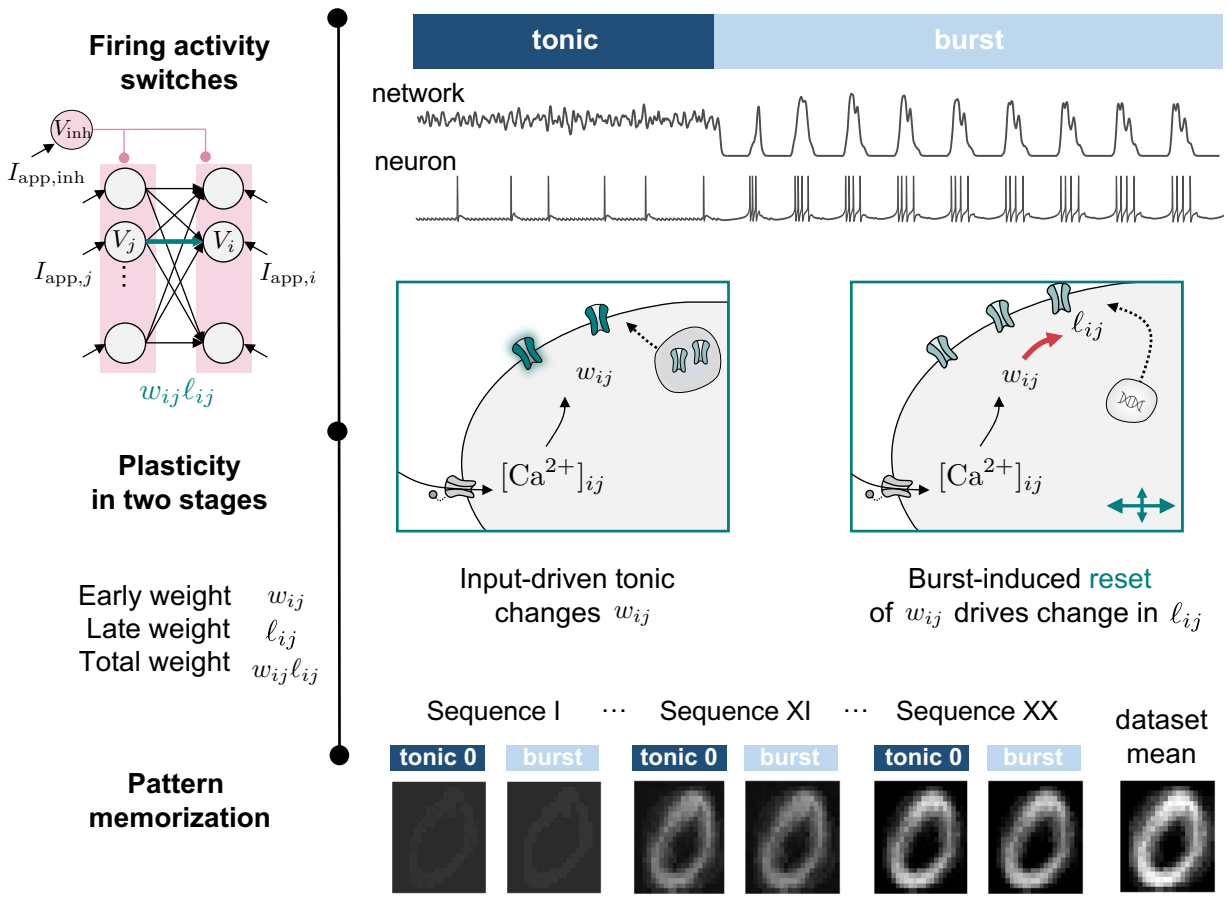
Acknowledgments
Kathleen Jacquerie was a Research Fellow of the Fonds de la Recherche Scientifique – FNRS. This work was supported by the Belgian Government through the Federal Public Service Policy and Support. Thanks to Juliette Ponnet, Nora Benghalem, Justine Magis, and Emmy Kellens who worked on this project during their Master's theses, and to Caroline Minne who focused on the synaptic weight reset.

Declaration of generative AI and AI-assisted technologies
ChatGPT was used to correct the spelling.

Author contributions
Conceptualization, K.J., G.D.;
Methodology, K.J., D.T., G.D.;
Software, K.J.;
Writing – Original Draft, K.J.;
Writing – Review and Editing, K.J., G.D., P.S., D.T.;
Visualization, K.J.;
Supervision, G.D., P.S., D.T.;
Funding Acquisition, K.J., G.D., P.S.

Keywords
synaptic plasticity, *calcium-dependent plasticity*, *late-phase plasticity*, *brain states*, *burst*, *neuromodulation*, *memory consolidation*

Graphical abstract



1 Summary

This study explores how alternating periods of tonic and burst firing phases in neural networks contribute to memory consolidation. We develop a biophysical neural network model that leverages these brain state fluctuations to investigate the interplay between neuronal firing shifts and synaptic plasticity. The model introduces a two-stage plasticity rule, combining conventional early-phase plasticity with a novel burst-driven late-phase plasticity. Using a learning protocol where tonic firing is associated with active learning periods and burst firing is associated with rest periods, we demonstrate that burst firing resets early-phase plasticity, enabling new memory formation during subsequent tonic-firing learning periods and memory consolidation through late-phase plasticity. We further confirm through a pattern recognition task that replacing burst firing with additional tonic firing or quiescent periods hinders memory consolidation. These findings suggest a mechanistic role for burst firing in memory consolidation and its importance for both computational models and experimental studies of learning.

2 Introduction

Biological neuronal networks undergo global changes in rhythmic activity that correspond to different brain states, which are defined by switches in neuronal excitability and orchestrated by neuromodulators (McGinley et al., 2015). Two such distinct brain states can be characterized by different global activity patterns: fast, low-amplitude oscillations and slow, large-amplitude oscillations, as seen in local field potential (LFP) or electroencephalogram (EEG) recordings. At the neuronal level, these states correspond to tonic firing – where action potentials are triggered by input signals – and endogenous collective bursting – a rapid succession of action potentials followed by a silent period. Endogenous bursting means that the bursts are generated by the neuron's own internal mechanisms, without relying on external synaptic input, and occur collectively in the network (McGinley et al., 2015; McCormick and Salkoff, 2015; Lee and Dan, 2012; Dastgheib et al., 2022). At the behavioral level, tonic firing has been shown to be an important regime during active learning (Sherman, 2001). Meanwhile, burst firing may correspond to states of rest (McGinley et al., 2015) and memory consolidation (Timofeev et al., 2002).

In each of these states, synaptic plasticity enables memory formation by adjusting connections between neurons based on their firing activity. Synaptic plasticity encompasses complex mechanisms occurring in two different phases (Citri and Malenka, 2008). Early-phase plasticity primarily involves three mechanisms: an increase in postsynaptic receptor efficacy, a rapid insertion of available receptors via exocytosis, and a restoration of postsynaptic receptor efficacy to baseline (Poirazi and Mel, 2001; Lamprecht and LeDoux, 2004; Magee and Grienberger, 2020). Late-phase plasticity occurs through *de novo* protein synthesis and structural changes in the synapse, also known as structural plasticity (Fauth and Tetzlaff, 2016).

So far, however, each of these phases has only been studied in isolation. The interaction between switches from tonic to burst firing and synaptic plasticity has not been deeply investigated in computational models for several reasons. First, it requires a network of neurons capable of robustly switching between different firing regimes in the presence of neuronal and synaptic variability. Second, a major challenge is the choice of synaptic plasticity rules. Previous modeling studies have captured early- and late-phase plasticity using highly detailed biophysical models (Smolen et al., 2006) or focusing solely on tonic firing activity using phenomenological rules (Luboeinski and Tetzlaff, 2021; Clopath et al., 2008; Zenke et al., 2015). However, these models are either too computationally intensive to scale to larger neural networks or are designed exclusively for tonic firing, making them unsuitable for our goal of integrating collective burst firing in a large network.

To tackle this challenge, we use a biophysical conductance-based neural network model capable of robustly transitioning between tonic and burst firing regimes (Jacquerie and Drion, 2021). As a key contribution, we develop a model of memory consolidation that combines early- and late-phase plasticity: we model the connection between neurons using a synaptic weight that is split into the product of two terms, the early-phase synaptic weight (early weight) and the late-phase synaptic weight (late weight). To drive change in the early weight, we use the early-phase calcium-based model introduced by Graupner and Brunel (2012), and propose a novel burst-driven late-phase plasticity mechanism: the change in the late weight is a function of the change in early weight, proportional to its negative derivative. This coupling models the transfer of memory from early-phase to late-phase plasticity using a single parameter. Such coupling of internal synaptic variables has previously been suggested by others, using more abstract models (Fusi et al., 2005; Benna and Fusi, 2016).

We test our model on a pattern recognition task where the network stores memory patterns during tonic firing and consolidates them during burst firing. Our results show enhanced memory consolidation due to the synergistic role of late-phase plasticity and the transition from tonic to burst firing, which improves both classification accuracy and generalization. Through ablation experiments, we demonstrate the importance of each component in our model: (1) we compare our model to a modified version where burst firing is replaced by additional tonic firing, replicating additional learning periods; (2) we block late-phase plasticity, allowing only early-phase plasticity during the transition from tonic to burst firing; (3) we replace burst firing with a quiescent period, a common alternative for modeling rest states. Finally, we investigate the impact of the parameter governing the timescale of late-phase plasticity, as well as the choice of the early-phase plasticity model on memory consolidation.

3 Results

3.1 Modeling robust transitions between tonic and burst firing in a biophysical network model

We simulate the transition from tonic to burst firing in a network of conductance-based model neurons (subsubsection 5.3.1), with the network architecture shown in Figures 1A-B. The network is composed of N presynaptic excitatory neurons projecting to M postsynaptic excitatory neurons with membrane potentials V_j and V_i, respectively. These neurons each receive GABA currents from one inhibitory neuron (V_{inh}), the activity of which controls whether the network is in a tonic firing or bursting regime. Each neuron also receives an individual applied current ($I_{\text{app,inh}}, I_{\text{app},j}, I_{\text{app},i}$) to set its firing rate for learning specified input patterns.

Figure 1A illustrates the transition between the two firing modes in a network of 50 presynaptic and 50 postsynaptic neurons: *input-driven tonic firing* and *collective burst firing*, controlled by the current $I_{\text{app,inh}}$ applied to the inhibitory neuron. This modulation in inhibitory levels mimics the variation of neuromodulators such as acetylcholine, dopamine, serotonin, or histamine (Tyree and Luis De Lecea, 2017), known to control brain state transitions (Hill and Tononi, 2005; Bazhenov et al., 2002; Olcese et al., 2010; Jacquerie and Drion, 2021; Drion et al., 2018).

During input-driven tonic firing, a depolarizing current $I_{\text{app,inh}}$ applied to the inhibitory neuron induces regular tonic spiking activity, as commonly observed in inhibitory pacemaking cortical neurons (McGinley et al., 2015). Consequently, this pacemaking inhibitory neuron sets the resting membrane voltage of all the excitatory neurons. Upon receiving an applied current ($I_{\text{app},j}, I_{\text{app},i}$), these excitatory neurons generate action potentials.

To trigger a network transition from input-driven tonic firing to collective burst firing, a hyperpolarizing current $I_{\text{app,inh}}$ is applied to the inhibitory neuron. This bursting activity – in which each neuron generates a rapid succession of action potentials followed by a quiescent period (Zeldenrust et al., 2018; Desroches et al., 2022; Fernandez and Luthi, 2019) – is independent of external input currents $I_{\text{app},j}$ and $I_{\text{app},i}$, and is caused by neuron intrinsic properties, such as T-type calcium channels that deinactivate when the neuron is hyperpolarized. The burst frequency, the intraburst frequency, and the number of spikes per burst are quantified in Figure SI 7. The term "collective" refers to the synchronization of individual bursting neurons despite heterogeneity in intrinsic properties and synap-

tic connections, resulting in slow large-amplitude population activity shown in the LFP oscillation, contrasting the fast small-amplitude observed in input-driven tonic firing (Figure 1A, LFP traces) (Jacquerie and Drion, 2021).

3.2 Modeling early- and late-phase synaptic plasticity

The synapses projecting from the presynaptic to the postsynaptic excitatory neurons undergo plasticity. In biological neurons, plasticity involves mechanisms in two phases: an early phase and a late phase, relying on different plasticity mechanisms and signaling pathways (Poirazi and Mel, 2001; Lamprecht and LeDoux, 2004; Fauth and Tetzlaff, 2016; Kotaleski and Blackwell, 2010; Feldman, 2020; Bading, 2013; Citri and Malenka, 2008; Berridge, 2014; Heidelberger et al., 2014; Golbert et al., 2017; Seibt and Frank, 2019). To model these two phases of plasticity between a presynaptic neuron j and a postsynaptic neuron i, we split the synaptic weight into two components: the early-phase synaptic

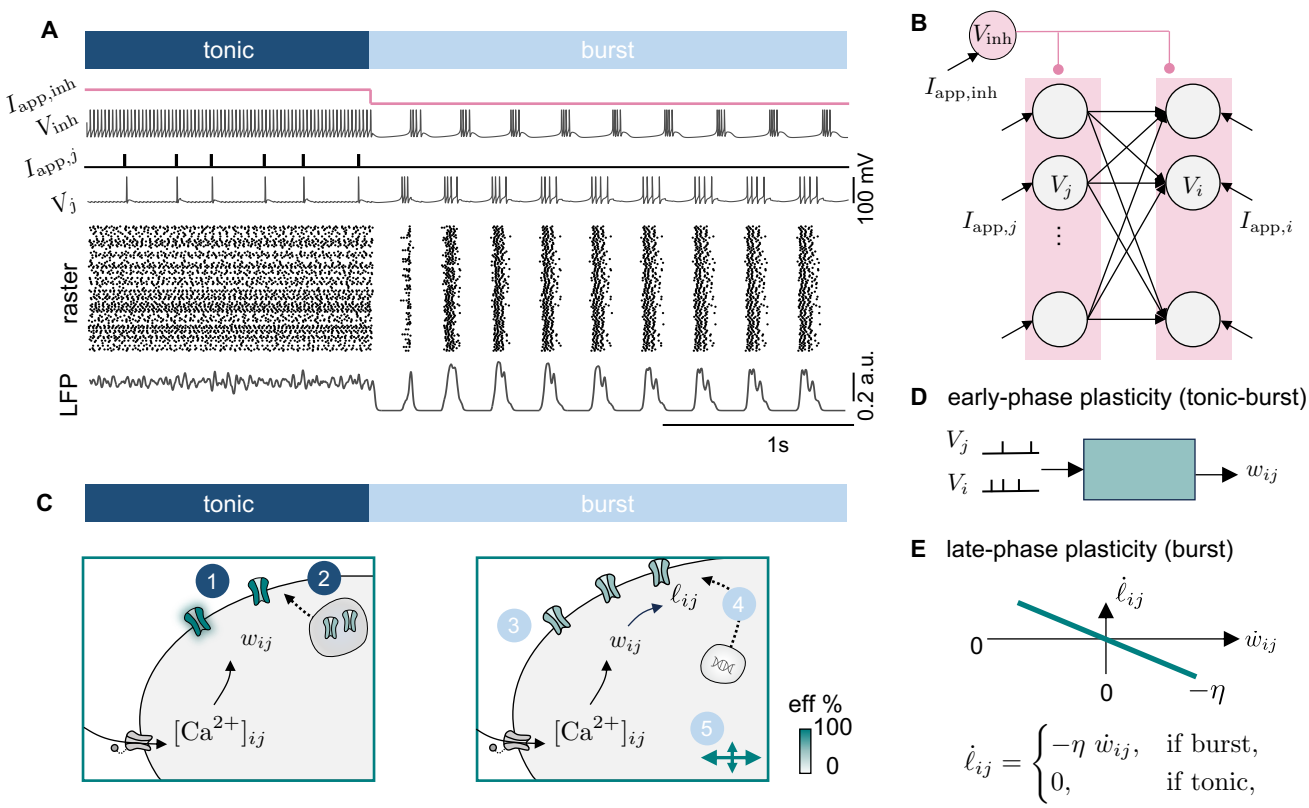


Figure 1: A biophysical network model of tonic and burst firing with a two-stage synaptic plasticity. A. A. Hyperpolarizing the current applied to the inhibitory neuron ($I_{\text{app,inh}}$) switches the network state from input-driving tonic firing to collective burst firing. Raster plot activity of the excitatory neurons and local field potential (LFP) traces are displayed. **B.** The network model consists of one inhibitory neuron (V_{inh}) projecting to all the excitatory neurons, connected in a feedforward configuration with N presynaptic neurons (V_j) to M postsynaptic neurons (V_i). An external current applied to the inhibitory neuron controls the network state ($I_{\text{app,inh}}$). Each excitatory neuron is subjected to an individual external current mimicking an external input. **C.** Schematic of possible synaptic mechanisms in a tonic firing: (1) increase in the efficacy (eff) of the postsynaptic receptors, (2) fast insertion of available postsynaptic receptors, and in a burst firing: (3) restoration of the receptor efficacy, (4) generation and insertion of new receptors, and (5) morphological changes of the synapse. From a modeling point of view, in tonic firing, calcium fluctuations resulting from presynaptic and postsynaptic input-driven activity trigger changes in the early weight (w_{ij}). Conversely, during burst firing, calcium fluctuations reflect collective activity that induces alterations in the early weight, activating downstream late-phase synaptic changes modeled by the late weight (ℓ_{ij}). **D.** The plasticity mechanisms operating in the early phase, active during both tonic and burst firing, regulate the early weight w_{ij} based on the pre- and postsynaptic activity (Graupner and Brunel, 2012). **E.** The late-phase plasticity, exclusively present during burst firing, governs the alteration of the late weight ℓ_{ij} in an opposite and proportional manner to the change observed in the early weight, with the late weight plasticity gain η.

weight (early weight, w_{ij}) and the late-phase synaptic weight (late weight, ℓ_{ij}), with the total synaptic weight defined as the product of the two. During tonic firing, the early weights are primarily governed by neuron correlations, as established in the literature (Abbott and Nelson, 2000). During burst firing, network activity supports the transfer of acquired learning into late-phase synaptic changes.

An increase in the early weight corresponds to enhanced postsynaptic receptor efficacy – for example by receptor phosphorylation or sensitivity increase (Lee, 2006)– and/or rapid insertion of available receptors via exocytosis (Bear et al., 2016; Seibt and Frank, 2019; Lamprecht and LeDoux, 2004) (Figure 1C, plasticity mechanisms 1 and 2). Existing calcium-based (Graupner and Brunel, 2012; Shouval et al., 2002) or time-dependent early-phase plasticity rules (Zenke and Gerstner, 2017) have captured these plasticity mechanisms by using the activity of the pre- and postsynaptic neurons to drive changes in the early weight (Figure 1D). Here, we use the calcium-based plasticity rule established in (Graupner and Brunel, 2012) for the majority of our simulations (Equation 2), although our results hold for a wide range of early-phase plasticity rules (subsection 3.6). The early-phase plasticity rule is active during both tonic and burst firing.

An increase in the late weight corresponds to protein synthesis or synaptic morphological changes (Pai et al., 2013; Citri and Malenka, 2008; Heidelberger et al., 2014; Golbert et al., 2017; Lamprecht and LeDoux, 2004) (Figure 1C, plasticity mechanisms 4 and 5). To model late-phase synaptic plasticity, previous studies have established models operating only during tonic activity, using either highly detailed biophysical equations representing protein actions (Smolen et al., 2006) or more abstract formulations (Zenke et al., 2015; Fauth and Van Rossum, 2019; Clopath et al., 2008). Here, we develop a simplified model such that the rate of change of the late weight is proportional to the negative rate of change of the early weight (Figure 1E):

$$\dot{\ell}_{ij} = \begin{cases} -\eta\ \dot{w}_{ij}, & \text{if burst,} \\ 0, & \text{if tonic.} \end{cases} \tag{1}$$

The parameter $\eta > 0$ represents the late-weight plasticity gain. Increasing η results in more pronounced modifications in the late weight, which can be interpreted as a greater quantity of synthesized proteins or more significant structural changes in the synapse. Despite its simplicity, this model encapsulates biological mechanisms behind the late-phase plasticity. The rise of calcium level activates calcium-calmodulin-dependent protein kinase II (CaMKII) or other members of the CaMK family that, in turn, activates cyclic cyclic adenosine monophosphate (cAMP) (Golbert et al., 2017; Lamprecht and LeDoux, 2004). This intracellular signaling molecule activates the production of Protein Kinase A that later enters the nucleus and activates transcription factors (Lamprecht and LeDoux, 2004). These can induce the expression of genes that encode new proteins. Altogether this process contributes to long-term morphological changes such as enlargement of the spine or insertion of new proteins.

This model relies on two hypotheses. The first suggests that there is a negative proportional relationship between the rate of change of the early and late weights. This is supported by the observation that learning acquired during tonic firing manifests as changes in synaptic efficacy, triggering downstream signaling pathways. The negative sign in the late-phase synaptic plasticity equation reflects the fact that the transfer of acquired learning during burst firing must occur while synaptic dynamic range is restored by a reset of synaptic efficacy (Figure 1C, plasticity mechanism 3). This restoration of synaptic modification range is required to avoid early saturation of potentiation and depression. Second, this rule is only active during burst firing. This hypothesis stems from the observation that the switch in neuronal activity from tonic to bursting is induced by neuromodulators, which may in parallel activate the late phase of synaptic plasticity.

3.3 Burst-driven late-phase plasticity enables memory consolidation

We first test our proposed model on a pattern recognition task. We use the classical MNIST image dataset (LeCun et al., 1998), which not only allows tractable model simulations but also enables an interpretable visualization of neuron learned synaptic weights (Figure 2). The task is learning to classify each of 10 digits based on an image of a handwritten version of that digit. Similar computational protocols have been explored in previous studies focusing on classification (Gurunathan and

Iyer, 2020; Zenke et al., 2015), including the MNIST dataset (Capone et al., 2019; Garg et al., 2022). The network is composed of 484 presynaptic neurons, each encoding a single pixel in an image (22×22 pixels, downsampled from the original size for computational efficiency), and 10 postsynaptic neurons, each coding for one of the corresponding digit classes ($N = 484$, $M = 10$).

The stimulation protocol is divided into 21 sequences (Figure 2, Roman numerals), with each sequence composed of 10 blocks. Each block consists of a learning period, followed by a rest period. The network performs supervised learning during input-driven tonic firing (Figure 2A, dark blue), where presynaptic neurons are driven by an external input such that their spiking frequency encodes pixel intensity, and the postsynaptic neuron corresponding to the presented digit class is driven to create a "one-hot" encoding of the class. These periods of active learning are interspersed with periods of rest, during which no inputs are shown to the network and the inhibitory neuron is hyperpolarized, which leads to bursting activity in the excitatory neurons (Figure 2A, light blue). As a control experiment (see below) we also consider a period of network quiescence (Figure 2A, lilac) in place of bursting (subsubsection 5.3.8). For sequences I through XX, in block i (for $i = 0, 1, \ldots, 9$), a randomly sampled image of digit i is presented to the network during the learning period. Sequence XXI is designed to test network sensitivity to noise (Figure 2A, grey), where presynaptic neurons, rather than being shown an image, are driven by random noise (subsubsection 5.3.8).

To visualize the learning progress, we track the synaptic weights associated with each digit, *i.e.*, the weights connecting all presynaptic neurons to the postsynaptic neuron that encodes that digit (Figure 2B-E). The value of each element is proportional to the total synaptic weight (product of early- and late-phase weights) at a given time, $(w\ell)_i \equiv w_{i:} \odot \ell_{i:}$, where $\odot$ is the element-wise product and ":" indicates all indices j. For visualization, these 484 synaptic weights are reshaped into a 22×22 matrix, with digits 0 and 9 shown in Figure 2B-E, and the remaining ones in Figure SI 8.

Our model, alternating tonic and burst firing states together with late phase synaptic plasticity ("tonic-burst plus late" model, TB+L) highlights the beneficial role of integrating burst firing periods during learning. We select 10-time points to display the weights, demonstrating the network evolution during the succession (Figure 2B). For sequence I, the weights for digits 0 and 9 are shown at the end of the tonic and burst states of block 0 (where digit 0 is learned). For sequence XI, the weights are displayed at the end of the tonic firing state where digits 0 and 9 are learned (tonic 0 and tonic 9) and their respective following burst firing states. For sequence XX, we show the last block. For the final sequence XXI, the weights are shown after the network was exposed to noise and a final burst firing state. First, the weight matrices strengthen their pixels associated with the digit from one sequence to another, indicating the network ability to learn. Second, burst firing facilitates memory consolidation: the weight matrices exhibit better contrast after the bursting state compared to its preceding tonic firing state (more details in section 3.7), even if no new inputs are shown during that period. Importantly, this memory consolidation occurs without requiring any reactivation of specific neuronal patterns associated with the stored information (Yaffe et al., 2014). Lastly, as demonstrated by the weight matrices in Sequence XXI, the network is resilient to noise.

3.4 Late-phase plasticity and burst firing are both necessary for memory consolidation

To illustrate the importance of alternating between periods of tonic and burst firing for memory consolidation, we replace each burst state with a tonic state ("tonic-tonic" model, TT). Consequently, the synapses are updated using another sample of the same digit instead of transitioning to burst firing (Figures 2C and SI 9). Although one might expect better learning because the network sees twice as many samples, replacing burst firing with additional tonic firing hinders memory consolidation. In the weight matrices in Sequences I–XX, the network learns only the current digit sample, erasing the memory of the previous one. Moreover, removal of the burst firing states renders the network vulnerable to noise, as shown by the weight matrices in sequence XXI.

To demonstrate the necessity of the late-phase plasticity, we block it by setting $\dot{\ell} = 0$ in the TB+L model ("tonic-burst" model, TB). This is analogous to experimental work in which inhibition of protein synthesis is shown to suppress long-term potentiation (Frey et al., 1993; Krug et al., 1984)

(subsubsection 6.2.2 for additional discussion). It results in a reset of the weights at each switch to collective bursting (Figures 2D and SI 10) – although the network is capable of learning, the acquired information vanishes during bursts. This configuration also makes the network vulnerable to noise.

Finally, rather than collective burst firing, periods of rest may be modeled by a quiescent network not activated by external inputs (Fauth and Van Rossum, 2019; González-Rueda et al., 2018). To demonstrate the necessity of bursting, we replace this firing state with quiescent neurons ("tonic-quiescent" model, TQ) by maintaining the depolarizing current applied to the inhibitory neuron (subsubsection 5.3.2). In this configuration, the network only retains the last presented digit in memory (Figures 2E, lilac and SI 11). However, the memory fades unless another training sample is shown, and the network remains susceptible to noise.

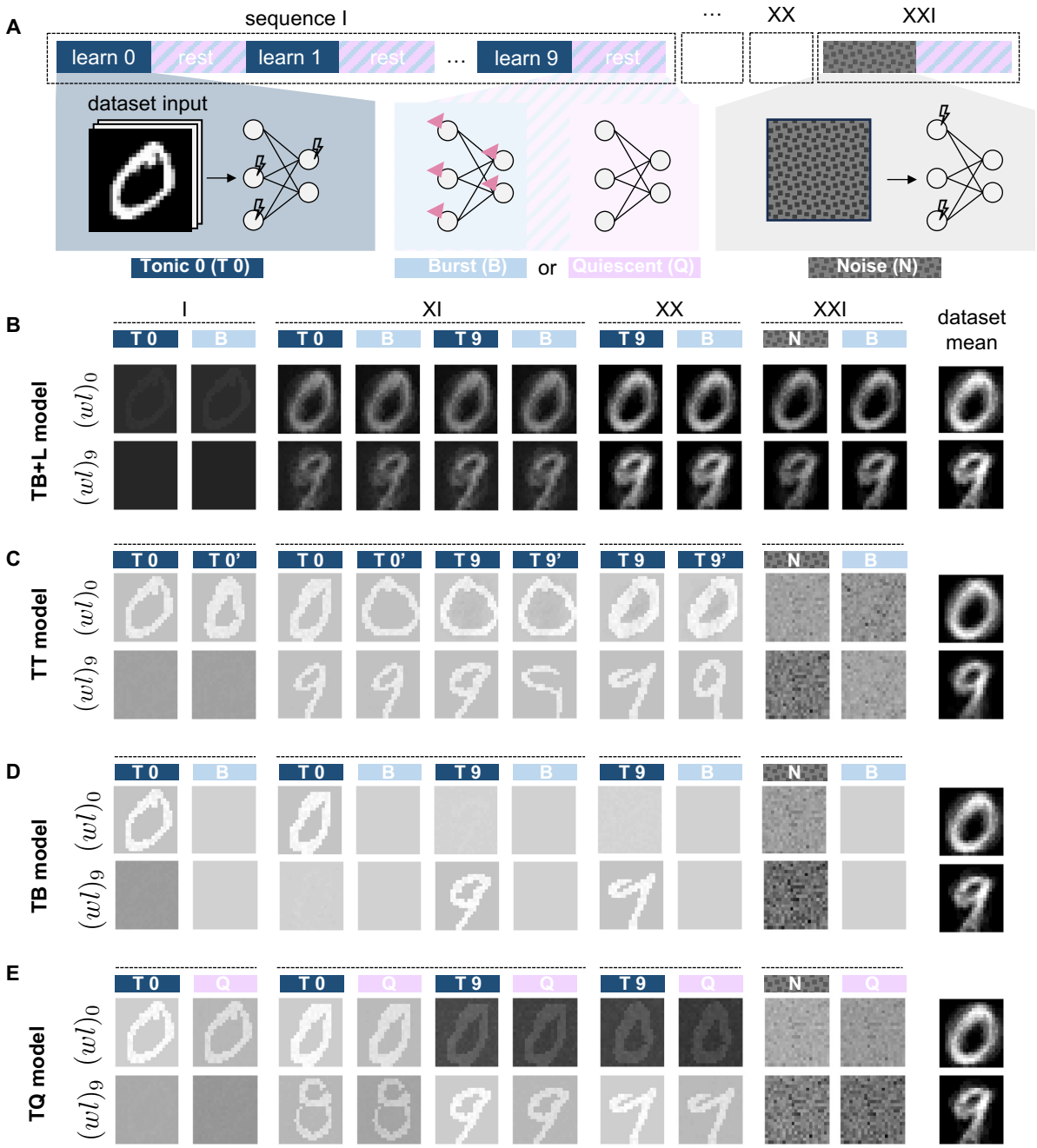


Figure 2: (Caption on the next page)

Figure 2: Burst-driven late-phase plasticity enables memory consolidation. A. A neural network is trained to classify digits from the MNIST dataset by interleaving periods of learning (input-driven tonic firing, shown by the lightning symbol) and resting (collective burst firing, shown by the pink triangle in the light blue box or quiescent neurons in the lilac box). It consists of 20 sequences (I-XX) of sequential training and one last sequence with only noise and rest (XXI). **B.** The weight matrices are shown for the output neurons associated with digits 0 and 9 at the transition from tonic to burst firing within one sequence (noted $(w\ell)_0$ and $(w\ell)_9$). The grey color scale is normalized between the minimum and maximum weight of the simulation. The network consolidates acquired learning during burst firing and it is robust to noise. The last column is the mean of the 20 dataset images for the digits 0 and 9 presented during the 20 sequences. **C.** Rest periods are replaced by additional tonic firing periods for sequences I to XX (T i'). The network only encodes the current digit, forgets previous learning, and is fragile to noise. The last column is the mean of the 40 dataset images presented during the 40 tonic firing states. **D.** Late-phase plasticity is blocked during burst firing periods ($\hat{\ell} = 0$). The network only encodes the digit just shown during the tonic state and then resets all its early weights. The network is fragile to noise. **E.** Burst firing periods are replaced by quiescent periods. The network just remembers the last representation of the digit and changes its representation each time a new sample of the same digit is presented. The memory fades throughout the simulation, and the network is fragile to noise.

3.5 Learning and generalization are improved by burst-driven late-phase plasticity

To quantify the learning process, we compute the correlation between the weights encoding each digit and the mean of the corresponding images in the training dataset over the course of learning (the mean of the 20 training images for digits 0 and 9 is shown in the last column of Figure 2B-E). In our model (TB+L) the correlation for a given digit increases with each state that the network is trained on. After just a few training sequences, each digit achieves a high correlation, remaining close to 1 even in the presence of noise (Figure 3A). In contrast, replacing rest periods with additional learning periods (TT model) significantly impacts the correlation evolution. The correlation for a given digit rises when the digit is presented in the sequence, but then declines when the subsequent digit is shown (Figure 3B). This confirms that the network fails to retain the information over a long period. When exposed to noise, the correlation drops for all digits, causing the network to forget its training. When late-phase plasticity is blocked (TB model), the correlation peaks only during the tonic state, dropping as it resets (Figure 3C). Similarly, replacing burst firing with quiescent neurons (TQ model) leads to an initial increase in correlation upon digit presentation, followed by a subsequent fade-out (Figure 3D).

To evaluate the learning performance, we measure the network ability to classify MNIST digits for each of these four models. At the end of each block, we present 20 samples of each digit 0 to 3. These samples are either samples seen during learning in Figure 2A (training set) or unseen samples (test set). Focusing on these 4 digits allows us to test whether the network can maintain its ability to recognize previously learned digits while simultaneously learning new patterns. For the purpose of evaluation, the network does not undergo plasticity during these testing periods. We evaluate the prediction determined by the neuron with the highest spike count, and accuracy is calculated as the percentage of correct predictions. To guarantee a fair comparison, we normalize the weights in models lacking late-phase plasticity by the maximum weight value from the TB+L model (more details in subsubsection 5.3.9 and see subsection 6.3 for unnormalized performance).

On the training set, the TB+L model demonstrates strong learning performance, with accuracy steadily increasing over time and surpassing that of other models (Figure 3E, dark teal trace). Without late-phase plasticity (TB, orange trace), the network forgets during burst states, as indicated by the sawtooth-shaped curve and the accuracy dropping close to 0%. This shows that burst firing without late-phase plasticity leads to a reset of the acquired learning (the mechanism is discussed in detail in subsection 3.6). Replacing burst-firing neurons with tonic-firing neurons (TT model, light teal) does not improve performance, as the curve remains below that of the TB+L model. Similarly, replacing burst-firing neurons with quiescent neurons (TQ model, red) does not enhance performance and leads to a repetitive learning pattern. At the start of each sequence, the model sequentially learns the digits 0 to 3 during the tonic-firing phase and maintains a low accuracy of around 25%, as it is only tested on samples from digits 0 to 3. As the sequence progresses, the network learns digits from 4 to 9, but it is still only tested on digits 0 to 3. Consequently, the accuracy rapidly drops to 0% as the network

fails to retain previously learned digits, remembering only the most recently learned ones. In contrast, by the end of the pattern recognition task, only the TB+L model remains robust to noise.

On the testing set, the TB+L model demonstrates generalization; the accuracy continuously increases over time and remains higher than the other models (Figure 3E, dark teal trace). The four models exhibit the same profile on the training set and on the testing set except that the accuracy is lower on unseen samples.

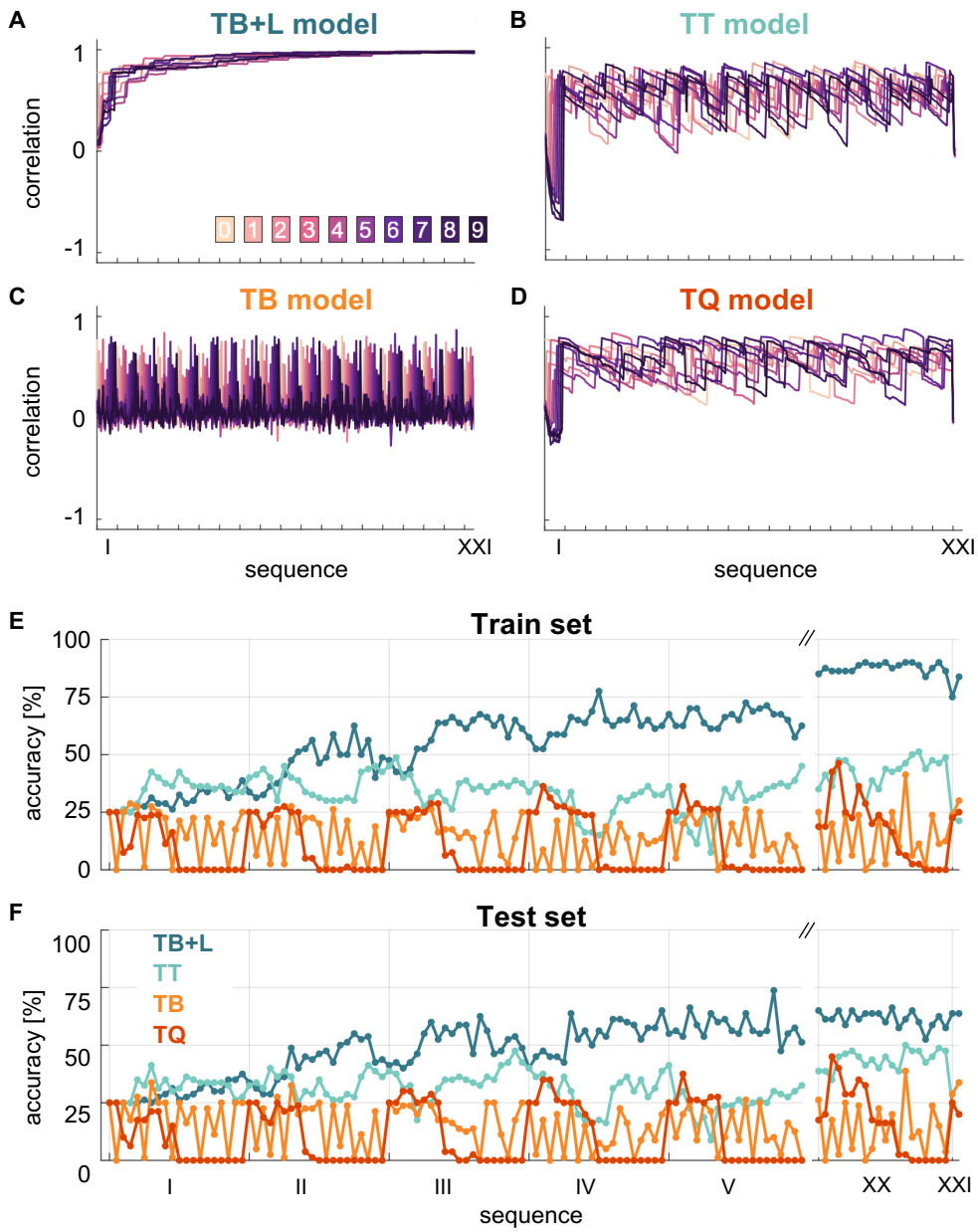


Figure 3: Burst-driven late-phase plasticity improves classification and generalization of learned digits. A-D. Evolution of the correlation between the mean image of the dataset and the resulting weight matrix at each state in each sequence for the different configurations; **A.** alternating tonic and burst firing with burst-driven late-phase plasticity (TB+L), **B.** tonic firing only (TT), **C.** alternating tonic and burst without late-phase plasticity (TB), and **D.** replacing burst firing with quiescence (TQ). Each color represents a different digit. **E-F.** Evolution of accuracy during the protocol (sequences I to V, XX, and XXI) from Figure 2, where the network classifies; **E.** 80 train samples (training set), **F.** or 80 unseen samples (testing set) of digits 0 to 3 across four configurations; TB+L (dark teal), TT (teal), TB (orange) and TQ (red) models. The weights for TT, TB, and TQ models are normalized by the maximum weight from the TB+L model.

3.6 Collective burst firing resets the early weights

To investigate how alternating between tonic and burst firing enable memory consolidation, we first examine the behavior of the early-phase synaptic plasticity. In particular, we explore the mechanisms by which switches to burst firing in the absence of late-phase plasticity reset the synaptic changes induced by early-phase plasticity during tonic firing (Figure 2D).

We consider the network architecture in Figure 1B, with 50 presynaptic neurons and 50 postsynaptic neurons ($N = M = 50$). Unlike the MNIST experiments which used a homogeneous network (subsubsection 5.3.8, here we use a heterogeneous network (subsubsection 5.3.4). We track the evolution of the early weights over the course of four switches from tonic to burst firing (Figure 4A).

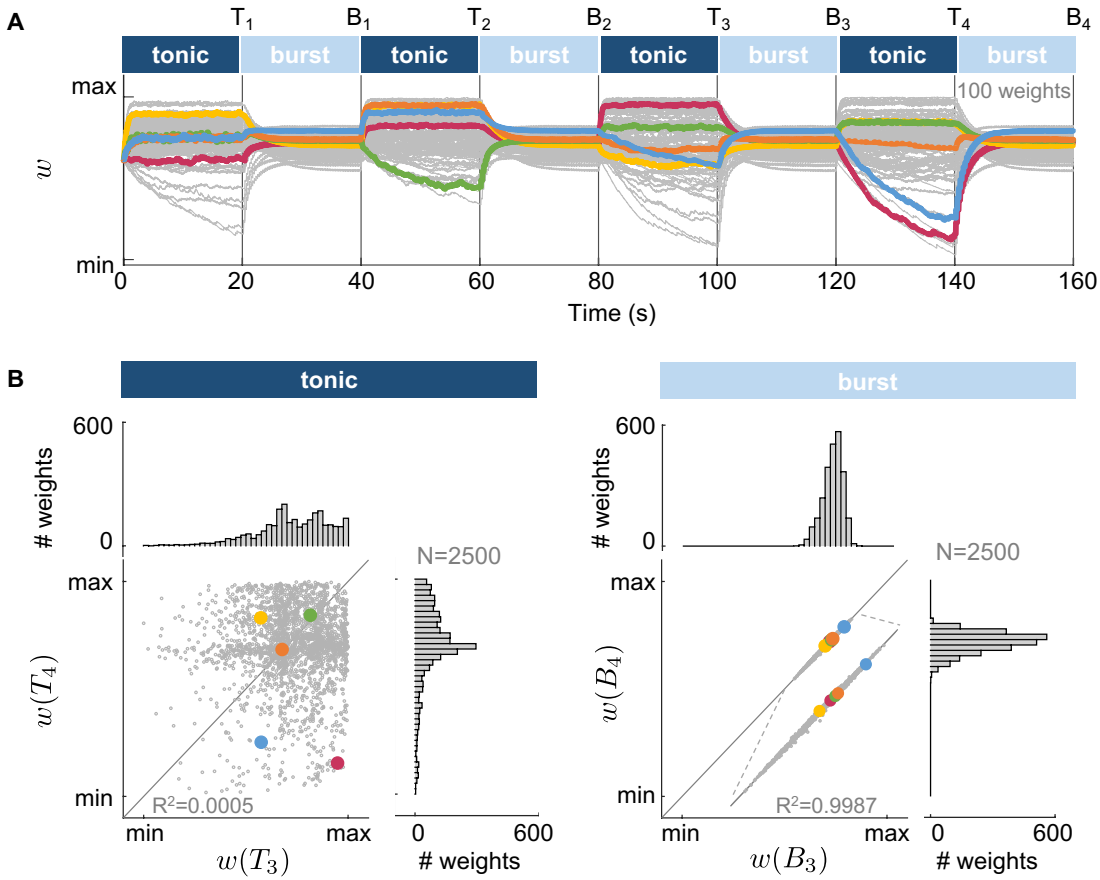


Figure 4: Collective burst firing resets the early weights. A. Evolution of the early weights during switches from input-driven tonic firing to collective burst firing. The network comprises 50 presynaptic neurons connected to 50 postsynaptic neurons for 100 traces among 2500 (gray lines) with 5 randomly highlighted lines (colored lines). T_k and B_k indicate the end of the k-th tonic and burst firing state, respectively. **B** Comparison of the early weights at the end of the third and fourth tonic firing states (dark blue) or the third and fourth burst firing states (light blue), normalized between the minimal and maximal values shown on a scatter plot. The histograms highlight the early weight distribution. The early weights reset to a similar steady state after each burst firing within a narrow range of values.

In the tonic firing period (Figure 4A, dark blue), each neuron receives a random pulse train input stimulation, the frequency selected uniformly at random between 0.1 and 50 Hz. The temporal evolution of the early weight $w_{ij}(t)$ is determined by the correlation between the activity of presynaptic neuron j and postsynaptic neuron i, which drives plasticity at those synapses. Therefore, at the end of each tonic firing period (time T_k, $k = 1, \ldots, 4$), the early weight $w_{ij}(T_k)$ converges to a value determined by the external inputs received. Because different neurons receive different inputs, early weights across the network settle at a wide range of values. Moreover, because input statistics differ across different periods of tonic firing, the early weights differ at the end of each one, as seen, for example, by plotting $w(T_4)$ against $w(T_3)$ (Figure 4B, left).

During burst firing (Figure 4A, light blue), the network is driven by an inhibitory current with no external input to the excitatory neurons. This drastically alters early weight dynamics: previously potentiated connections depress and vice versa (Jacquerie et al., 2022), with each early weight $w_{ij}(B_k)$ converging to a unique steady-state value within a narrow range, independent of its initial state (Figure 4B, right) – a phenomenon we call *burst-induced reset*. We derive an analytic expression for this steady-state value (Supplementary Information 6.4.3), which is well-matched to simulations (Figure SI 16D). This value depends on the parameters of the synaptic plasticity rule – such as the potentiation and depression rates – and the calcium concentration over time (resulting from the pre- and postsynaptic activity). During collective burst firing, all neurons within the network exhibit similar bursting activities, leading to a similar balance between potentiation and depression across neuron pairs (Figure SI 16C). Consequently, this homogeneity results in a narrow range of converging values.

Finally, we note that burst-induced reset is robust. Altering the applied current on inhibitory neurons tunes the network bursting properties but does not impact the presence of the reset (Figure SI 14). Furthermore, reset occurs across a wide range of synaptic plasticity models (Figure SI 17), including different variants of calcium-based rules (Graupner et al., 2016; Shouval et al., 2002; Deperrois and Graupner, 2020), phenomenological rules that utilize pairwise (Song et al., 2000) or triplet (Pfister and Gerstner, 2006) spike timing, weight-dependent rules (Van Rossum and Shippi, 2013; Gütig et al., 2003), and rules with parameters fit based on cortical (Sjöström et al., 2001) or hippocampal (Bi and Poo, 1998) plasticity induction protocols. We provide an analytic expression for the steady-state weights for a phenomenological model using spike time statistics (Supplementary Information 6.4.4), as well as a model with hard-bounds (Supplementary Information 6.4.5).

3.7 Burst-dependent up-selection and down-selection is controlled by the gain of the late-phase plasticity

We next evaluate the effect of the late-phase plasticity gain η (Equation 1) and the initial condition of the late weight $\ell_{ij}(0)$ on the dynamics of the synapse during burst firing states. We follow the computational experiments from González-Rueda et al. (2018), which enable quantifying memory consolidation using a single measure.

We consider a network of 100 presynaptic and 1 postsynaptic neuron ($N = 100$, $M = 1$, Figure 5A), which undergoes 10 alternating states of tonic and burst firing. During each state of tonic firing, 5 randomly selected presynaptic neurons are stimulated by a high-frequency pulse train ranging from 73 to 76 Hz (Figure 5B, dark purple), while the remaining 95 neurons receive stimulation from a low-frequency pulse train ranging from 0.1 to 5 Hz (Figure 5B, lilac). The output neuron is stimulated by a pulse train at an intermediate frequency, fixed at 25 Hz (purple). Synapses undergo early- and late-phase plasticity as introduced in section 3.2. During tonic firing, the five neurons firing at a high frequency potentiate their connection with the output neuron, while the rest tend to depress, as dictated by the early weight calcium-based rule.

To assess the network stored information across different states, we use the signal-to-noise ratio (SNR), defined as the maximum total weight divided by the mean of all total weights (Equation 3) (González-Rueda et al., 2018). Overall, SNR tends to increase over the course of the simulation (Figure 5Cii-v). However, for varying values of η and $\ell_{ij}(0)$, four different microscopic trends emerge across successive tonic and burst firing states (Figure 5Ci). In the first, which we term *up-selection*, the SNR after the burst firing state increases relative to the preceding tonic state (Figure 5Cii). Conversely, during *down-selection*, the SNR decreases during burst firing (Figure 5Ciii). Both mechanisms can coexist: the SNR can initially decrease during burst firing and then increase, in a *transition from down- to up-selection* (Figure 5Civ), as well as the reverse, starting with an increase and then a decrease, in a *transition from up- to down-selection* (Figure 5Cv).

The network can navigate within these four microscopic trends by varying the value of the gain η, tuned through the action of neuromodulators, which can influence protein synthesis pathways on a global scale (Brzosko et al., 2019; Pedrosa and Clopath, 2017; Nadim and Bucher, 2014; Pawlak and Kerr, 2008; Seol et al., 2007; Lisman et al., 2011). For example, dopamine or acetylcholine can

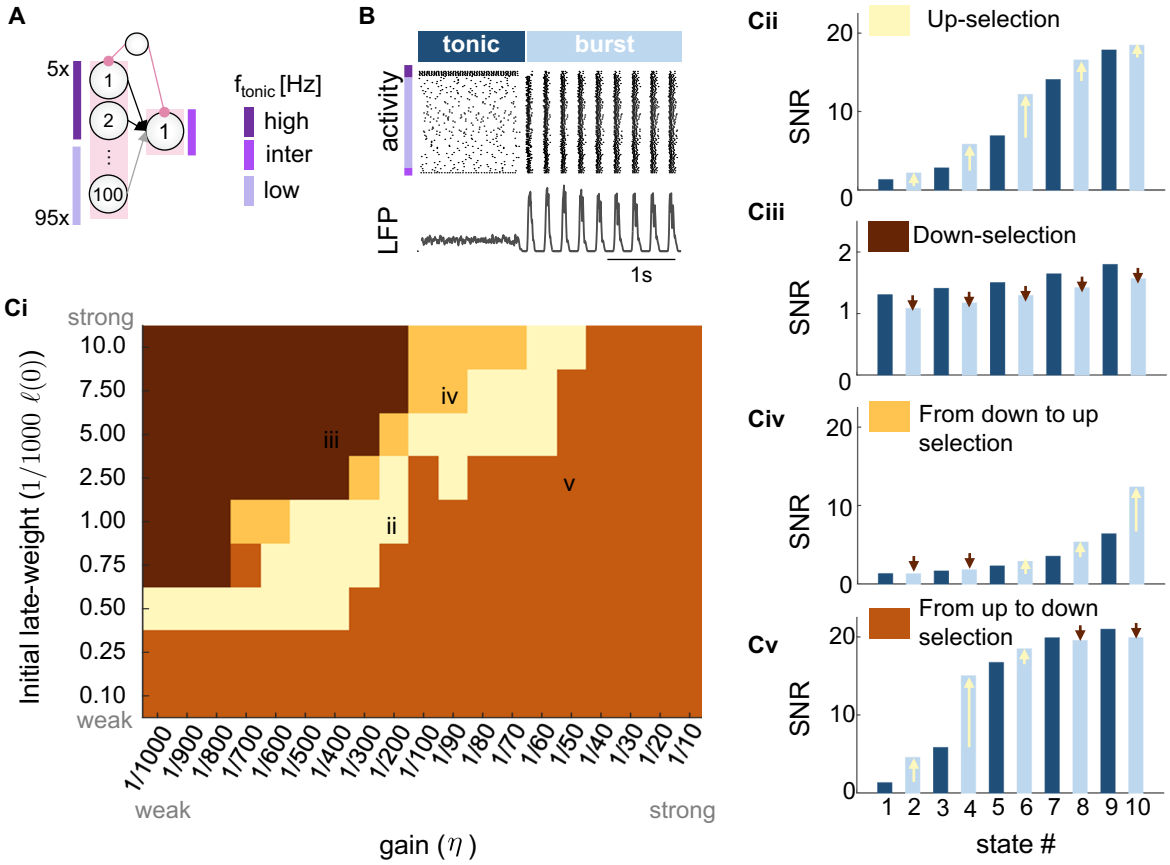


Figure 5: Dual mechanism of late-phase plasticity: Burst firing state can either increase or decrease the Signal-to-Noise Ratio (SNR), termed up-selection or down-selection. A. The effect of late-phase plasticity gain η is studied in a network with one inhibitory neuron projecting to 100 presynaptic neurons connected to 1 postsynaptic neuron during 10 states of 20 seconds, alternating between tonic and burst firing (inspired by (González-Rueda et al., 2018)). The neurons exhibit varying tonic firing frequencies (high, intermediate, or low - shades of purple). **B.** The raster plot over 4 seconds during the first switch from tonic to burst firing, along with the associated LFP. **C.** The evolution of the SNR at each successive tonic and burst firing state under different conditions. **Cii.** SNR increases from tonic to burst firing, indicating up-selection (light yellow, $\eta = 1/200$, $\ell_0 = 0.001$). **Ciii.** SNR decreases from tonic to burst firing, indicating down-selection (brown, $\eta = 1/400$, $\ell_0 = 0.005$). **Civ.** SNR transitions from down-selection to up-selection during the simulation (yellow, $\eta = 1/90$, $\ell_0 = 0.0075$). **Cv.** SNR transitions from up-selection to down-selection (orange $\eta = 1/50$, $\ell_0 = 0.0025$).

affect learning on the whole network (Pedrosa and Clopath, 2017). Additionally, these pathways can be locally modulated for specific synapses through synaptic-tagging and capture mechanisms (STC) (Lehr et al., 2022; Frey and Morris, 1997; Ding et al., 2022). According to this theory, synapses tagged after correlated pre- and postsynaptic activity experience enhanced protein synthesis, leading to synapse consolidation.

The effect of the gain, however, varies depending on the initial connectivity $\ell_{ij}(0)$. For low values, the microscopic trend remains the same regardless of the gain (Figure 5Ci, light brown color). This means that the network relies on burst firing to initially up-select the novel learning and then down-select it, which can be interpreted as a homeostatic technique. However, with intermediate connectivity, there is a range of gain values that promotes up-selection and thus consolidates memory during bursts. For very high connectivity, a weak gain will down-select, a mechanism compatible with synapse pruning that occurs during burst firing (Turrigiano and Nelson, 2004).

In summary, burst-driven late-phase plasticity can lead to both up-selection and down-selection. Up-selection acts as a mechanism for memory consolidation without needing to receive any input. This was also observed in the memorization task (subsection 3.3) through the increased contrast of the synaptic weights during bursting (Figure 2B). Conversely, down-selection prunes previously acquired

connections, allowing the network to retain and acquire memory while simultaneously reducing overall connectivity.

3.8 Different early-phase plasticity rules can lead to different learning strategies when learned patterns overlap

Finally, we evaluate the impact of the early-phase plasticity model on memory consolidation. Although the burst-induced reset phenomenon is robust across many early-phase plasticity models (subsubsection 6.4.6), the choice of model can affect the properties of memory consolidation, particularly in the case of overlapping patterns (*i.e.*, when two images from different categories share common pixels).

We compare two versions of the calcium-based model governing early-phase plasticity: the model proposed by Graupner and Brunel (2012) (Equation 2) and the one introduced by Graupner et al. (2016) (Equation 4). The 2012 model posits that at low calcium concentrations, the early weight follows a cubic function with two stable states, while the 2016 model suggests that the early change remains constant under low calcium concentrations. The numerical values of the late-phase plasticity parameters remain consistent across both models (Equation 1).

We consider a simplified pattern recognition task, using a network with 16 presynaptic neurons encoding a 4×4 pixel image and 2 postsynaptic neurons for classifying two patterns (Figure 6A, left, $N = 16, M = 2$). Following the supervised training paradigm in subsection 3.3, the network is trained on one of two different datasets: a pair of non-overlapping patterns (Figure 6B) or a pair of patterns which have overlapping pixels (Figure 6C, overlap indicated in dashed black lines).

The learning protocol, as depicted in Figure 6D, involves alternating between tonic and burst firing states. Each pattern consists of two bars, each containing 4 pixels, combining to form an 8-pixel pattern. During each tonic state, one bar of the pattern is presented. This reflects a learning scenario where different segments of a pattern (or class, like MNIST digits) are learned incrementally but still belong to the same class, despite individual pixel variations across samples. The exact sequence of stimuli is as follows: during the first tonic state, only the first bar (4 pixels) of the pattern is presented, and during the subsequent tonic state, the second bar is introduced. The two bars together form the complete 8-pixel pattern. This process is repeated for each new pattern. The burst state that follows each tonic state allows for the consolidation of the learned segment, preparing the network for learning the next bar during the subsequent tonic phase.

Figures 6Ei and Eii illustrate the evolution of the correlation between the weight matrices and the learned non-overlapping patterns (light orange and purple) for both the 2012 and 2016 models. The final weights and associated patterns are shown in Figures 6Eiii and Eiv, with sharper contrast and a more significant increase in correlation observed in the 2016 model compared to the 2012 model. To quantify this contrast, we calculate the average value of the 8 pixels forming one pattern and the average value of the other pixels in the weight matrix. We then compute the percentage difference between these two averages as a measure of contrast. For the 2012 model, the percentage difference is 17% for the first pattern and 20% for the second pattern. In the 2016 model, these values increase to 39% for both patterns.

Figures 6Fi-iv replicate these results for overlapping patterns. In this case, only the 2012 model successfully dissociates the patterns – the correlations slowly increase for both patterns throughout the protocol. The first pattern shows a 19% difference, and the second shows 21%. In contrast, with the 2016 model, the non-overlapping pixels are present in the final weight matrices, while the overlapping pixels are barely visible. Both patterns show a 10% difference. The correlation for one pattern increases when that pattern is presented, but the correlation for the other pattern decreases, reflecting that the overlapping pixels are being diminished. The evolution of the weight matrices throughout the protocol in both cases is shown in Figure SI 18.

This experiment highlights how the choice of early-phase plasticity models can lead to significantly different outcomes. The 2012 model effectively replicates memory consolidation, while the 2016 model reveals an intriguing and unexpected result: changing the early-phase plasticity model alters how overlapping patterns are learned. Specifically, the network inability to differentiate patterns based solely on overlapping pixels suggests that these pixels may not provide sufficient information for

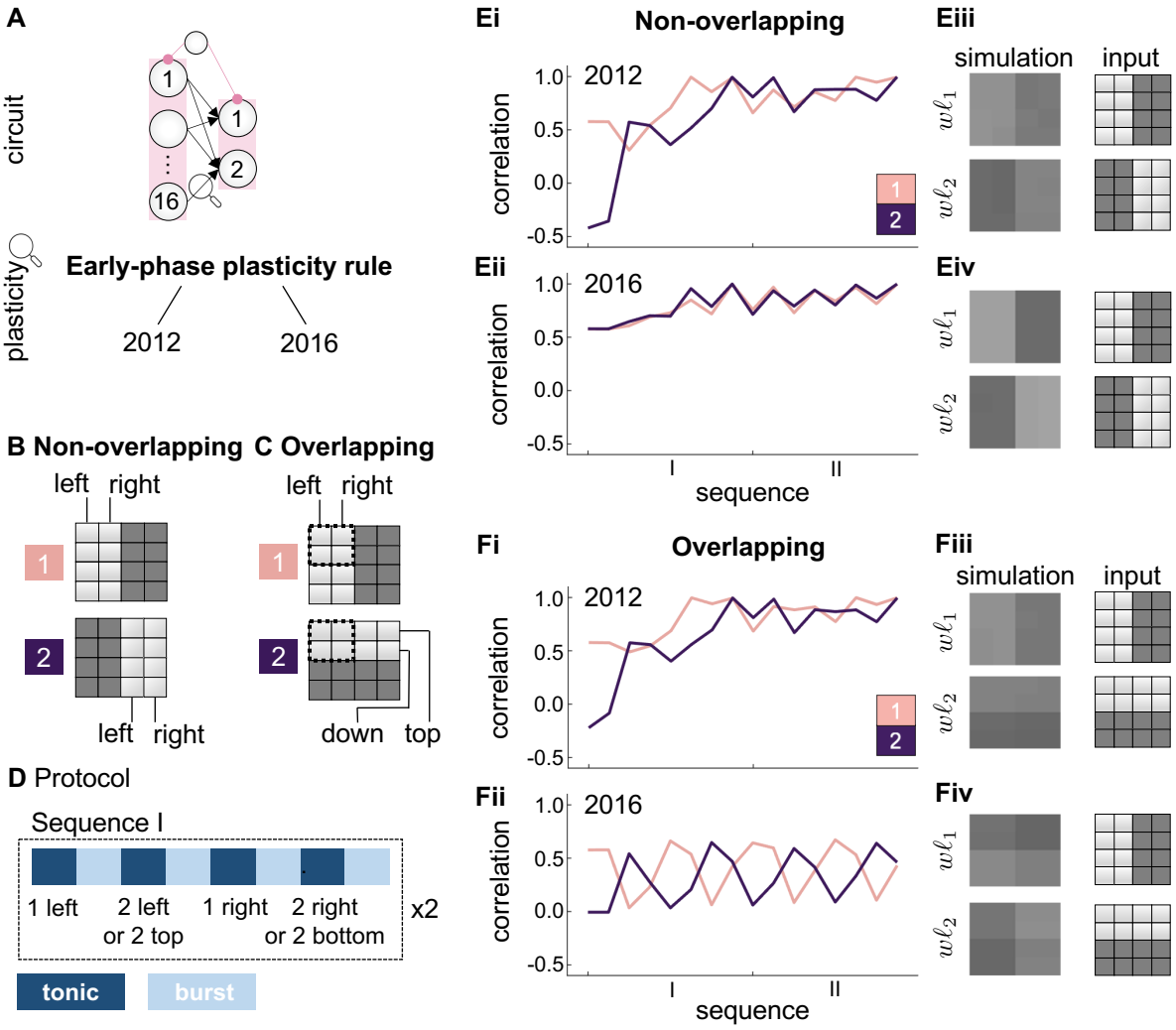


Figure 6: Different early-phase plasticity rules can result in different learning strategies when learned patterns overlap. A. The network consists of one inhibitory neuron synapsing GABA current to 16 presynaptic neurons connected to 2 postsynaptic neurons. The excitatory synapses are subject to plasticity based on the early-phase models proposed by (Graupner and Brunel, 2012) (2012 model) or (Graupner et al., 2016) (2016 model), combined with burst-driven late-phase plasticity. **B.** The task involves categorizing either non-overlapping patterns (#1, #2) or **C.** overlapping patterns (#1, #2) on a 4x4 grid image. Each pattern is composed of two bars (left, right, or top, down) **D.** The protocol is composed of two sequences. Each sequence is composed of 4 tonic firing states interleaved by 4 burst firing states. During each tonic firing state, the network is trained on one bar of a given pattern. **E.** Evolution of the correlation during both sequences for 2012 model(**Ei**) and 2016 model (**Eii**) for each pattern (orange and purple lines). **Eiii-iv** show the final weight matrices obtained for each pattern for respectively 2012 and 2016 models compared to the input pattern. Both early-phase rule fits support pattern categorization for non-overlapping patterns, as illustrated in the growing correlation and the final weight matrices ($w\ell$). **F.** The same graphs as **E** are shown for overlapping patterns. The patterns are distinctly learned according to the 2012 model, or the overlapping pixels are fewer than the unique pixels within the patterns according to the 2016 model.

accurate pattern memorization. This is akin to how humans might struggle to recognize an object when only shown a partial or fragmented view of it. For instance, when recognizing an animal from a piece of a picture, the brain can extrapolate to complete the whole representation, but at other times the fragment may not provide enough context. In biological systems, the brain likely employs multiple pathways to generate early-phase plasticity, leading to a diversity of memory consolidation results, depending on the input context and available cues. The novel late-phase plasticity introduced in this study does not interfere with these different pathways but rather regularizes learning. Our findings suggest that diverse synaptic plasticity rules account for different learning mechanisms, and these mechanisms are not mutually exclusive.

4 Discussion

Our brain exhibits a remarkable ability to transition between distinct states characterized by differences in neuronal firing patterns: neurons shift from input-driven tonic firing to collective burst firing, which has been shown, in some cases, to correspond to switching from learning to a resting state (Sherman, 2001). Concurrently, neurons undergo activity-dependent synaptic plasticity, known to occur in phases associated with distinct signaling pathways.

While significant efforts have been devoted to modeling the early phase of synaptic plasticity during tonic firing, our work focuses on the transitions between tonic firing and burst firing, and their impact on synaptic strengths within a biophysical neuronal network. We develop a straightforward burst-driven late-phase plasticity rule that complements well-established early-phase plasticity rules. Interleaving tonic firing and burst firing, coupled with this two-phase synaptic plasticity rule, enables robust memory consolidation.

4.1 Burst firing plays an important role in memory consolidation

Our model highlights a neuronal firing pattern signature that has not yet been fully investigated in synaptic plasticity research: collective burst firing within a network. These investigations have been sidelined for two main reasons. First, most experimental work has focused on studying synaptic plasticity under controlled firing conditions, where presynaptic and postsynaptic neurons fire in a highly regulated manner (Bi and Poo, 1998; Dudek and Bear, 1992; Inglebert et al., 2020; Sjöström et al., 2001). This controlled environment allows for a reliable exploration of the biophysics behind plasticity. In contrast, our model introduces burst firing activity, which is determined by intrinsic network and neuron parameters, presenting challenges in tracking burst firing characteristics such as the number of spikes and intraburst frequency as drivers of the synaptic change. Nevertheless, our computational framework provides a valuable tool for further research into burst firing dynamics. Second, burst firing has been identified as a signature of rest periods (Zagha, 2014; Sherman, 2001). However, in computational models, rest is often represented by quiescent neurons or low tonic activity, where neurons are not exposed to external inputs (Fauth and Van Rossum, 2019). Here, we propose that rest may involve more complex neuronal activity that plays a crucial role in shaping memory consolidation. Our model aims to illuminate the importance of investigating transitions between different firing activities and their implications for memory consolidation (Brokaw et al., 2016).

4.2 A simple burst-driven late-phase plasticity rule permits robust memory consolidation

In the literature, various late-phase plasticity rules exist, also referred to as structural plasticity (Fauth and Tetzlaff, 2016), ranging from highly detailed biophysical implementations that model different protein kinetics (Smolen et al., 2006), to less biologically detailed approaches (Deger et al., 2012) that utilize activity correlations to promote synapse removal. The first category is valuable for gaining a better understanding of the biological mechanisms involved in structural plasticity. However, it requires a larger number of equations, leading to an increased number of free parameters and computational complexity, especially as the network size grows. The second category provides a

useful framework for deriving analytical predictions but lacks biological interpretation. This approach designed for the tonic state is incompatible with a effective study of the effects of neuronal activity switches and capture the complex nature of bursting.

More recently, phenomenological rules have been proposed for late-phase plasticity (Zenke et al., 2015; Fauth and Van Rossum, 2019; Luboeinski and Tetzlaff, 2021), which also combine early-phase and late-phase plasticity (referred E-LTP and L-LTP in these articles). These rules exploit interesting phenomenological properties, such as weights with slow dynamics (Zenke et al., 2015) or the introduction of a probability of synapse removal that depends on the early weight (Fauth and Van Rossum, 2019). Such models have improved our understanding of synaptic plasticity mechanisms but, by relying on the value of the early weight itself to implement the late phase of plasticity (or maintenance and structural plasticity), they are not compatible with the burst-induced reset of the early weight investigated in this paper.

An alternative solution for memory consolidation is proposed by modeling the synaptic-tagging and capture (STC) hypothesis (Seibt and Frank, 2019). A potentiated synapse generates a local tag that activates and synthesizes plasticity-related-proteins (PRP) (Smolen et al., 2020). The maintenance of potentiation is achieved through the capture of PRP. However, implementing this hypothesis can be achieved by either describing the kinetics of different components, leading to a set of equations that becomes too complex for network use, or as an alternative, the synaptic weight is the *sum* of the early and late phase weights (Luboeinski and Tetzlaff, 2021; Clopath et al., 2008). The early weight follows traditional plasticity rules, while the late weight changes if the early weight surpasses a certain threshold. This mechanism is particularly effective in learning associated with tonic activity, where the early weight increases between correlated neurons, and the late weight further consolidates. However, considering this rule during bursting firing results in a reset of all the early weight once again dragging the late-phase weight in the same direction.

The burst-driven late-phase plasticity rule proposed in this paper stands out for two main reasons. First, the rule is simple, as it is only composed of one equation with one free parameter, yet it shows promising memory consolidation properties. Second, not only it is compatible with the burst-induced reset in the early weight observed during network bursting states, but it exploits the reset trace as the sole support of memory consolidation. This work demonstrates the potential of a simple model, which can be further refined by incorporating additional assumptions, such as activating the late-phase plasticity rule during tonic firing.

4.3 Burst-driven late-phase plasticity beyond biological networks

This work has implications beyond the field of synaptic plasticity and changes in rhythmic activity in a biophysical network model. It can draw inspiration for artificial spiking neural networks (SNNs). Current SNNs rely on continuous updates to the network based on input stimuli. However, in this project, we observe that during burst firing, the network can undergo plasticity without external stimuli, a process that might be beneficial to emulate in artificial systems. Our work suggests that incorporating such a burst state could allow networks to strengthen internal connections and consolidate learned patterns, without the need for constant external input. This mirrors concepts found in algorithms like wake-sleep, where the system alternates between learning new patterns and consolidating them without immediate external input (Hinton et al., 1995; Legenstein and Maass, 2007). Our proposed late-phase plasticity rule, using only the derivative of early weights and a proportional factor, makes it easily adaptable to artificial neural networks. Future work could integrate burst states and neuromodulator-induced switches to enhance memory consolidation and learning efficiency.

5 STAR METHOD

5.1 Key resources table

Reagent or **resource**	**Source**	**Identifier**
Julia	Julia programming	https://julialang.org
Matlab	Matlab and Matwhorks	https://matlab.mathworks.com
Custom code	This paper	https://github.com/KJacquerie/PhD

5.2 Resource availability

Lead contact

Further information and requests should be directed to and will be fulfilled by the lead contact, Kathleen Jacquerie (kathleen.jacquerie@gmail.com)

Material availability

NA

Data and code availability

All original data in this work was programmatically generated using the Julia programming language (Bezanson et al., 2017). Analyses were performed in Matlab. The code files are freely available at https://github.com/KJacquerie/PhD. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

5.3 Method details

5.3.1 Neuron and network model

All neurons are modeled using a single-compartment conductance-based using Hodgkin and Huxley formalism (Hodgkin and Huxley, 1952). The membrane voltage V of a neuron evolves according to the equation described by (Drion et al., 2018; Jacquerie and Drion, 2021):

$$C_{\mathrm{m}}\dot{V} = -I_{\mathrm{leak}} - \sum_{\mathrm{ion}\in\mathcal{I}} I_{\mathrm{ion}} - \sum_{p\in\mathcal{P}}\sum_{\mathrm{syn}\in\mathcal{S}} I_{\mathrm{syn},p} + I_{\mathrm{app}},$$

where C_{m} represents the membrane capacitance, I_{leak} is the leak current, I_{ion} are the intrinsic ionic currents with $\mathcal{I}$ the set of all ionic channels, $I_{\mathrm{syn},p}$ are the synaptic currents with $\mathcal{S}$ the set of all synaptic neurotransmitter types and $\mathcal{P}$ the set of all presynaptic neurons, and I_{app} denotes the applied current. Details about the ionic and synaptic currents and their associated dynamics and parameters are provided in the Supplementary Information 6.1.1.

The neuronal network comprises an inhibitory neuron ($\mathcal{N}_{\mathrm{inh}}$) projecting onto all excitatory neurons through $\mathrm{GABA_A}$ and $\mathrm{GABA_B}$ connections (refer to Figure 1B). Excitatory neurons are interconnected via a feedforward AMPA synapse, where the presynaptic neurons ($\mathcal{N}_{\mathrm{pre}}$) influence the postsynaptic neurons ($\mathcal{N}_{\mathrm{post}}$). The number of excitatory neurons varies across different computational experiments. The excitatory synaptic current perceived by the postsynaptic neuron i from presynaptic neuron j is characterized by:

$$I_{\mathrm{AMPA},ij} = w_{ij}\cdot\ell_{ij}\cdot s_{\mathrm{AMPA},j}\cdot(V_i - E_{\mathrm{AMPA}}),$$

where w_{ij} represents the early weight, ℓ_{ij} is the late weight, and their product $w_{ij}\cdot\ell_{ij}$ defines the synaptic weight. The variable $s_{\mathrm{AMPA},j}$ denotes the gating variable of the AMPA postsynaptic receptor (AMPAr), dynamically modulated by the presynaptic membrane voltage (V_j) and E_{AMPA} is the reversal potential of AMPAr. This model extends our previous representation of $\bar{g}_{\mathrm{AMPA},ij}$ as the maximal

conductance of the AMPA receptors to accommodate early and late phases of synaptic plasticity, as outlined in (Jacquerie and Drion, 2021).

5.3.2 Switch from tonic to burst firing

The inhibitory neuron dynamically influences the network activity. An external current applied to the inhibitory neuron ($I_{\text{app,inh}}$) models the effects of a neuromodulatory signal. A depolarizing current induces tonic firing activity in this inhibitory neuron, causing all excitatory cells to remain at rest. A sufficiently large external depolarizing pulse can evoke action potentials in the excitatory neurons. Conversely, a hyperpolarizing current applied to the inhibitory neuron switches the entire network into a synchronized collective bursting activity throughout the network (Zagha and McCormick, 2014; Jacquerie and Drion, 2021). In each computational experiment, the depolarizing current $I_{\text{app,inh}}$ is equal to $3\,\text{nA/cm}^2$ for tonic state and $-1.2\,\text{nA/cm}^2$ for burst state.

5.3.3 External applied pulse train current

In either learning, quiescent states, or noisy states, each excitatory neuron i is triggered by an applied current ($I_{\text{app},i}$) to make the neuron fire at a nominal frequency f_0. To generate this input-driven tonic activity, a neuron receives a pulse train current, where each pulse lasts 3 ms and has an amplitude that is independently sampled from a uniform distribution on an interval between $50\,\text{nA/cm}^2$ and $60\,\text{nA/cm}^2$. The interpulse intervals (*i.e.*, the time between two successive pulses) are independently sampled from a Normal distribution with a mean equal to $1/f_0$ and a standard deviation equal to $0.1/f_0$.

In burst firing states, each excitatory neuron i undergoes a noisy baseline such as the applied current ($I_{\text{app},i}$) randomly fluctuating between the interval 0 and $1\,\text{nA/cm}^2$.

5.3.4 Homogeneous and heterogeneous network

We define a *homogeneous* network (or circuit), where each neuron has the same intrinsic ion channel maximal conductances $\bar{g}_{\text{ion}} = \bar{g}^*_{\text{ion}}$ (numerical values are provided in subsubsection 5.3.4). We define a *heterogeneous* network (or circuit), where the maximal conductance $\bar{g}_{\text{ion}}$ for each neuron is randomly chosen within a $\pm 10\,\%$ interval around its nominal value $\bar{g}^*_{\text{ion}}$, such as $\bar{g}_{\text{ion}} = \bar{g}^*_{\text{ion}} \cdot (1+\epsilon)$ with $\epsilon \sim \text{Unif}(-0.1, 0.1)$.

5.3.5 Local field potential

The local field potential (LFP) measures the average behavior of interacting neurons. It reflects the collective excitatory synaptic activity received by the postsynaptic neuron population. The overall synaptic activity is measured by the mean of the individual synaptic currents:

$$\text{LFP}(t) = -\frac{1}{M}\sum_{j=1}^{M}\sum_{i=1}^{N} I_{\text{AMPA},ij}(t),$$

where M is the number of postsynaptic neurons and N is the number of presynaptic neurons (Jacquerie and Drion, 2021; Drion et al., 2018).

5.3.6 Synaptic plasticity

Early-phase synaptic plasticity The change in the early weight w_{ij} between a presynaptic neuron j and a postsynaptic neuron i is governed by the calcium-based model proposed by (Graupner and Brunel, 2012):

$$\tau_w \dot{w}_{ij} = -w_{ij}(1-w_{ij})(w^* - w_{ij}) - \gamma_{\text{d}} w_{ij}\Theta([\text{Ca}^{2+}]_{ij} - \theta_{\text{d}}) + \gamma_{\text{p}}(1-w_{ij})\Theta([\text{Ca}^{2+}]_{ij} - \theta_{\text{p}}). \quad (2)$$

Here, τ_w represents the time constant, w^* defines the stable state (equal to 0.5), γ_p is the potentiation rate, γ_d is the depression rate, θ_p is the potentiation threshold, and θ_d is the depression threshold. The function $\Theta(\cdot)$ is the Heaviside function, which returns 1 if $[\mathrm{Ca}^{2+}]_{ij} > \theta_\mathrm{p}$ and 0 otherwise.

The change in early weight depends on the calcium concentration $[\mathrm{Ca}^{2+}]_{ij}$, which is the sum of the calcium caused by the activity of the presynaptic neuron j and the activity of the postsynaptic neuron i. A pre- or postsynaptic spike translates into a calcium exponential decay. Further explanations are provided in section Supplementary Information 6.1.2. The calcium-based rule defined by (Graupner and Brunel, 2012) implements a soft-bound, where the perceived potentiation rate $\gamma_\mathrm{p}(1 - w_{ij})$ is smaller for high w_{ij} than for low w_{ij}. The same reasoning applies to the depression rate γ_d. Detailed parameter values are provided in Supplementary Information 6.1.3.

In Figures 4 and SI 18, the calcium-based model is modified according to the version of (Graupner et al., 2016). The model is similar except that the stable state is removed and the fitted plasticity parameters are adapted:

$$\tau_w \dot{w}_{ij} = \gamma_\mathrm{p}(1 - w_{ij})\Theta([\mathrm{Ca}^{2+}]_{ij} - \theta_\mathrm{p}) - \gamma_\mathrm{d} w_{ij}\Theta([\mathrm{Ca}^{2+}]_{ij} - \theta_\mathrm{d}).$$

Late-phase synaptic plasticity The late weight ℓ_{ij} between the presynaptic neuron j and the postsynaptic neuron i evolves according to Equation 1. The gain η is set to 1/200, except in Figure 5, where it is modified.

5.3.7 Computational experiment related to Figure 1

Figure 1C illustrates the activity in the feedforward network, comprising 50 presynaptic neurons connected to 50 postsynaptic neurons. The network is heterogeneous. The network is in tonic firing mode during 1.5 s (see subsubsection 5.3.2). Each excitatory neuron receives a train of current pulses with nominal frequencies f_0 randomly sampled from a uniform distribution between 0.1 and 50 Hz (see subsubsection 5.3.3). Then, the network switches to synchronized collective bursting during 2.5 s.

5.3.8 Computational experiment related to Figure 2

Figure 2A illustrates the simulation protocol for a learning task interleaved with periods of burst firing. The whole task lasts 6030 s (*i.e.*, 100 min and 30 s). The protocol is divided into 21 sequences (Roman numerals). In each sequence (I-XX), the network sequentially learns the digits 0 to 9. Each sequence lasts a total of 300 s and is composed of 10 blocks of tonic and burst firing. Each block lasts 30 s with a tonic or a burst firing state of 15 s each. The final sequence (XXI) lasts 30 s and consists of only one block. A cropped version of the MNIST dataset (LeCun et al., 1998) with 22×22 pixels is used. The initial dataset of 28×28 pixels has 3 pixels cropped on each side.

The network consists of 484 (22×22) presynaptic neurons representing individual image pixels and 10 postsynaptic neurons representing the digit classes (0 to 9). One inhibitory neuron projects onto all these neurons. The network is homogeneous to eliminate confounding factors while studying variations in the synaptic plasticity rule (parameters are available in section SI 6.1.1).

A learning state in the network corresponds to tonic activity (dark blue). Excitatory neurons are initially at resting membrane potential. Presynaptic neurons are activated by a train of current pulses with a nominal frequency reflecting the pixel color. The MNIST images are binarized such that a pixel with an intensity greater than 0.5 is considered a white pixel otherwise it is black. A white pixel is encoded by a neuron spiking at 45 Hz and a black pixel at 0.01 Hz. The postsynaptic neuron corresponding to the presented digit class spikes at 45 Hz, while the others spike at 0.0001 Hz. Each sequence is composed of 10 blocks. The learning states are sequentially organized from digit 0 to digit 9, with one sample digit randomly chosen from the MNIST dataset for each digit class.

A resting state corresponds to either collective bursting activity (dark blue) or quiescent activity (lilac) in the network. For the collective bursting activity, all excitatory neurons do not receive any input-driven pulse but only undergo baseline noise (subsubsection 5.3.3). For the quiescent activity, all neurons receive a pulse train with nominal frequencies independently sampled from a Normal distribution with a mean equal to 1 Hz and a standard deviation equal to 0.1 Hz..

A noisy state (gray) occurs during sequence XXI for 15 s. In this state, all excitatory neurons receive a pulse train with nominal frequencies independently sampled from a Normal distribution with a mean equal to 15 Hz and a standard deviation equal to 5 Hz.

Figure 2B illustrates the weight matrix of each digit. It corresponds to the total synaptic weights of the 482 presynaptic neurons onto the postsynaptic neuron associated with the considered digit, reshaped in a 22×22 grid matrix. The color of each pixel is proportional to the synaptic weight matrix $(w\ell)_i$ acquired at the end of each sequence (shown in Supplementary Information, Figure SI 8). For visualization purposes, synaptic weights are normalized so that the normalized synaptic weight values throughout the whole simulation span the interval $(0, 1)$, that is,

$$(\widetilde{w\ell})_i = \frac{(w\ell)_i - \min_{i,j,t}(w_{ij}(t)\ell_{ij}(t))}{\max_{i,j,t}(w_{ij}(t)\ell_{ij}(t)) - \min_{i,j,t}(w_{ij}(t)\ell_{ij}(t))},$$

where the maximum and minimum values are obtained by comparing the values across all pairs of presynaptic and postsynaptic neurons at the end of each state throughout the whole simulation. In Figure 2B, the weight matrices are shown for the states 1, 2 (in sequence I), 201, 202, 239, 240 (in sequence XI), 399, 400 (in sequence XX), 401, and 402 (in sequence XXI).

Figure 2C reproduces a similar protocol to Figure 2B, except that the bursting state (light blue) is replaced by an additional learning state for the same digit, using a different randomly chosen sample of the same digit. The weight matrices are obtained in the same manner, and the normalization is achieved by selecting the maximum and minimum values in this simulation. They differ from those in Figure 2B (see Supplementary Information 6.1.3 for the numeric values).

Figure 2D reproduces a similar protocol to Figure 2B, except that the late-phase plasticity is blocked. Only the early-phase plasticity is active during the bursting state.

Figure 2E reproduces a similar protocol to Figure 2B, except that the bursting state is replaced by a quiescent state.

5.3.9 Computational experiment related to Figure 3

Figures 3A-D illustrate the evolution of the two-dimensional correlation coefficient between the mean image of the dataset and the resultant normalized weight matrix at the end of each state, as obtained from the task depicted in Figures 2B-E. The mean image of the dataset is computed by averaging the pixel values across the 20 images per digit used to train the network (with one image per sequence) (see last column Figures 2B,D-E). For TT model, the mean image of the dataset employed in Figure 2C consists of 40 images instead of 20.

Figure 3E-F show the evolution of accuracy at each state of the learning task. Eighty samples are presented to the network, consisting of 20 samples from digit classes 0 to 3. Samples are either selected from the training set (Figure 3E) or the testing set (Figure 3F) and processed according to the methods described in subsubsection 5.3.8. Each sample is presented for 2 s. The external current applied to the inhibitory neuron is set to $0\,\mathrm{nA/cm^2}$ to modulate the network's overall excitability, mimicking a retrieval state. Presynaptic neurons are activated by a train of current pulses, with the frequency reflecting pixel intensity: white pixels are encoded by a neuron spiking at 55 Hz, and black pixels at 0.0001 Hz. Postsynaptic neurons are depolarized by $6.5\,\mathrm{nA/cm^2}$ and spike at 0.0001 Hz.

Plasticity is turned off, and synaptic weights are frozen at the values obtained at the end of each state (unnormalized weights). Since late-phase plasticity is only active in the TB+L model, we also consider normalized weights at a given state T^* for the TT, TB, and TQ models. The normalized weights are calculated as follows:

$$\widetilde{w_{ij} \cdot \ell_{ij}} = \max_{i,j} w_{ij}^{TB+L}(T^*)\ell_{ij}^{TB+L}(T^*) \frac{w_{ij}(T^*)\ell_{ij}(T^*) - \min_{i,j}(w_{ij}(T^*)\ell_{ij}(T^*))}{\max_{i,j}(w_{ij}(T^*)\ell_{ij}(T^*)) - \min_{i,j}(w_{ij}(T^*)\ell_{ij}(T^*)}),$$

where $\widetilde{w_{ij} \cdot \ell_{ij}}$ represents the normalized total weight, $\max_{i,j} w_{ij}^{TB+L}(T^*)\ell_{ij}^{TB+L}(T^*)$ is the scaling factor corresponding to the maximum weights at state T^* in the TB+L model, and the minimum and maximum values are taken across all presynaptic and postsynaptic neuron pairs at the end of state T^* for the respective model (TT, TB, or TQ).

The number of spikes generated by each output neuron is counted after a transient period of 0.5 s. The network prediction is determined by the output neuron with the highest spike count. Accuracy is defined as the percentage of correct predictions out of the 80 samples tested.

5.3.10 Computational experiment related to Figure 4

Figure 4A illustrates the evolution of the early weights within a heterogeneous network detailed in Figure 1C, consisting of 50 presynaptic neurons connected to 50 postsynaptic neurons, along with one inhibitory neuron projecting GABA current onto all of them. It comprises a total of 2500 weights. This computational experiment consists of 8 states interleaving tonic and burst firing, each lasting 20 s. During tonic firing, neurons are stimulated with a pulse train current, where the pulse frequency for each neuron is randomly chosen from a uniform distribution between 0.1 Hz and 50 Hz.

Figure 4B demonstrates the values of the normalized early weights at the end of the fourth tonic state (vertical axis) with the values of the normalized early weights at the end of the third tonic state (horizontal axis). The early weights are normalized using the formula

$$\tilde{w}_{ij}(t) = \frac{w_{ij}(t) - \min_{i,j,t}(w_{ij}(t))}{\max_{i,j,t}(w_{ij}(t)) - \min_{i,j,t}(w_{ij}(t))},$$

where $\max_{i,j,t}(w_{ij}(t))$ and $\min_{i,j,t}(w_{ij}(t))$ denote the maximum and minimum values obtained by comparing the values at the end of the third and fourth tonic states. The distribution of the normalized weights is represented for bins of width 0.025 between 0 and 1. Similarly, for burst firing (Figure 4C), the same data analysis protocol is used for the early weights at the end of the third burst state and the end of the fourth burst state. The scatter plots computed for the other states are shown in Figure SI 15.

The evolution of the mean and standard deviation of the early weights is shown in Figure SI 15.

5.3.11 Computational experiment related to Figure 5

Figure 5A illustrates a heterogeneous network comprising one inhibitory neuron connected to 101 excitatory neurons, where 100 presynaptic neurons synapse onto 1 postsynaptic neuron. The stimulation protocol consists of 10 states interleaving tonic firing and burst firing, each lasting 15 s.

For each tonic state, the first five neurons (dark purple) receive an applied current that consists of a train of pulses with a frequency independently sampled for each neuron from a uniform distribution between 73 and 76 Hz (the frequencies are resampled at the beginning of each tonic state). The other 95 presynaptic neurons (lilac) receive a pulse train with a frequency independently sampled for each neuron from a uniform distribution between 0.1 and 5 Hz. The postsynaptic neurons receive a pulse train at 25 Hz (purple). These values are selected to demonstrate strong and weak correlations, although other ranges of values also yield results. The pulse train is constructed as previously described. The raster plot and the LFP are computed for 4 s during the first state transition (Figure 5B). In Figure 5C, the signal-to-noise ratio is computed following the method proposed by (González-Rueda et al., 2018):

$$\mathrm{SNR} = \frac{\max_{i,j}(w_{ij}\ell_{ij})}{\mathrm{mean}_{i,j}(w_{ij}\ell_{ij})}. \tag{3}$$

Figures 5Ci highlights the trend followed by the SNR for different values of the late weight gain (η) and the initial value of the late weight (ℓ_0). Figures 5Cii to v are obtained for the following values: (ii) $\eta = 1/200$, $\ell_0 = 0.001$; (iii) $\eta = 1/400$, $\ell_0 = 0.005$; (iv) $\eta = 1/90$, $\ell_0 = 0.0075$; (v) $\eta = 1/50$, $\ell_0 = 0.0025$.

5.3.12 Computational experiment related to Figure 6

Figure 6A illustrates a homogeneous network comprising one inhibitory neuron connected to 18 excitatory neurons, where 16 presynaptic neurons synapse onto 2 postsynaptic neurons. Synapses undergo synaptic plasticity where the early weight is either driven by the early-phase plasticity fitted in the

calcium-based rule of (Graupner and Brunel, 2012) (2012 model) or (Graupner et al., 2016) (2016 model). Late-weight plasticity is active across both scenarios. Detailed parameters can be found in subsection 5.3.

The network is trained to learn two patterns, each corresponding to 8 active pixels out of the 16-pixel grid. The simulation protocol consists of two sequences of 4 tonic firing states interleaved with four burst firing states. During each tonic state, only 4 out of the 8 pixels are learned. It corresponds to creating samples for the same class. To achieve this, the corresponding input neurons are activated by a pulse train at 40 Hz, while the output neuron associated with the pattern class is triggered by a pulse train at 55 Hz. Inactive input pixels receive a pulse train at 0.01 Hz, and the second output neuron receives a pulse train at 0.0001 Hz. We alternate between learning one part of a pattern and, in the subsequent tonic firing state, one part of the other pattern. Correlation is computed as detailed in the methods section for Figure 2.

To compute the percentage difference, μ_{in} is defined as the mean of the pixels within the pattern, and μ_{out} as the mean of the pixels outside the pattern. In the non-overlapping case, for the first pattern, μ_{in} is the mean of the first two columns, and μ_{out} is the mean of the third and fourth columns. For the second pattern, these are reversed. In the overlapping case, μ_{in} and μ_{out} are calculated similarly, except for the second pattern, where μ_{in} is the mean of the first two rows, and μ_{out} is the mean of the third and fourth rows. The percentage difference is defined as $(\mu_{\text{in}} - \mu_{\text{out}})/((\mu_{\text{in}} + \mu_{\text{out}})/2)$.

References

Abbott LF, Nelson SB (2000) Synaptic plasticity: Taming the beast. *Nature Neuroscience* 3:1178–1183.

Bading H (2013) Nuclear calcium signalling in the regulation of brain function. *Nature Reviews Neuroscience* 14:593–608.

Bazhenov M, Timofeev I, Steriade M, Sejnowski TJ (2002) Model of thalamocortical slow-wave sleep oscillations and transitions to activated states. *Journal of Neuroscience* 22:8691–8704.

Bear MF, Connors BW, Paradiso MA (2016) *Neuroscience: exploring the brain* Fourth edition. Philadelphia : Wolters Kluwer Issue: International Edition.

Benna MK, Fusi S (2016) Computational principles of synaptic memory consolidation. *Nature Neuroscience* 19:1697–1706.

Berridge MJ (2014) Calcium regulation of neural rhythms, memory and alzheimer's disease. *Journal of Physiology* 592:281–293.

Bezanson J, Edelman A, Karpinski S, Shah VB (2017) Julia: A Fresh Approach to Numerical Computing. *SIAM Review* 59:65–98.

Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience* 18:10464–10472.

Bi Gq, Poo Mm (2001) Synaptic Modification by Correlated Activity: Hebb's Postulate Revisited. *Annual Review of Neuroscience* 24:139–166.

Brokaw K, Tishler W, Manceor S, Hamilton K, Gaulden A, Parr E, Wamsley EJ (2016) Resting state EEG correlates of memory consolidation. *Neurobiology of Learning and Memory* 130:17–25.

Brzosko Z, Mierau SB, Paulsen O (2019) Neuromodulation of Spike-Timing-Dependent Plasticity: Past, Present, and Future. *Neuron* 103:563–581 Publisher: Elsevier Inc.

Capone C, Pastorelli E, Golosio B, Paolucci PS (2019) Sleep-like slow oscillations improve visual classification through synaptic homeostasis and memory association in a thalamo-cortical model. *Scientific Reports* 9:1–11.

Chen JY, Lonjers P, Lee C, Chistiakova M, Volgushev M, Bazhenov M (2013) Heterosynaptic plasticity prevents runaway synaptic dynamics. *Journal of Neuroscience* 33:15915–15929.

Citri A, Malenka RC (2008) Synaptic plasticity: Multiple forms, functions, and mechanisms. *Neuropsychopharmacology* 33:18–41.

Clopath C, Ziegler L, Vasilaki E, Büsing L, Gerstner W (2008) Tag-trigger-consolidation: A model of early and late long-term-potentiation and depression. *PLoS Computational Biology* 4.

Dastgheib M, Kulanayagam A, Dringenberg HC (2022) Is the role of sleep in memory consolidation overrated? *Neuroscience & Biobehavioral Reviews* 140:104799.

Deger M, Helias M, Rotter S, Diesmann M (2012) Spike-Timing Dependence of Structural Plasticity Explains Cooperative Synapse Formation in the Neocortex. *PLoS Computational Biology* 8:e1002689.

Deperrois N, Graupner M (2020) Short-term depression and long-term plasticity together tune sensitive range of synaptic plasticity. *PLoS Computational Biology* 16:1–25.

Desroches M, Rinzel J, Rodrigues S (2022) Classification of bursting patterns: A tale of two ducks. *PLOS Computational Biology* 18:e1009752.

Destexhe A, Mainen ZF, Sejnowski TJ (1994) Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *Journal of Computational Neuroscience* 1:195–230.

Ding Y, Wang Y, Cao L (2022) A Simplified Plasticity Model Based on Synaptic Tagging and Capture Theory: Simplified STC. *Frontiers in Computational Neuroscience* 15:798418.

Drion G, Dethier J, Franci A, Sepulchre R (2018) Switchable slow cellular conductances determine robustness and tunability of network states. *PLoS Computational Biology* 14:1–20.

Dudek SM, Bear MF (1992) Homosynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. *Proceedings of the National Academy of Sciences* 89:4363–4367.

Fauth M, Tetzlaff C (2016) Opposing Effects of Neuronal Activity on Structural Plasticity. *Frontiers in Neuroanatomy* 10.

Fauth MJ, Van Rossum MC (2019) Self-organized reactivation maintains and reinforces memories despite synaptic turnover. *eLife* 8:1–19.

Feldman DE (2020) Spike timing–dependent plasticity. *Neural Circuit and Cognitive Development* pp. 127–141 Publisher: Elsevier Inc.

Fernandez LM, Luthi A (2019) Sleep Spindles: Mechanisms and Functions. *Physiological Reviews* 100:805–868.

Frey U, Huang YY, Kandel ER (1993) Effects of cAMP Simulate a Late Stage of LTP in Hippocampal CA1 Neurons. *Science* 260:1661–1664.

Frey U, Morris RGM (1997) Synaptic tagging and long-term potentiation. *Nature* 385:533–536.

Fusi S, Drew PJ, Abbott L (2005) Cascade Models of Synaptically Stored Memories. *Neuron* 45:599–611.

Garg N, Balafrej I, Stewart TC, Portal JM, Bocquet M, Querlioz D, Drouin D, Rouat J, Beilliard Y, Alibart F (2022) Voltage-dependent synaptic plasticity: Unsupervised probabilistic Hebbian plasticity rule based on neurons membrane potential. *Frontiers in Neuroscience* 16 Publisher: arXiv Version Number: 5.

Gerstner W, Kistler WM (2002) Mathematical formulations of Hebbian learning. *Biological Cybernetics* 87:404–415.

Golbert DC, Souza AC, Almeida-Filho DG, Ribeiro S (2017) Sleep, Synaptic Plasticity, and Memory In Byrne JH, editor, *Learning and Memory: A Comprehensive Reference*, pp. 539–562. Academic Press, Oxford, second edition.

González-Rueda A, Pedrosa V, Feord RC, Clopath C, Paulsen O (2018) Activity-Dependent Downscaling of Subthreshold Synaptic Inputs during Slow-Wave-Sleep-like Activity In Vivo. *Neuron* 97:1244–1252.e5.

Graupner M, Brunel N (2012) Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proceedings of the National Academy of Sciences of the United States of America* 109:3991–3996.

Graupner M, Wallisch P, Ostojic S (2016) Natural firing patterns imply low sensitivity of synaptic plasticity to spike timing compared with firing rate. *Journal of Neuroscience* 36:11238–11258.

Gurunathan A, Iyer LR (2020) Spurious learning in networks with Spike Driven Synaptic Plasticity In *International Conference on Neuromorphic Systems 2020*, pp. 1–8, Oak Ridge TN USA. ACM.

Gütig R, Aharonov R, Rotter S, Sompolinsky H (2003) Learning Input Correlations through Nonlinear Temporally Asymmetric Hebbian Plasticity. *The Journal of Neuroscience* 23:3697–3714.

Heidelberger R, Shouval H, Zucker RS, Byrne JH (2014) Synaptic Plasticity In *From Molecules to Networks: An Introduction to Cellular and Molecular Neuroscience: Third Edition*, pp. 533–561. Elsevier Inc.

Hill S, Tononi G (2005) Modeling sleep and wakefulness in the thalamocortical system. *Journal of Neurophysiology* 93:1671–1698.

Hinton GE, Dayan P, Frey BJ, Neal RM (1995) The "Wake-Sleep" Algorithm for Unsupervised Neural Networks. *Science* 268:1158–1161.

Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology* 117:500–544.

Inglebert Y, Aljadeff J, Brunel N, Debanne D (2020) Synaptic plasticity rules with physiological calcium levels. *Proceedings of the National Academy of Sciences* 117:33639–33648.

Jacquerie K, Drion G (2021) Robust switches in thalamic network activity require a timescale separation between sodium and T-type calcium channel activations. *PLOS Computational Biology* 17:e1008997.

Jacquerie K, Minne C, Ponnet J, Benghalem N, Sacré P, Drion G (2022) Switches to rhythmic brain activity lead to a plasticity-induced reset in synaptic weights preprint, Neuroscience.

Kotaleski JH, Blackwell KT (2010) Modelling the molecular mechanisms of synaptic plasticity using systems biology approaches. *Nature Reviews Neuroscience* 11:239–251.

Krug M, Lössner B, Ott T (1984) Anisomycin blocks the late phase of long-term potentiation in the dentate gyrus of freely moving rats. *Brain Research Bulletin* 13:39–42.

Lamprecht R, LeDoux J (2004) Structural plasticity and memory. *Nature Reviews Neuroscience* 5:45–54.

LeCun Y, Cortes C, Burges C (1998) MNIST handwritten digit database.

Lee HK (2006) Synaptic plasticity and phosphorylation. *Pharmacology & Therapeutics* 112:810–832.

Lee SH, Dan Y (2012) Neuromodulation of Brain States. *Neuron* 76:209–222 Publisher: Elsevier Inc.

Legenstein R, Maass W (2007) Edge of chaos and prediction of computational performance for neural circuit models. *Neural Networks* 20:323–334.

Legenstein RA, Maass W (2005) A Criterion for the Convergence of Learning with Spike Timing Dependent Plasticity In *NIPS 2005*, p. 8, Vancouver, Canada.

Lehr AB, Luboeinski J, Tetzlaff C (2022) Neuromodulator-dependent synaptic tagging and capture retroactively controls neural coding in spiking neural networks. *Scientific Reports* 12:17772.

Lisman J, Grace AA, Duzel E (2011) A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends in Neurosciences* 34:536–547.

Luboeinski J, Tetzlaff C (2021) Memory consolidation and improvement by synaptic tagging and capture in recurrent neural networks. *Communications Biology* 4:275.

Magee JC, Grienberger C (2020) Synaptic Plasticity Forms and Functions. *Annual Review of Neuroscience* 43:95–117 Publisher: Annual Reviews.

McCormick DA, Salkoff DB (2015) Brain State Dependent Activity in the Cortex and Thalamus. *Current Opinion in Neurobiology* 31:133–140.

McGinley MJ, Vinck M, Reimer J, Batista-Brito R, Zagha E, Cadwell CR, Tolias AS, Cardin JA, McCormick DA (2015) Waking State: Rapid Variations Modulate Neural and Behavioral Responses. *Neuron* 87:1143–1161 Publisher: Elsevier.

Morrison A, Diesmann M, Gerstner W (2008) Phenomenological models of synaptic plasticity based on spike timing. *Biological Cybernetics* 98:459–478.

Nadim F, Bucher D (2014) Neuromodulation of neurons and synapses. *Current Opinion in Neurobiology* 29:48–56.

Olcese U, Esser SK, Tononi G (2010) Sleep and synaptic renormalization: A computational study. *Journal of Neurophysiology* 104:3476–3493.

Pai TP, Chen CC, Lin HH, Chin AL, Lai JSY, Lee PT, Tully T, Chiang AS (2013) *Drosophila* ORB protein in two mushroom body output neurons is necessary for long-term memory formation. *Proceedings of the National Academy of Sciences* 110:7898–7903.

Pawlak V, Kerr JND (2008) Dopamine Receptor Activation Is Required for Corticostriatal Spike-Timing-Dependent Plasticity. *Journal of Neuroscience* 28:2435–2446.

Pedrosa V, Clopath C (2017) The role of neuromodulators in cortical plasticity. A computational perspective. *Frontiers in Synaptic Neuroscience* 8:38–38.

Pfister JP, Gerstner W (2006) Triplets of spikes in a model of spike timing-dependent plasticity. *Journal of Neuroscience* 26:9673–9682.

Poirazi P, Mel BW (2001) Impact of Active Dendrites and Structural Plasticity on the Memory Capacity of Neural Tissue. *Neuron* 29:779–796.

Rubin J, Lee DD, Sompolinsky H (2001) Equilibrium Properties of Temporally Asymmetric Hebbian Plasticity. *Physical Review Letters* 86:364–367.

Seibt J, Frank MG (2019) Primed to sleep: The dynamics of synaptic plasticity across brain states. *Frontiers in Systems Neuroscience* 13:1–19.

Seol GH, Ziburkus J, Huang S, Song L, Kim IT, Takamiya K, Huganir RL, Lee HK, Kirkwood A (2007) Neuromodulators Control the Polarity of Spike-Timing-Dependent Synaptic Plasticity. *Neuron* 55:919–929.

Sherman SM (2001) Tonic and burst firing: dual modes of thalamocortical relay. *Trends in neurosciences* 24:122–6.

Shouval HZ, Bear MF, Cooper LN (2002) A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proceedings of the National Academy of Sciences of the United States of America* 99:10831–10836.

Sjöström PJ, Turrigiano GG, Nelson SB (2001) Rate, Timing, and Cooperativity Jointly Determine Cortical Synaptic Plasticity. *Neuron* 32:1149–1164.

Smolen P, Baxter DA, Byrne JH (2006) A Model of the Roles of Essential Kinases in the Induction and Expression of Late Long-Term Potentiation. *Biophysical Journal* 90:2760–2775.

Smolen P, Baxter DA, Byrne JH (2020) Comparing Theories for the Maintenance of Late LTP and Long-Term Memory: Computational Analysis of the Roles of Kinase Feedback Pathways and Synaptic Reactivation. *Frontiers in Computational Neuroscience* 14:569349.

Song S, Miller KD, Abbott LF (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience* 3:919–926.

Timofeev I, Grenier F, Bazhenov M, Houweling AR, Sejnowski TJ, Steriade M (2002) Short-and medium-term plasticity associated with augmenting responses in cortical slabs and spindles in intact cortex of cats in vivo. *Journal of Physiology* 542:583–598.

Turrigiano GG, Nelson SB (2004) Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience* 5:97–107.

Tyree S, Luis De Lecea (2017) Optogenetic Investigation of Arousal Circuits. *International Journal of Molecular Sciences* 18:1773.

Van Rossum MC, Bi GQ, Turrigiano GG (2000) Stable Hebbian learning from spike timing-dependent plasticity. *Journal of Neuroscience* 20:8812–8821.

Van Rossum MC, Shippi M (2013) Computational modelling of memory retention from synapse to behaviour. *Journal of Statistical Mechanics: Theory and Experiment* 2013.

Wang HX, Gerkin RC, Nauen DW, Bi GQ (2005) Coactivation and timing-dependent integration of synaptic potentiation and depression. *Nature Neuroscience* 8:187–193.

Yaffe RB, Kerr MSD, Damera S, Sarma SV, Inati SK, Zaghloul KA (2014) Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proceedings of the National Academy of Sciences* 111:18727–18732.

Zagha EM D (2014) Neural Control of Brain State. *Current Opinion in NeurobiologyBiology* 0:178–186 ISBN: 6176321972 _eprint: NIHMS150003.

Zagha E, McCormick DA (2014) Neural control of brain state. *Current Opinion in Neurobiology* 29:178–186 Publisher: Elsevier Ltd.

Zeldenrust F, Wadman WJ, Englitz B (2018) Neural Coding With Bursts—Current State and Future Perspectives. *Frontiers in Computational Neuroscience* 12:48.

Zenke F, Agnes EJ, Gerstner W (2015) Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature Communications* 6.

Zenke F, Gerstner W (2017) Hebbian plasticity requires compensatory processes on multiple timescales. *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.