The trade-off between temporal precision and effect amplitude of inhibitory plasticity regulation determines separability of learned representations

Patricia Rubisch[1,2*, §], **Melanie I Stefan**[2], **Matthias H Hennig**[1]

***For correspondence:** patricia.rubisch@medicalschool-berlin.de (p.rubisch@sms.ed.ac.uk)

Present address: [§]Department of Human Medicine, Medical School Berlin, Germany

[1]Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh; [2]Department of Human Medicine, Medical School Berlin

Synaptic plasticity, the process by which synapses change in an activity-dependent manner, is assumed to be the basis of learning. Experimental evidence demonstrates that activity originating from other synapses in close proximity to an observed one can influence the outcome of plasticity including activity from inhibitory synapses. Under the assumption that the regulatory effect of inhibition is mediated by hyperpolarisation, we identify a trade-off between temporal precision and effect amplitude due to the treatment of postsynaptic activity in three different voltage-dependent plasticity models. Generally, inhibitory regulation of plasticity enhances the competition between lateral neurons driving the development of functionally relevant connectivity structures in recurrent excitatory-inhibitory networks. Thus, all models show signs of the ability to perform Independent Component Analysis (ICA) and lead to receptive field development. Models which are highly sensitive to local synaptic information tend to result in a higher degree of separation between learned features. This work stresses the importance of considering inhibition in plasticity research as well as indicates that learned representations are influenced by the available information at a synaptic site.

Introduction

Long Term Potentiation (LTP) and Long Term Depression (LTD), forms of activity-dependent plasticity, are the foundational processes by which an organism learns. The standard plasticity protocols which describe the change of synaptic strength with regard to pre-and postsynaptic activity such as Spike-Timing-Dependent Plasticity (STDP) and the frequency-effects (***Abbott and Nelson, 2000***; ***Sjöström et al., 2001***; ***Froemke and Dan, 2002***; ***Ngezahayo et al., 2000***; ***Dudek and Bear, 1993***) are often performed in mediums containing GABA antagonists or blockers. Therefore, they exclude any effects mediated by inhibitory activity. However, ***Steele and Mauk*** (***1999***) showed that inhibition influences and changes the established experimental results of the standard protocols. Furthermore, ***Hayama et al.*** (***2013***) and ***Paille et al.*** (***2013***) reported regulation of STDP when GABA is present at the synapse. ***Hayama et al.*** (***2013***) reports a promotion of LTD when GABAergic activity is delivered prior to a pre-post spike pair. Moreover, ***Paille et al.*** (***2013***) demonstrates the reversal of the causal relationship of standard STDP by inhibitory activity.

Models of synaptic plasticity allow us to explore the functional role of inhibition in plasticity regulation within networks of neurons. Rate- and Spike-timing-dependent models find that inhibition promotes functional development and stability (***Miehl and Gjorgjieva, 2022***; ***Rossbroich and Zenke, 2023***; ***Agnes and Vogels, 2024***). While they present a compact description of plasticity and in the case of rate models can derive analytical solutions, they require a-priori assumptions about the re-

lationship between E/I balance and plasticity updates. Moreover, they do not account for the effect of single inhibitory spikes in STDP experiments as described in the literature. We hypothesise that these effects can be accounted for by voltage-dependent plasticity models. In voltage-dependent plasticity models, plasticity induction is dependent on the relationship between the membrane potential and (arbitrary) plasticity thresholds. Thus, any hyperpolarisation of the postsynaptic membrane potential caused by inhibitory activity directly impacts the plasticity updates. Therefore, they provide the opportunity to investigate inhibitory regulation of plasticity without specifying the interaction explicitly. Moreover, a detailed study of inhibitory regulation in voltage-dependent plasticity models is currently missing.

We address this gap by studying three models with varying sensitivity to the instantaneous membrane potential: the model introduced in ***Clopath et al.*** (***2010***) (CM), its variation as presented in ***Meissner-Bernard et al.*** (***2020***) (MB) and the novel Voltage-Dependent Pathway (VDP) model. The models use different combinations of pre-and/or postsynaptic activity filters and thresholding to compute whether the synapse strengthens or weakens. We find that the differential treatment of the postsynaptic activity leads to differences in sensitivity to membrane potential fluctuations and restricts the effective timescale of the inhibitory regulation. While all models allow for inhibition to control plasticity, the network-driven competition retains hallmarks of the effective time scales causing developed functionally relevant structures such as receptive fields to reflect this.

Results

Pre- and postsynaptic filtering in voltage-dependent plasticity models

The family of voltage-dependent plasticity models assume that plasticity updates are driven by the behaviour of the postsynaptic membrane potential and thus the models are not only sensitive to postsynaptic spike timing but also take into account events which might not lead to spike emission.

Here we compare three models of this type which are based on the following assumptions: They implement Hebbian plasticity where the sign and amplitude of the update are dependent on the difference between the postsynaptic trace(s) and the plasticity thresholds $\theta^{+,-}$. The induction mechanism includes separate parameters and variables for the potentiation- and depression-inducing conditions which are indicated by the superscript + and − respectively. In the following pre- and postsynaptic traces are used to refer to the low-pass filtered presynaptic spike train $\bar{s}$ and the postsynaptic membrane potential $\bar{v}^{+,-}$. ***Meissner-Bernard et al.*** (***2020***) (MB) exclusively uses filtered activity to compute synaptic plasticity updates Δw as defined by Equation 1.

$$\Delta w = \eta \left(a^{+} \bar{s}(t) [\bar{v}^{+}(t) - \theta^{+}]_{+} - a^{-} \bar{s}(t) [\bar{v}^{-}(t) - \theta^{-}]_{+} \right) \tag{1}$$

The scaling parameters $\eta, a^{+,-}$, are the learning rate which scales the size of the weight update and the amplitudes of LTP and LTD.

In contrast, the model by ***Clopath et al.*** (***2010***) (CM) only computes LTD updates at the time of presynaptic spiking. In contrast, potentiation can occur with a longer latency between the pre- and postsynaptic event. However, the instantaneous postsynaptic membrane potential needs to be above the potentiation threshold. These requirements are formalised as :

$$\Delta w = \eta \left(a^{+} \bar{s}(t) [\bar{v}^{+}(t) - \theta^{-}]_{+} [v(t) - \theta^{+}]_{+} - a^{-} s(t) [\bar{v}^{-}(t) - \theta^{-}]_{+} \right) \tag{2}$$

Third, we introduce a new variation of a voltage-dependent plasticity model, the Voltage-Dependent Pathway mode (VDP). Similarly to the established models, it employs thresholds to determine whether the potentiation or/and depression conditions are fulfilled. However, if the instantaneous membrane potential is above the potentiation/depression threshold, the respective pathway $p^{+,-}$ is activated. The difference between their relative onset and activation determines the overall synaptic weight change. Therefore, the VDP update equation is formalised as:

$$\Delta w = \eta \left(a^{+} \bar{s}(t) p_{+}(t) - a^{-} \bar{s}(t) p^{-}(t) \right) \tag{3}$$

Model	update	presynaptic	postsynaptic
MB (***Meissner-Bernard et al., 2020***)	LTP & LTD	low-pass filtered	threshold on low-pass filtered
CM (***Clopath et al., 2010***)	LTP	low-pass filtered	threshold on low-pass filtered & instantaneous
	LTD	spike time (instantaneous)	low-pass filtered
VDP	LTP & LTD	low-pass filtered	threshold on instantaneous followed by down-stream dynamics

Table 1. Summary of the treatment of the presynaptic and postsynaptic activity by the different voltage-dependent models in the LTP and LTD induction conditions. All models assume individual thresholds for LTP and LTD induction for which $\theta^+ > \theta^-$.

with $p^{+,-}(t)$ following bi-exponential functions with saturating activation as $\sigma(v - \theta^{+,-})$. For details, see Methods. The main differences between these models are summarised in Table 1.

Of special interest are the differences in the postsynaptic induction component between the models. Experimental data has shown that the requirements for plasticity induction can vary between neurons. The VDP assumption follows the experimental literature reporting that local depolarisations alone are sufficient to induce plasticity (***Brandalise et al., 2016***; ***Lisman and Spruston, 2005***; ***Golding et al., 2002***; ***Stiefel et al., 2005***; ***Hardie and Spruston, 2009***). Moreover, the addition of the pathway activity is motivated by the down-stream processes which lead to the insertion and exocytosis of AMPA receptors which are the driving force behind LTD and LTP (***Colbran and Brown, 2004***; ***Mateos-Aparicio and Rodríguez-Moreno, 2020***; ***Kandel et al., 2013***). In contrast, models like MB and CM assume that prolonged depolarisations are necessary for plasticity induction following the fundamental results of ***Ngezahayo et al.*** (***2000***). These assumptions are also consistent with the observation that backpropagating Action Potentials (bAPs) can drive and control plasticity updates (***Magee and Johnston, 1997***; ***Christie et al., 1996***; ***Sjöström and Häusser, 2006***; ***Fuenzalida et al., 2010***). As a result, both MB and CM need prolonged depolarisations or a postsynaptic action potential for plasticity induction.

Postsynaptic filtering extends the effective regulation time window of inhibition

Paille et al. (***2013***) and ***Hayama et al.*** (***2013***) investigated the influence of inhibition on plasticity induction during standard STDP protocols using isolated pre-post pairings. They report qualitatively different results with the former one showing a flip of the Hebbian STDP to non-Hebbian STDP and the latter a strengthening of LTD. Nonetheless, both demonstrate the importance of considering all inputs received by the postsynaptic neuron. We hypothesise that inhibition exerts its regulatory function directly through hyperpolarisation of the postsynaptic membrane potential, mediated by the concomitant effects on local calcium levels. We used an STDP protocol with precisely timed inhibitory inputs to assess the relationship between local membrane potential, inhibitory timing and predicted plasticity updates. While all models predict a relative strengthening of LTD by inhibitory spiking prior to the postsynaptic event, the strength and effective time window depend on the model's induction mechanism and its assumptions on the postsynaptic activity. Figure 1B-C shows that the filtering of the membrane potential in the MB and CM model prolongs the time window in which inhibitory activity affects plasticity, reducing LTP when inhibition follows a postsynaptic spike. In the MB model, this leads to a complete reversal of LTP to LTD during inhibitory presynaptic activity (Fig. 1A). In the VDP, both LTP and LTD amplitudes decrease, with a stronger reduction of LTP.

One key difference between the models is that only in the VDP the effect of inhibition is independent of the duration of the inhibitory event (duration of current injected). In contrast, longer inhibitory inputs cause stronger attenuation of LTP and LTD in the MB and CM models (Fig. 1D).

The induction mechanism determines subthreshold fluctuation sensitivity

The results above illustrate how isolated spike pairs are affected by inhibition, and the use of a filtered or unfiltered membrane potential in plasticity updates explains the difference between

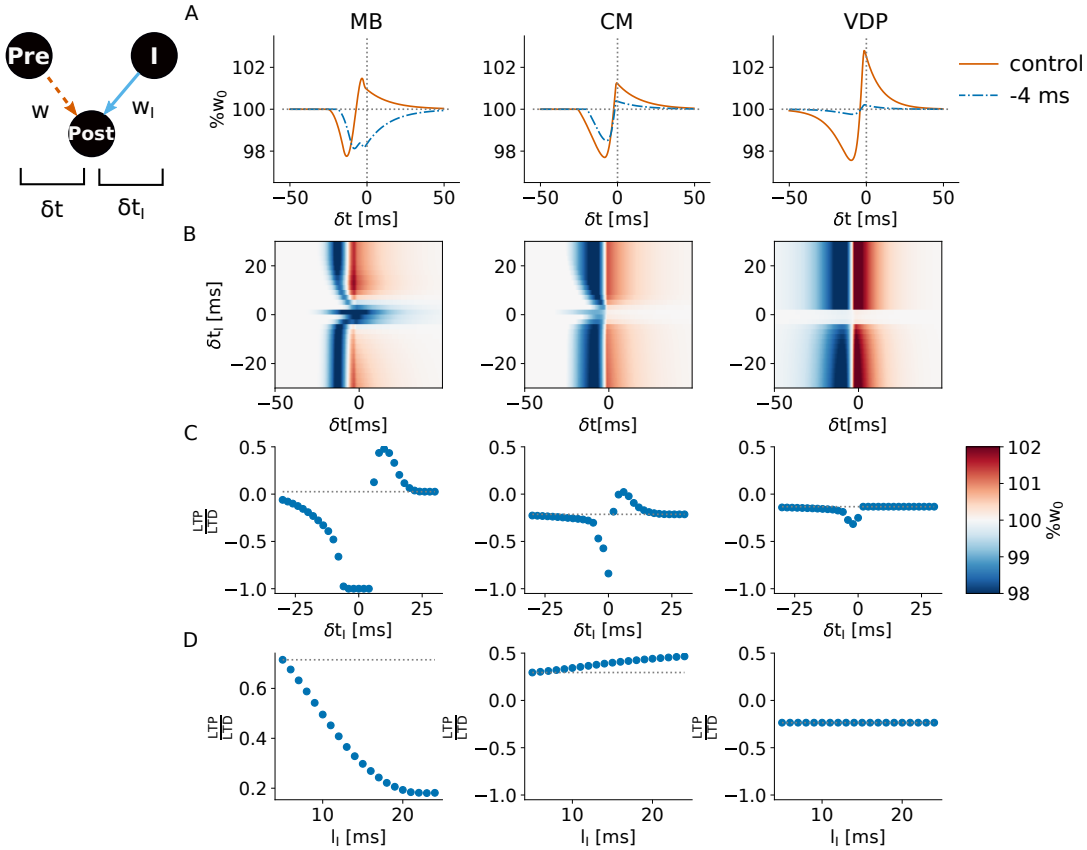

Figure 1. Precise inhibitory spike timing regulates the LTP-to-LTD ratio in STDP experiments. A standard STDP protocol is extended by an inhibitory presynaptic neuron as indicated by the circuit diagram. The inhibitory spike time δt_I is measured relative to the postsynaptic spike. Models are indicated by their abbreviations on top. **A:** Example STDP-kernels of the standard protocol without inhibition (orange) and with inhibitory spiking at $\delta t_I = -4$ms. **B**: STDP kernels over several different inhibitory spike timings shown as a heatmap. Blue areas indicate LTD while red denotes LTP. **C**: The LTP-LTD ratio of the STDP kernel depends on the inhibitory spike timing. Models which filter (MB and CM) the postsynaptic trace before thresholding show a broader temporal effect including positive inhibitory spike timings. **D**: The length of the hyperpolarisation does not influence plasticity models which only consider the instantaneous membrane potential. The postsynaptic and inhibitory spike is replaced with a depolarisation protocol imitating a rebound spike with different lengths of the hyperpolarising period I_i. The scatter plot shows the LTP-LTD ratio of the resulting STDP kernels.

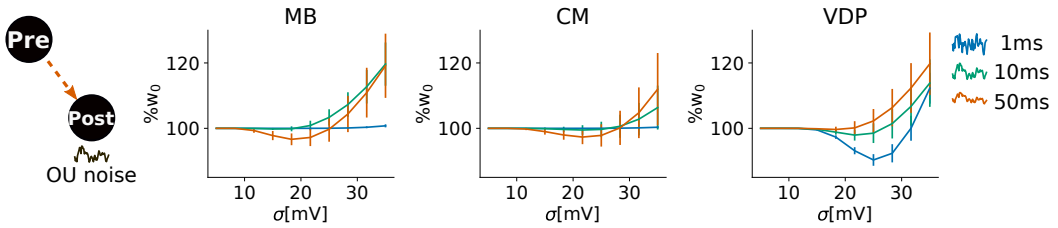

Figure 2. The induction principle of the plasticity model decides its sensitivity to fast fluctuations in the postsynaptic membrane potential. Instead of recording the membrane potential, v is replaced by an Ornstein-Uhlenbeck (OU) noise. The simulation is run for 10s. The presented results are averaged and computed over 100 instantiations. To control the fluctuation speed the time constant τ of the OU noise is varied (colours) while increasing the standard deviation σ results in growing fluctuation amplitude. Plasticity models are indicated by their abbreviation as in previous figures.

the models. As filtering the membrane potential suppresses fast fluctuations, we now compare how the models react to inputs during more realistic, fluctuation-driven spiking activity. To this end, we inject Ornstein-Uhlenbeck (OU) noise to simulate the combined input from many neurons, and change the time constant τ and the amplitude σ of the noise.

The results show that the VDP has a higher sensitivity to fast fluctuations than the other models (Fig. 2). While the MB and CM models allow inhibition to have a prolonged effective window, this reduces their sensitivity to fast events. In these models, only prolonged postsynaptic depolarisations cause weight changes due to the filtering of the membrane potential.

Recurrent network activity and stability

The results so far suggest that the way signals are integrated in the plasticity models qualitatively influences weight changes, and thus the connectivity and stability of recurrent networks. Specifically, in a recurrently connected network, a neuron's activity is the sum of the recurrent excitatory and inhibitory synaptic input. A shift in the excitation/inhibition (EI) ratio systematically changes the firing rate as well as the membrane potential statistics. At constant postsynaptic output rates, the mean membrane potential increases with increasing inhibitory synaptic input while its variance decreases (Fig. 3B-C).

Due to the different fluctuation sensitivity and temporal window of inhibitory regulation, the different models therefore predict distinct behaviours for different EI ratios (Fig. 3D). Excitation-dominated input causes high postsynaptic firing rates. The CM and VDP both predict weight increases under these conditions, a frequently discussed problem of Hebbian plasticity. In contrast, the MB model predicts LTD in the same condition. LTD for excitatory-dominated synaptic input at high rates pushes the activity towards lower rates, where the model predicts LTP. As a consequence, the synaptic weights will stabilise at an intermediate EI ratio, which in turn can stabilise a recurrent network.

Another stable point shared by all models is reached by inhibitory-dominated input, resulting in low firing rates. In this case, plasticity updates are rare and LTD-dominated, so the weights are, on average, reduced. This silencing of synaptic connectivity can lead to weight pruning.

The inhibitory synapses display a similar behaviour for low postsynaptic activity, where LTD dominates. In contrast, at high activity levels and low inhibitory firing rates, all models show stable inhibitory synapses with balanced LTP and LTD.

For stable network dynamics, the weight changes at intermediate firing frequencies are most important. Specifically, all models have points at which the mean synaptic weight is 0(intersections with the 100% line), and at these points, the EI ratio determines whether these can be stable. Comparing the models shows that the MB model and the VDP show a pattern supporting stability. At a constant postsynaptic rate, excitatory-dominated input induces LTD while inhibition-dominated input induces LTP. This causes the activity to change between inhibition-dominated activity, which

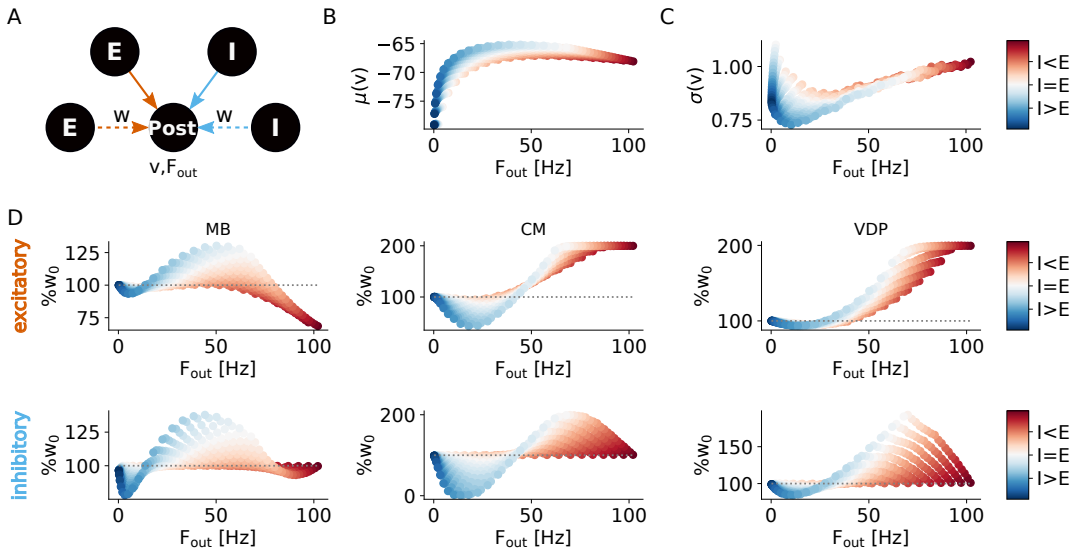

Figure 3. The fluctuations of the sub-threshold membrane potential influence plasticity predictions in the VDP. **A** A postsynaptic neuron is driven by 100 excitatory and 100 inhibitory Poisson neurons with firing rates $F_E, F_I \in [0, 100)$Hz (5Hz resolution). We measure the postsynaptic membrane potential, and frequency as well as the weight change of an excitatory (orange) and inhibitory (blue) synapse. The presynaptic neurons fire regularly with the same frequency as the corresponding populations. **B**: The mean membrane potential increases when the synaptic input is inhibition-dominated. **C**: The standard deviation of the membrane potential measuring the fluctuation amplitude is increasing when the synaptic input is dominated by excitation. **D**: Plasticity predictions of the plasticity models of the inhibitory and excitatory weights. Model identity is indicated by the abbreviations on top. All results are averaged over 10s simulation time.

is pushed towards excitatory-dominated activity by LTP and vice versa, leading, on average, to stable network activity. The CM model, in contrast, shows the opposite pattern which suggests this model is inherently less stable.

Hallmarks of time-restricted competition in developed functional connectivity

The analysis so far has established that the MB and VDP models can stabilise recurrent networks. Next, we asked whether these plasticity models also differ in terms of the learned representations when presented with structured input activity. To this end, we presented fully plastic recurrent EI networks with increasingly complex inputs and analysed the developed representations. The results demonstrate how the inhibitory plasticity regulation in the different models shapes the competition between lateral neurons and determines the features of the input representations connecting them to known properties of the visual system.

Competition in neural networks is necessary in order to develop a connectivity structure which can meaningfully encode input features and separate different sources of input. It is needed throughout neural development e.g. for developing ocular separation. Ocular separation describes the differentiation into left-eye and right-eye input-receiving neurons in the visual cortex before eye opening (***Huberman et al., 2008***; ***Ackman et al., 2012***). Therefore the neurons learn to separate two distinct input sources. The visual input during this developmental phase is highly correlated within one eye while uncorrelated between. To mimic this input structure, we defined two input populations (blue/red) which overlaps in activity are only circumstantial due to background activity (see Fig. 4A). As in the visual cortex, the competition mediated by lateral inhibition in the recurrent network causes the input weights to differentiate along the input correlation structure. The excitatory neurons develop preferences for one input source which can be measured by the relative dominance of their respective connections, see Fig. 4B-C. Due to the difference in temporal window and amplitude strength of the inhibitory regulation of the plasticity models, the characteristics

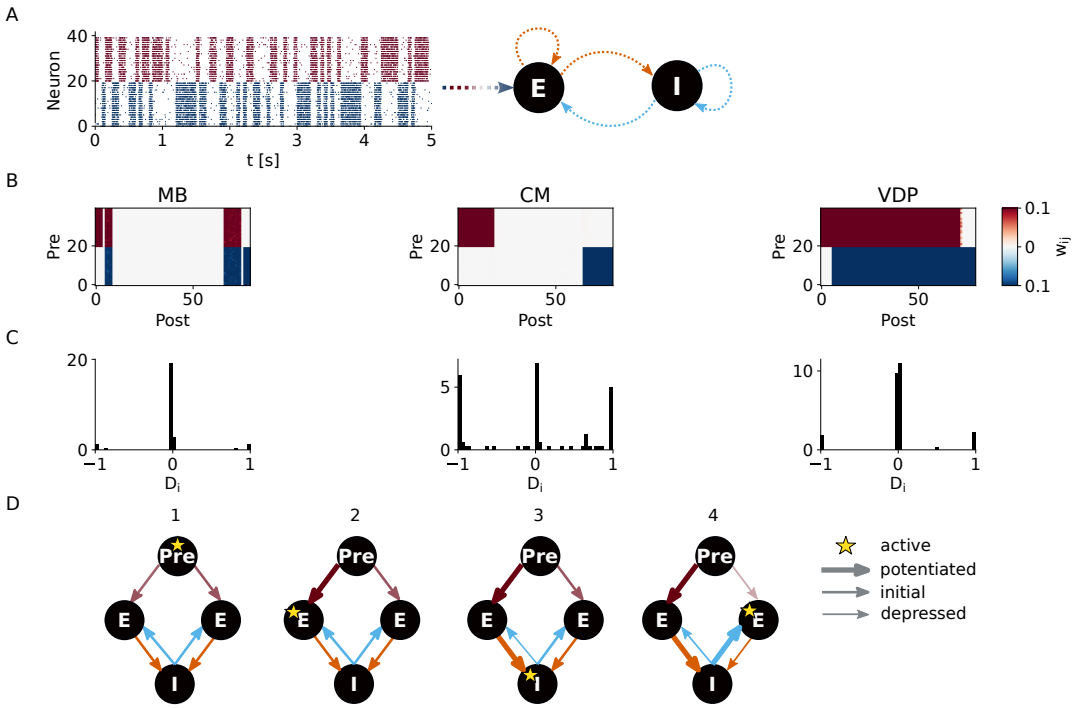

Figure 4. Voltage-dependent plasticity models can learn representations according to the cross-correlation pattern of the input. **A**: A fully plastic recurrent EI network is presented with ocular separation input. Two input populations (red and blue) do not correlate but highly correlate within their input group. **B**: Input-weight matrix and **C** dominance scores of the excitatory neurons showing developed by respective plasticity models. **D**: Illustration of the recurrently-driven and inhibition-enhanced competition between lateral neurons in a recurrent EI network.

of the connectivity patterns developed by the networks vary. While the MB model shows undifferentiated neurons with both minimal and maximal connections, the CM only develops silent undifferentiated neurons. In contrast, the VDP with the highest temporal regulation develops a large amount of undifferentiated but maximally connected excitatory neurons. As previously mentioned the differentiation is driven by the competition between lateral neurons which general principle is illustrated in Fig. 4D. Depending on the temporal window and amplitude of the inhibitory regulation, the depression of the input weight at timepoint 4 and the latency between the events of 2 to 4 differs between plasticity models. Due to the prolonged temporal window of the MB and CM models extending to positive time points, synaptic weights are suppressed by inhibitory regulation disregarding the temporal order of pre and postsynaptic activity. Therefore, on average more synaptic weights undergo (stronger) inhibition-driven depression leading to the development of silent neurons. In the VDP, inhibitory regulation is only effective in a short temporal window which is restricted by the onset of postsynaptic activity. This enforces the causal nature of Hebbian plasticity. However, due to the short temporal window and small amplitude, on average potentiation dominates the updates leading to a larger amount of neurons with maximal connectivity to both input populations.

Brito and Gerstner (***2016***) demonstrate that non-linear Hebbian plasticity implements a general principle for receptive field development and unifies the biologically motivated plasticity models with principled approaches such as Independent Component Analysis (ICA). The non-linearity specifically denotes the frequency-plasticity relationship. Thus, the voltage-dependent plasticity models also implement non-linear Hebbian plasticity. As a consequence, all models should be able to perform ICA. To test this, the networks are presented with the bars-and-stripes or Földiak's task (***Savin et al., 2010***; ***Földiak, 1990***). Crosses are randomly sampled from independent bars and

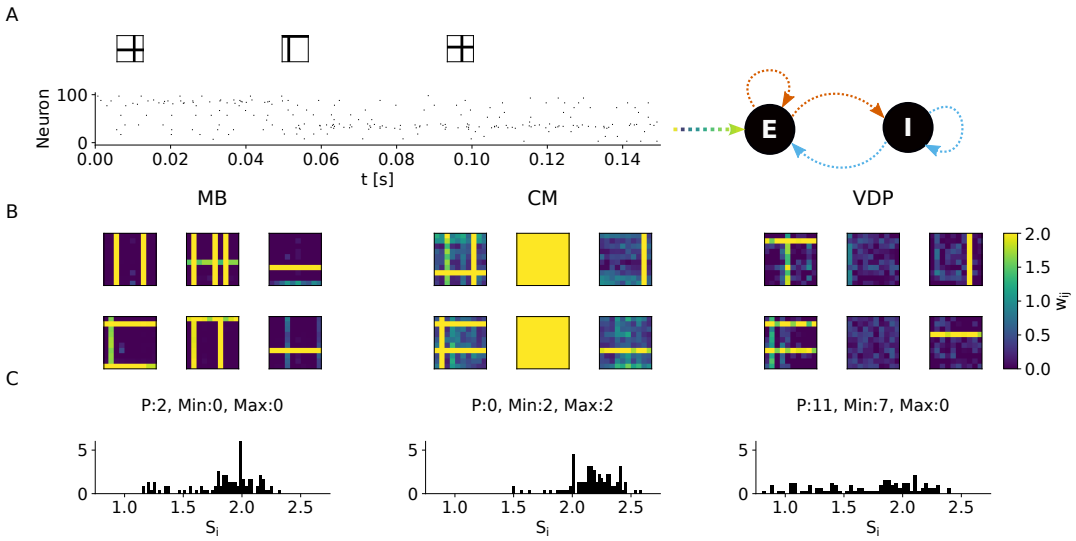

Figure 5. The excitatory population develops independent components when presented with temporally correlated signals. **A**: An EI network is presented with the bars-and-stripes task (***Földiak, 1990***). **B**: Learned receptive fields. **C**: Separation scores of the excitatory neurons. A neuron which encodes a single bar or stripe will have $S_i = 1.1$. A cell which encodes a cross or two bars will have $S_i = 2$. These scores are computed on the cleaned-up synaptic weight patterns to eliminate the influence of the pattern in near-0 synaptic weights. All weights $w_{ij} < \mu(w)$ are set to 0 for clean-up. The scores on top are the count of all neurons with near-perfect scores $1.0 < S_i < 1.2$ (P), the number of neurons which do not receive input $\mu(w_i) < \frac{\mu(w)}{3}$ (Min) and neurons which receive maximal input from all presynaptic neurons $\mu(w_i > 1.8$(Max). Here $\mu(w)$ denotes the mean of all input weights and $\mu(w_i)$ the mean weight of the receptive field of neuron i.

stripes (Fig. 5A). In contrast, to the ocular separation input whose sources were identifiable due to their correlation pattern, the independent sources are no longer uncorrelated. A network which successfully performs ICA on this input develops receptive fields resembling the independent components, one neuron's receptive field encoding a bar or a stripe, not the presented cross. Similar to the previous experiments, all models develop such receptive fields but differ strongly in their statistics (Fig. 5B-C). The MB network develops predominantly multi-factor cells, showing not only one bar or stripe but up to four independent factors encoded by a single excitatory neuron. Interestingly, these cells do not necessarily resemble the presented input but might consist of two stripes instead of a cross. The CM network is the only network which develops "copying" cells (denoted by Max in Fig. 5C). These cells do not show any discernible pattern in their receptive fields due to all their synapses saturating at the maximal weight are always active. This network as well as the VDP network also develops silent neurons, which do not receive any input due to their weights being uniformly 0 (, denoted by Min in Fig. 5C). The separation score S_i measured the quality and quantity of the receptive fields. A score of $S_i = 1.1$ indicates a receptive field which exactly encodes one independent component disregarding whether this is a bar or stripe. The histograms plotted in Figure 5D show that the VDP develops the most single-factor cells and overall heterogeneous receptive fields. It does not show a cluster around $S_i = 2$ indicating the absence of a large amount of multi-factor cells and a better-separated input representation. The tight temporal window of the inhibitory regulation in the VDP results in a higher degree of representational separability as well as shows hallmarks of sparse coding by pruning synaptic weights causing silent neurons.

The competition which drives ocular separation, as well as input factorisation also drives the receptive field development in V1. The receptive fields which resemble oriented edge-detectors have also been predicted by ICA and sparse coding on naturalistic input (***Van Hateren and van der Schaaf, 1998***; ***Bell and Sejnowski, 1997***; ***Brito and Gerstner, 2016***; ***Zylberberg et al., 2011***; ***Zhu and Rozell, 2013***). Therefore, they are a well-known example of input factorisation in the visual system.

To investigate whether the previously observed pattern of differences presits in the characteristics of the developed receptive fields, we present the recurrent EI network with CIFAR10 input patches (***Krizhevsky et al., 2009***). Resembling On- and Off cells in the retina, the input consists of two separate populations, an On- and an Off-population. Neurons in the On-population fire proportionally to the positive part of the input while the value of the negative pixels scales the Off-neuron's firing rate (Fig. 6A). To compute the receptive fields of each excitatory neuron, the differences between the Off-weights and On-weights are calculated. It is important to note that homogeneous receptive fields with $w_i = 0 \forall i$ such as those present in the MB network (see Fig. 6B) do not necessarily indicate a silent neuron as in the previous experiments. These neurons have equally strong connections to the On and Off-input with perfectly overlapping patterns. All networks develop a mixture of (uniform) mono-polar and oriented receptive fields (Fig. 6B). Mono-polar receptive fields are characterised by a small variance and $\mu \in [-1, 1]$ (Fig. 6C). They are essentially copying the input from either the On- or Off-population. The MB network develops the largest amount of mono-polar and 0-receptive fields which leads to clusters in the distributions of μ and σ. While the CM model has a larger amount of mono-polar receptive fields than the VDP network, neither of their distributions shows clear clusters. The heterogeneous distribution of the statistics of the receptive fields developed by the VDP network indicates that it learned the most diverse receptive fields overall. The biases in the orientations of the receptive fields align with the bias present in the input patches. To measure the orientation and relative quantity of oriented receptive fields, the receptive fields along the input and a randomly sampled control are convolved with oriented Gabor-patches. The absolute responses are scaled by the maximal response in order to ensure equal scale for better comparison of the response profiles of each condition. The input data shows a peak response at ca. 40°. While the CM and VDP network show the same peak response, the MB peak response is slightly shifted to the left (Fig. 6D). The response difference for vertical and horizontal orientations is the slowest in the receptive fields developed by the VDP and therefore closest to the pattern present in the input data. The MB receptive fields show the biggest difference between the response for horizontal and vertical orientations. However, all networks show a slight preference for horizontal orientations (0°) which resembles the response pattern of the input data. A small deviation from the overall input response pattern is present in the CM response pattern. Instead of linearly decreasing the response to progressively vertically oriented gratings (45°-90°), it shows a stronger slope in its responses around 70°.

Summarised, the biases and characteristics of the developed connectivity structures in all experiments suggest that the tight temporal inhibitory plasticity regulation in the VDP favours the separability of input representations which are simultaneously better-tuned to the input biases. Therefore, the trade-off between temporal precision and effect amplitude directly affects the characteristics of the internal representations of a recurrent network driven by structured activity.

Discussion

Taken together, these results indicate that inhibitory regulation of plasticity in fully plastic recurrent networks is sufficient for the development of functional-relevant connectivity patterns. Inhibition in voltage-dependent plasticity models enhanced recurrently-driven competition due to the direct hyperpolarisation of the membrane potential. The detailed exploration of three voltage-dependent plasticity models revealed a trade-off between the temporal precision of the inhibitory regulation effect and its amplitude. The stronger the induction mechanism filters the postsynaptic membrane potential, the broader the temporal window in which inhibitory spiking affects the synaptic weight update. This can even result in a reversal from LTP to LTD for pre-post pairs if the considered presynaptic activity trace is the filtered spike train and not instantaneous.

Functional experiments have revealed that a narrower temporal window which does not extend to positive spike pairings is favourable for input representation learning. However, all tested induction mechanisms are able to develop receptive fields when presented with naturalistic input.

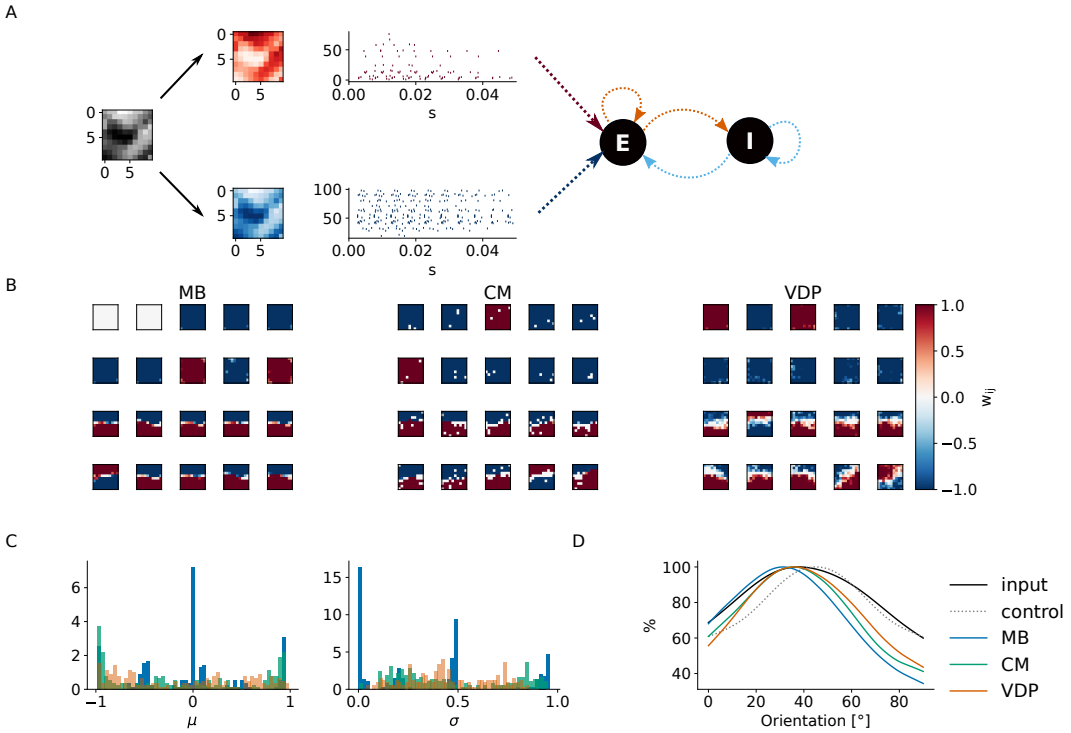

Figure 6. The induction mechanisms develop characteristic receptive fields. **A**: A fully-plastic recurrent EI network is presented with CIFAR10 input patches which are represented by the firing rates of an On- and Off-population. **B**: The developed receptive fields show specific characteristics depending on the applied plasticity model. Shown are 10 receptive fields with the smallest standard deviation and 10 receptive fields with the highest standard deviation. Blue indicates a strong connection to the Off-population while red indicates a strong input from the On-population. **C**: Statistics of the receptive fields. The histogram on the left shows distribution of the mean of the receptive field strength and the one on the right shows the standard deviation. **D**: Responsiveness of oriented Gabor-patches to the input data (black), random patches (grey-dotted), receptive fields of the MB (blue), CM (green) and VDP-network (orange). The colour scheme for the receptive fields is consistent between **C** and **D**.

Biologically, this can indicate that the learned representations become more refined with dendritic distance. The models presented in ***Clopath et al.*** (***2010***) and ***Meissner-Bernard et al.*** (***2020***) more strongly constrain the conditions under which plasticity events can occur. By filtering the membrane potential they emphasise the importance of the preceding activity. Therefore, temporally distant events can influence plasticity induction as demonstrated by the wider effective window of inhibitory regulation. When considering the spatial extent of a neuron and its dendrites, this also transfers to the locality in a simplified neuron model. Filtering the membrane potential is often used to model the influence of bAPs at distant dendritic locations. The stronger the filter, the further the modelled dendritic location ***Clopath et al.*** (***2010***); ***Meissner-Bernard et al.*** (***2020***). Therefore, these models include global information such as a bAP. As the influence of a bAP attenuates with increasing distance if not actively maintained, at distant (apical) dendrites, only the local information of their received synaptic currents are available. Thus, the VDP is likely to more accurately model synaptic behaviour where plasticity is driven by the local, instantaneous information. In contrast, the MB and CM model account for the global events which influence local activity at proximal dendritic sites.

However, neurons and synapses are highly heterogeneous: from genetic markers to protein and channel sub-type composition and local distribution. As previously discussed, it is more likely that a neuron has synaptic sites where the learning dynamics are more closely governed by any of the presented models. Furthermore, we only considered synapses which in their standard configuration reproduce the standard causal STDP kernel while several different STDP kernel forms have been identified experimentally (***Abbott and Nelson, 2000***). The implications for the inhibitory regulation effect when considering different STDP kernels or looking into a more detailed model considering channel compositions are yet to be investigated. It is highly likely that such an endeavour would also reveal the basis for the observations described by ***Paille et al.*** (***2013***) as the reversal from LTD to LTP is fundamentally incompatible with the mechanisms of voltage-dependent plasticity models which follow the standard observation that the LTD threshold is below the LTP threshold ***Ngezahayo et al.*** (***2000***).

In summary, this work investigates the role of inhibition in voltage-dependent plasticity models. Inhibitory regulation of plasticity varies depending on the treatment of the postsynaptic membrane potential. Models which filter the membrane potential allow inhibition to promote depression within a wider temporal range with a stronger effect amplitude. By studying its implication in recurrent competition between lateral neurons of the same afferent pathway, we identified that this temporal precision-effect amplitude trade-off comes at the cost of the precision of the learned representations. While all voltage-dependent models show hallmarks of ICA and can develop functional-relevant representation from naturalistic input, models using the instantaneous membrane potential tend to learn highly separate features which are more closely tuned to the input. This work stresses the importance of considering inhibition and the local EI balance when describing the plasticity pattern at a synaptic site which warrants further investigation.

Methods

Neuron model

All experiments are performed using a Leaky-Integrate and Fire (LIF) neuron with conductance-based synapses. They follow the equation:

$$\tau \frac{dv}{dt} = -(v - E_L) - g_I \cdot (v - E_I) - g_E \cdot (v - E_E) \tag{4}$$

$$\frac{dg_x}{dt} = -\frac{g_x}{\tau_{syn}} + \sum_j s_j \cdot w_j \cdot g_x^{max} \tag{5}$$

with $x = [E, I]$ and E_L denoting the resting potential of the neuron. g_I and g_E are the synaptic conductances of the inhibitory and excitatory synapses. Whenever a spike arrives at the synapse g_x grows proportionally to the current synaptic weight w_x. We use g_x^{max} to denote the maximal

conductance. We generally assume the notation of w_{ij} denoting the synapse connection between presynaptic neuron j and postsynaptic neuron i. Therefore, j in Equation 5 indicates that the synaptic current is summed over all presynaptic neurons.

Voltage-dependent plasticity models

The CM and BM models have been implemented according to their definitions in their corresponding studies (***Clopath et al., 2010***; ***Meissner-Bernard et al., 2020***). It is important to note that we did not include any homeostatic scaling in these models to exclusively study the effects on Hebbian learning. The novel model, the Voltage-Dependent Pathway (VDP) model, while following a similar update equation introduces a novel step by including separate pathway dynamics. These are defined as bi-exponential differential equations governed by:

$$\tau_f \frac{p}{dt} = -p + o \cdot p_r \tag{6}$$

$$\tau_r \frac{p_r}{dt} = -p_r + [[\sigma(v - \theta, k) - p]_+ - p_r]_+ \tag{7}$$

$$o = \frac{\tau_r}{\tau_f}^{\frac{\tau_f}{\tau_r - \tau_f}} \tag{8}$$

All parameters are pathway specific. This means that the potentiation and depression pathway vary in their time scale and thresholds. The slope of the sigmoid function is set to $k = 0.4$ for both. An example of the overall induction mechanism of the VDP is given in Figure 7 of the Appendix.

All models consider the presynaptic trace $\bar{s}$ in their induction mechanism. This is a low-pass filtered spike train which is computed by:

$$\tau \frac{d\bar{s}}{dt} = -\bar{s} + s(t) \tag{9}$$

$s(t)$ is the presynaptic spike train, which is 1 if a spike is emitted by the neuron and 0 otherwise. Similarly, the postsynaptic trace $\bar{v}$ is governed by:

$$\tau \frac{d\bar{v}}{dt} = -\bar{v} + v, \tag{10}$$

which low-pass filters the membrane potential. Superscripts indicate a potentiation and depression-specific filtering time constant τ.

Ornstein-Uhlenbeck noise for testing fluctuation sensitivity

To obtain the results presented in Figure 2, we replace the neuron model computing the subthreshold membrane potential with an Ornstein-Uhlenbeck noise. Therefore in these simulations, the membrane potential is described by the following differential equation:

$$\frac{dv}{dt} = -(v - E_L)/\tau + \sigma \delta W$$

δW is a Wiener process. For simulations, the seed of the random process implementing the Wiener process is set to ensure reproducible results. The speed and amplitude of fluctuations are changed by varying τ and σ.

Measures of Separation and Specialisation

All functional network consists of $N = N_E + N_I = N_E + 0.2 * N_E$ neurons where all synapses including the recurrent connection are plastic. To visualise the specialisation of the input weights, scores over all weights belonging to the receptive field of excitatory i are computed.

The dominance score D_i of the ocular separation experiment is computed as follows:

$$D_i = \frac{\sum_{j=0}^{0.5N} w_{ij} - \sum_{j=0.5N}^{N} w_{ij}}{\sum_{j=0}^{N} w_{ij}} \tag{11}$$

This compares the strengths of connectivity to the two input populations. $D_i < 0$ indicates a preference for the blue or left population, while $D_i > 0$ indicates that the input to this neuron is dominated by inputs from the red or right population.

In order to qualitatively and quantitatively describe the receptive fields in the bars and stripes task, the separation score S_i is computed per receptive field r_i, where r_i denotes all weights $w_{ij} \forall j$ of neuron i.

$$S_i = \sum_v c_S(r_i, 0°) \cdot max(c_S(r_i, 0°)) + \sum_h c_S(r_i, 90°) \cdot max(c_S(r_i, 90°)) \tag{12}$$

This measures the cosine-similarity c_S between the receptive field r_i and the ideal receptive fields of all bars $0°$ and the ideal receptive fields of all stripes $90°$. We scale the respective sums of cosine similarities with the maximal similarity to disambiguate between partial representations and ensure that $S_i = 1.1$ for only single factor representations.

References

Abbott, L. F. and Nelson, S. B. (2000). Synaptic plasticity: taming the beast. *Nature neuroscience*, 3(11):1178–1183.

Ackman, J. B., Burbridge, T. J., and Crair, M. C. (2012). Retinal waves coordinate patterned activity throughout the developing visual system. *Nature*, 490(7419):219–225.

Agnes, E. J. and Vogels, T. P. (2024). Co-dependent excitatory and inhibitory plasticity accounts for quick, stable and long-lasting memories in biological networks. *Nature Neuroscience*, 27(5):964–974.

Bell, A. J. and Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research*, 37(23):3327–3338.

Brandalise, F., Carta, S., Helmchen, F., Lisman, J., and Gerber, U. (2016). Dendritic nmda spikes are necessary for timing-dependent associative ltp in ca3 pyramidal cells. *Nature Communications*, 7(1):13480.

Brito, C. S. and Gerstner, W. (2016). Nonlinear hebbian learning as a unifying principle in receptive field formation. *PLoS computational biology*, 12(9):e1005070.

Christie, B. R., Magee, J. C., and Johnston, D. (1996). The role of dendritic action potentials and ca2+ influx in the induction of homosynaptic long-term depression in hippocampal ca1 pyramidal neurons. *Learning & memory*, 3(2-3):160–169.

Clopath, C., Büsing, L., Vasilaki, E., and Gerstner, W. (2010). Connectivity reflects coding: a model of voltage-based stdp with homeostasis. *Nature neuroscience*, 13(3):344–352.

Colbran, R. J. and Brown, A. M. (2004). Calcium/calmodulin-dependent protein kinase ii and synaptic plasticity. *Current opinion in neurobiology*, 14(3):318–327.

Dudek, S. M. and Bear, M. F. (1993). Bidirectional long-term modification of synaptic effectiveness in the adult and immature hippocampus. *Journal of Neuroscience*, 13(7):2910–2918.

Földiak, P. (1990). Forming sparse representations by local anti-hebbian learning. *Biological cybernetics*, 64(2):165–170.

Froemke, R. C. and Dan, Y. (2002). Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature*, 416(6879):433–438.

Fuenzalida, M., Fernandez de Sevilla, D., Couve, A., and Buno, W. (2010). Role of ampa and nmda receptors and back-propagating action potentials in spike timing–dependent plasticity. *Journal of neurophysiology*, 103(1):47–54.

Golding, N. L., Staff, N. P., and Spruston, N. (2002). Dendritic spikes as a mechanism for cooperative long-term potentiation. *Nature*, 418(6895):326–331.

Hardie, J. and Spruston, N. (2009). Synaptic depolarization is more effective than back-propagating action potentials during induction of associative long-term potentiation in hippocampal pyramidal neurons. *Journal of Neuroscience*, 29(10):3233–3241.

Hayama, T., Noguchi, J., Watanabe, S., Takahashi, N., Hayashi-Takagi, A., Ellis-Davies, G. C., Matsuzaki, M., and Kasai, H. (2013). Gaba promotes the competitive selection of dendritic spines by controlling local ca2+ signaling. *Nature neuroscience*, 16(10):1409–1416.

Huberman, A. D., Feller, M. B., and Chapman, B. (2008). Mechanisms underlying development of visual maps and receptive fields. *Annuual Review of Neuroscience*, 31:479–509.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S., Hudspeth, A. J., and Mack, S. (2013). *Principles of neural science*, volume 5. McGraw-hill New York.

Krizhevsky, A., Hinton, G., et al. (2009). Learning multiple layers of features from tiny images.

Lisman, J. and Spruston, N. (2005). Postsynaptic depolarization requirements for ltp and ltd: a critique of spike timing-dependent plasticity. *Nature Neuroscience*, 8(7):839–841.

Magee, J. C. and Johnston, D. (1997). A synaptically controlled, associative signal for hebbian plasticity in hippocampal neurons. *Science*, 275(5297):209–213.

Mateos-Aparicio, P. and Rodríguez-Moreno, A. (2020). Calcium dynamics and synaptic plasticity. *Calcium Signaling*, pages 965–984.

Meissner-Bernard, C., Tsai, M. C., Logiaco, L., and Gerstner, W. (2020). Dendritic voltage recordings explain paradoxical synaptic plasticity: a modeling study. *Frontiers in synaptic neuroscience*, 12:585539.

Miehl, C. and Gjorgjieva, J. (2022). Stability and learning in excitatory synapses by nonlinear inhibitory plasticity. *PLOS Computational Biology*, 18(12):e1010682.

Ngezahayo, A., Schachner, M., and Artola, A. (2000). Synaptic activity modulates the induction of bidirectional synaptic changes in adult mouse hippocampus. *Journal of Neuroscience*, 20(7):2451–2458.

Paille, V., Fino, E., Du, K., Morera-Herreras, T., Perez, S., Kotaleski, J. H., and Venance, L. (2013). Gabaergic circuits control spike-timing-dependent plasticity. *Journal of Neuroscience*, 33(22):9353–9363.

Rossbroich, J. and Zenke, F. (2023). Dis-inhibitory neuronal circuits can control the sign of synaptic plasticity. *arXiv preprint arXiv:2310.19614*.

Savin, C., Joshi, P., and Triesch, J. (2010). Independent component analysis in spiking neurons. *PLoS computational biology*, 6(4):e1000757.

Sjöström, P. J. and Häusser, M. (2006). A cooperative switch determines the sign of synaptic plasticity in distal dendrites of neocortical pyramidal neurons. *Neuron*, 51(2):227–238.

Sjöström, P. J., Turrigiano, G. G., and Nelson, S. B. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32(6):1149–1164.

Steele, P. M. and Mauk, M. D. (1999). Inhibitory control of ltp and ltd: stability of synapse strength. *Journal of neurophysiology*, 81(4):1559–1566.

Stiefel, K. M., Tennigkeit, F., and Singer, W. (2005). Synaptic plasticity in the absence of backpropagating spikes of layer ii inputs to layer v pyramidal cells in rat visual cortex. *European Journal of Neuroscience*, 21(9):2605–2610.

Van Hateren, J. H. and van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1394):359–366.

Zhu, M. and Rozell, C. J. (2013). Visual nonclassical receptive field effects emerge from sparse coding in a dynamical system. *PLoS Computational Biology*, 9(8):e1003191.

Zylberberg, J., Murphy, J. T., and DeWeese, M. R. (2011). A sparse coding model with synaptically local plasticity and spiking neurons can account for the diverse shapes of v1 simple cell receptive fields. *PLoS Computational Biology*, 7(10):e1002250.

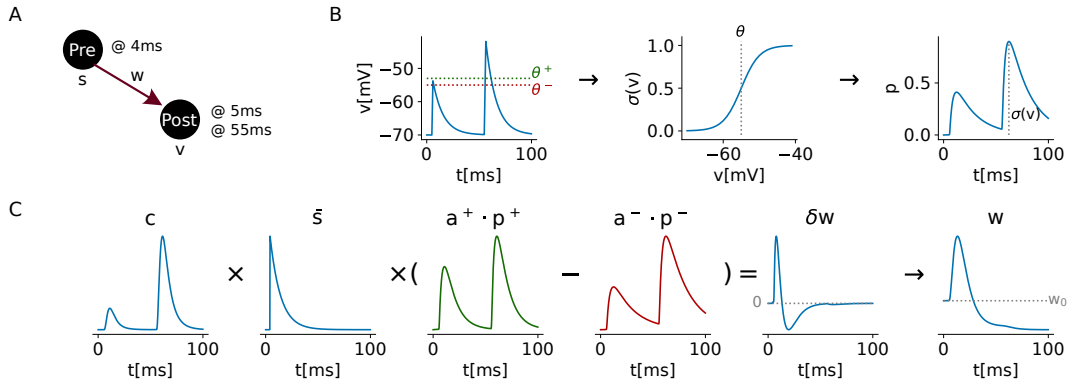

Figure 7. Illustration of the VDP model. **A**: Stimulation schema. The presynaptic spike train s and the postsynaptic membrane potential v are measured to compute the synaptic weight update. The time points denote spike and postsynaptic stimulation time to elicit the activation of the traces in B and C. **B**: Visualisation of the pathway induction process. The instantaneous membrane potential (left) is applied to a sigmoid function (middle) as a soft threshold with the middle points θ^+ and θ^- for the potentiation and depression pathway. The result of the sigmoid operation is used as the maximum amplitude or offset for the biexponential pathway activation p (right). **C**: Illustration of the computation and functions in the synaptic weight update equation 3. The component is a scaling factor mimicking calcium concentration dependent on the potentiation and depression pathway. This has been omitted in the equations to allow a direct comparison of the models. The scaling does not qualitatively influence the inhibitory effect. From left to right: The calcium level c is multiplied with the presynaptic trace $\bar{s}$ and the difference between the potentiation p^+ and depression pathway p^- in order to calculate the momentary weight w. Due to the amplitude parameter $a_+ < 1$, the illustrative example results overall in LTD, despite the weight increasing at first after the induction mechanism.

Appendix

To test the validity of the VDP, we tested whether core results of the experimental literature focusing on excitatory synaptic plasticity can be reproduced. The results which show that the VDP can be considered a valid synaptic plasticity model are shown in Figure 8.

Furthermore, by changing the relative dynamics of the potentiation and depression pathway the VDP can reproduce different STDP kernel shapes as described in ***Abbott and Nelson*** (***2000***). To characterise the different kernel shapes we measure the shape s defined as :

$$s = \frac{\sum_{i=0}^{t} \delta w^{\delta i} - \delta w^{-\delta i}}{\sum_{i=-t}^{t} |\delta w^{\delta i}|} \tag{13}$$

and the LTP/LTD ratio r defined as :

$$r = \frac{\sum_{i=-t}^{t} (\delta w^{\delta i} \cdot \delta w^{\delta i} > 0) - \sum_{i=-t}^{t} (\delta w^{\delta i} \cdot \delta w^{\delta i} < 0)}{\sum_{i=-t}^{t} |\delta w^{\delta i}|} \tag{14}$$

Figure 9 shows four distinct types of STDP kernel shapes along side their s and r scores.

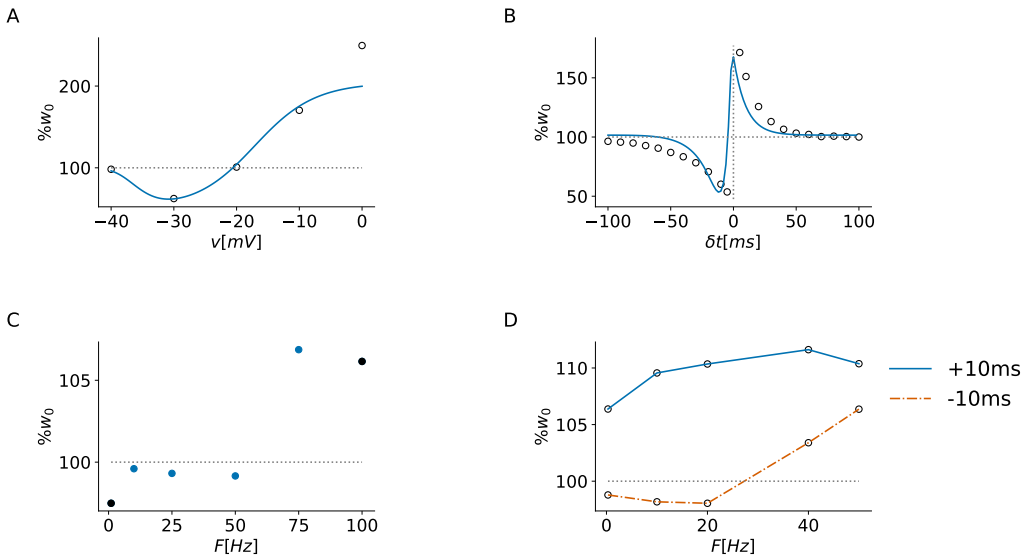

Figure 8. Predictions of the VDP for the reproduction of experimental protocols. **A**: The voltage dependency follows the relationship as described by ***Ngezahayo et al.*** (***2000***) closely. **B**: The VDP can reproduce the standard STDP kernel as measured by ***Froemke and Dan*** (***2002***). The open-faced circles indicate the experimental measurements. **C**: Low Frequency Stimulation (LFS) and Theta Burst Stimulation effects (TBS) are reproduced. The solid circles indicate conditions measured in ***Dudek and Bear*** (***1993***). **C**: The qualitative effect of the Pairing Frequency as described in ***Sjöström et al.*** (***2001***) can be reproduced by the VDP. Open-faced circled here indicate the conditions tested by ***Sjöström et al.*** (***2001***). They do not represent their measurements, as fitting the data quantitatively is limited by the description of the experimental protocol and used neuron model.

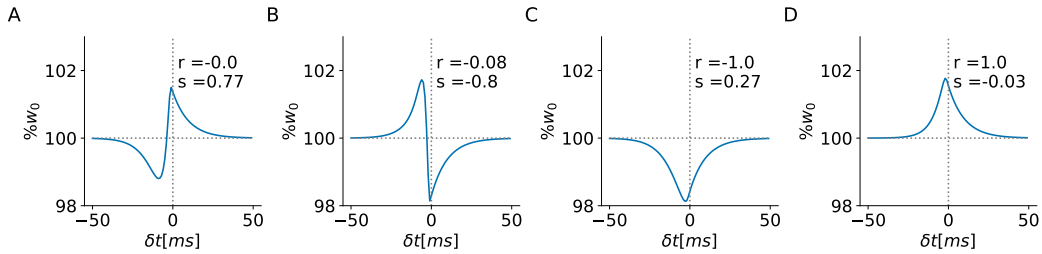

Figure 9. Different STDP kernel shapes reproduced by the VDP. The differences can be measured by calculating the kernel shape s and LTP/LTD ratio r. **A**: Balanced Hebbian STDP kernel with r = 0.0 and s = 0.77. **B**: Balanced anti-Hebbian STDP kernel with r = -0.08 and s = -0.8. **C**: LTD-dominated nearly-symmetric kernel. r = -1.0, s = 0.27. **D**: LTP-dominated nearly symmetric STDP kernel. r = 1.0 and s = -0.03.