The role of temporal sensitivity of synaptic plasticity in representation learning

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. Recent experiments demonstrated that synaptic plasticity not only depends on excitatory activity but also on the rate and the timing of inhibitory events. Hypothesising that the regulatory effect of inhibition is mediated by membrane potential hyperpolarisation, we identify a trade-off between temporal precision and magnitude of the inhibitory regulation of plasticity in three different voltage-dependent plasticity models. We show that in recurrent networks all three models develop receptive fields from naturalistic input, yet the temporal precision of inhibitory regulation is critical for the separation and heterogeneity of the resulting representation.

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***) found a promotion of LTD when GABAergic activity is delivered prior to a pre-post spike pair. Moreover, ***Paille et al.*** (***2013***) demonstrated the reversal of the causal relationship of standard STDP by inhibitory activity. A direct link between inhibition and the regulation of plasticity at excitatory synapses is reported by ***Person and Raman*** (***2010***). They found that hyperpolarisation due to inhibition deactivates L-type calcium channels which provides an LTP induction signal.

Models of synaptic plasticity allow us to explore the functional role of inhibition in plasticity regulation within networks of neurons. In rate- and Spike-timing-dependent models inhibition and its influence on plasticity promotes functional development and stability (***Miehl and Gjorgjieva, 2022***; ***Rossbroich and Zenke, 2023***; ***Agnes and Vogels, 2024***). However, they require a-priori assumptions about the relationship between E/I balance and plasticity updates. We hypothesise that inhibitory regulation effects are already accounted for in voltage-dependent plasticity models avoiding any assumptions on the mechanisms themselves. 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 different 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 leading to the development of functionally relevant structures, the heterogeneity and separability of representations are determined by the temporal precision of the regulation.

Results

Pre- and postsynaptic filtering in voltage-dependent plasticity models

In voltage-dependent plasticity models, we assume that the direction of plasticity updates depends on the postsynaptic membrane potential. Thus the models are not only sensitive to postsynaptic spike timing, but also predict plasticity updates in absence of postsynaptic spikes (***Clopath and Gerstner, 2010***; ***Meissner-Bernard et al., 2020***; ***Clopath et al., 2010***).

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 has separate parameters and variables for the potentiation- and depression-inducing conditions, 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 $\bar{v}^{+,-}$ (see Methods for details). ***Meissner-Bernard et al.*** (***2020***) (MB) exclusively uses filtered activity to compute synaptic plasticity updates Δw as follows:

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

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) computes LTD updates only at the time of presynaptic spiking while potentiation can occur with a longer latency between the pre- and postsynaptic event. However, for an update to occur the instantaneous postsynaptic membrane potential has to exceed the potentiation threshold. These requirements are formalised as:

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

Thirdly, we introduce a new variation of a voltage-dependent plasticity model, the Voltage-Dependent Pathway mode (VDP). As the two other models, it employs thresholds to determine whether the potentiation and/or depression conditions are fulfilled. However, if the instantaneous membrane potential is above the potentiation/depression threshold, the respective pathway $p^{+,-}$ is activated thereby adding another dynamical component. The difference between their respective dynamics and activation amplitude determines the overall synaptic weight change. The VDP update equation is:

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

with $p^{+,-}(t)$ following bi-exponential functions with saturating activation $p^{\infty} \lesssim \sigma\left(v - \theta^{+,-}\right)$ if the membrane potential is static. For further 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,***

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 $\theta^+ > \theta^-$.

2005; ***Golding et al., 2002***; ***Stiefel et al., 2005***; ***Hardie and Spruston, 2009***). Moreover, the addition of the pathway activity is motivated by the downstream processes which lead to the insertion and endocytosis of AMPA receptors which are the driving force behind LTP and LTD, respectively(***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 reported qualitatively different results. While the former found a flip of Hebbian STDP to non-Hebbian STDP, the latter demonstrated a strengthening of LTD. Nonetheless, both reports underline 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 plasticity predictions for isolated spike pairs are affected by inhibition. The differences between the models can be explained by the absence or presence of a postsynaptic membrane potential filter. Low-pass filtered as used in the MB and CM model suppresses fast fluctuations by reducing the high-frequency components in the signal. This can be explicitly shown by replacing the discrete current injections to produce spikes with a more realistic, fluctuation-driven postsynaptic membrane potential. To this end, we inject Ornstein-Uhlenbeck (OU) noise to simulate the combined input from many neurons. The fluctuation speed and amplitude are manipulated by changing the time constant τ and the amplitude σ of the noise respec-

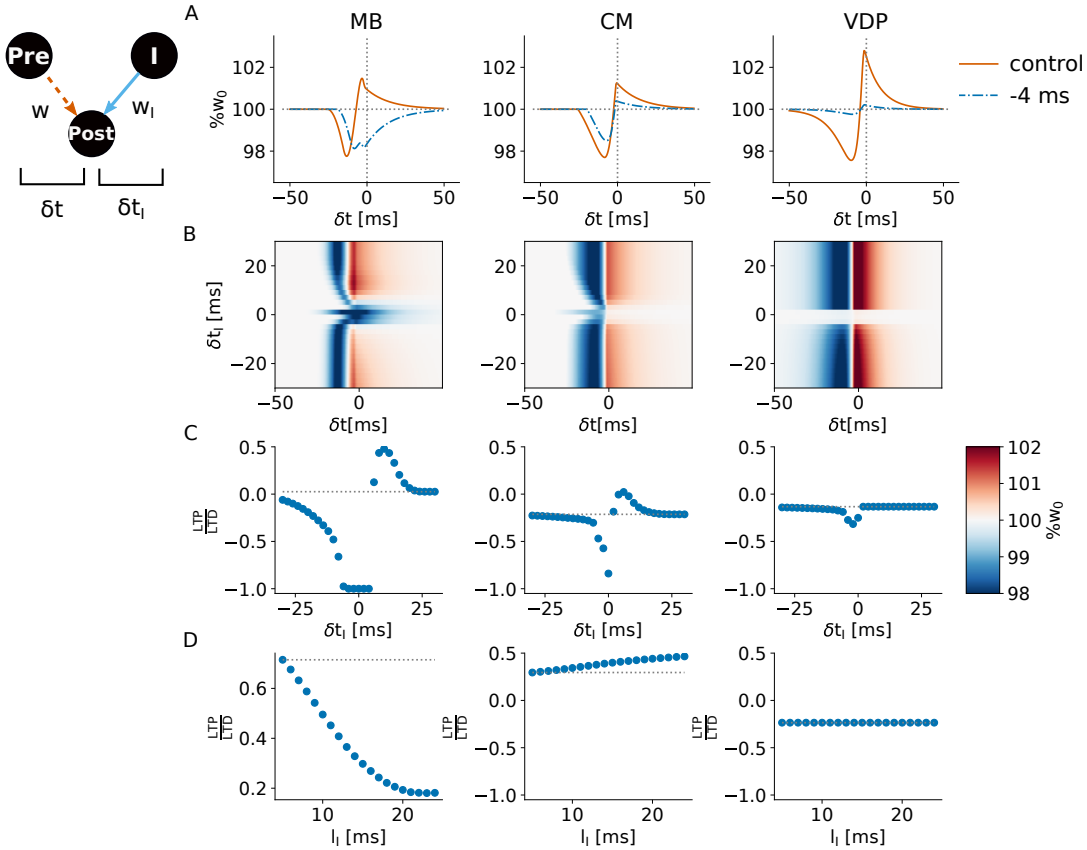

Figure 1. Precise inhibitory spike timing regulates the LTP-to-LTD ratio in STDP experiments. A standard STDP protocol is extended by an inhibitory presynaptic neuron as indicated by the circuit diagram. The inhibitory spike time δt_I is measured relative to the postsynaptic spike. Models are indicated by the 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.

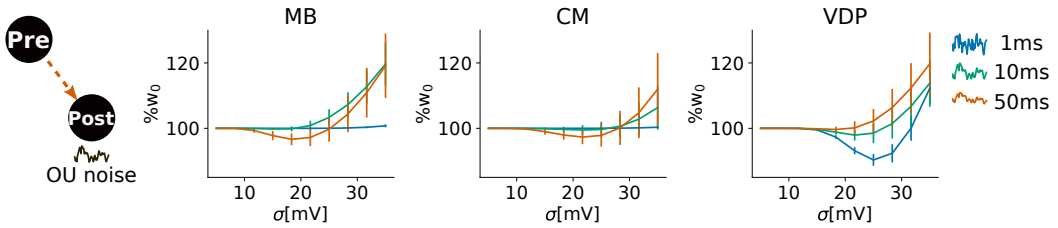

Figure 2. The induction principle of the plasticity model decides its sensitivity to fast fluctuations in the postsynaptic membrane potential. Instead of recording the membrane potential, υ is replaced by an Ornstein-Uhlenbeck (OU) noise. The simulation is run for 10s. The presented results are averaged 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.

tively.

The results confirm the proposed effects of the postsynaptic filter: the VDP has a higher sensitivity to fast fluctuations than the other models (Fig. 2). Furthermore, models which filter the membrane potential need prolonged depolarisations to induce plasticity. These can either occur due to slow fluctuations or large amplitude. This is why the minimal amplitude for which the MB and CM model predict synaptic weight changes shifts towards larger values with increasing fluctuation speed.

Synaptic EI balance and weight stability

The Ornstein-Uhlenbeck noise experiment demonstrates that the speed of membrane potential fluctuations influences the models' predictions. However, that experiment does not consider what synaptic input causes the membrane potential to change as well as ignores any active spiking of the neuron. In order to investigate how the synaptic plasticity and postsynaptic activity co-varies dependent on the excitation/inhibition (EI) balance of the synaptic input, we simulated a Leaky Integrate-and-Fire (LIF) neuron receiving input from excitatory and inhibitory Poisson neurons (Figure 3A). This experiment enables us to examine whether the firing rate dependency follows a linear or non-linear behaviour and might lead to pathological weight depression/potentiation as present in other Hebbian learning rules (***Litwin-Kumar and Doiron, 2014***; ***Keck et al., 2017***).

The postsynaptic frequency is driven by excitatory input. Therefore, excitatory-dominated synaptic input (dark red) causes high firing rates and inhibition-dominated input (dark blue) results in low firing. With increasing firing rate the mean membrane potential and its variance increase as well. However, the dependency on the synaptic input composition is counter-intuitive. At constant firing rates, the smaller the EI balance (I > E), the higher the mean membrane potential, while the variance decreases, see Figure 3B and C. These results can also be obtained when simulating a Hodgkin-Huxley neuron (***Hodgkin and Huxley, 1952***) as indicated by the results in the Appendix.

We are particularly interested in the interaction between EI balance and LTP/LTD update. In general, potentiation of excitatory synapses increases the contribution of excitation and increases the EI balance while LTD increases the inhibitory contribution and shifts the EI balance towards stronger inhibitory-dominated input compositions. This relationship is flipped for inhibitory synapses. LTP increases the inhibitory contribution and reduces the EI balance while LTD shifts the EI balance towards excitation-dominated input by weakening the inhibitory input.

All models predict pathological behaviours in the low and/or high-frequency range in which the EI balance is either minimal (open-faced stars in Figure 3D) or maximal (solid stars in Figure 3D). At low firing rates, LTD updates dominate for both excitatory and inhibitory synapses. At excitatory synapses, this causes a negative feedback loop: LTD of the synapse decreases the received excitatory input, decreasing the postsynaptic firing which in turn causes stronger LTD. Essentially, the excitatory synaptic weight is removed. Due to the opposite effect of LTD at inhibitory synapses

on the EI balance, this negative feedback loop is not as effective at inhibitory connection as at excitatory synapses. Inhibitory LTD increases the EI balance. If excitatory and inhibitory input are balanced (E=I), the mean synaptic update is 0 (indicated by the horizontal dotted line). Therefore, inhibitory synapses can in principle stabilize at small firing rates without being 0 given balanced synaptic input.

The counterpart to this negative feedback loop leads to exploding weights at high firing rates, but can only be present in the development of excitatory synapses. Both the CM and VDP models show this behaviour. Any input which leads to LTP at higher firing rates progressively strengthens the excitatory input. Therefore, postsynaptic firing rates will increase, which promotes high-amplitude LTP updates further strengthening the excitatory contribution. Consequently, the weights grow infinitely large without a pre-defined upper bound.

These pathological behaviours have been previously described for Hebbian learning rules and require additional non-Hebbian mechanisms e.g. homeostatic plasticity to stabilise recurrent networks (***Zenke et al., 2013***; ***Keck et al., 2017***; ***Zenke and Gerstner, 2017***; ***Litwin-Kumar and Doiron, 2014***). Our focus lies on the isolated effects of the Hebbian term of the voltage-dependent plasticity models. This is why we excluded the homeostatic terms of the models introduced in ***Clopath et al.*** (***2010***); ***Meissner-Bernard et al.*** (***2020***).

The MB model does not predict this positive feedback loop for excitatory weight development. At high frequencies, excitatory synapses undergo LTD. Instead of increasing the EI balance, the excitatory contribution weakens. This progressively reduces the postsynaptic firing rate and shifts the synaptic weight towards a regime in which it can stabilise as described in more detail below. Nonetheless, it is important to note, that an increase in the potentiation amplitude parameter a^+ leads to an overall increase of LTP updates resulting in an overall prediction of LTP for all firing rates. Therefore, the MB can in principle produce the positive feedback loop present for excitatory weight development. However, we do not consider those settings functional due to the absence of LTD.

As in all Hebbian learning rules, the predicted weight updates are dependent on the presynaptic activity. For inhibitory synapses at high postsynaptic firing rates, the presynaptic activity is low due to I « E. Thus, synaptic updates are rare increasing their probability to be stable. Furthermore, the MB model predicts inhibitory LTD at high firing rates. This increases the EI balance and increases the firing rate. Updates become rarer and the synapse stabilises despite the high frequency. In contrast, the CM and VDP predict LTP for high frequencies. The EI balance decreases and the postsynaptic firing rate reduces increasing update amplitude and frequency (Figure 3D bottom).

We are particularly interested in answering whether synapses can be (dynamically) stable when the input composition changes (E>I to I>E) due to plasticity events at intermediate firing rates. In the following, it is assumed for simplicity that the excitatory and inhibitory synapses develop independently of each other. We will give a short interpretation of co-development and its implication for weight stability at the end. Furthermore, the frequency ranges we discuss concern those in which a change in EI balance at constant firing rates leads to a change in the sign of the predicted synaptic update. Therefore, we want to know at which points plasticity changes from LTD to LTP. These regions are highlighted in Figure 3D.

The MB excitatory weight predictions show the broadest frequency range for which changes in synaptic EI balance can lead to a dynamically stable synapse. Despite the narrower frequency range, the VDP results are characterised by a similar relationship between EI balance and predicted weight update. Excitation-dominated synaptic input results in LTD updates. This leads to a reduction of excitation and pushes the EI balance towards stronger inhibition. In contrast, inhibition-dominated input leads predominantly to LTP. This strengthens the excitatory contribution and pushes the input composition back towards I<E. Therefore, the synaptic weight can fluctuate between these states and show quasi-stable behaviour over longer timescales despite changes in the received synaptic input.

The CM results are characterised by the opposite relationship between shifts in synaptic EI

balance and LTP/D predictions. Excitation-dominated synaptic input leads to LTP updates. The excitatory contribution increases, pushing the EI balance towards stronger excitation-domination. This leads to positive feedback between LTP and EI balance as previously described. Therefore, the excitatory weight starts to grow without compensation. In contrast, for synaptic input which is characterized by I>E, the model predicts LTD and strengthens the inhibition-dominance of the synaptic input, following the pattern of the previously described negative feedback loop. Instead of pushing towards converging weight update predictions with changing EI balance as predicted by the MB and VDP models, the CM predicts diverging weight updates pushing the synaptic weight away from a stable point. This stable point is the intersection of the weight predictions with the 100% indicator line when I=E.

Due to the opposite relationship between EI balance and inhibitory LTP/D, the frequency ranges for quasi-stable behaviour and conditions leading to it are significantly broader in all models' predictions. Additionally, the inhibitory weights have a higher chance to stabilise at small but non-zero weights. The predicted LTD decreases the inhibition received by the postsynaptic neuron causing the EI balance to increase. Thus, successive LTD updates have progressively smaller amplitudes. If the decrease in inhibition is accompanied by an increase in the postsynaptic frequency, the decrease in LTD amplitude is further promoted in all models.

Furthermore, inhibitory LTP at intermediate firing rates reduces the EI balance. An additional reduction in postsynaptic firing rates shifts the synaptic conditions towards LTD predictions. This however increases the EI balance leading to higher firing rates and LTP which leads to quasi-stable behaviour of the inhibitory synapses in all models.

When considering the co-development of inhibitory and excitatory synapses, the flipped relationship between EI balance and LTP/LTD effects can compensate for each other. This happens when excitatory synapses undergo LTD and decrease the excitatory contribution while simultaneously inhibitory synapses undergo LTD and decrease the inhibitory contribution. The result is similar for the opposite case for excitatory LTP and inhibitory LTP. If both weight changes result in a similar amplitude change in EI balance of the synaptic input, they cancel each other out. As a consequence, both inhibitory and excitatory synapses would co-fluctuate between these two states. Due to the previously described result for the excitatory weight development depending on the EI balance, such interactions are possible in the MB and VDP models. Due to the diverging excitatory weight development in the CM model, this behaviour is further accelerated by the inhibitory plasticity.

An interesting final observation is that the inhibitory LTP at high frequencies might be able to counteract the pathological weight development in the CM and VDP. Decreasing the EI balance decreases the excitatory LTP amplitude while increasing the inhibitory LTP amplitude. Whether this pattern is capable of pushing the postsynaptic activity and synaptic EI balance received by the neuron towards a stable point, however, is dependent on the relative synaptic strengths and conductances of excitation and inhibition.

Overall, these results indicate that the MB and VDP models predict regions in which excitatory and inhibitory synapses can co-stabilise due to the relationship between activity and synaptic EI balance. In contrast, the CM shows diverging patterns for excitatory synapses. Therefore, it is more likely that they develop strongly binary patterns with either 0 weights or maximal weights. However, the postsynaptic frequency dependency in all models is non-linear which is a prerequisite for the development of functionally relevant structures such as receptive field formation (***Brito and Gerstner, 2016***).

Temporal sensitivity determines plasticity predictions during presentation onset

The analysis so far has established that the models predict different plasticity outcomes for differences in E-I balance. Next, we asked whether these differences are also present in the weight development during the presentation of structured input activity. The structure of naturalistic stimuli is complex, and they can often be decomposed into sources or causal factors, a process referred

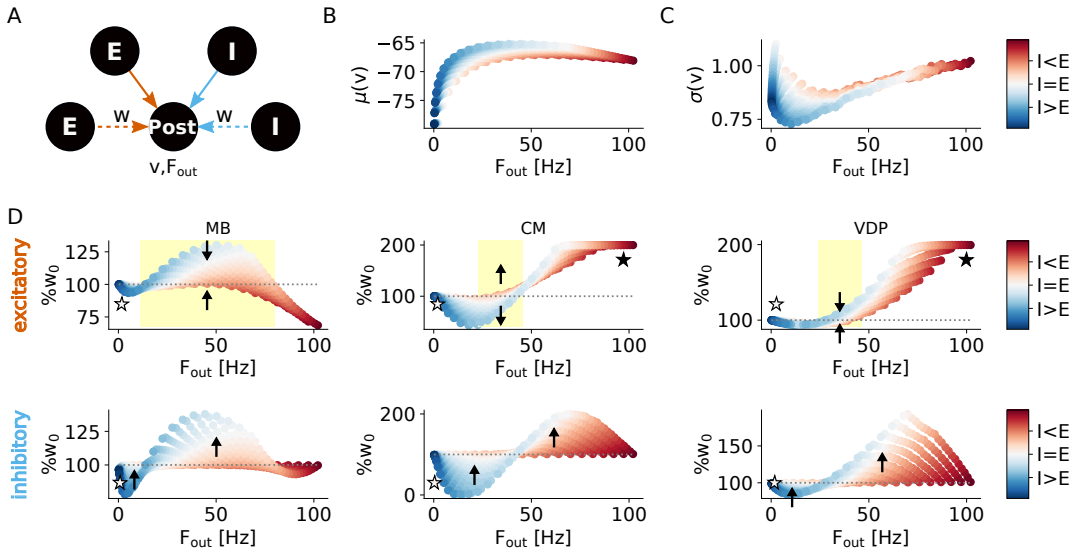

Figure 3. The fluctuations of the sub-threshold membrane potential influence plasticity predictions in the VDP. **A** A postsynaptic neuron is driven by 100 excitatory and 100 inhibitory Poisson neurons with firing rates $F_E, F_I \in [0, 100)$Hz (5Hz resolution). We measure the postsynaptic membrane 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. The open-faced star indicates the regime of pathological behaviour at low postsynaptic frequencies (negative plasticity-activity feedback). The black star indicated the regime of pathological behaviour at high postsynaptic frequencies (positive plasticity-activity feedback). Yellow-backed regions indicate intermediate frequency prediction of special interest. Arrows indicate the direction of EI balance shift due to the predicted synaptic plasticity update.

to as causal modelling (***Dayan and Abbott, 2005***). In order to gain an understanding of how the temporal sensitivity and inhibitory regulation affect the ability of the models to develop causal representations, we constructed a toy problem with three distinct inputs which overlap in their activity pattern. They activate three postsynaptic populations which receive pre-specified inhibitory input. Only the postsynaptic population which is supposed to encode the currently non-active input receives inhibition. Figure 4A and B show an illustration of the set-up and activity. We are specifically interested in the difference between the weight update predictions between the encoding weights (solid arrows in Figure 4A and data shown in C) and the non-encoding weights (indicated by the dotted arrows, data shown in D).

We consider two key factors when comparing the weight development between plasticity models: the amplitude difference between encoding and non-encoding weight change prediction and the update prediction surrounding stimulus changes. The difference in non-encoding and encoding weight change prediction is the most apparent. The MB and CM models both predict that the maximal amplitude of the weight change for the non-encoding weights is twice that of the encoding weights. In contrast, in the VDP model, the non-encoding weight change amplitudes are slightly smaller than those predicted for the encoding synapses. Another distinct difference between the models is the time point at which the peak of the non-encoding weight updates is reached relative to the presentation onset. The MB and CM predict maximal weight change at the time of stimulus change with a slow increase and decrease which roughly spans one stimulus presentation. For the VDP in contrast, the predicted updates of the non-encoding weights reach their maximum shortly after the stimulus changes and quickly relaxes back to their limit. This indicates that the non-encoding weights undergo smaller plasticity updates in the VDP circuit in comparison to the MB and CM models.

Similar patterns are present in the prediction for the weight change of the encoding weights. Both MB and CM show slower decreases indicating a slower reaction to the changing conditions. Interestingly, the MB and VDP predictions indicate target update sizes which are approached during a single presentation window, while the CM only demonstrates one such limit level. Moreover, in both the MB and VDP model the synapses connecting to the inhibited population relax towards the lower limit which is approximately 0 in the VDP circuit. The MB lower limit might be 0 as well, however, the relaxation speed is too slow to reach the limit. In contrast to this, the CM model does not react with a decrease in weight update size to inhibition. The synapses connecting to the inhibited population show an increase in LTP amplitude after inhibition onset, while the synapses which connect to the population which has not previously or currently received inhibition stay at the limit level and do not show any significant deviations from this level of weight change. Therefore, the MB and VDP actively react to the inhibition with the expected behaviour: The population receiving inhibition is undergoing less LTP. In contrast, the CM increases the encoding weights during active inhibition.

In summary, the MB model reacts slowly to a presentation change and actively reduces the weight growth of inhibited neurons. The CM reacts to changes in the activity pattern regardless of whether it is due to inhibition onset or termination. However, the latter elicits faster changes in the weight predictions with a higher amplitude. The VDP is the only model which promotes the growth of the encoding weights of the active input while quickly reducing and pausing plasticity of the non-active input encoding synapses. Furthermore, it reacts quickly to the changes in activity. Overall these patterns indicate that the overlap of developed representation or encoding learned via the MB and CM is larger. This is caused by the broader effective time window of inhibition which is a result of the postsynaptic filtering. However, the CM develops representation by actively promoting LTP after presentation onset. This can lead to qualitatively different representations in comparison to the composition-dependent promotion of LTP and LTD in the MB and VDP models.

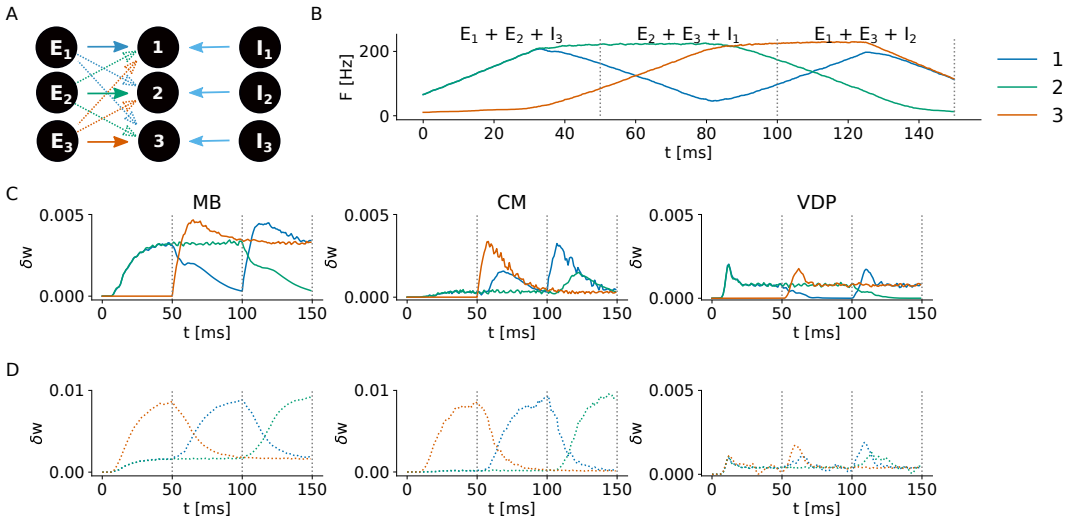

Figure 4. The higher temporal precision of the inhibitory regulation increases the VDP's sensitivity and reaction speed to changes in the received synaptic input. **A**: Three distinct excitatory Poisson neuron populations (N=10) drive neural activity of 3 LIF neurons. Each LIF neuron receives inhibition at a predefined time from its dedicated inhibitory Poisson neuron population (N=10). Solid arrows indicate encoding weights which connect a specific Poisson population with the neuron which is inhibited when the input population is in-active. Dotted lines indicate non-encoding weights. All results are means of 100 independent runs. **B**: Postsynaptic firing rate of the excitatory neurons and the activation pattern of the input and inhibitory populations. **C** Synaptic update over time of the encoding weights. **D**: Synaptic updates of the non-encoding weights.

Temporal precision of inhibitory regulation shapes learned representation

The results so far have been an analysis of the inhibitory regulation. With the help of a toy experiment, we develop a hypothesis about the structure of the representation when learning with a specific plasticity model. In order to test their validity, we demonstrate functional structure development in recurrent networks and connect these to well-known properties of the visual cortex.

Brito and Gerstner (***2016***) demonstrated that non-linear Hebbian plasticity implements a general principle for receptive field development and unified the biologically motivated plasticity models with principled approaches such as Independent Component Analysis (ICA). The non-linearity specifically refers to the frequency dependency. Thus, the voltage-dependent plasticity models implement non-linear Hebbian plasticity. As a consequence, all models should be able to perform ICA. To test this, 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 stripes (Fig. 5A). 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. These cells do not show any discernible pattern in their receptive fields due to all their synapses saturating at the maximal weight and are always active. It also develops neuron with the opposite connectivity, uniform synapses with 0 weights. Neurons with this connectivity are predominantly silent as their activity is not driven by feedforward input but only recurrent activity. They are also present in the network whose synapses learn via the VDP model.

The separation score S_i measured the quality and quantity of the receptive fields. A score above the green line indicates a receptive field which encodes one independent component disregarding

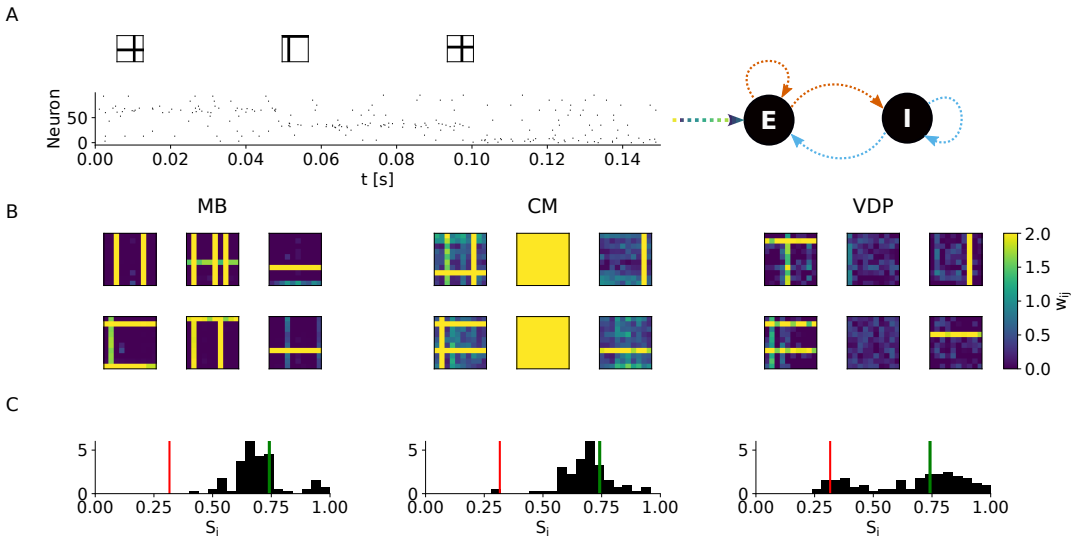

Figure 5. The excitatory population develops independent components when presented with temporally correlated signals. **A**: An EI network is presented with the bars-and-stripes task (***Földiak, 1990***). **B**: Learned receptive fields. **C**: Separation scores of the excitatory neurons. A neuron which encodes a single bar perfectly has $S_i = 1.0$. The green line indicates the upper bound for a multi-factor cell. The red line indicates the corresponding lower bound. Any score below this characterises random receptive fields with no discernable pattern. These scores are computed with cleaned-up receptive fields to reduce the contribution of spurious patterns present in near-0 weights. All weights $w_{ij} < \mu(w)$ are set to 0. The mean is computed over all receptive fields of the network.

whether this is a bar or stripe. Scores below the red line indicate a random receptive field. Multi-factor cells score in between the two. The histograms plotted in Figure 5D show that the VDP develops the highest number of single-factor cells and overall heterogeneous receptive fields. The separation score distribution is bimodal with one peak at single-factor scores and a second one at the transition from random receptive field scores to multi-factor cells. The tight temporal window of the inhibitory regulation in the VDP results in a higher degree of representational separability which agrees with our previous results. Furthermore, the pruning of synaptic weights causing a subset of neurons to be silent shows hallmarks of sparse coding.

The competition and inhibitory effects which drive input factorisation also drive the receptive field development in V1. 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 persists 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 show clearly discernable clusters like the MB network results. In direct comparison, the statistics of the VDP synapses cover the space more uniformly. This indicates a higher degree of heterogeneity of the receptive fields. Therefore, the VDP network developed the most diverse representations.

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° with a small bias for horizontal gratings (0°). While the CM and VDP orientation curve peaks match the input response peak, the MB peak response is slightly shifted to the left (Fig. 6D). The response difference for vertical and horizontal orientations is the smallest in the receptive fields developed by the VDP, followed by the bias presented in the CM receptive fields while the MB shows the strongest bias difference. All of them however match the relatively higher response to horizontal edges. A unique characteristic of the CM response curve is the prediction of a slope change at around 70°. All other data including the input response curve are characterised by a linear decrease in the response to progressively vertically oriented gratings (60°-90°). Taking together the representation features developed by the VDP match the input biases the best in comparison to the other two models.

In summary, 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 developed representations of a recurrent network driven by structured activity.

Discussion

These results suggest that the inhibitory regulation through hyperpolarisation of the membrane potential is a plausible model for developing functional connectivity patterns. A detailed analysis of three voltage-dependent plasticity models highlights a trade-off between the timing precision and strength of the inhibitory regulation. A stronger filtering of the postsynaptic membrane potential by the induction mechanism results in a wider time window where inhibitory spiking impacts synaptic weight updates. If the model also filters the presynaptic spike train, reducing the amplitude of the LTP update can even switch from LTP to LTD for pre-post spike pairs.

Functional experiments have shown a correlation between the duration of the regulation effect and representation separation. When inhibitory regulation is effective over a longer duration, representation learning relies on the average of postsynaptic activity over time. As a result, change points are less effective at driving representation apart leading to overlaps and greater similarities. Therefore, the model with the smallest regulation amplitude but shortest temporal window best captures crucial differences between inputs, leading to more distinct and accurate representations. By reducing the influence of past activity, this model minimises overlaps and improves the separation of learned features, enhancing the overall quality of the representation.

Biologically, these results suggest that synapses located on distal dendrites can rely more on the local, fast-changing information (***Branco and Häusser, 2011***; ***Oviedo and Reyes, 2005***; ***Magee et al., 1998***), if plasticity has higher temporal sensitivity, such as proposed in the VDP. This implies that distal synapses might develop more precise and distinct representations, as they are less influenced by longer temporal integration. In contrast, at proximal synapses, the voltage signal corresponds to information integrated over longer periods of time and over many synapses. For this case our simulations suggest plasticity will yield more overlapping representations. As the models

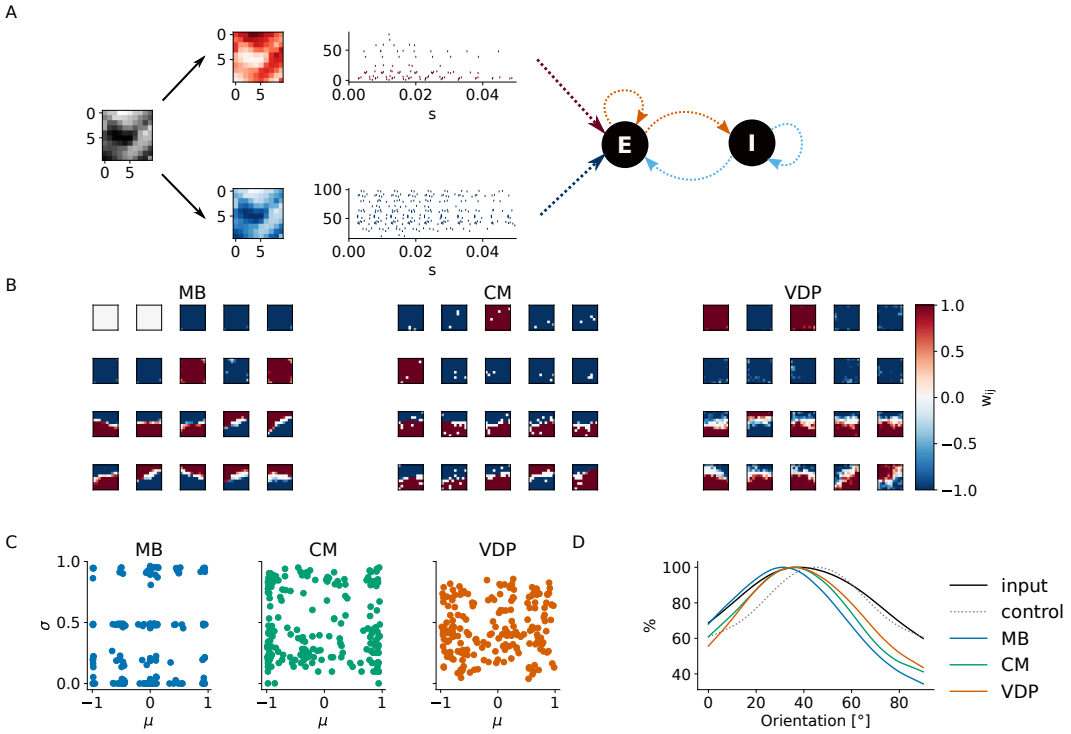

Figure 6. The three models develop characteristic receptive fields. **A**: A fully-plastic recurrent EI network is presented with CIFAR10 input patches which are represented by the firing rates of an On- and Off-population. **B**: The developed receptive fields show specific characteristics depending 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**: Mean and standard deviation of the receptive fields. **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**.

by ***Clopath et al.*** (***2010***) and ***Meissner-Bernard et al.*** (***2020***) do not have a high temporal sensitivity, they are predicted to behave similarly at distal and proximal synapses. In contrast, the VDP predicts a difference in representation learning at distal and proximal locations. This prediction is testable in biophysical neuron models, and differences in distal and proximal tuning properties should be detectable in imaging experiments.

Neurons and synapses are highly heterogeneous: from genetic markers to protein and channel sub-type composition and local distribution (***Cizeron et al., 2020***; ***Zhu et al., 2020***; ***Magee et al., 1998***). Thus it is unlikely that a single rule can account for all behaviours. Here we only considered synapses reproducing the standard causal STDP kernel, while several different STDP kernel forms have been identified experimentally (***Abbott and Nelson, 2000***). These models cannot reproduce effects such as the the reversal from LTD to LTP described by ***Paille et al.*** (***2013***), as it is incompatible with voltage-dependent plasticity models where the LTD threshold is always below the LTP threshold (***Ngezahayo et al., 2000***).

While we analyse and discuss the stability of a single, isolated synapse, this work does not explicitly investigate network stability. All functional simulations have been run for an extensive time to ensure that the representations are stable despite ongoing plasticity. However, the plasticity models only model Hebbian principles and do not include homeostatic components. The interaction between Hebbian and homeostatic mechanisms acting on biologically plausible time scales thus remains a question for future investigations.

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 the learned representation 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 characteristics of ICA and can develop receptive fields from naturalistic input, models using the instantaneous membrane potential tend to learn highly separate and heterogeneous features which are more closely tuned to the input. This work predicts that synapses whose plasticity is driven by local, fast-fluctuating information like those on distal dendrites develop causal representations accurately representing the input composition due to the precisely timed inhibitory plasticity regulation. Investigating the relationship and interplay between local EI balance and LTP/LTD at different dendritic sites warrants further investigations to accurately describe the functional role of inhibitory regulation of plasticity in biological neurons.

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.

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 postsynaptic neuron i.

$$S_i = max(c_S(r_i, [0°, 90°])) \tag{11}$$

This measures the maximal 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 $90°$. Before computing S_i, we clean the receptive fields r_i by discarding any weights which are smaller than the overall average postsynaptic connection. This decreases the contribution of residual patterns in non-functional,

small synapses. We take the maximum because it accurately describes whether the receptive field represents a single factor. When the receptive field encodes a single factor the cosine similarity c_S to this factor is 1.0, while the similarity to all remaining factors is smaller than 1. The presence of multiple factors in the receptive fields, disregarding whether in the form of a cross or two non-intersecting factors, decreases the cosine similarities between the receptive field and all single factors. Therefore we can compute an upper bound for multi-factor presence of two factors, which simultaneously provides a lower bound for a single-factor representation. At the same time, we can also determine the upper bound similarity score for non-informative random fields.

Code availability

Code to reproduce all presented experimental results is available at: https://github.com/prubisch/inhibitory_plasticity_regulation_in_voltage_dependent_plasticity_models. The BRIAN2 package (***Stimberg et al., 2019***) was used for all simulations.

Acknowledgements

P.R. was supported by EPSRC DTA (PhD scholarship). We would like to thank the members of the ′Neurons and Systems′ group for helpful discussions and comments.

Author contributions

Conceptualization, P.R. and M.H.H.; methodology, P.R.; investigation, P.R.; writing. P.R., M.I.S., M.H.H.; supervision, M.H.H.

Declaration of interests

All authors declare no competing interests.

References

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

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.

Branco, T. and Häusser, M. (2011). Synaptic integration gradients in single cortical pyramidal cell dendrites. *Neuron*, 69(5):885–892.

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.

Cizeron, M., Qiu, Z., Koniaris, B., Gokhale, R., Komiyama, N. H., Fransén, E., and Grant, S. G. (2020). A brainwide atlas of synapses across the mouse life span. *Science*, 369(6501):270–275.

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.

Clopath, C. and Gerstner, W. (2010). Voltage and spike timing interact in stdp–a unified model. *Frontiers in synaptic neuroscience*, 2:25.

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.

Dayan, P. and Abbott, L. F. (2005). *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT press.

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.

Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500–544.

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.

Keck, T., Toyoizumi, T., Chen, L., Doiron, B., Feldman, D. E., Fox, K., Gerstner, W., Haydon, P. G., Hübener, M., Lee, H.-K., et al. (2017). Integrating hebbian and homeostatic plasticity: the current state of the field and future research directions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1715):20160158.

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.

Litwin-Kumar, A. and Doiron, B. (2014). Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nature communications*, 5(1):5319.

Magee, J., Hoffman, D., Colbert, C., and Johnston, D. (1998). Electrical and calcium signaling in dendrites of hippocampal pyramidal neurons. *Annual review of physiology*, 60(1):327–346.

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.

Oviedo, H. and Reyes, A. D. (2005). Variation of input-output properties along the somatodendritic axis of pyramidal neurons. *Journal of Neuroscience*, 25(20):4985–4995.

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.

Person, A. L. and Raman, I. M. (2010). Deactivation of l-type ca current by inhibition controls ltp at excitatory synapses in the cerebellar nuclei. *Neuron*, 66(4):550–559.

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.

Stimberg, M., Brette, R., and Goodman, D. F. (2019). Brian 2, an intuitive and efficient neural simulator. *elife*, 8:e47314.

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.

Zenke, F. and Gerstner, W. (2017). Hebbian plasticity requires compensatory processes on multiple timescales. *Philosophical transactions of the royal society B: biological sciences*, 372(1715):20160259.

Zenke, F., Hennequin, G., and Gerstner, W. (2013). Synaptic plasticity in neural networks needs homeostasis with a fast rate detector. *PLoS computational biology*, 9(11):e1003330.

Zhu, F., Collins, M. O., Harmse, J., Choudhary, J. S., Grant, S. G., and Komiyama, N. H. (2020). Cell-type-specific visualisation and biochemical isolation of endogenous synaptic proteins in mice. *European Journal of Neuroscience*, 51(3):793–805.

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.

Appendix

Illustration of the Voltage-Dependent Pathway model

The VDP models the downstream processes of plasticity induction as a simplified pathway-dynamic which is activated by depolarisation of the postsynaptic membrane. Therefore, the overall weight update from a single pairing can only be determined after several milliseconds when the pathway activations for LTP and LTD have both returned to 0. It emphasizes the temporal competition between LTP and LTD pathway activation as the central mechanism determining the sign and amplitude of the synaptic change. The illustration provided in Figure 7 visualises the mechanism used in the VDP.

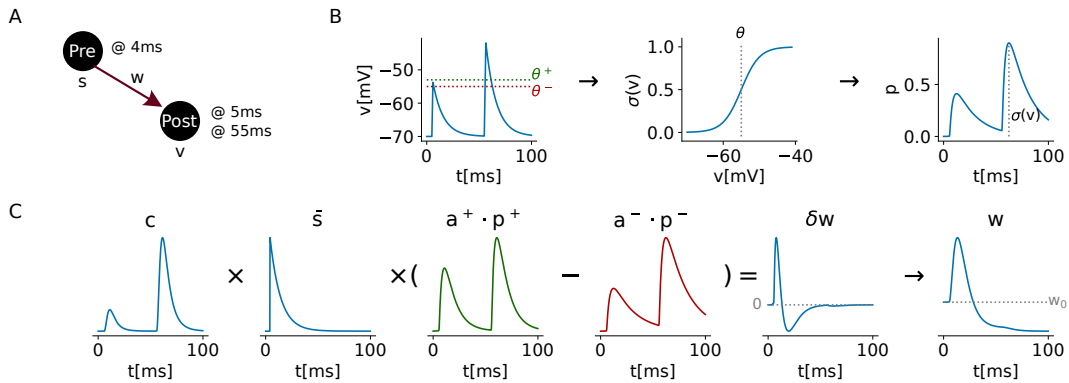

Figure 7. Illustration of the VDP model. **A**: Stimulation schema. The presynaptic spike train s and the postsynaptic membrane potential v are measured to compute the synaptic weight update. The time points denote spike and postsynaptic stimulation time to elicit the activation of the traces in B and 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.

Reproduction of central experimental results with the VDP

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.

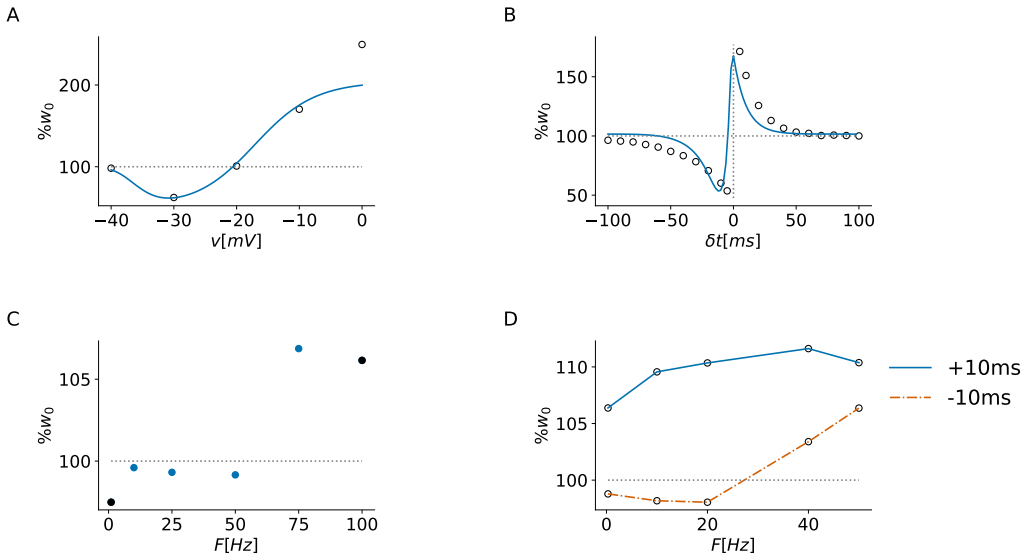

Figure 8. Predictions of the VDP for several experimental protocols. **A**: The voltage dependency follows the relationship as described by ***Ngezahayo et al.*** (***2000***) closely. **B**: The VDP can reproduce the standard STDP kernel as measured by ***Froemke and Dan*** (***2002***). The open-faced 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 neuron model used.

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 :

$$s = \frac{\sum_{i=0}^{t} \delta w^{\delta i} - \delta w^{-\delta i}}{\sum_{i=-t}^{t} |\delta w^{\delta i}|} \tag{12}$$

and the LTP/LTD ratio r is defined as :

$$r = \frac{\sum_{i=-t}^{t} (\delta w^{\delta i} \cdot \delta w^{\delta i} > 0) - \sum_{i=-t}^{t} (\delta w^{\delta i} \cdot \delta w^{\delta i} < 0)}{\sum_{i=-t}^{t} |\delta w^{\delta i}|} \tag{13}$$

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

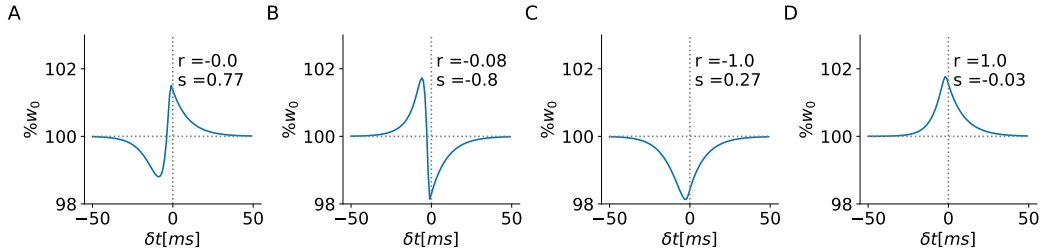

Figure 9. Different STDP kernel shapes reproduced by the VDP. The differences can be measured by calculating the kernel shape s and LTP/LTD ratio r. **A**: Balanced Hebbian STDP kernel with r = 0.0 and s = 0.77. **B**: Balanced anti-Hebbian STDP kernel with r = -0.08 and s = -0.8. **C**: LTD-dominated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

EI-balance and membrane potential statistics in a Hodgkin-Huxley regular spiking neuron

All simulations use a LIF neuron as described in Section Methods. As this is a simplified neuron model without an explicit spike shape estimation, it is central to test whether the relationship between EI balance and membrane potential statistics holds in neuron models which approximate the spike shape. The results of the same simulation with a regular spiking Hodgin-Huxley model (***Hodgkin and Huxley, 1952***) are illustrated in Figure 10. It confirms that inhibitory dominated synaptic input results in a higher mean membrane potential at similar firing rates in comparison to excitatory-dominated synaptic input. However, the dip in the variance of the membrane potential at low firing rates is not present in this simulation, indicating that this is an artefact caused by the cut

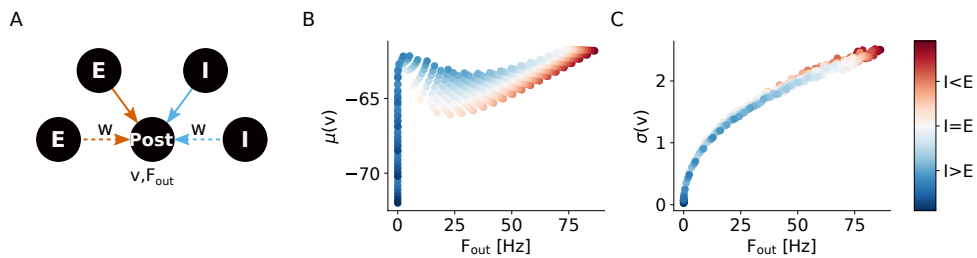

Figure 10. Different STDP kernel shapes reproduced by the VDP. The differences can be measured by calculating the kernel shape s and LTP/LTD ratio r. **A**: Balanced Hebbian STDP kernel with r = 0.0 and s = 0.77. **B**: Balanced anti-Hebbian STDP kernel with r = -0.08 and s = -0.8. **C**: LTD-dominated nearly-symmetric kernel. r = -1.0, s = 0.27. **D**: LTP-dominated nearly symmetric STDP kernel. r = 1.0 and s = -0.03.

Influence of time constants on the ICA mechanism

A central question which arises, when comparing different synaptic plasticity models is whether similar results can be achieved simply by adapting model parameters. In this work we argue that the temporal extent of the inhibitory regulation as well as the sensitivity is inherent to models. To demonstrate that this is true for the functional mechanism of developing separated representation of presented input pattern, we vary the time constants and measure how they influence the slope of weight change of the mechanistic simulation illustrated in Figure 4. Due to the differences in the nature of the postsynaptic temporal parameters, the results in Figure 11 only show the effect of the postsynaptic parameters of the MB and CM as these are directly comparable. They also demonstrate that the presynaptic temporal constant has the biggest effect by allowing a longer effective pairing window in these models. The VDP has slower postsynaptic dynamics which is why the presynaptic constant has a less strong effect in this model.

Within-network receptive field similarity

Figure 12A shows further results on the clustering of the receptive field statistics. To investigate the robustness of the clusters, the statistics were randomly shuffled (grey data). The similarity between receptive fields of the same network and its relationship to the variance of the receptive fields, Figure 12B-D, show the gradual change of receptive fields in the CM and VDP network, while the MB network builds clusters of receptive field with similar structures which is also captured by the clustering in their descriptive statistics.

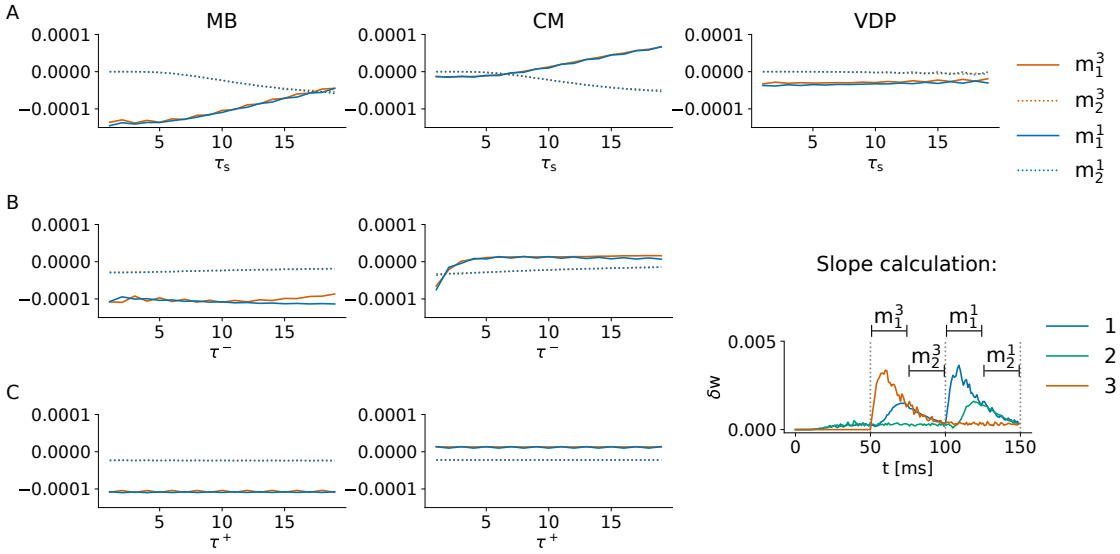

Figure 11. Different STDP kernel shapes reproduced by the VDP. The differences can be measured by calculating the kernel shape s and LTP/LTD ratio r. **A**: Balanced Hebbian STDP kernel with r = 0.0 and s = 0.77. **B**: Balanced anti-Hebbian STDP kernel with r = -0.08 and s = -0.8. **C**: LTD-dominated nearly-symmetric kernel. r = -1.0, s = 0.27. **D**: LTP-dominated nearly symmetric STDP kernel. r = 1.0 and s = -0.03.

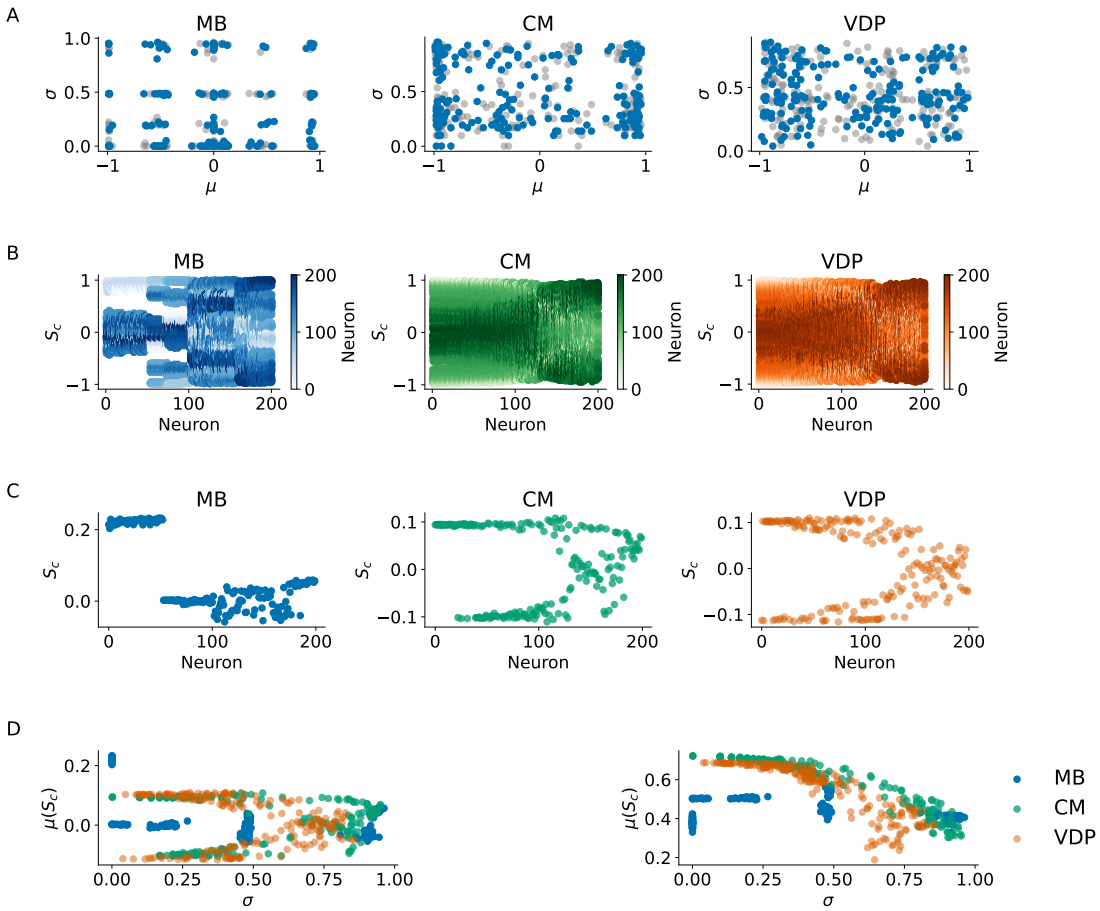

Figure 12. Clustering and further similarity analysis of the receptive fields. **A**:Original pairings of mean μ and variance σ (blue) and reshuffled pairs (gray). **B**: Cosine similarity $S_c(i, j)$ of the neurons i and j receptive field. **C**: Mean cosine similarity $\mu(S_c(i, j))$ or neuron i. **D**: Relationship between the standard deviation σ and $\mu(S_c(i, j))$ (left) and $\mu|(S_c(i, j))|$ (right).