

Pre- and postsynaptically expressed spike-timing-dependent plasticity contribute differentially to neuronal learning

Beatriz E. P. Mizusaki^{1,2,3}, Sally S. Y. Li¹, Rui Ponte Costa^{3,4,5}, P. Jesper Sjöström^{1,*},

1 Centre for Research in Neuroscience, Brain Repair and Integrative Neuroscience Programme, Departments of Medicine, Neurology and Neurosurgery, The Research Institute of the McGill University Health Centre, Montreal General Hospital, Montreal, Quebec, Canada

2 Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

3 Computational Neuroscience Unit, Department of Computer Science, SCEEM, Faculty of Engineering, University of Bristol, Bristol, United Kingdom

4 Department of Physiology, University of Bern, Bern, Switzerland

5 Centre for Neural Circuits and Behaviour, Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, United Kingdom

* jesper.sjostrom@mcgill.ca

Abstract

A plethora of experimental studies have shown that long-term synaptic plasticity can be expressed pre- or postsynaptically depending on a range of factors such as developmental stage, synapse type, and activity patterns. The functional consequences of this diversity are not clear, although it is understood that whereas postsynaptic expression of plasticity predominantly affects synaptic response amplitude, presynaptic expression alters both synaptic response amplitude and short-term dynamics. In most models of neuronal learning, long-term synaptic plasticity is implemented as changes in connective weights. The consideration of long-term plasticity as a fixed change in amplitude corresponds more closely to post- than to presynaptic expression, which means theoretical outcomes based on this choice of implementation may have a postsynaptic bias. To explore the functional implications of the diversity of expression of long-term synaptic plasticity, we adapted a model of long-term plasticity, more specifically spike-timing-dependent plasticity (STDP), such that it was expressed either independently pre- or postsynaptically, or in a mixture of both ways. We compared pair-based standard STDP models and a biologically tuned triplet STDP model, and investigated the outcomes in a minimal setting, using two different learning schemes: in the first, inputs were triggered at different latencies, and in the second a subset of inputs were temporally correlated. We found that presynaptic changes adjusted the speed of learning, while postsynaptic expression was more efficient at regulating spike timing and frequency. When combining both expression loci, postsynaptic changes amplified the response range, while presynaptic plasticity allowed control over postsynaptic firing rates, potentially providing a form of activity homeostasis. Our findings highlight how the seemingly innocuous choice of implementing synaptic plasticity by single weight modification may unwittingly introduce a postsynaptic bias in modelling outcomes. We conclude that pre- and postsynaptically expressed plasticity are not interchangeable, but enable complimentary functions.

Author summary

Differences between functional properties of pre- or postsynaptically expressed long-term plasticity have not yet been explored in much detail. In this paper, we used minimalist models of STDP with different expression loci, in search of fundamental functional consequences. Biologically, presynaptic expression acts mostly on neurotransmitter release, thereby altering short-term synaptic dynamics, whereas postsynaptic expression affects mainly synaptic gain. We compared models where plasticity was expressed only presynaptically or postsynaptically, or in both ways. We found that postsynaptic plasticity had a bigger impact over response times, while both pre- and postsynaptic plasticity were similarly capable of detecting correlated inputs. A model with biologically tuned expression of plasticity also completed these tasks over a range of frequencies. Also, postsynaptic spiking frequency was not directly affected by presynaptic plasticity of short-term plasticity alone, however in combination with a postsynaptic component, it helped restrain positive feedback, contributing to activity homeostasis. In conclusion, expression locus may determine affinity for distinct coding schemes while also contributing to keep activity within bounds. Our findings highlight the importance of carefully implementing expression of plasticity in biological modelling, since the locus of expression may affect functional outcomes in simulations.

Introduction

Learning and memory in the brain, as well as refinement of neuronal circuits and the development of receptive fields, are widely attributed to long-term synaptic plasticity [1]. While this notion is not yet formally experimentally proven [2], it has in recent years received strong experimental support in several brain regions, in particular the amygdala [3] and the cerebellum [4]. The notion that synaptic plasticity underlies memory is typically attributed to Hebb [5], but it is in actuality an idea that extends considerably farther back in time, e.g. to Ramon y Cajal and William James [6].

After the discovery by Bliss and Lomo [7] of the electrophysiological counterpart of Hebb's postulate, now known as long-term potentiation (LTP), much effort has been focused on establishing the induction and expression mechanisms of long-term plasticity. In the 1990s, this led to a heated debate on the precise locus of expression of LTP, with some arguing for postsynaptic expression, whereas others were in favour of a presynaptic locus of LTP [8](Fig. 1A). Beginning in the early 2000's, this controversy was gradually resolved by the realisation that plasticity depends critically on several factors, notably animal age, induction protocol, and precise brain region [9–11]. Indeed, this resolution has now been developed to the point that it is currently widely accepted that specific interneuron types have dramatically different forms of long-term plasticity [12,13], meaning that long-term plasticity in fact depends on the particular synapse type [14]. In retrospect, it is probably not all that surprising that LTP in different circuits is expressed either pre- or postsynaptically, or both, given the diversity of computational functions of different synapses [15]. Nevertheless, the precise functional benefits of having LTP be expressed on one side of the synapse or the other have remained quite poorly explored, with only a handful of classical theoretical papers addressing this point [16–21].

Going back several decades, a multitude of highly influential computer models of neocortical learning and development have been proposed, some of them focusing on aspects such as the dependence of induction on firing rates [22–24], while others have emphasised the role of the relative millisecond timing of spikes in connected cells [25–27], and some yet have included both [28]. Irrespective of whether timing, rate, or other factors are used to determine the outcome of plasticity in theoretical

models, it has virtually always been the case that – with a few notable exceptions [18,19,21] – the expression of plasticity itself has been regarded as a simple change in the magnitude of synaptic inputs between neurons of the network. As a minimal description this is of course a perfectly reasonable approach, as it is a parsimonious assumption that induction of long-term plasticity manifests itself in the alteration of connectivity weights.

However, the expression of plasticity is not always well modelled by this sole change of instantaneous magnitude. This is because presynaptically expressed plasticity leads to changes in synaptic dynamics, whereas postsynaptic expression does not (Fig.1B). For instance, during high-frequency bursting, as the readily releasable pool of vesicles in a synaptic bouton runs out, leading to short-term depression of synaptic efficacy [29], while at other synapse types short-term facilitation dominates [30]. Such short-term plasticity is important from a functional point of view because it acts as a filter of the information that is transmitted by a synapse [31–33]. Short-term depressing connections are more likely to elicit postsynaptic spikes due to brief non-sustained epochs of activity, whereas facilitating synapses require that presynaptic activity be sustained for some period of time to elicit postsynaptic spikes. In other words, short-term facilitating connections act as high-pass filtering burst detectors [34,35], while short-term depression provides low-pass filtering inputs more suitable for correlation detection and automatic gain-control [36–38]. As a corollary, it follows that presynaptic expression of plasticity may change the computational properties of a given synaptic connection. In this case, increasing the probability of release by the induction of LTP led to more prominent short-term depression due to depletion of the readily-releasable pool depletion, and as a consequence to a bias towards correlation detection at the expense of burst detection [39,40].

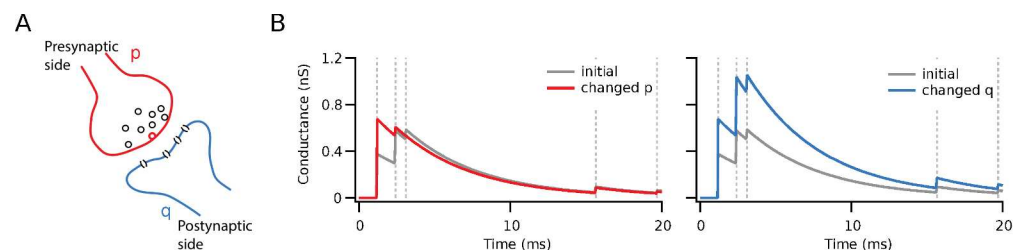


Fig 1. The postsynaptic response to the same stimulus after plasticity depends on the locus of expression. (A) Representation of pre- (red) and postsynaptic (blue) sides of a synapse, with probability of vesicle release p , and quantal amplitude q , i.e. the amplitude of postsynaptic response to a single vesicle. (B) Example of the difference between pre- and postsynaptic expression at inputs onto a cell. The identical initial response is illustrated in grey, while the potentiated responses are coloured red or blue. The amplitude of the first response after learning was set to be the same after pre- (red) and postsynaptic (blue) potentiation. With postsynaptic potentiation, the gain was increased by the same amount for all responses in the high-frequency burst. With presynaptic potentiation, however, the efficacy of the response train was redistributed toward its beginning, enhancing the first response but not the last.

Experimentally, it is long known that the induction of neocortical long-term plasticity may alter short-term depression [16,41]. Although the functional consequences of short-term plasticity itself are as outlined above quite well described [39,42], the theoretical implications of *changes* in short-term plasticity due to the induction of long-term plasticity are not well described. Yet, a majority of theoretical studies of long-term plasticity assumes that synaptic amplitude but not

synaptic dynamics are altered by synaptic learning rules. One of the motivations of our present study is the observation that this seemingly innocuous assumption may not be neutral, but may in effect introduce a bias, because changing synaptic weight in theoretical models of long-term plasticity is equivalent to assuming that synaptic plasticity is solely postsynaptically expressed. This begs the question: What are the functional implications of pre- versus postsynaptically expressed long-term plasticity? Providing answers to this central issue is important for understanding brain functioning, as well as for knowing when weight-only changes in computer modelling is warranted, and when it is not.

Here, we use computational modelling to explore and describe the consequences of expressing plasticity pre- or postsynaptically in a single neuron under two simple paradigms (Fig. 2). One paradigm explores the postsynaptic response in relation to a repeated time-locked stimulus [26, 43, 44], while the other investigates the neuron's ability to detect a correlated stimulus [45–47]. Initially, we compare and contrast relatively artificial scenarios, for which the locus of expression is either solely presynaptic, solely postsynaptic, or equally divided between both sides. We then move on to investigating the functional impact in a biologically realistic model with separate pre- and postsynaptic components that were tuned to experimental data from connections between neocortical layer-5 pyramidal cells [19]. We report that presynaptically expressed plasticity adjust the speed of learning, while postsynaptic expression is more efficient at regulating spike timing and frequency. We conclude that pre- and postsynaptically expressed plasticity enable different complimentary functions and are not equivalent.

Results

Postsynaptically expressed plasticity is readily implemented as a simple change in synaptic gain, by adjusting the quantal amplitude, q . The impact of postsynaptic expression is therefore relatively unambiguous, since it scales all postsynaptic responses the same way. For example, in the case of repeated measures of presynaptic stimulation, the standard deviation and the mean of synaptic responses scale the same, so the coefficient of variation remains the same [48], which means synaptic noise levels remain the same after postsynaptically expressed plasticity.

Presynaptic plasticity, however, has at least two different distinct types of impact on a synapse. First, in the reliability of transmission due to stochastic vesicle release. Assuming release is binomially distributed, increasing the probability of release, p , increases the mean of synaptic responses while keeping the standard deviation roughly the same, which means the coefficient of variation is decreased after presynaptic LTP [48]. Second, increasing the probability of release depletes the pool of readily releasable pool of vesicles more rapidly. Therefore, synaptic short-term dynamics are necessarily changed by presynaptically expressed long-term plasticity, resulting in functional differences.

To limit the scope of the study, we focus on early forms of plasticity for which we have detailed experimental data [19]. We thus do not consider the possibility that the number of release sites, n , may change, as it does in late, protein-synthesis dependent forms of plasticity [49].

To distinguish between the two distinct types of impact of presynaptic plasticity, we model them separately. We start with presynaptic expression modelled as direct changes in the probability of vesicle release and compare that to postsynaptic expression. Subsequent to that, we model presynaptic expression as changes in short-term plasticity and compare that postsynaptic expression. This way, we aim to systematically tease apart any different contributions of the two distinct impacts of

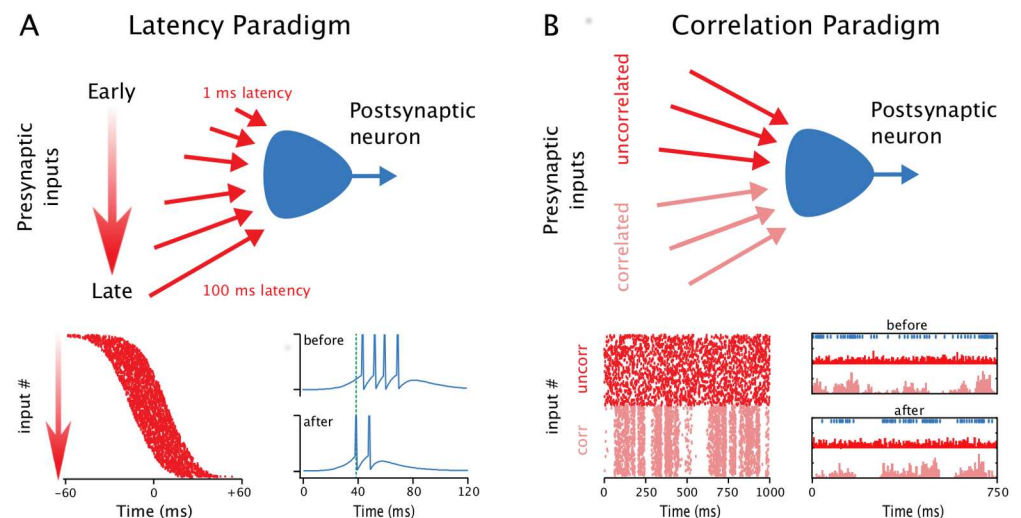


Fig 2. Two different STDP learning paradigms were explored. (A) Inputs arriving with a gradient of early to late timings resulted in reduced latency of the postsynaptic spiking response after STDP, as previously described [26]. In each trial, the postsynaptic neuron repeatedly received a brief volley of stimuli, between which short-term plasticity variables were allowed to return to their initial resting values. Bottom, left: Each presynaptic spike (raster dots) arrived with a different delay in the volley. Bottom, right: After a period of learning, the postsynaptic spiking response (blue) was shortened and started earlier. The learning task is thus to reduce the latency and to shorten the duration of the postsynaptic spiking response [26]. **(B)** Correlated inputs were selectively potentiated by STDP, as previously described [45]: The postsynaptic neuron received persistent stimulation, with half of the inputs having correlated activity, while the rest were uncorrelated. Bottom, left: Raster plot illustrating the correlated (corr) and uncorrelated (uncorr) input spiking. Bottom, right: After learning, the postsynaptic spiking (blue raster at top) was more correlated with the correlated inputs (pink histograms) than with the uncorrelated inputs (red histograms), indicating that the former drove postsynaptic activity. The learning task is thus to select for inputs that are correlated at the expense of those that are not [45]. In both paradigms, STDP is modelled with the same parameters, to permit comparison (e.g. timing window $\tau = 20$ ms, see Methods).

presynaptically expressed plasticity.

113

Presynaptic expression modelled as changes in stochastic release

114

For the first set of simulations, we considered the probability of release (p_j) and the quantal amplitude (q_j), i.e. pre- and postsynaptic quantities, respectively. In each simulation, plasticity was expressed exclusively presynaptically, exclusively postsynaptically, or equally divided between both sides at connections onto a single-compartment point neuron (see Methods). Here, changes in p_j were explored in terms of their impact on stochastic release.

115
116
117
118
119
120

In the latency paradigm, in which a volley of stimuli arrives at the postsynaptic neuron with varying delays (Fig. 2A), plasticity resulted in the shortening of the time to respond — the latency — of the postsynaptic neuron, as well as a temporal sharpening of the response, with fewer spikes and shorter inter-spike intervals [26]. The average latency reduction (Fig. 3A, B), as well as the overall distribution of synaptic weights, decrease of postsynaptic activity duration and increase of postsynaptic firing

121
122
123
124
125
126

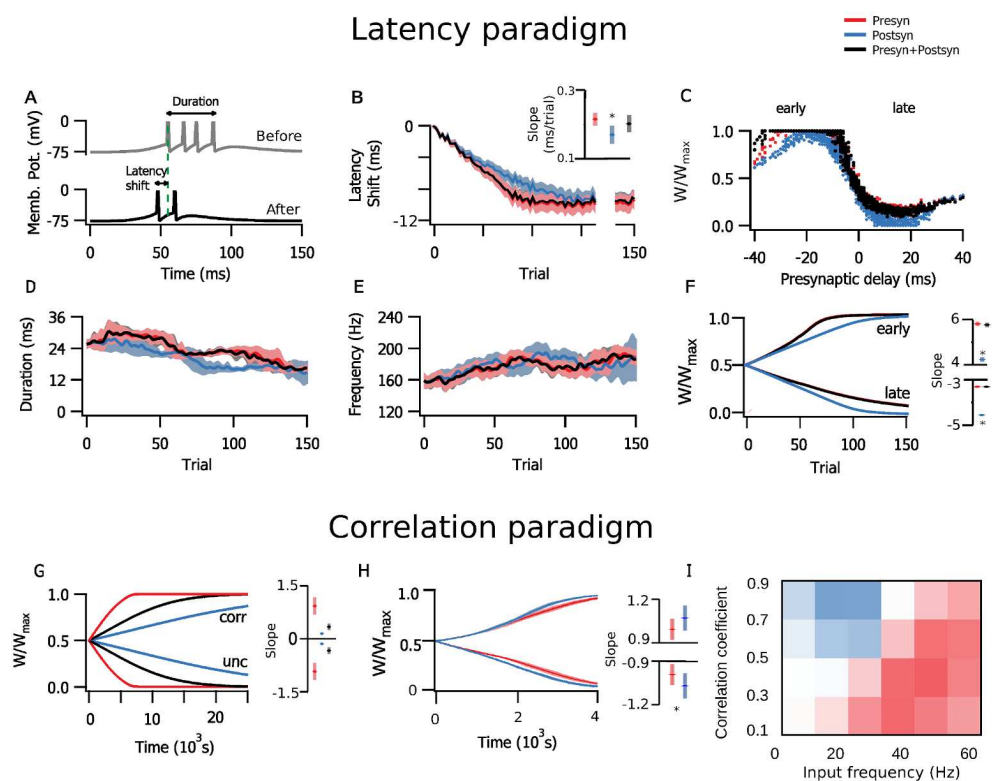


Fig 3. Presynaptically expressed plasticity typically promoted faster learning. (A-F) Simulations in the latency paradigm (see Fig. 2A). (A) Example traces of the postsynaptic membrane potential before (grey) and after (black) plasticity. Initial latency of response is marked by a green dashed line. (B) STDP-mediated learning shortened postsynaptic latency to spike, as previously shown [26]. The graphs here and below are colour-coded: only presynaptic plasticity (red), only postsynaptic plasticity (blue), or both pre- and postsynaptic plasticity (black) are implemented. Note how learning with presynaptic plasticity appears faster. Inset: presynaptic plasticity was faster than postsynaptic plasticity alone (t-test, p-value = 0.008). (C) Synaptic weight distribution after 150 trials, normalized and sorted by the fixed presynaptic delay. (D, E) Postsynaptic response duration (i.e., the interval between first and last spike in each trial) and the burst frequency did not differ across different expression loci. (F) Time evolution of average synaptic weight among early and late presynaptic inputs (i.e., input cells that spiked in the first or the second half of the stimulus) show how post-only expression (blue) is relatively slower. (G-I) Simulations in the correlation paradigm (see Fig. 2B). (G) Potentiation and depression of the average synaptic weight among correlated inputs was relatively faster in the presynaptic case. (H) However, for inputs with a very high correlation ($c > 0.9$), learning was faster with postsynaptic expression. This indicated that which form of plasticity led to faster learning depended on the details of the input firing patterns. (I) We explored this finding further in simulations where all inputs were correlated, but half expressed plasticity presynaptically, and the other postsynaptically. The map shows which side potentiated faster, indicating that postsynaptic expression (blue) won out for a relatively small parameter space where input firing frequency was low and correlations quite high.

frequency (Figs. 3C, D, E) did not differ appreciably with the locus of plasticity. In comparison to the purely postsynaptic case, simulations with presynaptic plasticity

127
128

presented a smaller variance of the latency shift (Fig. 3B). Potentiation also developed faster with presynaptic expression (Fig. 3F). This can be framed as a consequence of potentiation requiring glutamate release [50], so that in a more reliable synapse, with a high p value, there is a greater propensity for potentiation. Conversely, depression was slower with presynaptically expressed plasticity, again because lowered probability of release effectively also led to less plasticity (Fig. 3F).

Next, we explored the correlation paradigm, in which plasticity selectively potentiates correlated inputs (Fig. 2A,B) ([45]). Here, all plasticity implementations detected the input correlations. However, presynaptically expressed plasticity generally promoted faster learning, e.g. synaptic weights evolved more rapidly (Fig. 3G), similar to what we found above for the latency paradigm. However, there were exceptions to this general observation — postsynaptic plasticity was faster for strongly correlated input firing at certain input frequencies (Fig. 3H).

We wanted to explore this exception in more detail. One implication of the presynaptic side potentiating faster than the postsynaptic side is that, in a scenario in which there is competition due to e.g. limited resources, inputs with presynaptic plasticity would be expected to overcome inputs with postsynaptic plasticity. To explore this possibility, we ran simulations where all of inputs were correlated, but half of them expressed plasticity only presynaptically, and the other half only postsynaptically. Because of normalization, these two input populations competed, so that one potentiated at the expense of the other, which depressed. With this approach, we systematically explored the correlation-frequency space. We found that postsynaptic expression won for very highly correlated inputs for sufficiently low input frequencies (Fig. 3I), a scenario that corresponds to fewer inputs spiking synchronously.

Presynaptic expression modelled as changes in short-term plasticity

We next explored the effects of altering short-term dynamics (see Methods). This adds another aspect of presynaptically expressed plasticity, since short-term plasticity takes into account the history of presynaptic activity. In this scenario, presynaptic changes redistribute synaptic resources used over a limited time period, instead of bringing about an average increase or decrease [16, 41]. Even if the amplitude of an individual EPSP were affected equally by pre- and by postsynaptically expressed plasticity, the total input from a burst would still differ dramatically depending on the site of expression (Fig. 1B).

In the simulations with the timed input configuration, results differed considerably depending on the specific locus of plasticity in the latency configuration. Postsynaptic expression alone provided the largest latency reduction, and also achieved it faster than the other plasticity implementations (Fig. 4A, B). The results of presynaptic expression were also more subtle compared to the mixed setting with both pre- and postsynaptic expression, for which results may vary between extremes according to the ratio of pre- and postsynaptic expression. Effects of postsynaptic plasticity over response duration and intraburst frequency (Figs. 4C and 4D) were also more marked, as expected from a higher integrated input (Fig. 1B). The simulations with both sides changing appeared closer to either the presynaptic case (duration, Fig. 4C) or the postsynaptic case (frequency, Fig. 4D). Here, changes in p had a relatively greater influence on response duration, while changes in q had greater impact on the response frequency. Nevertheless, synaptic efficacy was still potentiated faster and depressed slower in the presynaptic case (Fig. 4E). This was similar to the above stochastic release implementation of presynaptically expressed plasticity, although it was less

pronounced. This means that even if the rate of learning was effectively faster, presynaptic expression affected latency less rapidly than postsynaptic expression did (Fig. 4F).

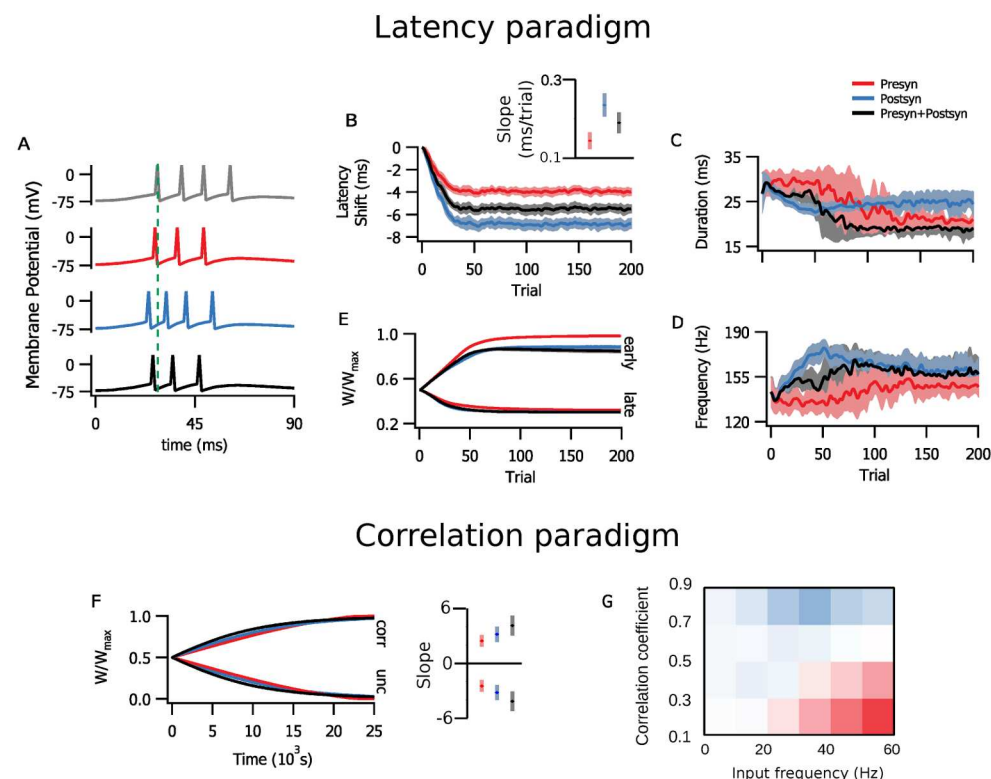


Fig 4. Altering short-term plasticity was less efficient at reducing postsynaptic latency. (A-E) Simulations in the latency paradigm (see Fig. 2A) are colour-coded: red denotes presynaptic plasticity alone, blue postsynaptic plasticity alone, and black combined pre- and postsynaptic plasticity. **(A)** Example traces of postsynaptic activity before (grey) and after plasticity (coloured). Initial response latency is illustrated by the vertical dashed line. **(B)** Latency reduction was both faster and more marked for postsynaptic (blue) than for presynaptic (red) or combined (black) plasticity. Inset: The slope of latency reduction was steeper when postsynaptic expression was involved (t-tests: between p and q, p-value < 10^{-6} ; between p and both, p-value = 0.0008, between q and both p-value = 0.003) **(C)** Combined and presynaptic plasticity reduced response duration more than with postsynaptic expression alone. **(D)** Burst frequency was similarly increased with all three forms of plasticity, although rate change was faster with postsynaptic plasticity. **(E)** Time course of average synaptic weights for early (left) and late (right) inputs. **(F, G)** Simulations in the correlation paradigm (see Fig. 2B) **(F)** Time course of average synaptic weights for correlated (left, "corr") and uncorrelated (right, "unc") inputs were largely indistinguishable. **(G)** As with Fig. 3I, the map of competition between input populations with pre- or postsynaptically expressed plasticity indicated a less marked differentiation.

On the other hand, under STP modulation, plasticity rates in the correlated inputs paradigm evolved differently compared to the above stochastic release implementation (Fig. 4G), even though the overall effect on the covariance between pre- and postsynaptic activity was similar (not shown). In the simulations where long-term

plasticity affected short-term plasticity, the rate of change was slightly faster with postsynaptic than with presynaptic plasticity. It thus appears that computational advantages could be tailored to a functional task at hand by recruiting pre- or postsynaptic plasticity differentially.

Comparisons with a biologically tuned model

The above minimalist toy models had the advantage that they provided full control of several key parameters. However, the relevance of the findings for the intact brain were unclear. To address this shortcoming, we explored the biological plausibility in a model [19] (see Methods) that was fitted to long-term synaptic plasticity data obtained from connections between rodent visual cortex layer-5 pyramidal neurons [41, 51, 52]. We could thus to some extent verify whether the results obtained with the minimal models hold in a more complex, data-driven context. We want to clarify upfront that in this model, LTP is expressed both pre- and postsynaptically, whereas LTD is solely postsynaptically expressed. This asymmetry may seem odd, but it is derived from experimental data, and we have previously found that this arrangement provides certain computational advantages [19].

We first explored the latency paradigm (Fig. 2A). To avoid disrupting the parameter tuning, instead of normalising the total synaptic change on each side, we kept the data-derived ratios and blocked either pre- or postsynaptic changes. Even so, we found that both pre- and postsynaptic plasticity components independently led to the shortening of postsynaptic latency (Fig. 5A-C). As with the above simplistic modelling scenarios, postsynaptic changes appeared to affect spike timing more. Thus, when both pre- and postsynaptic plasticity were active, the presence of postsynaptic potentiation further reduced the latency compared to presynaptic plasticity alone (Fig. 5B).

Following the experimental results [41, 52], postsynaptic plasticity in the tuned model lacked the capacity to depress. As a consequence, postsynaptic plasticity led to inflated postsynaptic frequency and duration when implemented alone (Fig. 5D, E). However, the inclusion of presynaptic LTD was enough to produce a temporally sharpened response of shorter duration. With postsynaptic plasticity, the dynamics developed faster (Fig. 5F), as result of a positive-feedback loop due to increased postsynaptic firing frequency (compare Fig. 5A).

In the correlation paradigm (Fig. 2A), groups of correlated and uncorrelated inputs clustered (Fig. 6A) without the need for added competition through weight normalization [46, 53]. This only occurred when both pre- and postsynaptic plasticity components were implemented, and was not achieved through other models with physiologically compatible parameters [47].

To better understand the robustness of this property, we quantified the capacity of separation between correlated and uncorrelated populations with a linear separator. It was trained to classify inputs as correlated or uncorrelated according to the average and variance of p values (Fig. 6C). The presynaptic frequency range for optimal separation was between 50 and 80 Hz (Fig. 6C). At the other end of the range, it was bounded by the STDP correlation time scale of $\tau = 20$ ms (see Methods), meaning interspike intervals longer than 20 ms could not represent the minimal interval of correlation. At the upper end of the range, the high presynaptic frequency yielded overall potentiation that included uncorrelated inputs, limiting the separation from the more potentiated correlated population (see appendix).

In the same way as in the latency paradigm (Fig. 5D), postsynaptic potentiation increased postsynaptic firing rate (Fig. 6B). However, presynaptic plasticity alone produced no such effect. In combination with postsynaptic plasticity, presynaptic

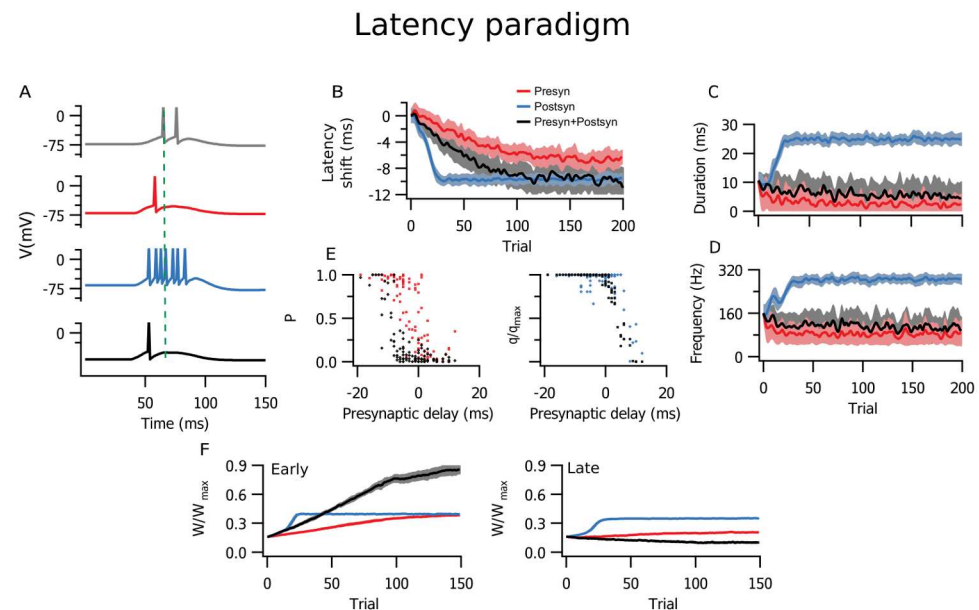


Fig 5. A biologically tuned model verified key findings obtained with minimalist models. (A) Sample traces of postsynaptic activity before (grey) and after only presynaptic (red), only postsynaptic (blue), or both pre- and postsynaptic learning (black). The initial response latency is indicated by the green dashed line. (B) The postsynaptic response latency was shortened by learning, although both faster and more efficiently with postsynaptic learning. (C) Distribution of pre- (p) and postsynaptic efficacies (q) after 200 learning trials. (D, E) Changes in duration and burst frequency of postsynaptic activity mirrored those obtained with the stochastic minimalist models (Fig. 4C, D). (F) Average synaptic weight of early (left) and late (right) presynaptic inputs evolved in distinct manners, however (compare e.g. Fig. 4).

plasticity provided a degree of output control, as its introduction helped to maintain a lower postsynaptic firing frequency even as q saturated (Fig. 6B).

Discussion

In recent years, it has become clear that diversity in LTP expression is both ubiquitous and considerable, depending on factors such as animal age, induction protocol, and precise brain region [9–11, 15]. In this work, we explored possible functional properties of either pre- or postsynaptic locus of plasticity expression, and found that even in a single neuron scenario overall dynamics may be affected by it. This is an important feature to be considered, as many theoretical studies have focused on induction but not many in the expression of plasticity. Plasticity has in the typical phenomenological model been implemented by default as a straightforward change in synaptic weight [26, 54, 55], although there are a few notable exceptions [16–18, 56, 57]. In other words, in the absence of better information, a standard assumption has been that that locus of expression does not matter appreciably for the modelling scenario at hand. Our findings challenge this standard assumption, highlighting how it may introduce a bias. For example, over-representation of postsynaptic expression may exaggerate the capacity to learn spike timing (e.g., Figs. 4B, 5B).

We investigated two different learning paradigms, one with differently timed inputs, in which postsynaptic latency to spike was used as a measure of learning (Fig. 2A),

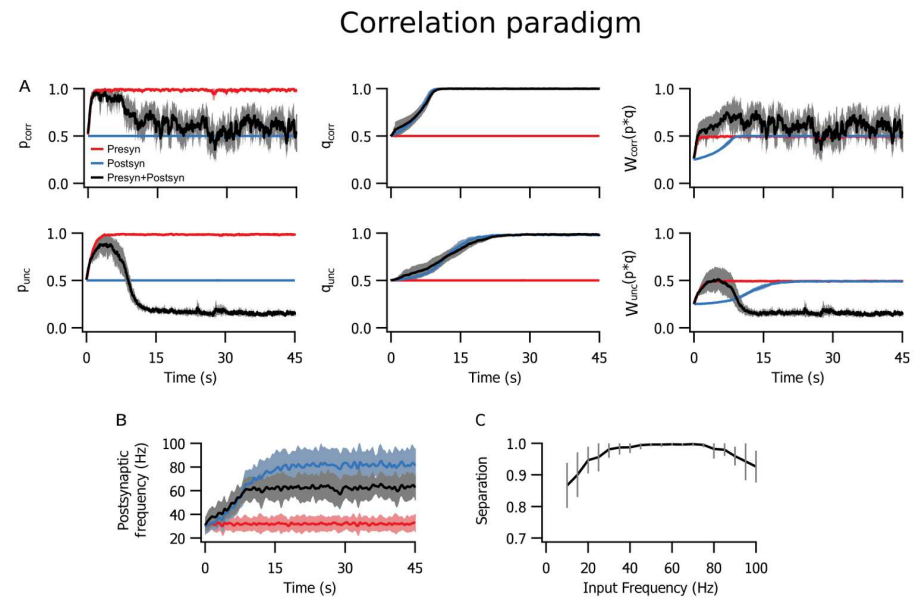


Fig 6. The biologically tuned model clustered inputs with correlated and uncorrelated activity. (A) Normalized averages for presynaptic (p), postsynaptic (q) and combined pre- and postsynaptic (W) plasticity of correlated (corr) and uncorrelated (unc) inputs show that meaningful learning and segregation of inputs occurred when both pre- and postsynaptic learning mechanisms were engaged. (B) The postsynaptic spiking frequency increased when postsynaptic plasticity was engaged (blue and black), but not with presynaptic-only learning (red). (C) The separation between correlated and uncorrelated inputs was optimal for presynaptic frequencies in the range 50 and 80 Hz.

and another under constant stimulation, where a subset of inputs were correlated and potentiated together (Fig. 2B). We first worked with simplified conceptual STDP models and later with a more realistic, biologically tuned model in which pre- and postsynaptic components were tuned to connections between neocortical layer 5 pyramidal cells [19].

Pre- and postsynaptic expression favour different coding schemes

Our study showed that the locus of expression of plasticity determined affinity for different coding schemes. Presynaptic plasticity expressed as the regulation of release probability alone did not result in any differences over average postsynaptic activity measurements compared to postsynaptic expression. However, in the presence of short-term plasticity, presynaptic expression of long-term plasticity had a smaller impact on the spike latency in comparison to postsynaptic expression (e.g., Fig. 4B). This was because, as synaptic response amplitude grew, fewer inputs were needed to evoke a postsynaptic spike. With presynaptic expression, however, the spike still depended on the sum of a larger number of inputs. However, weight changes developed faster with presynaptic plasticity, thereby increasing the speed of learning. This effect, however, was not present in the correlation paradigm, where both pre- and postsynaptically expressed cases performed similarly.

Presynaptically expressed plasticity alone was not ideally suited for rate coding, because it did not impact the average summed input effectively. As a consequence, postsynaptic firing frequency remained relatively unchanged after presynaptically expressed plasticity (e.g., Figs. 4D, 5D, 6B). Presynaptic plasticity thus appeared to

act as a limiter or a form of homeostasis for postsynaptic activity, in agreement with previously published interpretations [40]. The flip-side of this stabilizing feature of changes in short-term plasticity [58] is in other words the loss of ability to rate code well. An important cautionary take-home message from this observation is that the default implementation of plasticity as purely postsynaptic may thus lead to an erroneous overestimation of the impact on postsynaptic firing rates.

It is also interesting to think about the role of presynaptic plasticity if it is not very useful in the context of usual 'coding' frameworks. Frequently the effect of unreliability of single synapses is considered to be simply of noise or energy economy [59]. However, one can in fact consider this unreliability as a representation of uncertainty over a synaptic weight compared to its optimal value [60,61]. It would then be plausible to consider pre-synaptic plasticity as an uncertainty tuning over the posterior distribution in a probabilistic inference framework [62].

A biologically tuned model corroborated the toy model predictions

The same basic properties were observed in the biologically tuned model with simultaneous pre- and postsynaptic plasticity. Learning was dramatically affected by postsynaptic plasticity, while the presynaptic side appeared to act more on the rate of learning and on weight dynamics. It is possible that these results could be modified according to the ratio of pre- versus postsynaptic forms of plasticity, to optimize for the computational task at hand. It is noteworthy that the biologically tuned model was also capable of separating groups of correlated and uncorrelated inputs without the need for a hard competitive mechanism.

Experimental tests of model predictions

Since it is possible to specifically block pre- or postsynaptic STDP pharmacologically [41,52], several of our findings related to the locus of expression of plasticity are possible to directly test experimentally. For example, at connections between neocortical layer-5 pyramidal cells, it is possible to block nitric oxide signalling to abolish pre- but not postsynaptic expression of LTP [52]. It is also possible to use GluN2B-specific blockers such as ifenprodil or Ro25-6581 to block presynaptic NMDA receptors necessary for presynaptically expressed LTD without affecting postsynaptic NMDA receptors that are needed for LTP [41,63]. As a proxy for learning rate, one could explore *in vitro* how blockade of different forms of plasticity expression impacts the number of pairings required for plasticity, or alternatively how the magnitude of plasticity is affected for a given number of pairings [52,55]. *In vivo*, the impact on cortical receptive fields could similarly be explored. For example, we predict that receptive field discriminability is poorer when presynaptic LTP is abolished by nitric oxide signalling blockade [19].

Conclusions

Here, we have challenged the standard assumption that modelling synaptic plasticity as a weight change is neutral and unbiased. We found that even in a simple feed-forward scenario, the locus of expression may have considerable impact on learning outcome. We expect that these effect will only be greater in recurrent networks, where presynaptic plasticity at loops and re-entrant pathways will exacerbate the effects of changes in synaptic dynamics due to alterations of the accumulated difference. This additional level of complexity may in particular complicate very large recurrent network models [64,65].

As our collective understanding of the expression of long-term plasticity has improved, it has become clear that the long-held notion that plasticity is expressed predominantly postsynaptically is erroneous [9–11]. Since presynaptic expression is still relatively poorly studied, our understanding of long-term presynaptic plasticity in health and disease needs to be generally improved [66]. Specifically, our study highlights the need for more detailed modelling of the role of the site of expression. It is clear that it has implications on the relevant form of information coding, be it spike-, rate-based, or probabilistic. In modelling long-term plasticity, correctly implementing changes in weight is thus a matter of gravity.

Methods

Neuron model

All of the simulations consisted of one postsynaptic neuron receiving a number of presynaptic Poisson inputs. In the first section, we used a simple leaky integrate-and-fire model defined by

$$\tau_V \frac{dV}{dt} = E_v - V(t) - g_e(t)(E_e - V(t)), \quad (1)$$

in which the membrane potential V decayed exponentially with a time constant of $\tau_V = 20\text{ms}$ to the resting value of $E_v = -74\text{ mV}$, and the threshold for an action potential was $V_{th} = -54\text{ mV}$. After each spike it was reset at $V_0 = -60\text{ mV}$ with a refractory period of 1 ms.

Inputs were accounted as conductance-based excitatory contributions with reversal potential $E_e = 0\text{ mV}$, amplitude q_j , summed after the l^{th} spike of presynaptic neuron j , that decayed exponentially with a time constant of $\tau_g = 5\text{ms}$:

$$\frac{dg_e}{dt} = -\frac{g_e}{\tau_g} + \sum_{j,l} q_j \delta(t - t_j^l) . \quad (2)$$

In the the last section, we used the adaptive exponential integrate-and-fire model [67] to reduce unrealistic bursting and to comply with the biological tuning [19]:

$$\frac{dV}{dt} = \frac{1}{C} [g_L(E_L - V) + g_L \Delta_T e^{\left(\frac{V - V_T}{\Delta_T}\right)} - g_e V - z] , \quad (3)$$

$$\tau_W \frac{dz}{dt} = c_z(V - E_L) - z . \quad (4)$$

The corresponding parameters for a pyramidal neuron were $C = 281\text{ pF}$, $g_L = 30\text{ nS}$, $E_L = -70.6\text{ mV}$, $\Delta_T = 2\text{mV}$, $c = 4\text{nS}$, $\tau_W = 144\text{ms}$. Spiking threshold was $V_T = -50.4\text{ mV}$, and after each spike V was reset to the resting potential E_L while z increased by the quantity $b = 0.0805\text{ nA}$ (as in [67]).

Stimulation paradigms

The postsynaptic neuron received either one of two stimulus configurations. The first one was based on [26] and is referred to as the Latency Peduction (Fig. 2A). In every 375-ms-long trial, the postsynaptic cell received a volley of Poisson inputs that arrived with a specific delay, normally distributed around a time reference, for each specific presynaptic neuron. Each input lasted for 25 ms with a spiking frequency of 100 Hz. We measured the time to spike of the first postsynaptic spike in response to a bout of stimuli using the mean of the presynaptic delay distribution as a reference point. For

clarity, in the Results, curves that represent latency shift, intraburst frequency or burst duration were smoothed using a moving average filter with a window of three points.

The second type of stimulation paradigm was based on [45] and is referred to as the Correlation Paradigm (Fig. 2B). This configuration consisted of continuous Poisson inputs with fixed frequency. However, half of the inputs had correlated fluctuations of activity, with a time window of $\tau_{corr} = 20$ ms, while the other half was uncorrelated. Correlations were implemented as in [68].

Additive STDP model

For the majority of the simulations we opted to implement STDP with the simple additive model proposed by Song and Abbott [26]:

$$\tau_{STDP} \frac{dW_{ij}}{dt} = \sum_k \sum_l F(t_i^k - t_j^l)$$

$$F(t_i - t_j) = \begin{cases} c_{pot}, & t_i > t_j \\ c_{dep}, & t_i < t_j \end{cases} \quad (5)$$

Each increment to the synaptic weights W_{ij} (since there was only one postsynaptic cell, we consider $W_j = W_{ij}$ throughout this paper) was computed after a pair of pre- and postsynaptic spikes, and the parameters were set to $\tau_{STDP} = 20$ ms, $c_{pot} = 0.005$, and $c_{dep} = -0.00525$. We separated the synaptic weight W_j as a product between pre- and postsynaptic counterparts, probability of release $P_j = (0, 1]$ and quantal amplitude $q_j = (0, q_{max}]$ respectively, so that $W_j = q_j P_j$. The probability of release was simulated in two different ways, one equivalent to regulating the probability of stochastic interactions and the other via short-term plasticity.

When the weight convergence rates were compared, we had to ensure that $\Delta W = W^f - W^i$ per time step was the same for all simulations. Therefore, we normalised the changes so that if only q was changed:

$$\Delta W^q = P^i(q^f - q^i) = P^i \Delta q \quad , \quad (6)$$

and if only P was changed,

$$\Delta W^P = q^i \Delta P \quad . \quad (7)$$

The initial value of all simulations was the same for P and q , so in these cases $\Delta P = \Delta q \equiv \Delta$. This amount was equally divided between P and q when both were changed simultaneously:

$$\begin{aligned} \Delta W^{Pq} &= P^f q^f - P^i q^i = (P^i + \Delta P^{Pq})(q^i + \Delta q^{Pq}) - P^i q^i \\ &= (P^i + \Delta P^{Pq})(q^i + \Delta P^{Pq}) - P^i q^i \quad , \end{aligned} \quad (8)$$

so that

$$\Delta^{Pq} = -\frac{1}{2} \left[(P^i + q^i) - \sqrt{(P^i + q^i)^2 + 4P^i \Delta} \right] \quad . \quad (9)$$

The largest possible change for P or q separately was $\Delta_{tot} = 1 - q^i$. To keep the same range of W for changing P and q simultaneously, we limited the maximal values P and q in this case at $q^{max} = P^{max} = \sqrt{q^i}$.

Biologically tuned STDP model

We compared the results of the straightforward additive model to a slightly more complex STDP model that acts separately over pre- and postsynaptic factors [19]. Parameters were fitted to experimental data from connections between pyramidal cells from layer 5 of V1 [41, 51, 52]. The equations for pre- and postsynaptic changes followed:

$$\Delta q_j = c_+ x_{j+}(t) y_-(t - \epsilon) Y(t) , \quad (10)$$

$$\Delta P_j = -d_- y_-(t) y_+(t) X_j(t) + d_+ x_{j+}(t - \epsilon) y_+(t) X_j(t) . \quad (11)$$

where $X_j(t) = \sum_l \delta(t - t_j^l)$ is increased at each spike from the presynaptic neuron j and $Y(t) = \sum_k \delta(t - t_i^k)$ at each spike from the postsynaptic neuron i . ϵ is to emphasise that ΔW was calculated before x_{j+} and y_- were updated, upon the arrival of a new spike. y_+ and y_- are postsynaptic traces,

$$\frac{dy_+}{dt} = -\frac{y_+}{\tau_{y_+}} + Y , \quad (12)$$

$$\frac{dy_-}{dt} = -\frac{y_-}{\tau_{y_-}} + Y , \quad (13)$$

with decay times τ_{y_+} and τ_{y_-} respectively, and x_{j+} was a presynaptic trace with decay time τ_{x_+} :

$$\frac{dx_{j+}}{dt} = -\frac{x_{j+}}{\tau_{x_+}} + X_j . \quad (14)$$

The parameter values were taken from [19]: $d_- = 0.1771$, $\tau_{y_-} = 32.7\text{ms}$, $d_+ = 0.15480$, $c_+ = 0.0618$, $\tau_{y_+} = 230.2\text{ms}$ and $\tau_{x_+} = 66.6\text{ms}$. To avoid manipulation of the fitting, weight changes were not normalised in this case.

In the last section, we used a linear least squares separator to classify presynaptic inputs according to synaptic weight average and variance.

Presynaptic factor

Presynaptic control of the probability of release per stimulus was implemented either as a Markovian process or as short-term plasticity. In the former case, probability (P_j) of stochastic neurotransmitter vesicle release followed a binomial distribution. Based on the findings reported by [69], each presynaptic neuron had $N = 5$ release sites that functioned independently. In the second we considered a dynamic modulation of the EPSPs through STP. The probability P_j was decomposed into the product of instantaneous probability of release $p_j(t)$ and availability of local resources $r_j(t)$, resulting in the following synaptic efficacy:

$$W_j(t) = q_j p_j(t) r_j(t) . \quad (15)$$

In the latter case, the dynamics of $p_j(t)$ and $r_j(t)$ followed the model proposed by Tsodyks and Markram [70]:

$$\frac{dr_j(t)}{dt} = \frac{1 - r_j(t)}{\tau_D} - p_j(t) r_j(t) X_j(t) , \quad (16)$$

$$\frac{dp_j(t)}{dt} = \frac{P_j - p_j(t)}{\tau_F} + P_j [1 - p_j(t)] X_j(t) . \quad (17)$$

Depression and facilitation time constants, $\tau_D = 200$ ms and $\tau_F = 50$ ms respectively, were chosen as representative values for connections between pyramidal neurons [71]. The resulting short-term plasticity could be either depressing, if $P_j > P_C$, or facilitating, if $P_j < P_C$. For the values of τ_D and τ_F used, $P_C \approx 0.3$.

Supporting information

S1 Appendix — Rate Model

Using a simple firing rate model with linear response, we were able to illustrate how synaptic plasticity could separate correlated and uncorrelated inputs without competition between the two populations. Considering a neuron receiving independent Poisson inputs with fixed firing rate (pooled into a single average input $I(t)$), we found that the system tends to a specific non-zero average value for P , denoted P^* below. We converted the biophysically tuned model (eqs. 10 and 11) to a firing rate representation with time-averaged values:

$$\langle dq \rangle = c^+ \tau_{x+} \tau_{y-} I \nu^2, \quad (18)$$

$$\langle dP \rangle = \nu I \tau_{y+} (d_+ \tau_{x+} I - d_- \tau_{y-} \nu). \quad (19)$$

Postsynaptic output ν was then considered as a simple firing rate model with linear relation to average input I , weighted by average synaptic efficacy (eq. 15):

$$\nu = \alpha + \beta q r p I. \quad (20)$$

To determine α and β values that corresponded to the simulated neurons (for fixed values of q and P), we fitted to data from simulations without plasticity. Since I was fixed, we could also consider stationary values for $r(t)$ and $p(t)$, \bar{r} and \bar{p} , from eqs. 16 and 17:

$$\bar{r} = \frac{1 + P I \tau_F}{1 + P I \tau_F + P I \tau_D (1 + I \tau_F)}, \quad (21)$$

$$\bar{p} = \frac{P(1 + I \tau_F)}{1 + P I \tau_F}. \quad (22)$$

We thus have $\langle dq \rangle (P, q, I)$ and $\langle dP \rangle (P, q, I)$ in the LTP equations 18 and 19:

$$\nu \approx \alpha + \frac{\beta q I P (1 + I \tau_F)}{1 + P I \tau_F + P I \tau_D (1 + I \tau_F)}. \quad (23)$$

We plotted $dP \times dq$ as a vector field (Fig. 7A), which shows how P tended to the specific value P^* , which corresponds to the average value of P for uncorrelated inputs. Integrate-and-fire simulation averages also converged to this point (black line, Fig. 7A). This is in contrast to correlated inputs, which potentiated more (Fig. 6). The value P^* was relatively stable with frequency, but saturated at a limited frequency value (Fig. 7B), effectively limiting the range of possible separation between the correlated and the uncorrelated populations.

Data availability statement

All files will be available on GitHub.

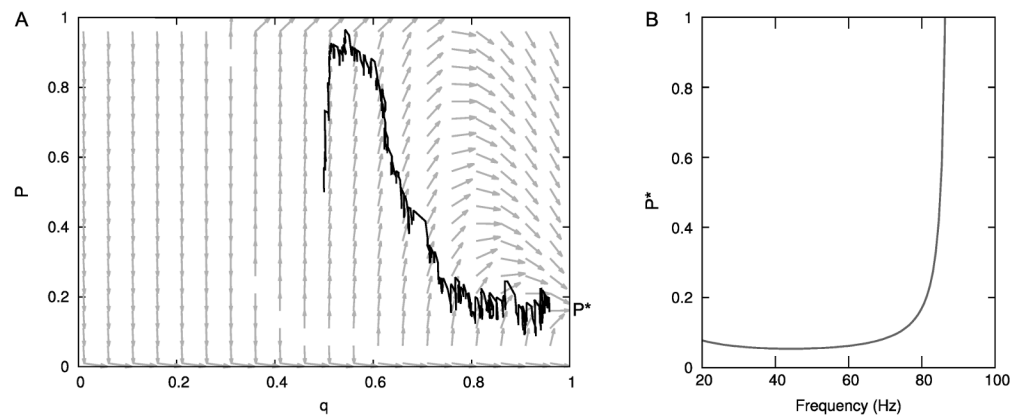


Fig 7. Plasticity separated correlated and uncorrelated inputs up to a limiting frequency. (A) Vector field ($p \times q$) representing the rate model for uncorrelated synaptic inputs only. The black line shows corresponding integrate-and-fire simulation averages for uncorrelated inputs (compare Fig. 6A). Note how both models converge to the same fixed point, indicated with the label P^* . **(B)** The point of convergence P^* was relatively stable with respect to presynaptic frequency up to a limiting frequency of around 85 Hz, where it saturated. Since correlated inputs tended to saturate, this shows an effective upper frequency limit to the clustering of correlated and uncorrelated inputs.

Acknowledgements

We thank Alanna Watt and Mark van Rossum for suggestions, help, and useful discussions.

Author contributions

Conceptualization: Beatriz E.P. Mizusaki, Rui P. Costa, P. Jesper Sjöström.
Formal analysis: Beatriz E.P. Mizusaki, Sally S.Y. Li.
Funding acquisition: Beatriz E.P. Mizusaki, Sally S.Y. Li, Rui P. Costa, P. Jesper Sjöström.
Investigation: Beatriz E.P. Mizusaki, Sally S.Y. Li.
Methodology: Beatriz E.P. Mizusaki, Rui P. Costa.
Project administration: P. Jesper Sjöström.
Resources: Rui P. Costa, P. Jesper Sjöström.
Software: Beatriz E.P. Mizusaki, Sally S.Y. Li, Rui P. Costa.
Supervision: P. Jesper Sjöström.
Validation: Beatriz E.P. Mizusaki.
Visualization: Beatriz E.P. Mizusaki.
Writing – original draft: Beatriz E.P. Mizusaki, Rui P. Costa, P. Jesper Sjöström.
Writing – review & editing: Beatriz E.P. Mizusaki, Rui P. Costa, P. Jesper Sjöström.

Funding

This work was funded by CNPq 202183/2015-7 (BEPM), Canada Summer Jobs (SSYL), EPSRC EP/F500385/1 (RPC), BBSRC BB/F529254/1 (RPC), Fundação para a Ciência e a Tecnologia SFRH/BD/60301/2009 (RPC), CFI LOF 28331 (PJS),

CIHR OG 126137 (PJS), CIHR PG 156223 (PJS), CIHR NIA 288936 (PJS), FRQS
CB Sr 254033 (PJS), NSERC DG 418546-2 (PJS), NSERC DG 2017-04730 (PJS), and
NSERC DAS 2017-507818 (PJS). The funders played no role in the study design, data
collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors have declared that no competing interests exist.

References

1. Maheux J, Froemke RC, Sjöström PJ. Functional plasticity at dendritic synapses. In: Stuart G, Spruston N, Häusser M, editors. *Dendrites*. Oxford University Press; 2015. p. 465–498.
2. Stevens CF. A Million Dollar Question: Does LTP = Memory? *Neuron*. 1998;20(1):1–2.
3. Nabavi S, Fox R, Proulx CD, Lin JY, Tsien RY, Malinow R. Engineering a memory with LTD and LTP. *Nature*. 2014;511:348–352.
4. Thompson RF. In Search of Memory Traces. *Annual Review of Psychology*. 2005;56:1–23.
5. Hebb DO. *The organization of Behavior*. Wiley, New York; 1949.
6. Markram H, Gerstner W, Sjöström PJ. A history of spike-timing-dependent plasticity. *Frontiers in Synaptic Neuroscience*. 2011;3(4):1–24.
7. Bliss TVP, Lomo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology*. 1973;232(2):331–356.
8. Bear MF, Malenka RC. Synaptic plasticity: LTP and LTD. *Current Opinion in Neurobiology*. 1994;4(3):389–399.
9. Malenka RC, Bear MF. LTP and LTD: an embarrassment of riches. *Neuron*. 2004;44:5–21.
10. MacDougall MJ, Fine A. The expression of long-term potentiation: reconciling the preists and the postivists. *Philosophical Transactions of the Royal Society B*. 2013;369:20130135.
11. Padamsey Z, Emptage N. Two sides to long-term potentiation: a view towards reconciliation. *Philosophical Transactions of the Royal Society B*. 2013;369:20130154.
12. Kullmann DM, Lamsa KP. Long-term synaptic plasticity in hippocampal interneurons. *Nature Reviews Neuroscience*. 2007;8:687–699.
13. Lamsa KP, Kullmann DM, Woodin MA. Spike-timing dependent plasticity in inhibitory circuits. *Frontiers in Synaptic Neuroscience*. 2010;2(8):1–8.
14. Larsen RS, Sjöström PJ. Synapse-type-specific plasticity in local circuits. *Current Opinion in Neurobiology*. 2015;35:127–135.

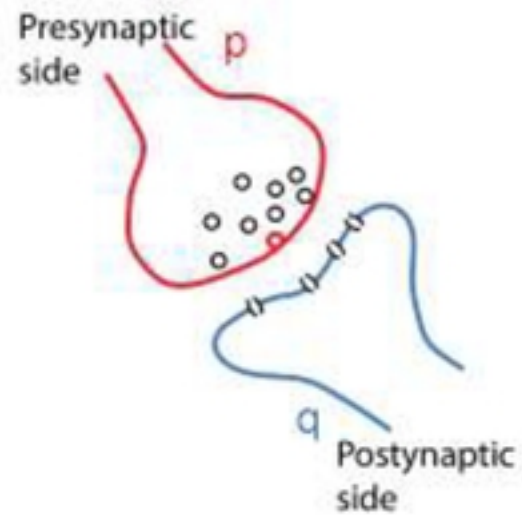
15. Kullmann DM. The Mother of All Battles 20 years on: is LTP expressed pre- or postsynaptically? *The Journal of Physiology*. 2012;590(10):2213–2216. 505 506
16. Markram H, Tsodyks M. Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*. 1996;382:807–810. 507 508
17. Lisman JE. Bursts as a unit of neural information: making unreliable synapses reliable. *Trends in Neurosciences*. 1997;20(1):38–43. 509 510
18. Senn W, Markram H, Tsodyks M. An Algorithm for Modifying Neurotransmitter Release Probability Based on Pre- and Postsynaptic Spike Timing. *Neural Computation*. 2001;13(1):35–67. 511 512 513
19. Costa RP, Froemke RC, Sjöström PJ, van Rossum MCW. Unified pre- and postsynaptic long-term plasticity enables reliable and flexible learning. *eLife*. 2015;4:e09457. 514 515 516
20. Cui Y, Prokin I, Xu H, Delord B, Genet S, Venance L, et al. Endocannabinoid dynamics gate spike-timing dependent depression and potentiation. *eLife*. 2016;5:e13185. 517 518 519
21. Costa RP, Padamsey Z, D'Amour JA, Emptage NJ, Froemke RC, Vogels TP. Synaptic transmission optimization predicts expression loci of long-term plasticity. *Neuron*. 2017;96(1):177–189. 520 521 522
22. Bienenstock EL, Cooper LN, Munro PW. Theory for the development of neuron selectivity, orientation specificity and binocular interaction in visual cortex. *The Journal of Neuroscience*. 1982;2(1):32–48. 523 524 525
23. Linsker R. Self-organization in a perceptual network. *Computer*. 1988;21:105–117. 526 527
24. von der Malsburg C. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*. 1973;14(2):85–100. 528 529
25. Gerstner W, Kempter R, Van Hemmen JL, Wagner H. A neuronal learning rule for sub-millisecond temporal coding. *Nature*. 1996;383:76–78. 530 531
26. Song S, Miller KD, Abbott LF. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience*. 2000;3(9):919–926. 532 533 534
27. van Rossum MCW, Bi Gq, Turrigiano GG. Stable Hebbian learning from spike timing-dependent plasticity. *The Journal of Neuroscience*. 2000;20(23):8812–8821. 535 536 537
28. Clopath C, Busing L, Vasilaki E, Gerstner W. Connectivity reflects coding: a model of voltage-based STDP with homeostasis. *Nat Neurosci*. 2010;13(3):344–52. doi:nn.2479 [pii] 10.1038/nn.2479. 538 539 540
29. Zucker RS, Regehr WG. Short-Term Synaptic Plasticity. *Annual Review of Physiology*. 2002;64:355–405. 541 542
30. Blackman AV, Abrahamsson T, Costa RP, Lalanne T, Sjöström PJ. Target-cell-specific short-term plasticity in local circuits. *Frontiers in Synaptic Neuroscience*. 2013;55:11. 543 544 545

31. Dittman JS, Kreitzer AC, Regehr WG. Interplay between Facilitation, Depression, and Residual Calcium at Three Presynaptic Terminals. *The Journal of Neuroscience*. 2000;20(4):1374–1385. 546
32. Fortune ES, Rose GJ. Short-term synaptic plasticity as a temporal filter. *Trends in Neurosciences*. 2001;24(7):381–385. 549
33. Fuhrmann G, Segev I, Markram H, Tsodyks M. Coding of Temporal Information by Activity-Dependent Synapses. *Journal of Neurophysiology*. 2002;87(1):140–148. 551
34. Maass W, Zador AM. Dynamic stochastic synapses as computational units. *Neural Computation*. 1999;11:903–917. 554
35. Matveev V, Wang XJ. Differential Short-term Synaptic Plasticity and Transmission of Complex Spike Trains: to Depress or to Facilitate? *Cerebral Cortex*. 2000;10(11):1143–1153. 557
36. Markram H, Wang Y, Tsodyks M. Differential signaling via the same axon of neocortical pyramidal neurons. *PNAS USA*. 1998;95:5323–5328. 559
37. Abbott LF, Varela JA, Sen K, Nelson SB. Synaptic Depression and Cortical Gain Control. *Science*. 1997;275:220–224. 561
38. Rosenbaum R, Rubin J, Doiron B. Short Term Synaptic Depression Imposes a Frequency Dependent Filter on Synaptic Information Transfer. *PLoS Computational Biology*. 2012;8(6):e1002557. 563
39. Abbott LF, Regehr WG. Synaptic computation. *Nature*. 2004;431:796–803. 566
40. Costa RP, Mizusaki BEP, Sjöström PJ, van Rossum MCW. Functional consequences of pre- and postsynaptic expression of synaptic plasticity. *Philosophical Transactions of the Royal Society B*. 2017;372:20160153. 567
41. Sjöström PJ, Turrigiano GG, Nelson SB. Neocortical LTD via Coincident Activation of Presynaptic NMDA and Cannabinoid Receptors. *Neuron*. 2003;39:641–654. 570
42. Deng PY, Klyachko VA. The diverse functions of short-term plasticity components in synaptic computations. *Communicative & integrative biology*. 2011;4(5):543–548. 573
43. Gawne TJ, Kjaer TW, Richmond BJ. Latency: another potential code for feature binding in striate cortex. *Journal of Neurophysiology*. 1996;76(2):1356–1360. 576
44. Gollisch T, Meister M. Rapid Neural Coding in the retina with relative Spike latencies. *Science*. 2008;319:1108–1111. 579
45. Song S, Abbott LF. Cortical Development and Remapping through Spike Timing-Dependent Plasticity. *Neuron*. 2001;32:339–350. 581
46. Gütiğ R, Aharonov R, Rotter S, Sompolinsky H. Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *The Journal of Neuroscience*. 2003;23(9):3697–3714. 583
47. Babadi B, Abbott LF. Stability and Competition in Multi-spike Models of Spike-Timing Dependent Plasticity. *PLoS Computational Biology*. 2016;12(3):e1004750. 586

48. Brock JA, Thomazeau A, Watanabe A, Li SSY, Sjöström PJ. A Practical Guide to Using CV Analysis for Determining the Locus of Synaptic Plasticity. *Frontiers in Synaptic Neuroscience*. 2020;12:11. 589-591
49. Kandel ER. The Molecular Biology of Memory Storage: A Dialogue Between Genes and Synapses. *Science*. 2001;294:1030–1038. 592-593
50. Sjöström PJ, Rancz EA, Roth A, Häusser M. Dendritic Excitability and Synaptic Plasticity. *Physiological Reviews*. 2008;88:769–840. 594-595
51. Sjöström PJ, Turrigiano GG, Nelson SB. Rate, timing and cooperativity jointly determine cortical synaptic plasticity. *Neuron*. 2001;32:1149–1164. 596-597
52. Sjöström PJ, Turrigiano GG, Nelson SB. Multiple forms of long-term plasticity at unitary neocortical layer 5 synapses. *Neuropharmacology*. 2007;57:176–184. 598-599
53. Miller KD. Synaptic economics: competition and cooperation in synaptic plasticity. *Neuron*. 1996;17:371–374. 600-601
54. Pfister JP, Gerstner W. Triplets of spikes in a model of Spike timing-dependent plasticity. *The Journal of Neuroscience*. 2006;26(38):9673–9682. 602-603
55. Froemke RC, Tsay IA, Raad M, Long JD, Dan Y. Contribution of Individual Spikes in Burst-Induced Long-Term Synaptic Modification. *Journal of Neurophysiology*. 2006;95:1620–1629. 604-606
56. Seung HS. Learning in Spiking Neural Networks by reinforcement of stochastic synaptic Transmission. *Neuron*. 2003;40:1063–1073. 607-608
57. Carvalho TP, Buonomano DV. A novel learning rule for long-term plasticity of short-term synaptic plasticity enhances temporal processing. *Frontiers in Integrative Neuroscience*. 2011;5:20. 609-611
58. Abbott LF, Nelson SB. Synaptic plasticity: taming the beast. *Nature Neuroscience*. 2000;3:1178–1183. 612-613
59. Harris J, Jolivet R, Attwell D. Synaptic Energy Use and Supply. *Neuron*. 2012;75(5):762–777. 614-615
60. Höyer PO, Hyvärinen A. Interpreting neural response variability as Monte Carlo sampling of the posterior. *Advances in Neural Information Processing Systems*. 2003; p. 293–300. 616-618
61. Aitchison L, Latham PE. Synaptic sampling: A connection between PSP variability and uncertainty explains neurophysiological observations; 2015. <https://arxiv.org/abs/1505.04544>. 619-621
62. Aitchison L, Jegminat J, Menendez JA, Pfister JP, Pouget A, Latham PE. Synaptic plasticity as Bayesian inference. *Nature Neuroscience*. 2021;24:565–571. 622-623
63. Abrahamsson T, Chou CYC, Li SY, Mancino A, Costa RP, Brock JA, et al. Differential Regulation of Evoked and Spontaneous release by Presynaptic NMDA Receptors. *Neuron*. 2017;96:839–855. 624-626
64. Markram H, Muller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA, et al. Reconstruction and Simulation of Neocortical Microcircuitry. *Cell*. 2015;163:456–492. 627-629

65. Merolla PA, Arthur JV, Alvarez-Icaza R, Cassidy AS, Sawada J, Akopyan F, et al. A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*. 2014;345:668–673. 630
631
632
66. Monday HR, Castillo PE. Closing the gap: long-term presynaptic plasticity in brain function and disease. *Current Opinion in Neurobiology*. 2017;45:106–112. 633
634
67. Brette R, Gerstner W. Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity. *Journal of Neurophysiology*. 2005;94:3637–3642. 635
636
637
68. Brette R. Generation of correlated spike trains. *Neural computation*. 2009;21(1):188–215. doi:10.1162/neco.2008.12-07-657. 638
639
69. Markram H, Lübke J, Frotscher M, Roth A, Sakmann B. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *Journal of Physiology*. 1997;500:409–440. 640
641
642
70. Tsodyks MV, Markram H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *PNAS USA*. 1997;94:719–723. 643
644
645
71. Costa RP, Sjöström PJ, van Rossum MCW. Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits. *Frontiers in Computational Neuroscience*. 2013;7:75. 646
647
648

A



B

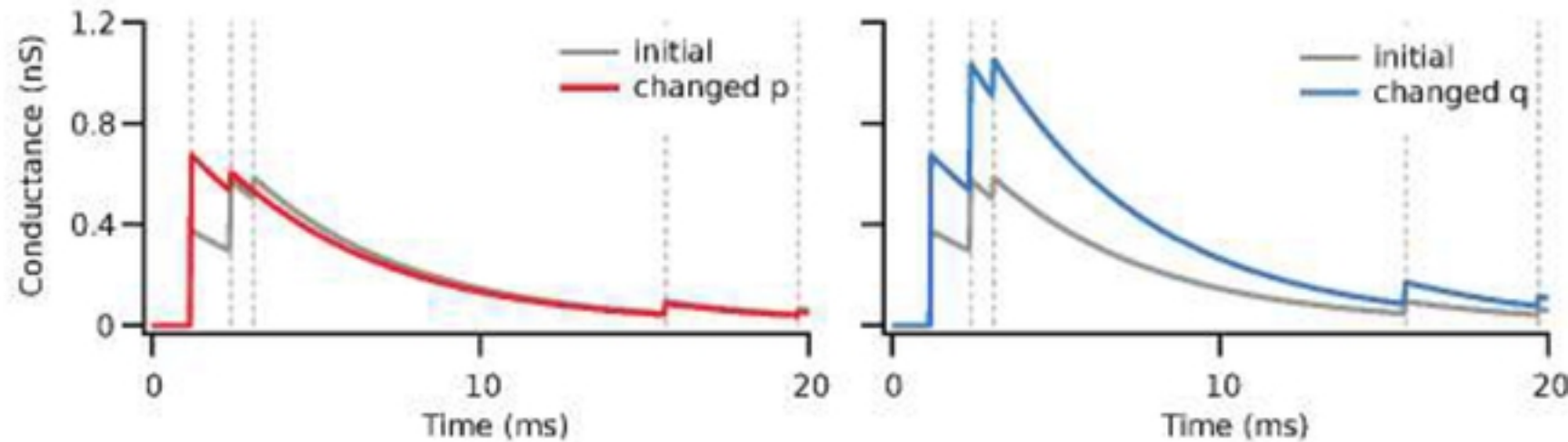
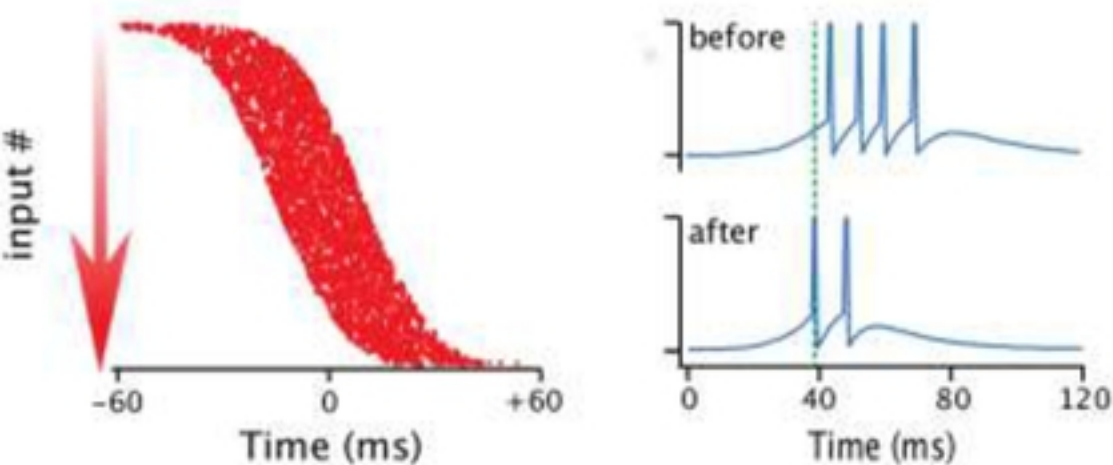
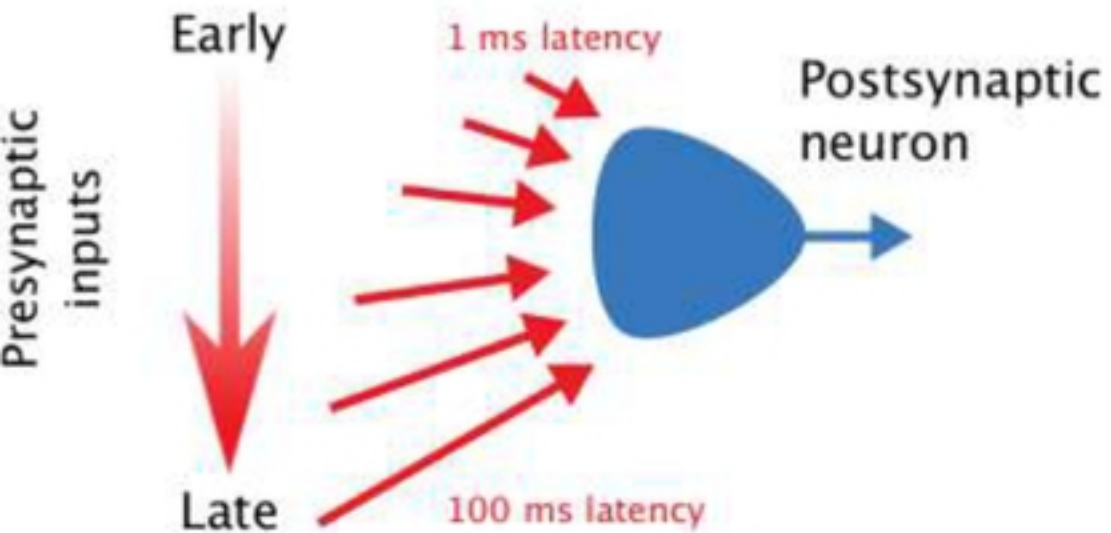


Fig 1

A

Latency Paradigm



B

Correlation Paradigm

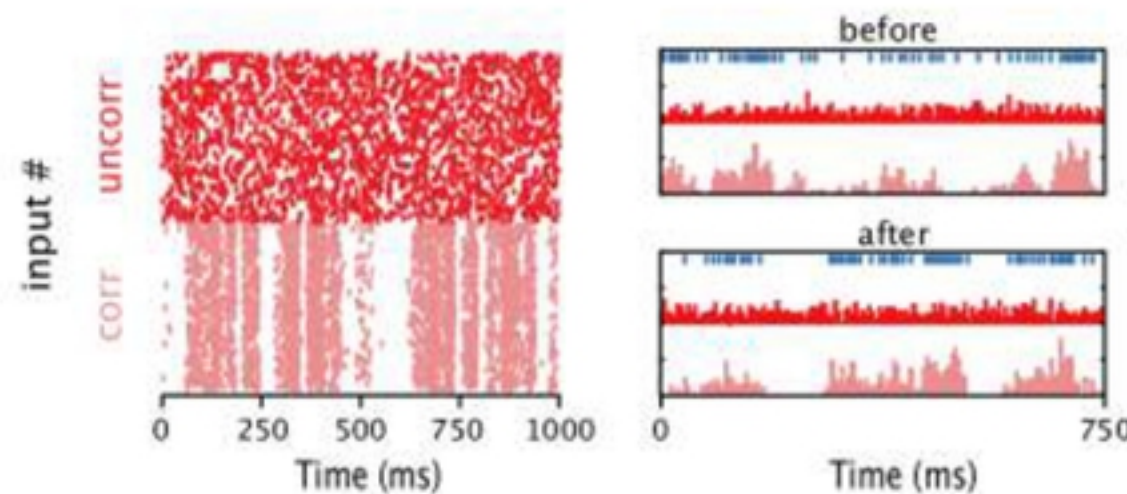
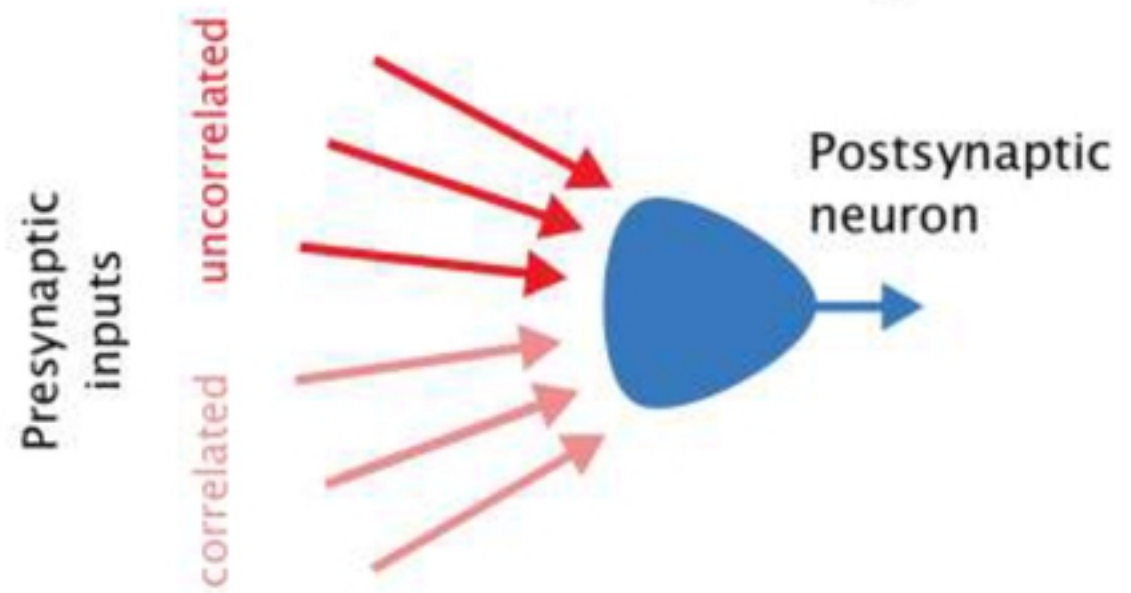
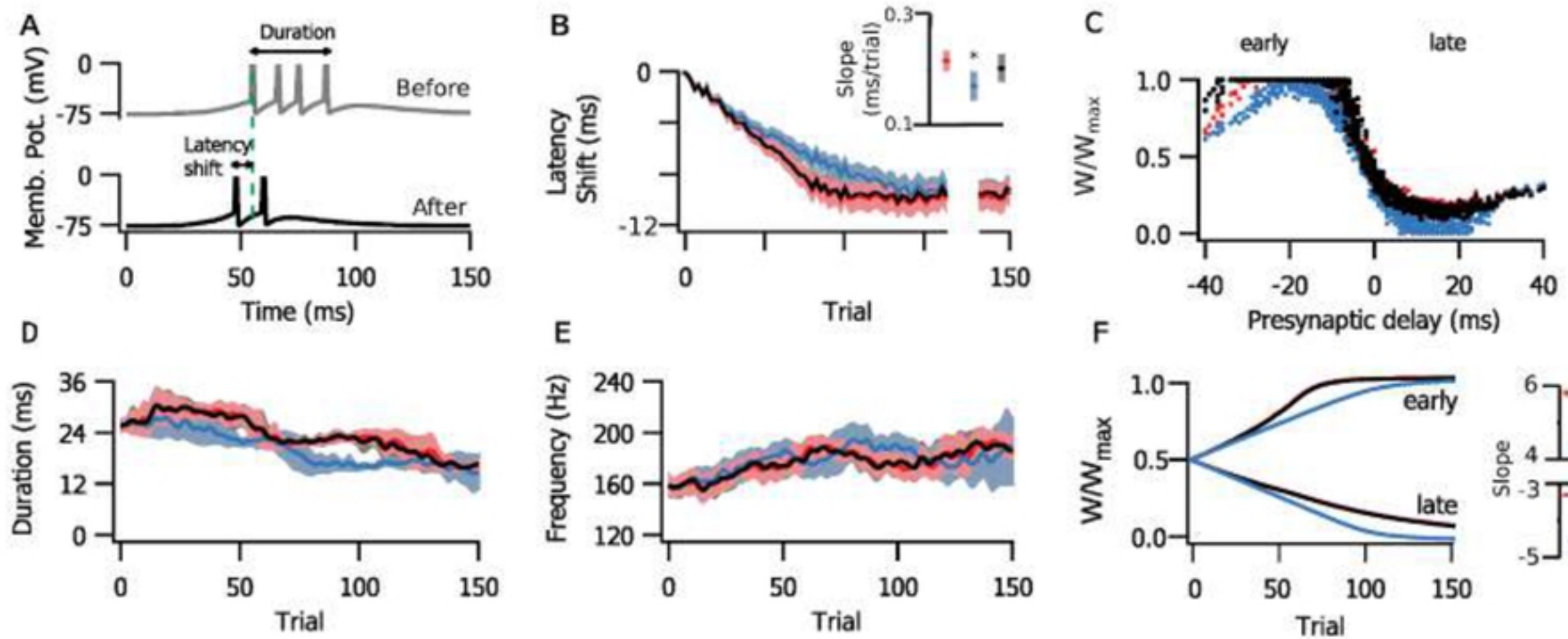


Fig 2

Latency paradigm



Correlation paradigm

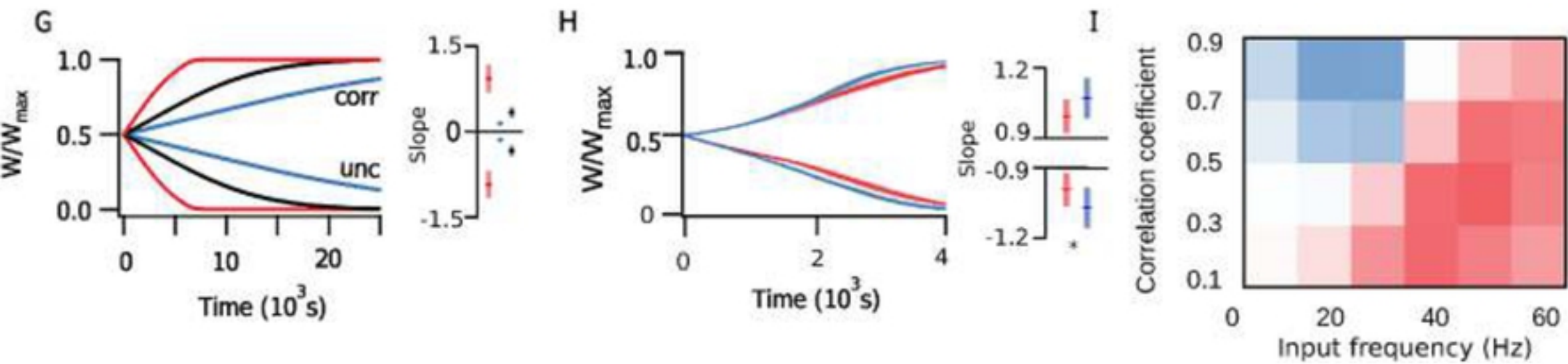
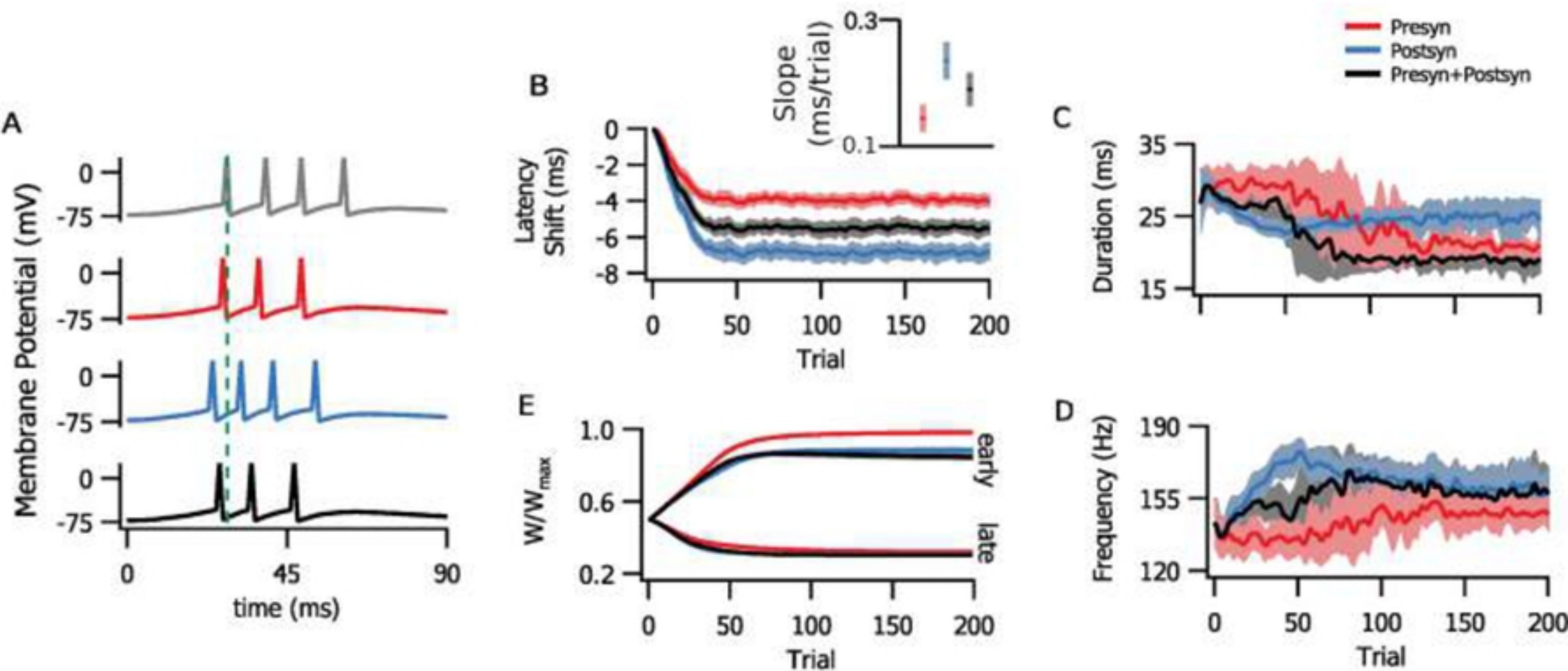


Fig 3

Latency paradigm



Correlation paradigm

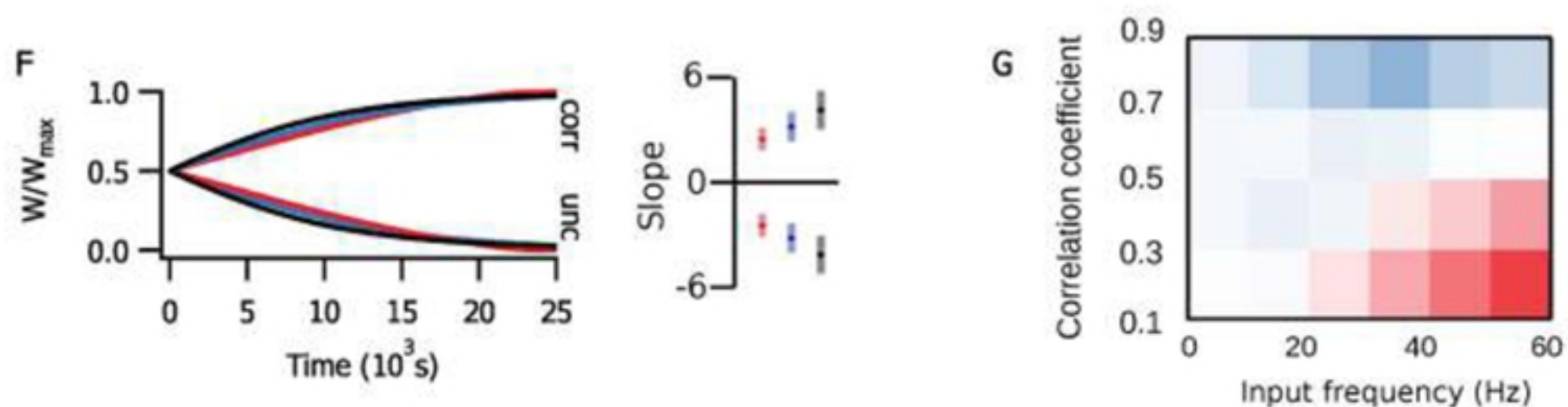
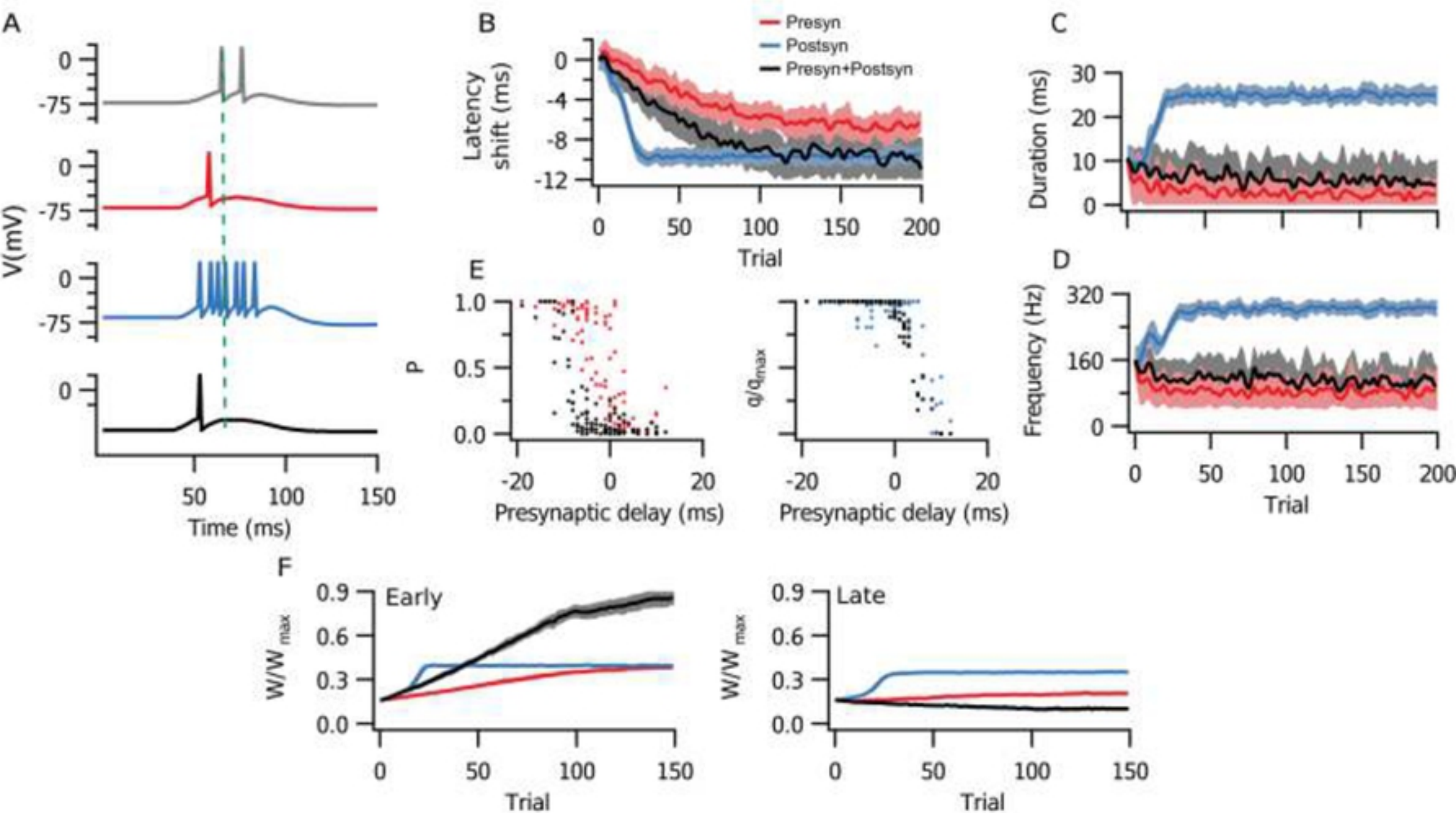


Fig 4

Latency paradigm



Correlation paradigm

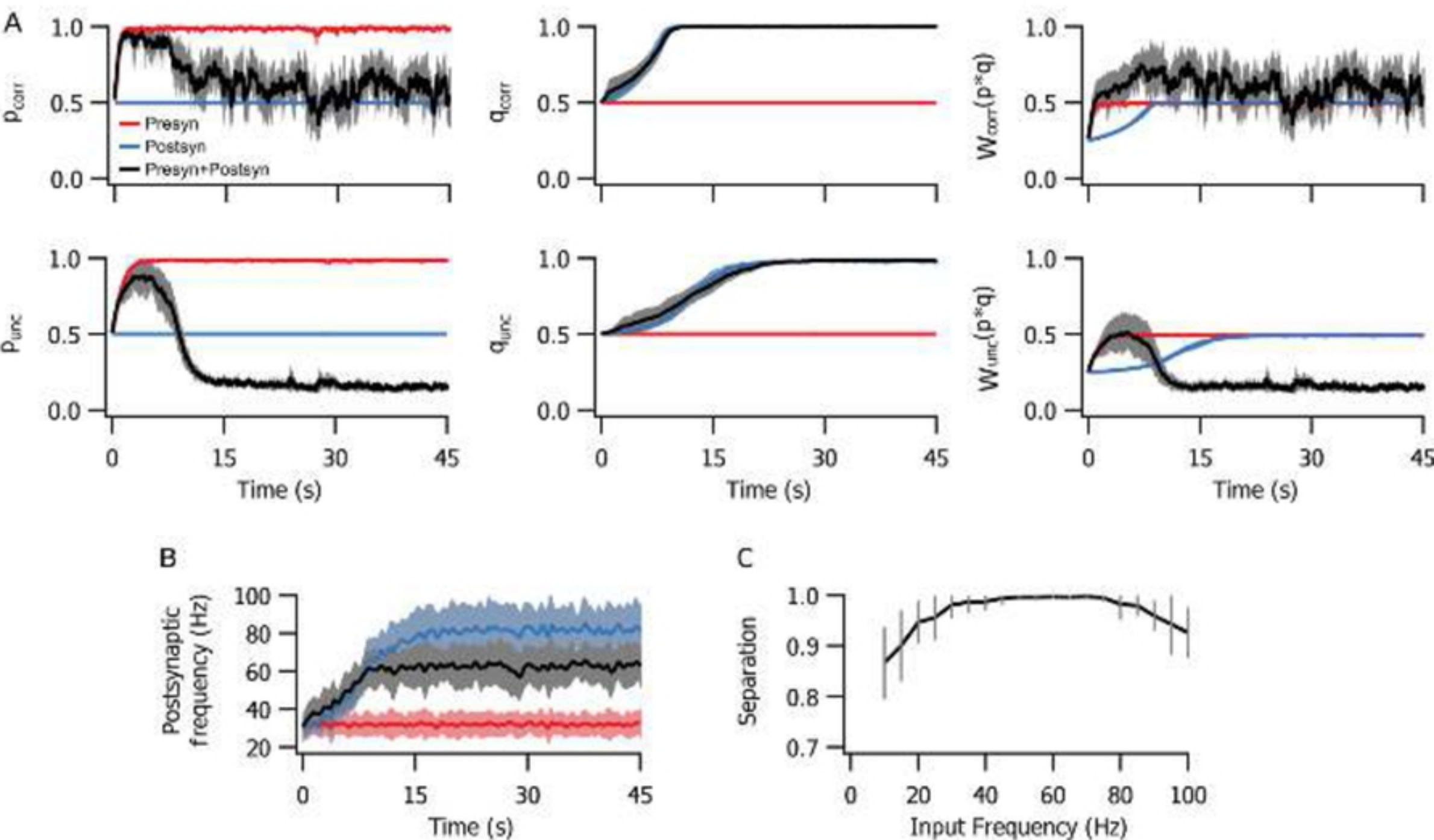
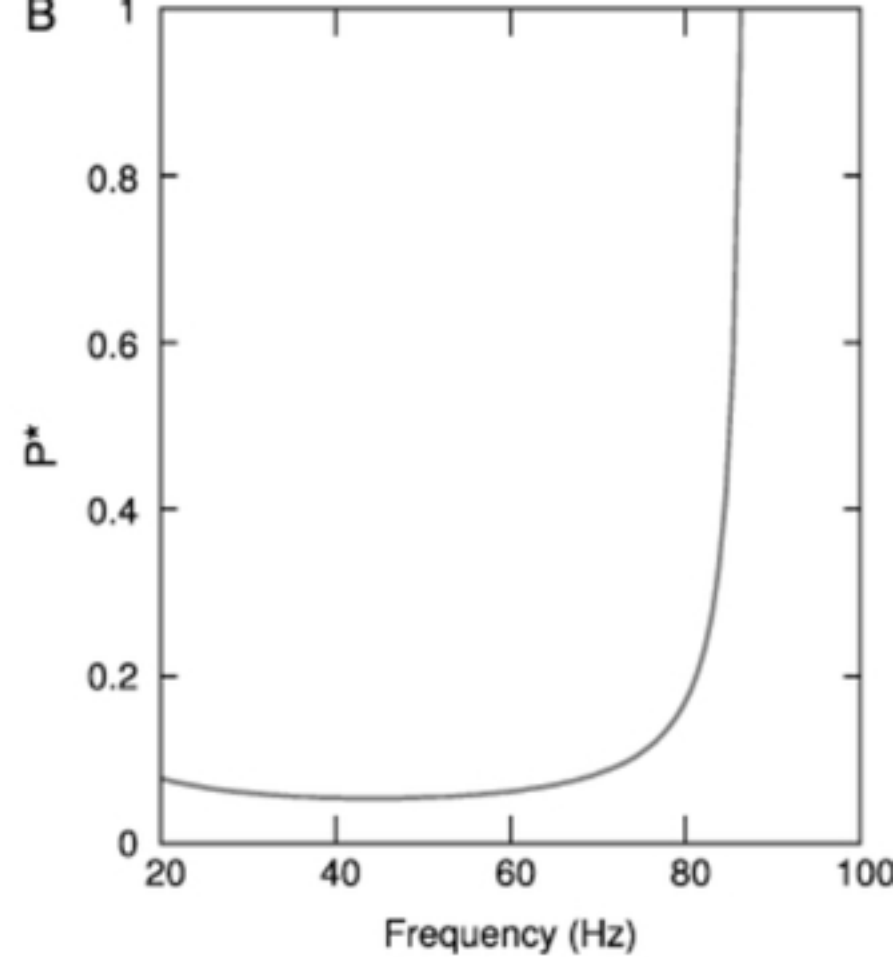
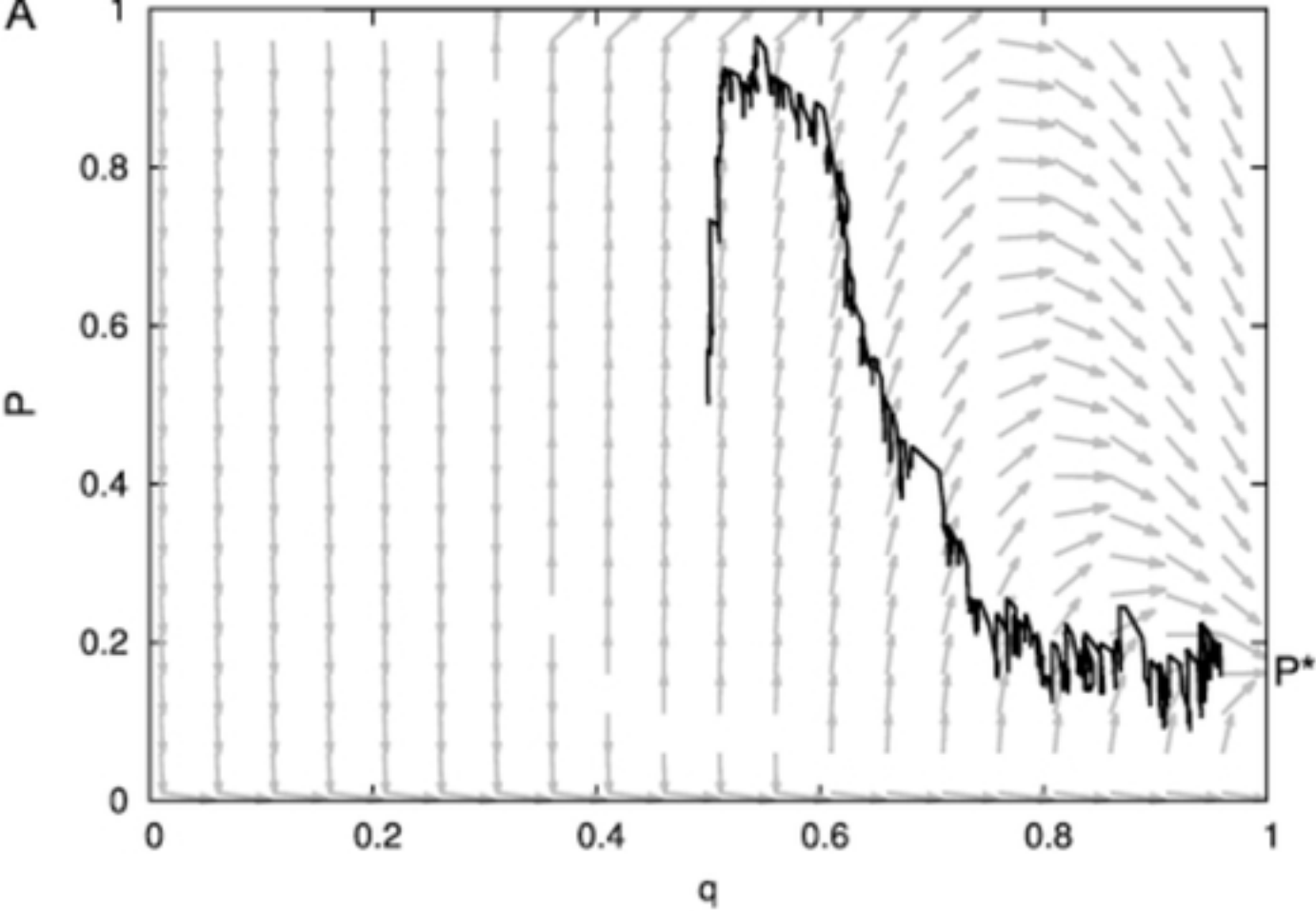


Fig 6



Supp Fig S1