

SHORT-TERM SYNAPTIC PLASTICITY MAKES NEURONS SENSITIVE TO THE DISTRIBUTION OF PRESYNAPTIC POPULATION FIRING RATES

Luiz Tauffer

School of Computer Science and Communication
KTH Royal Institute of Technology
Stockholm, Sweden
Bernstein Center Freiburg, University of Freiburg
Freiburg, Germany
tauffer@kth.se

Arvind Kumar

School of Computer Science and Communication
KTH Royal Institute of Technology
Stockholm, Sweden
arvkumar@kth.se

July 18, 2019

ABSTRACT

The ability to discriminate spikes that encode a particular stimulus from spikes produced by background activity is essential for reliable information processing in the brain. We describe how synaptic short-term plasticity (STP) modulates the output of presynaptic populations as a function of the distribution of the spiking activity and find a strong relationship between STP features and sparseness of the population code, which could solve the discrimination problem. Furthermore, we show that feedforward excitation followed by inhibition (FF-EI), combined with target-dependent STP, promote substantial increase in the signal gain even for considerable deviations from the optimal conditions, granting robustness to this mechanism. A simulated neuron driven by a spiking FF-EI network is reliably modulated as predicted by a rate analysis and inherits the ability to differentiate sparse signals from dense background activity changes of the same magnitude, even at very low signal-to-noise conditions. We propose that the STP-based distribution discrimination is likely a latent function in several regions such as the cerebellum and the hippocampus.

Introduction

The brain is a highly noisy system. At the cellular level, the neurons are unreliable in eliciting spikes and synapses are unreliable in transmitting the spikes to the postsynaptic neurons. At the network level, the connectivity and balance of excitation and inhibition gives rise to fluctuations in the background activity (Brunel, 2000; Kumar et al., 2008) which can be as high the mean stimulus response (Arieli et al., 1996). In such a noisy environment, a neuron is faced with two crucial tasks: (1) discriminating between a stimulus and fluctuations in the background activity and (2) discriminating between inputs of the same intensity (number of spikes in a period of time) but with different arrangements.

If synapses were static (that is, when the postsynaptic conductances do not depend on the immediate spike history) both these tasks cannot be accomplished, unless synapses are specifically tuned to do so. For instance, the identification of specific spiking patterns, filtering out presumed noise sequences, can be accomplished by precise tuning of synaptic weights (Gütig and Sompolinsky, 2006). This solution, however, relies on training synaptic weights using a certain supervised learning rule and even then it could only work for a specific set of spike timing sequences. Active dendrites (with voltage dependent ionic conductance) can also work as pattern detectors (Hawkins and Ahmad, 2016), but this mechanism would only work for signals constrained to locally clustered synapses. Therefore, despite being relevant for the understanding of signal processing in the brain, the mechanisms by which neural ensembles solve the activity discrimination problem have remained elusive.

Here, we show that short-term plasticity (STP) of synapses provides an effective and general mechanism to solve the two aforementioned tasks. STP refers to the observation that synaptic strength changes on spike-by-spike basis,

depending on the timing of previous spikes (Stevens and Wang, 1995; Zucker and Regehr, 2002). STP arises because neurotransmitter release dynamics is history dependent. STP is believed to play several important roles in neural information processing (Fuhrmann et al., 2002; Abbott and Regehr, 2004; Rotman et al., 2011; Scott et al., 2012; Jackman and Regehr, 2017). Relevant to the two aforementioned tasks, synapses that express STP (dynamic synapses) have been suggested to help a neuron learn to identify input spike sequences (Buonomano, 2000; Rotman and Klyachko, 2013). But this solution also relies on a supervised learning rule and is constrained to specific temporal sequences.

An obvious consequence of STP is that the effective postsynaptic conductances depend on the firing rates of individual presynaptic neurons. This suggests that postsynaptic targets of populations with dynamic synapses could distinguish between different input distributions even without supervised learning. To demonstrate this feature of STP, we measured the response of postsynaptic neurons for a weak stimulus with amplitude one order of magnitude smaller than the background activity. By systematically changing the distribution of rates over the presynaptic neuron ensemble, we found that weak signals can be differentiated from the noisy fluctuations if the signal is appropriately distributed over the input ensemble. The optimal distribution that maximizes the discriminability depends on the nature of STP. We found that, for facilitatory synapses, sparse codes give better discrimination between a weak signal and dense background changes of the same intensity. By contrast, for depressing synapses, sparse codes result in very negative gains in relation to dense background changes of the same intensity. We also investigated the feedforward networks with excitation and disinhibitory inhibition, with the target-dependent STP, and found that this arrangement allows for extra robustness for this mechanism.

Finally, we demonstrate how STP can endow a postsynaptic neuron with the ability to differentiate sparsely encoded activity from dense activity of the same magnitude, a function that would be especially important at very low signal-to-noise regimes. Thus, our results reveal that the nature of STP may also constrain the nature of neural population code.

Results

Here we are interested in a mechanism by which a neuronal network or a single postsynaptic neuron receiving multiple inputs may distinguish between different spike patterns with the same intensity (e.g. the same number of spikes). To illustrate this situation, consider a simple example of a sequence of 7 spikes arriving from cells presynaptic to a common readout neuron, where all the spikes can arrive from a single presynaptic neuron (Figure 1A) or from different neurons (Figure 1B). Static synapses evoke exactly the same postsynaptic conductance (PSC) sequence for the two distinct activity distributions (black lines), making both indistinguishable for a readout neuron. However, when synapses are dynamic, short-term facilitation (STF, blue line) enhances the PSC amplitudes compared to the static synapses (cf. Figure 1A,B, bottom traces). Short-term depression (STD, red line) results in a weaker response as compared to the static synapses (cf. Figure 1A,B, bottom traces). If the incoming spikes are distributed along different synapses, the sequence of PSCs is identical for all types of synaptic dynamics (cf. Figure 1A,B, bottom traces).

In vivo neural coding is certainly more complex than the above example. However, this simple example suggests that in the case of a neuron receiving synaptic inputs via thousands of noisy synapses, STP could be a mechanism to differentiate between an evoked signal from the background activity fluctuations of the same amplitude, provided the former is encoded as a specific pattern that can exploit the STP properties of the synapses. In the following, we describe how well dynamic synapses could endow feedforward circuits with such activity distribution discrimination properties in low signal-to-noise regimes (Figure 1C).

Optimal activity distribution with dynamic synapses

We implemented dynamic synapses with the rate-based TM model (Tsodyks et al., 1998) (equation 4). In this model, the instantaneous proportion of resources released (PRR) depend on the vesicle release probability and available vesicle resources. These variables depend on the spike history and, therefore, on the spike rate when the spikes are arriving in a Poisson manner. To quantify how the available vesicle resources change over time for a given spike rate, we defined the total amount of resources (Q^s) a synapse releases over a time period of T_s (equation 6; Figure 2B). Given the STP properties and a stationary baseline firing rate r_{bas} , Q^s reaches a stationary state (Q_{bas}^s). However, any transient change in the firing rate will alter the PRR (Figure 2A) and Q_{ext}^s .

We found that the extra amount of resources released by a single synapse (Q_{ext}^s) varied in a nonlinear fashion as a function of the transient rate increase r_{ext} (equation 6, Figure 2B). The slope of Q_{ext}^s (Figure 2B *inset*) for depressing synapses is monotonically decreasing, indicating that any increase in the firing rate in those synapses will produce sublinear increase in Q_{ext}^s , whereas for facilitatory synapses the slope initially increases, indicating that increases in the firing rate of those synapses, up to some point, will produce supralinear increase in Q_{ext}^s .

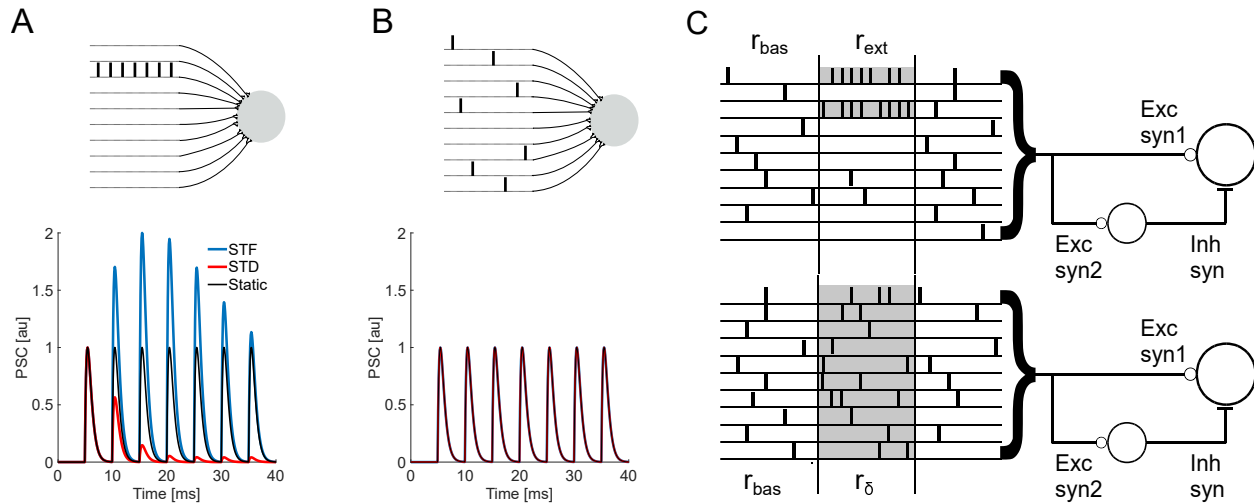


Figure 1: Distribution of the spiking activity over presynaptic neurons and short-term-plasticity. **A** Top: A neuron receives inputs from 10 presynaptic neurons. Of these only one of the neurons elicits 7 spikes. Bottom: The PSC generated by 7 consecutive spikes for three different types of synapses (static – black; facilitatory – blue; depressing – red). The PSCs are different for each of these three types of synapses. **B** Top: A neuron receives inputs from 10 presynaptic neurons. In this scenario each neuron elicited a single spike. But the postsynaptic neurons received 7 spikes as in the panel A. Bottom: The PSC generated by a sequence of 7 consecutive spikes arriving at the same time as in panel A coming from three different types of synapses (static – black; facilitatory – blue; depressing – red). The PSCs are identical for each of these three types of synapses (lines overlapped). **C** Feedforward excitation/inhibition (FF-EI) configuration and two distributions of an extra spike rate R_{ext} . Top: the extra rate is distributed into a few presynaptic neurons (gray), with each chosen unit increasing its rate by $r_{ext} = R_{ext}/N_{ext}$. Bottom: the extra rate is distributed homogeneously throughout the population of N units, with each unit increasing its rate by $r_{\delta} = R_{ext}/N$.

In the brain, neurons typically receive inputs from a large ensemble of presynaptic neurons. In the ongoing activity state these neurons spike at a low-baseline firing rate (r_{bas}) and in the event-related activity state, firing rate of a subset of presynaptic neurons is transiently increased. Therefore, it is important to understand how the total synaptic resources released (Q^p) changes as the firing rate of a fraction of neurons is transiently increased. The calculation of Q^p is the sum of individual Q^s (equation 7). As is evident from the equation 7, Q^p depends on number of neurons whose firing rate is altered. We distribute a fixed event-related population rate increase R_{ext} into varied number of chosen synapses N_{ext} , each of these chosen synapses increasing its firing rate by r_{ext} , that is, $R_{ext} = N_{ext} \times r_{ext}$, and report the changes in Q_{ext}^p .

We found that when synapses show short-term facilitation, for fixed values of R_{ext} , Q^p in response to external input (Q_{ext}^p) varied in a non-monotonic fashion as a function of N_{ext} (Figure 2C left). The total synaptic resources released during the event-related activity is a sum of synaptic resources released because of the ongoing activity (Q_{bas}^p) and the transient increase caused by the event (Q_{ext}^p). Q_{bas}^p is independent of N_{ext} , therefore, the change in synaptic resources released by a transient firing rate increase ($Q_{ext}^p - Q_{bas}^p$) also varied in a non-monotonic fashion, as a function of N_{ext} (Figure 2C left). By contrast, for short-term depressing synapses the change in synaptic resources released during the event-related activity varied in a monotonic fashion (Figure 2C right). For both facilitatory and depressing synapses, the change in synaptic resources released converged to their respective Q_{δ}^p when the input rate R_{ext} was distributed over all the input neurons such that $r_{ext} = r_{\delta} = R_{ext}/N$.

These results suggest that, when synapses are facilitatory, total synaptic resources released during a event-related activity state is maximized when event-related spiking activity is confined to a small number of synapses. Q_{ext}^p was smaller than Q_{δ}^p when R_{ext} was distributed into a small subset of presynaptic neurons, because those chosen neurons spiked at very high rates and the synapses ran out of vesicle resources rapidly. When the event-related input was distributed over all the presynaptic neurons, the Q_{ext}^p also decreased because in such a scenario $r_{ext} = r_{\delta} = R_{ext}/N$ was too small to fully exploit the benefits of synaptic facilitation. In contrast to the facilitatory synapses, for depressing synapses it was more beneficial to distribute the event-related spiking activity over the whole input ensemble in order to maximize the amount total synaptic resources released. That is, in this scenario r_{ext} was small enough to avoid any losses in vesicle release because of synaptic depression.

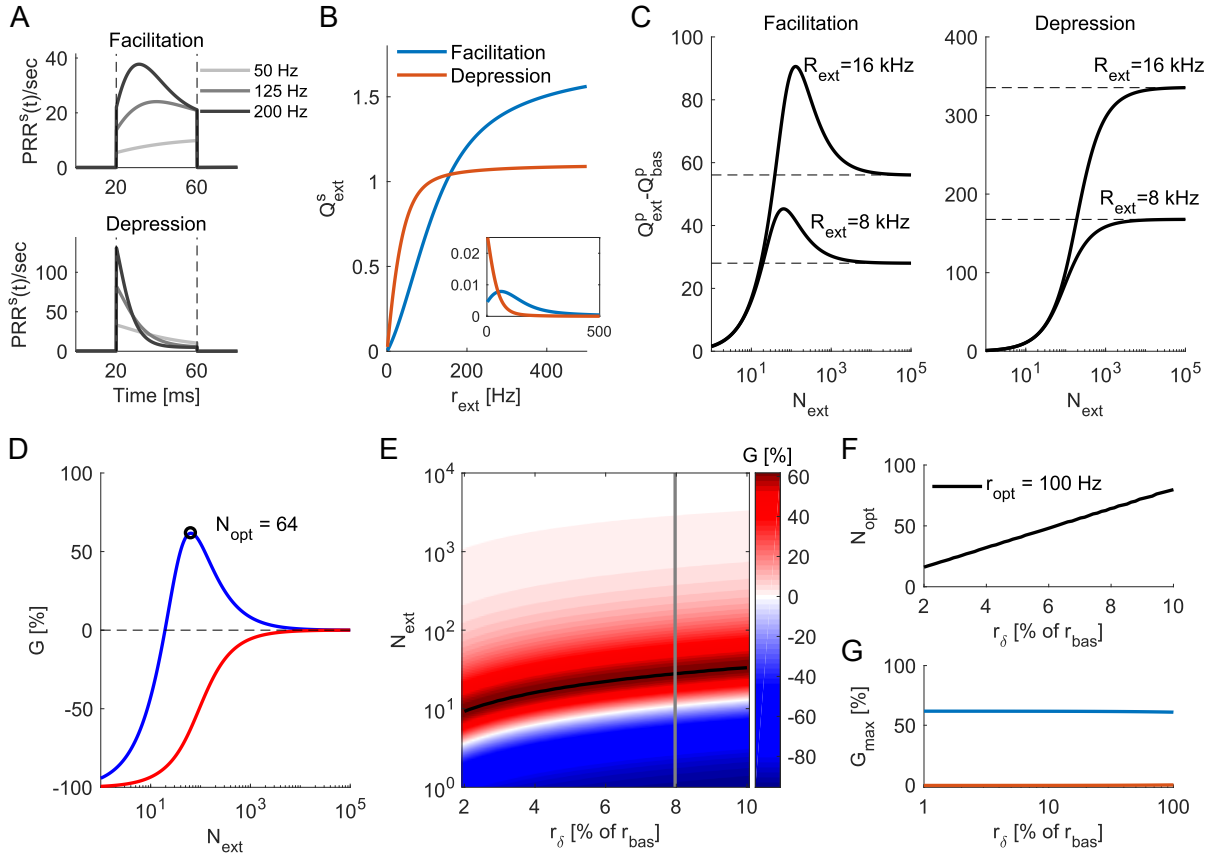


Figure 2: **A** Temporal profile of the PRR for a facilitatory (*top*, $U = 0.1$, $\tau_f = 200ms$ and $\tau_{rec} = 50ms$) and depressing (*bottom*, $U = 0.7$, $\tau_f = 50ms$ and $\tau_{rec} = 200ms$) synapses with increased rates during a period of $T_s = 40ms$. **B** The amount of resources released by a single synapse, Q^s . This was obtained by integrating $PRR^s(t)$ over T_s (area under the curves in **A**, equation 6). Q_{ext}^s for depressing synapses saturates at lower firing rates than facilitatory synapses. The inset shows the derivative of Q^s and highlights the nonlinearities in Q^s , with depressing synapses showing monotonically decreasing slopes (decreasing release rate) and facilitatory synapses showing an initial region of increasing slopes (increasing release rate) with respect to r_{ext} . **C** The extra proportion of released resources, $Q_{ext}^p - Q_{bas}^p$ as a function of the number of presynaptic neurons (N_{ext}) whose firing rate increases by two different values of R_{ext} . Dashed lines mark the value achieved when $N_{ext} = N$ i.e. the dense distribution case. A population of facilitatory synapses (left) maximizes its release with low N_{ext} , while a population of depressing synapses (right) maximizes its release with $N_{ext} = N$. **D** The gain (G , equation 8) as a function of N_{ext} for a fixed R_{ext} . The N_{ext} that maximizes G , for this particular extra rate, is $N_{opt} = 64$ for facilitatory synapses and $N_{opt} = N$ for depressing synapses. Notice that if the extra rate is allocated in even fewer input units, G can be negative. **E** G surface for a facilitatory synapse as a function of r_δ and N_{ext} . The black line marks the maximum values of G i.e. N_{opt} for each r_δ . The gain curves at panel **D**, where $r_\delta = 8\%$ of r_{bas} , is marked with a gray line for reference. **F** The relationship between N_{opt} and r_δ is linear. At maximal gain (G_{max}), the firing rate of the event-related neurons (N_{ext}) is r_{opt} (100Hz for this specific example). **G** G_{max} for the two STP regimes shown in panel **B**. For a low signal-to-basal ratio $r_\delta/r_{bas} < 1$, the gain can be considered independent from the stimulus intensity r_δ .

109 Activity distribution-dependent Gain

110 To further quantify the effect of distribution of event-related activity over the input ensemble (that is, how neurons
111 increase their rate in the event-related phase), we defined the distribution gain G as the proportional change in Q_{ext}^p
112 in relation to Q_δ^p (equation 8). We found that Q_δ^p is approximately a linear function of r_δ for a wide range of scenarios (see
113 Materials and Methods) and, because of that, with the dense distribution of the inputs (when all the input neurons change
114 their input rate by a small amount r_δ in the event-related activity state), even STP type synapses behave approximately
115 as static synapses. Therefore, G can be understood as a gain over a dense distribution or a gain over static synapses. For

facilitatory synapses, similar to Q_{ext}^p , G also followed a non-monotonic curve as a function of N_{ext} , with a single peak at N_{opt} (Figure 2D, blue line). By contrast, depressing synapses resulted in negative gains to every distribution, except for $N_{ext} = N$ where $G = 0\%$ (Figure 2D, red line).

Next, we estimate N_{opt} and G for a range of extra input intensities (Figure 2E, for facilitatory synapses). For these calculations, we parameterized the extra input R_{ext} as a fraction of the baseline firing rate R_{bas} (correspondingly, r_δ as % of r_{bas} , see Materials and Methods). We found that, for facilitatory synapses, N_{opt} increased linearly with the extra input intensity (Figure 2G), resulting in an optimal encoding rate r_{opt} which was independent of the input intensity. For depressing synapses, the optimal distribution $N_{opt} = N$ did not change with R_{ext} , making the optimal encoding rate $r_{opt} = r_\delta$.

Because the presynaptic neurons are assumed to be Poisson processes, an advantage of parametrizing R_{ext} in terms of fraction of R_{bas} is that it directly translates to signal-to-noise ratio. For the example shown in Figure 2G, we found that STP could amplify the presynaptic output for weak signals (which were less than 10% of the baseline activity) by up to 60% if the extra rate was distributed over N_{opt} synapses as opposed to N synapses. For low signal-to-noise ratios ($r_\delta < r_{bas}$), the gain at the optimal distribution G_{max} was approximately constant and always positive for facilitatory synapses, while depressing synapses keep $G_{max} = 0$ at $N_{opt} = N$ (Figure 2G). Finally, we analytically show that the independence of r_{opt} and G_{max} from the extra rate intensity is true for a wide range of basal rates and STP types (see Materials and Methods).

These results suggest that when synapses are facilitatory, the input should be distributed sparsely (or sparse code, that is, only a small set of neurons change their firing rate in the event-related state) to maximize the synaptic resources released at the downstream neuron. By contrast, when synapses are depressing, the input should be distributed densely (or dense code, that is, all the neurons change their firing rate in the event-related state) to maximize the synaptic resources released at the downstream neuron. Thus, for sparse population activity, facilitatory synapses get optimally utilized and depressing synapses get subutilized.

Effects of STP parameters on optimal rate and gain

Next, we investigated how N_{opt} , r_{opt} and G_{max} vary with STP parameters. To this end we systematically changed synapses from facilitatory to depressing by varying the set of parameters $\{U, \tau_f, \tau_{rec}\}$. We found that r_{opt} decayed exponentially as the synapses become more depressing (Figure 3A). This follows from the fact that facilitatory synapses profit from high firing rates and depressing synapses avoid negative gains at lower rates.

The maximum gain G_{max} also decreased exponentially as synapses were systematically changed from facilitatory to depressing (3B). We found that the relationship between gain and optimal rate was linear from mildly to strongly facilitatory synapses (3C), with larger basal rates constraining the optimal conditions to lower rates with lower gains.

Interestingly, increasing the basal firing rate r_{bas} substantially reduced r_{opt} and G_{max} . This is surprising because, at such low values of spiking rates, STP effects are hardly perceivable in traditional paired-pulse ratio analyses. The high value of G_{max} , when the system operates at low r_{bas} , is because of the many synapses taking advantage of the nonlinearities in their individual Q^s . That is, even if there was a small change in Q^s , because of the large number of input synapses, the collective effect was much higher. Therefore, increased baseline activity strongly impaired the capacity to exploit these nonlinearities and, ultimately, the distinction of activity distributions (sparse vs. dense).

Relationship between facilitatory synapses and sparse coding

We quantified the optimal distribution of an evoked neural signal by OD (see Materials and Methods). High OD ($OD \rightarrow 1$) indicates a dense distribution in which many neurons spike to encode the extra activity, whereas low OD ($OD \rightarrow 0$) implies a sparse distribution. We found that OD changed abruptly from sparse to dense as synapses were changed from facilitatory to depressing (Figure 3D). Facilitatory synapses yielded maximum response for sparse while depressing synapses yielded maximum response (avoid negative gains) for dense distributions. The transition point from sparse to dense OD did not depend on the stimulus duration. However, the basal rate strongly modified the transition point, with higher r_{bas} allowing only strongly facilitatory synapses to take advantage of sparse distributions. This configuration remained independent of the stimulus intensity as long as the circuit operates at low signal-to-noise conditions ($r_\delta/r_{bas} < 1$, Figure 2G).

In the above, we characterized the synapses as facilitatory and depressing without being specific about the model parameters. Next, we systematically varied each of the STP parameters and measured the OD for maximum gain. We found that the transition region was primarily governed by the facilitation factor U ($U \sim 0.45$), with a weak dependence on τ_{rec} and τ_f (Figure 3E). The relative contribution of τ_{rec} and τ_f became more relevant at higher T_s (Supplementary Figure 1).

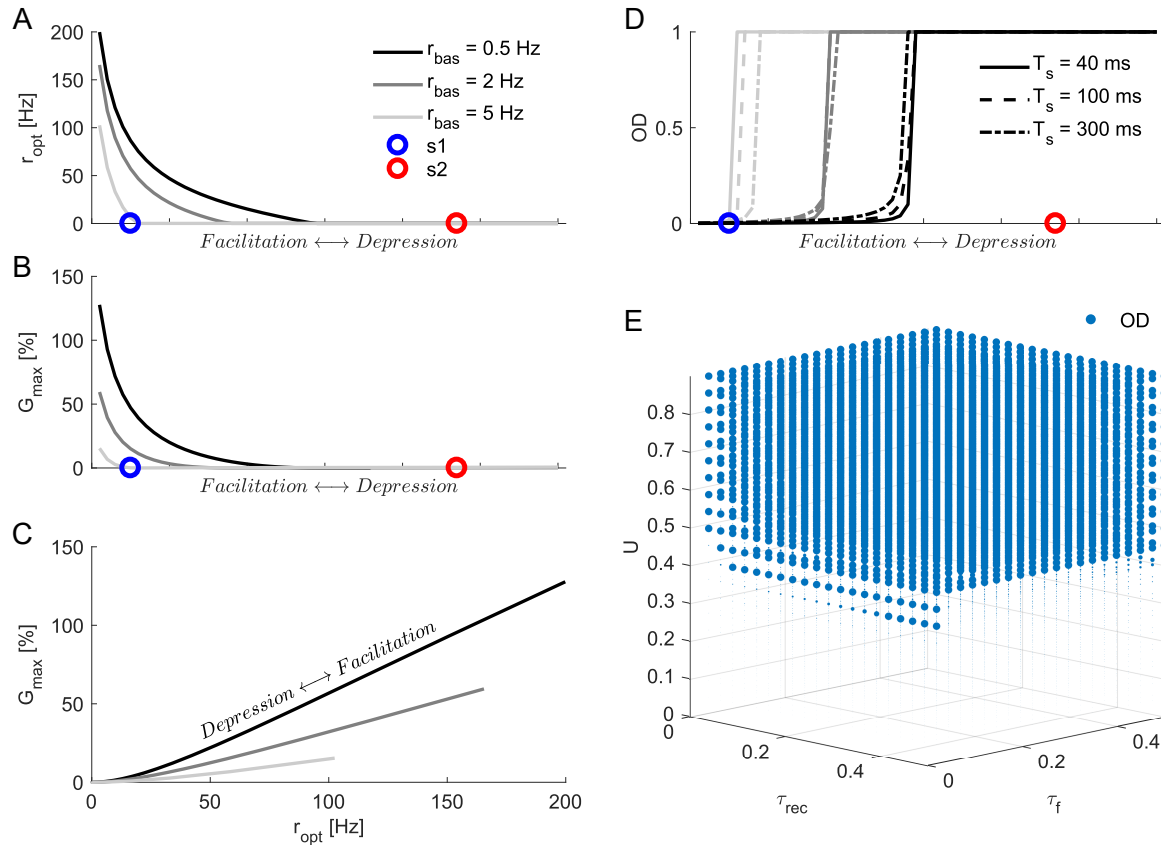


Figure 3: Effects of STP attributes on maximum gain of the neural population. Here the STP parameters U , τ_{rec} and τ_{fac} were varied in the following range: $U : 0.05 \rightarrow 0.9$, $\tau_{rec} : 20 \rightarrow 500ms$ and $\tau_f : 500 \rightarrow 20ms$. **A** The optimum frequency r_{opt} as a function of STP properties that gradually and monotonically change the synapse from facilitatory to depressing. r_{opt} is high for facilitatory synapses and low for depressing synapses. r_{opt} monotonically decreases as synapses change from facilitatory to depressing. As the basal rate is increased, r_{opt} decreased for all types of synapses. The circle markers show r_{opt} for a facilitatory (s1, blue) and a depressing (s2, red) synapse used later in the study. **B** G_{max} as a function of STP properties that gradually and monotonically change the synapse from facilitatory to depressing. **C** The relationship between G_{max} and r_{opt} . Notice the approximately linear relationship for facilitatory synapses, with the slope steadily decreasing with increasing r_{bas} . **D** Optimal distribution of rate over N_{opt} presynaptic neurons that maximize the gain G . The change from sparse to dense optimal distribution (0 to 1 OD , respectively) is abrupt and occurs approximately at the same STP region for all T_s . However, the transition point where the input distribution changes from sparse to dense code is strongly modulated by r_{bas} – higher basal rates allows for sparse code only for more facilitatory synapses. **E** Optimal distribution of rate as a function of the three key model parameters (U , τ_{rec} and τ_{fac}). The variable U is the most influential in defining the optimal encoding distribution, with $U \sim 0.45$ defining the OD transition point for $T_s = 40ms$ and $r_{bas} = 0.5Hz$. Marker sizes represent OD values, with large ones for $OD = 1$ and small ones for $OD \approx 0$.

These results clearly highlight the importance of the stationary basal rate in how well the synaptic gain modulation operates, as only low r_{bas} allows for significant gains. Importantly, the switch-like behavior of the optimal distribution indicates that, for a given population code, there is a robust range of STP attributes that could produce positive gains. This transition point seems to be relatively independent of the signal duration but is strongly affected by r_{bas} . Finally, having a low initial release probability (defined in the model by a low U) seems to be the preeminent feature in defining the optimal OD . Next, we investigate how fine tuning of parameters could affect the optimal conditions.

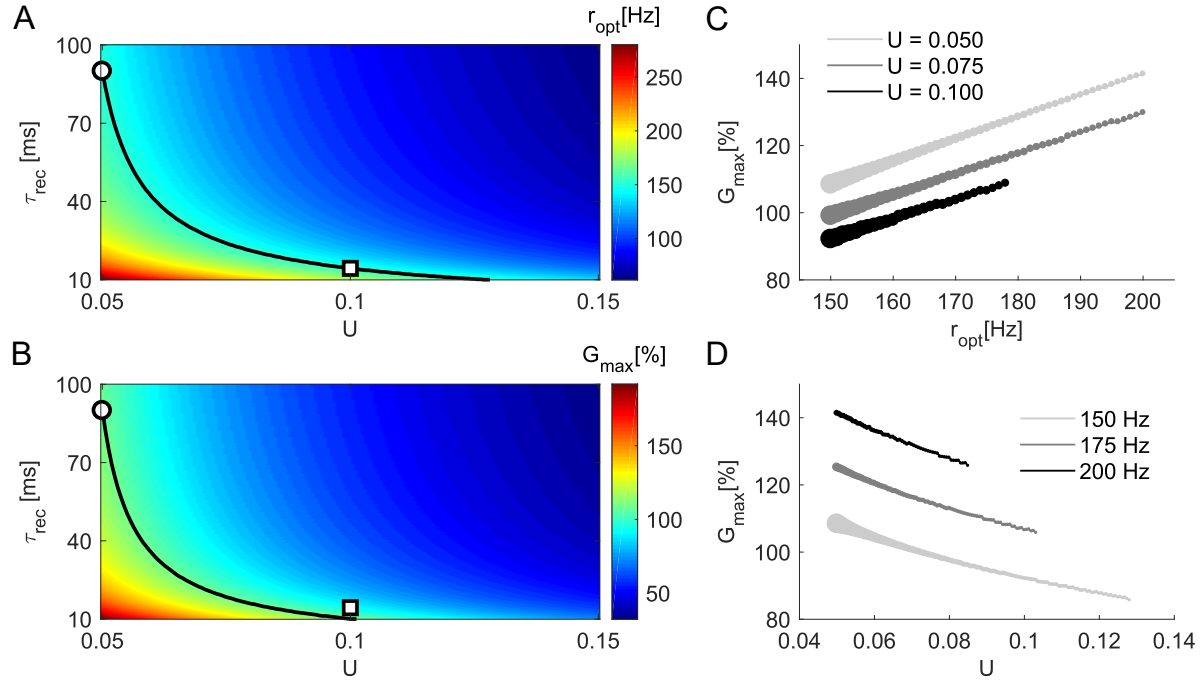


Figure 4: Effects of resources recovery time τ_{rec} and facilitation factor U on G_{max} and r_{opt} for $T_s = 40ms$, for facilitatory synapses. **A** r_{opt} as a function of facilitation factor (U) and recovery time constant (τ_{rec}). The r_{opt} surface shows that a given optimum encoding rate can be matched by different combinations of synaptic parameters. For example, the iso-frequency curve of $150Hz$ (black line) is achieved with $U = 0.05$ and $\tau_{rec} = 90ms$ (\circ) or with $U = 0.1$ and $\tau_{rec} = 15ms$ (\square). **B** G_{max} as a function of facilitation factor (U) and recovery time constant (τ_{rec}). Same maximal gain can be observed for many different combinations of U and τ_{rec} . The black line shows the contour for $G_{max} = 109\%$. The two configurations with same r_{opt} marked in panel A have distinct gains ($\circ = 109\%$, $\square = 92\%$). **C** We fix U and vary τ_{rec} (circle sizes) to match r_{opt} (x axis), then observe the gain. Larger values of U systematically produce smaller gains. Recovery time has a lower boundary $\tau_{rec} = 10ms$. **D** G_{max} as a function of U for three different values of r_{opt} . Larger values of U require smaller values of τ_{rec} (circle sizes) to match the same r_{opt} , but as a consequence the gain decreases as we increase U .

Effects of different sources of enhancement on G_{max}

The enhancement of the output at facilitatory synapses could, in principle, have many causes (Valera et al., 2012; Thanawala and Regehr, 2013; Jackman and Regehr, 2017). Using the TM model (equation 4), we phenomenologically accounted for two important sources: a low initial release probability which sequentially increases with each incoming spike (Jackman et al., 2016) and fast replenishment of readily available resources (Crowley et al., 2007). The first characteristic is mimicked by a low facilitation factor U , which determines the initial release probability after a long quiescent period and the proportional increase in it after each spike. The second mechanism is captured by a fast recovery time constant τ_{rec} .

We systematically varied U and τ_{rec} and measured G_{max} and r_{opt} . We found that several different combinations of U and τ_{rec} resulted in the same optimal distribution gain and rate. However, when we changed U and τ_{rec} while keeping the r_{opt} fixed, G_{max} could no longer be kept constant and vice versa. For instance, the two parameter sets $\{U = 0.05, \tau_{rec} = 90ms\}$ and $\{U = 0.1, \tau_{rec} = 15ms\}$ gave $r_{opt} = 150Hz$ (Figure 4A), but the first parameter set gave $G_{max} = 109\%$ and the second parameter set gave $G_{max} = 92\%$ (Figure 4B). Holding U fixed and choosing τ_{rec} to match with different r_{opt} showed that G_{max} consistently dropped for higher U (Figure 4C,D).

These results indicate that, in terms of maximum gain G_{max} , the fine tuning of intracellular mechanisms that work to steadily increase a low initial release probability might be more important than fast vesicle replenishment mechanisms. This remains true for larger T_s (Supplementary Figure 2).

In summary, these results show that a set of presynaptic STP parameters generates a gain surface G that, in principle, could be tuned to match presynaptic population activity characteristics. The optimum rate and the maximum gain are

independent of the stimulus intensity for a low signal-to-noise ratio, with facilitatory synapses yielding high gains for sparse distributions while depressing synapses avoid negative gains only with dense distributions. For low basal activity ($r_{bas} = 0.5Hz$) and short duration integration window ($T_s = 40ms$) conditions, the parameter U is the principal determinant of the optimal distribution. Furthermore, lower U yields a higher gains than lower τ_{rec} when the optimal encoding rate is kept constant.

Feedforward inhibition and heterogeneous STP

In the above we ignored the fact that presynaptic STP can be target-dependent (Reyes et al., 1998; Markram et al., 1998; Rozov et al., 2001; Sun et al., 2005; Pelkey and McBain, 2007; Bao et al., 2010; Blackman et al., 2013; Larsen and Sjöström, 2015; Éltés et al., 2017) and the spike trains coming from the same axon can be modulated by different short-term dynamics at different synapses. In the following, we describe the effects of such heterogeneity in a feedforward excitation and inhibition (FF-EI) motif (Figure 1C), an ubiquitous circuit motif across the brain (Klyachko and Stevens, 2006; Dean et al., 2009; Isaacson and Scanziani, 2011; Wilson et al., 2012; Jiang et al., 2015; Grangeray-Vilmin et al., 2018).

We extend our previous analysis to a scenario in which the presynaptic population makes synaptic contacts not only with a readout neuron, but also with the local inhibitory population which projects to the readout neuron creating the FF-EI motif. Both, the readout neuron and the inhibitory group receive the same spike trains via two different types of synapses, $s1$ and $s2$ (Figure 1C). Because the presynaptic population activity is the same for both synapses ($r_{bas} = 0.5Hz$, $T_s = 40ms$), the differences in gain (G) are governed by the STP properties of the two synapses. Figure 5A shows G for a facilitatory ($s1$, $U = 0.1$, $\tau_f = 200ms$, $\tau_{rec} = 50ms$) and a depressing ($s2$, $U = 0.7$, $\tau_f = 50ms$, $\tau_{rec} = 200ms$) synapse.

In the case of a FF-EI network, those two synapse types may be associated with the two branches, for example $s1$ to the feedforward excitation (FFE) branch (targeting a principal neuron) and $s2$ to the feedforward inhibition (FFI) branch (targeting local interneurons which eventually project to principal neurons) (Figure 5B inset). In this arrangement, the combined gain is determined by the two branches $G^{com} = G^{s1} - G^{s2}$. We found that the combined gain of the FF-EI circuit also varied non-monotonically as a function of N_{ext} and peaked at N_{opt}^{com} which corresponded to the combined optimum encoding rate r_{opt}^{com} (Figure 5B,C). Note that the combined maximum gain of the FF-EI circuit is larger than the gain obtained via the FFE branch with facilitatory synapses alone (Figure 2C). This substantial increase is a consequence of the strictly negative profile of G^{s2} . When the extra input is distributed in N_{opt}^{com} units (sparse coding), the depressing branch of the FF-EI drove the local inhibitory group with weaker strength than a scenario in which $N_{ext} = N$ (dense coding). Therefore, with sparse distribution of the input the readout neuron experienced strong excitation from the FFE branch and weak inhibition from FFI branch.

Similar to the behavior of facilitatory synapses, in the FF-EI network N_{opt}^{com} increased linearly as a function of r_{δ} , maintaining a constant optimal encoding rate r_{opt}^{comb} (Figure 2D, top). We also observed that r_{opt}^{comb} was larger than r_{opt}^{s1} ($N_{opt}^{com} < N_{opt}^{s1}$), making the isolated gain of $s1$ suboptimal. However, this can be compensated by putting $s2$ into a very negative gain region (Figure 5D bottom, red dashed line), with a sparse distribution of the inputs. We show analytically that r_{opt}^{com} and G_{max}^{com} are independent of the extra rate for a wide range of conditions (see Materials and Methods).

We extended this analysis to a large range of $\{s1, s2\}$ STP combinations by gradually changing the set of parameters $\{U, \tau_f, \tau_{rec}\}$ (Figure 5F). We found that G_{max}^{com} increased monotonically when we made the synapse $s1$ more facilitatory or when we made the synapse $s2$ more depressing. The anti-diagonal (where $s1 = s2$) marked the region of zero gain and any point above it ($s2$ more facilitatory than $s1$) resulted in $G_{max}^{com} = 0$, whereas any point below it ($s1$ more facilitatory than $s2$) resulted in $G_{max}^{com} > 0$. As expected, sparse distribution of the extra rate resulted in very high gain when $s1$ is highly facilitatory and $s2$ highly depressing.

Effects of basal activity on the FF-EI network

Next, we investigated the effects of the stationary basal activity at the combined optimal conditions of a FF-EI network. We found that the optimal rate and optimal gain both decreased as r_{bas} was increased (Figure 5G). Separation of the individual contributions of $s1$ and $s2$ branches revealed that this decrease was primarily because of a reduction in the gain of facilitatory synapses ($s1$) whereas the strong negative gain of depressing synapses ($s2$) remained approximately unaltered. This suggests that a population of facilitatory synapses will rapidly lose its activity distribution discrimination capacity as the baseline firing rate is increased, whereas a population of depressing synapses can preserve this capability even at larger basal rates.

Thus, these results show that a FF-EI network with target-dependent STP can make the discrimination of sparse inputs more robust than what could be achieved by the feedforward excitation alone. This can be achieved when the excitatory

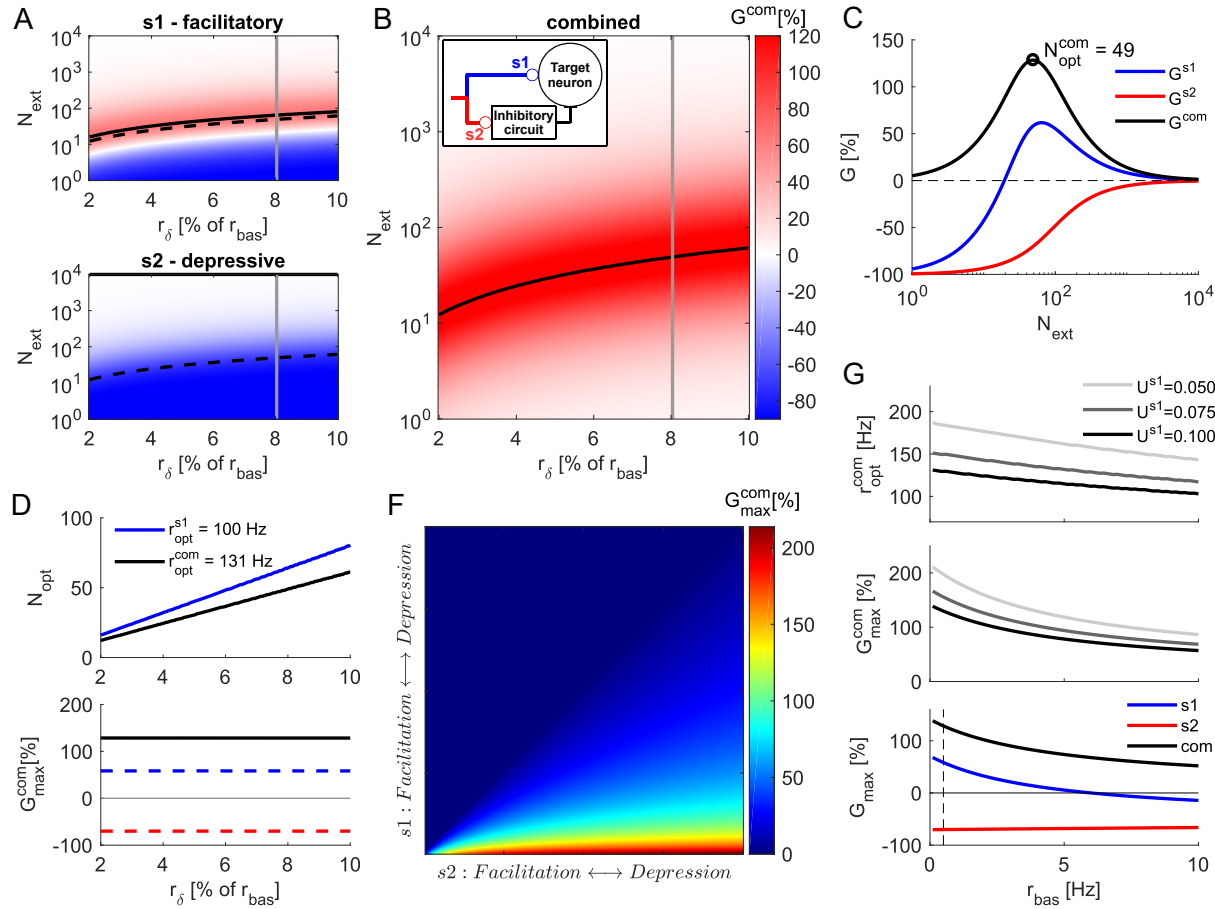


Figure 5: Combined optimal distribution of input in a feedforward excitation-inhibition circuit with target-dependent STP. **A** G as a function of r_δ and N_{ext} for a facilitatory synapse ($s1$, top) and for a depressing synapse ($s2$, bottom). This is similar to the Figure 2E. **B** The combined gain ($G^{com} = G^{s1} - G^{s2}$) of the FF-EI circuit as a function of r_δ and N_{ext} obtained by combining the gains of the feedforward excitation and inhibition branches. The black line marks the N_{opt}^{com} for every stimulus intensity r_δ and is represented with dashed black lines in panel A. In-box: schematic of the FF-EI circuit. **C** Gain as a function of N_{ext} for $r_\delta = 8\%$ of r_{bas} (gray lines in panels A and B). G^{com} inherits the non-monotonicity from G^{s1} (blue, compare with Figure 2D). The gain for a depressing synapse is negative (G^{s2} , red) for every $N_{ext} < N$. **D** top N_{opt} as a function of r_δ produces iso-frequency lines (compare with Figure 2F). r_{opt}^{com} is markedly larger than r_{opt}^{s1} . bottom G_{max}^{com} is independent of r_δ (for $r_\delta < r_{bas}$, compare with Figure 2G). Gain for both synapse types at r_{opt}^{com} (dashed lines). The small decrease in synaptic gain for $s1$ is compensated by putting $s2$ in a very negative gain region. **E** G_{max}^{com} surface for different combinations of STP characteristics of $s1$ and $s2$. Notice that G_{max}^{com} steadily increases for $s1 \rightarrow Fac$ or $s2 \rightarrow Dep$. **F** Effects of ongoing basal activity r_{bas} on optimal conditions for $T_s = 40ms$. Increasing basal activity decreases the combined optimum rate. Results for 3 different U at the facilitatory synapse (top). Increasing basal activity consistently decreases G_{max}^{com} (bottom). G_{max}^{com} decay happens mostly due to decay of the positive gain at the facilitatory synapse $s1$ (blue), while the negative gain at the depressing synapse $s2$ is kept negative and change only slightly (red). Dashed vertical line marks the basal activity used for most part of our analysis, $r_{bas} = 0.5Hz$, where both branches contribute significantly to increase the combined gain.

branch is facilitatory while the activation of the inhibitory branch is depressing (by placing $s1$ and $s2$ at the region below the anti-diagonal on Figure 5F).

Sparse code identification by a postsynaptic neuron model

The ability of STP to amplify the output of a presynaptic population would be functionally relevant only if this amplification is transferred to the postsynaptic side. We tested the postsynaptic effects of the STP based modulation of

the presynaptic activity distribution by simulating an integrate and fire (IF) neuron model (equation 17) as a readout device for a FF-EI circuit (Figure 6A). We simulate a presynaptic population with characteristics similar to the cerebellar molecular layer, a massively feedforward system with properties much alike the ones we have described so far (Ito, 2006).

Specifically, the readout neuron received input from 160,000 presynaptic neurons. The presynaptic background activity was modeled as independent and homogeneous Poisson spike trains with average firing of $r_{bas} = 0.5Hz$ ($R_{bas} = 80kHz$). In addition, the population of presynaptic neurons increased their firing rate ($R_{ext} = 1 \dots 10\%$ of R_{bas}) during a brief time window ($T_s = 40ms$) to mimic a event-related activity. The extra presynaptic activity was either confined to a small set of presynaptic neurons ($N_{ext} = N_{opt, sparse}$) or distributed over a large number of neurons ($N_{ext} = 16,000, dense$). The excitatory synapses onto the readout neuron ($s1$) were facilitatory and the STP parameters for each synapse were drawn from a Gaussian distribution ($s1, U : mean = .1, s.d. = .02, \tau_{rec} : mean = 50, s.d. = 10ms$ and $\tau_f : mean = 200, s.d. = 40ms$). The feedforward inhibitory activity was modeled as a Poisson process whose firing rate (λ_i , equation 19) was linearly dependent on the excitatory input of depressing synapses, whose STP parameters for each synapse were drawn from a Gaussian distribution ($s2, U : mean = 0.7, s.d. = .14, \tau_{rec} : mean = 200, s.d. = 40ms$, and $\tau_f : mean = 50, s.d. = 10ms$).

The distribution of the input had a noticeable effect in the output of the target neuron, as shown by the peristimulus time histogram (Figure 6B). While the dense distribution elicited transients at the beginning and ending of the stimulus period because of the inhibition slow time constant, the sparse code elicited a sustained elevated firing rate response throughout the stimulus period. The stimulus induced membrane potential responses for the two types of input patterns (dense and sparse) were also similar to the firing rate responses (Figure 6C). By interchangeably setting $s1$ and $s2$ to static, we identified that both branches contributed significantly to keep the mean membrane potential high in the presence of extra sparse input.

The contribution of each branch becomes clear at the average change in the total excitatory and inhibitory conductances of the readout neuron. When both synapses were dynamic and the stimulus was sparse (Figure 6D, *leftmost*), the average excitation was larger (because of synaptic facilitation) and the average inhibition was lower (because of synaptic depression) than the average changes caused by a stimulus of the same intensity but with dense distribution (Figure 6D, *rightmost*). Note how, with dynamic synapses and dense distribution of the stimulus, the conductance changes matched the expected change for static synapses (dashed line). When we kept the stimulus distribution sparse, but interchangeably set $s1$ and $s2$ to static, the conductance trace related to the static branch reached the same value as for the dense distribution and the system was left with the gain produced at the dynamic branch. Dense distributions, therefore, do not exploit the STP nonlinearities and the synapses behave approximately as static, as predicted.

Next, we systematically changed N_{ext} as percentages of N_{opt} ($N_{ext} = 1, 10, 25, 50, 100, 200, 400, 1000\%$ of N_{opt} , black circles in Figure 6E) and found that both the mean membrane potential and the average spike count during the stimulus period followed profiles that closely matched the predicted G^{com} curve (Figure 6E). This result confirms that the modulation of the proportion of released resources from the presynaptic population is faithfully translated into postsynaptic variables (gain estimated as the presynaptic population and membrane potential and spike rate measured on the postsynaptic neuron side). Furthermore, this result also highlights the robustness of this mechanism – even with considerable deviations from the optimum encoding distribution ($N_{ext} = 50\%$ or $N_{ext} = 200\%$ of N_{opt} , marked as the first black points at left and right from $N_{ext} = N_{opt}$), the evoked responses remained reasonably close to the optimal.

To further assess how individual realizations of the sparse input could be distinguished from a dense input of the same intensity, we sampled the output spike count of the readout neuron for a period of $40ms$ during the ongoing basal activity just before the stimulus and during the $40ms$ stimulus period for both sparse and dense distributions (Figure 6F). We used the Bhattacharyya coefficient (BC) as a measure of overlap between these sample distributions and $1 - BC$ as a measure of difference (Figure 6G). The dense input had almost complete overlap with the basal condition. On the other hand, the sparse input produced increasingly different response distributions from both the dense input and basal condition, with almost complete separation at $r_{\delta} = 10\%$ of r_{bas} .

Taken together, these results illustrate the potential role of dynamic synapses in amplification of sparse signals at the presynaptic side (Q^p, G), even when such signal intensity is just a small fraction of the ongoing basal activity and, therefore, likely to be buried in proportionally large noise fluctuations. In addition, for a dense distribution of the input, the system can preserve short periods ($\sim 10ms$) of increased (decreased) spike probability right after stimulus onset (offset) due to delayed inhibition, which is a known characteristic of FF-EI motifs and might serve as indication of global background rate changes.

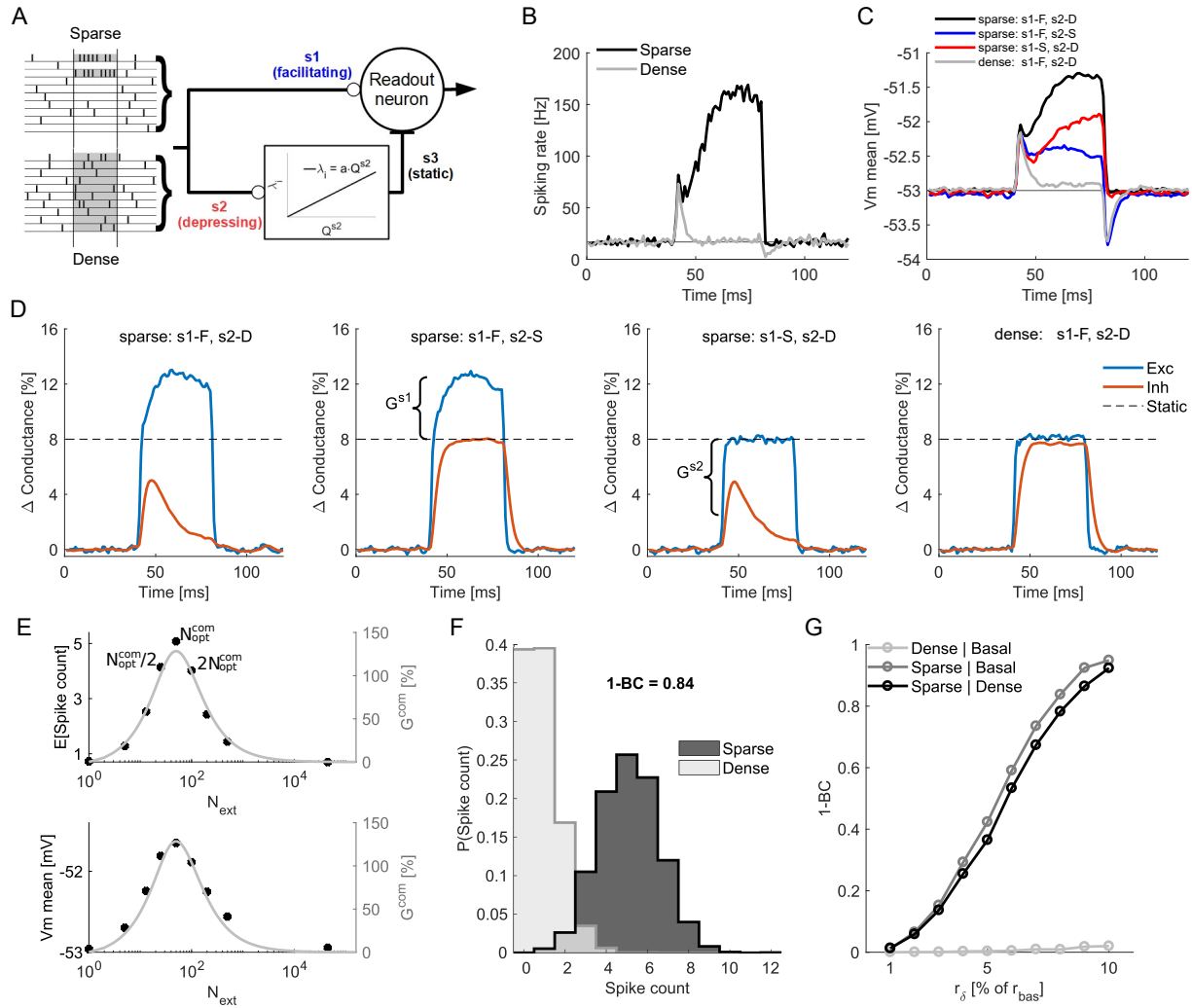


Figure 6: Transfer of STP gain from presynaptic side to postsynaptic neuron. **(A) Right** Schematic representation of a readout neuron receiving feedforward excitation from $N = 160,000$ neurons and feedforward inhibitory input. The inhibitory group was driven by the *PRR* of synapses of type s_2 according to the linear function shown in the rectangular box. **(A) Left** Sparse and dense input patterns are schematically shown. The basal rate was set to $r_{bas} = 0.5 \text{ Hz}$. The population temporally ($T_s = 40 \text{ ms}$ marked in gray) increases its firing rate by $R_{ext} = 1...10\%$ of R_{bas} in two different configurations: sparse (Top, $N_{ext} = N_{opt}$) or dense distribution (Bottom, $N_{ext} = 16,000$). For panels B, C, D and E, $R_{ext} = 8\%$ of R_{bas} . **(B)** PSTH of the spiking rate (3000 realizations) of readout neuron receiving sparse (black) or dense (gray) distribution of input activity. **(C)** Mean membrane potential for sparse input (black) and dense input (gray) when the synapse s_1 was facilitatory and synapse s_2 was depressing. Red trace: Membrane potential when the synapse s_1 was facilitatory and synapses s_2 was static. Blue trace: Membrane potential when the synapse s_1 was static and synapses s_2 was depressing. The red and blue traces show the contributions from synaptic facilitation (on the feedforward excitation branch) and depression (on the feedforward inhibitory branch) to the neuron response. **(D)** Changes in the total excitatory and inhibitory conductances for the four configurations of synapses (as show in the panel C). The dashed line marks the conductance changes for the static synapses condition. **(E)** Effect of varying N_{ext} as a proportion of N_{opt}^{com} on the expected spike count (top, black circles) and the mean membrane potential (bottom, black circles) during the event-related activity period. Both profiles match the combined gain curve (gray line, compare with Figure 5C), with peak at $N_{ext} = N_{opt}^{com}$. **(F)** Probability distribution of output spike counts within T_s . The sparse distribution increases substantially the elicited number of spikes at the readout neuron when compared to the dense distribution. **(G)** Separation (1-BC) between spike count distributions as a function of r_δ . The sparse distribution produced increasingly substantial separation when compared to basal (dark gray) and dense distribution (black) whereas the separation was always small when comparing dense distribution with basal activity (light gray).

Continuous rate distribution

Thus far, we have considered a binary distribution of the extra rate: a fraction of presynaptic cells increased their rate by r_{ext} or not at all. Although some bursting networks (e.g. cerebellar parallel fibers) do seem to operate in a quasi-binary fashion (burst or no-burst), it is important to extend the analysis to continuous distributions, which most parts of the brain seem to operate under. We do this by assuming that the distribution of event-related neural firing rates follows a Gamma distribution, which allows us parameterized control of the sparseness of the neural code (with the mean of the distribution) and of the distribution shape (with the skewness and kurtosis):

$$r_{ext} \sim \text{Gamma}(k, \theta), \quad (1)$$

where k is the shape parameter and θ is the scale parameter. For $k = 1$ it is equivalent to an exponential distribution and, for increasing values of k , it becomes a right-skewed distribution, with the skewness approaching zero for higher values of k (becoming approximately Gaussian). For each shape parameter, we controlled the mean of the distribution by varying the scale parameter, because for a Gamma distributed r_{ext} the expected value is

$$E[r_{ext}] = k\theta. \quad (2)$$

For the above specified distribution of extra rates and a given presynaptic set of STP parameters, the expected amount of resources released by a population is

$$E[Q_{ext}^p] = \int_{r_\delta}^{\infty} \frac{\text{Gamma}(r_{ext}; k, \theta)}{r_{ext}} Q^s(r_{ext}) dr_{ext}, \quad (3)$$

which we solved numerically for two synapse types ($s1$ -facilitatory and $s2$ -depressing) and a range of rate distributions (Figure 7B). The distribution gain G for $E[Q_{ext}^p]$ was then calculated in relation to the dense case, where $N_{ext} = N$ and $r_{ext} = r_\delta$ (Figure 7A).

We found that, similar to the binary distribution case, the gain for facilitatory synapses followed a non-monotonic curve as a function of θ (for a fixed k), with negative values at high θ (overly sparse distribution), a single peak at the optimal θ choice and convergence to 0 at low θ (dense distribution). By contrast, depressing synapses showed negative gains, monotonically converging to zero at low θ . The combined gain reached high values when $s1$ synapses were in very positive and $s2$ synapses were in very negative operating regions (Figure 7C).

Interestingly, not only the gain magnitudes were very similar to the ones obtained with binary distributions (compare colorbars of Figure 5A,B and Figure 7A), but also with continuously distributed rates the points of maximum gain were obtained at high mean rates (in relation to r_δ) and, therefore, representative of sparse distributions of the population activity. For increasing values of k , the skewness of these distributions approached zero (i.e. became closer to a Gaussian) and the mean r_{ext} of the optimal θ approaches the r_{opt} obtained by binary distributions. These results further highlight the relevance of the activity distribution-dependent gain modulation in presynaptic populations with STP.

Discussion

Our results suggest a close relationship between short-term synaptic plasticity and the nature of the population code that, within physiological values, can endow a postsynaptic neuron with the ability to discriminate between weak signals and background activity fluctuations of the same amplitude.

Relevance to specific brain circuits

We have shown that STP can enhance the effective input when (1) stimulus is sparse, temporally bursty and (2) feedforward excitatory synapses on the principal cells are facilitatory and feedforward excitatory synapses on local fast-spiking, inhibitory interneurons are depressing. These two conditions are fulfilled in several brain regions.

In the cerebellum, glomeruli in the granular layer actively sparsify the multimodal input from mossy fibers into relatively few simultaneously bursting parallel fibers (PF) (Billings et al., 2014) projecting to Purkinje cells (PuC). A single PuC might sample from hundreds of thousands of PFs (Tyrrell and Willshaw, 1992; Ito, 2006). In behaving animals, PF present two stereotypical activity patterns, a noisy basal state with rates lower than 1 Hz during long periods interleaved by short duration ($\sim 40\text{ ms}$), high frequency (usually $> 100\text{ Hz}$) bursts carrying sensory-motor information (Chadderton et al., 2004; Jorntell and Ekerot, 2006; van Beugen et al., 2013). Given the large number of PFs impinging on a PuC,

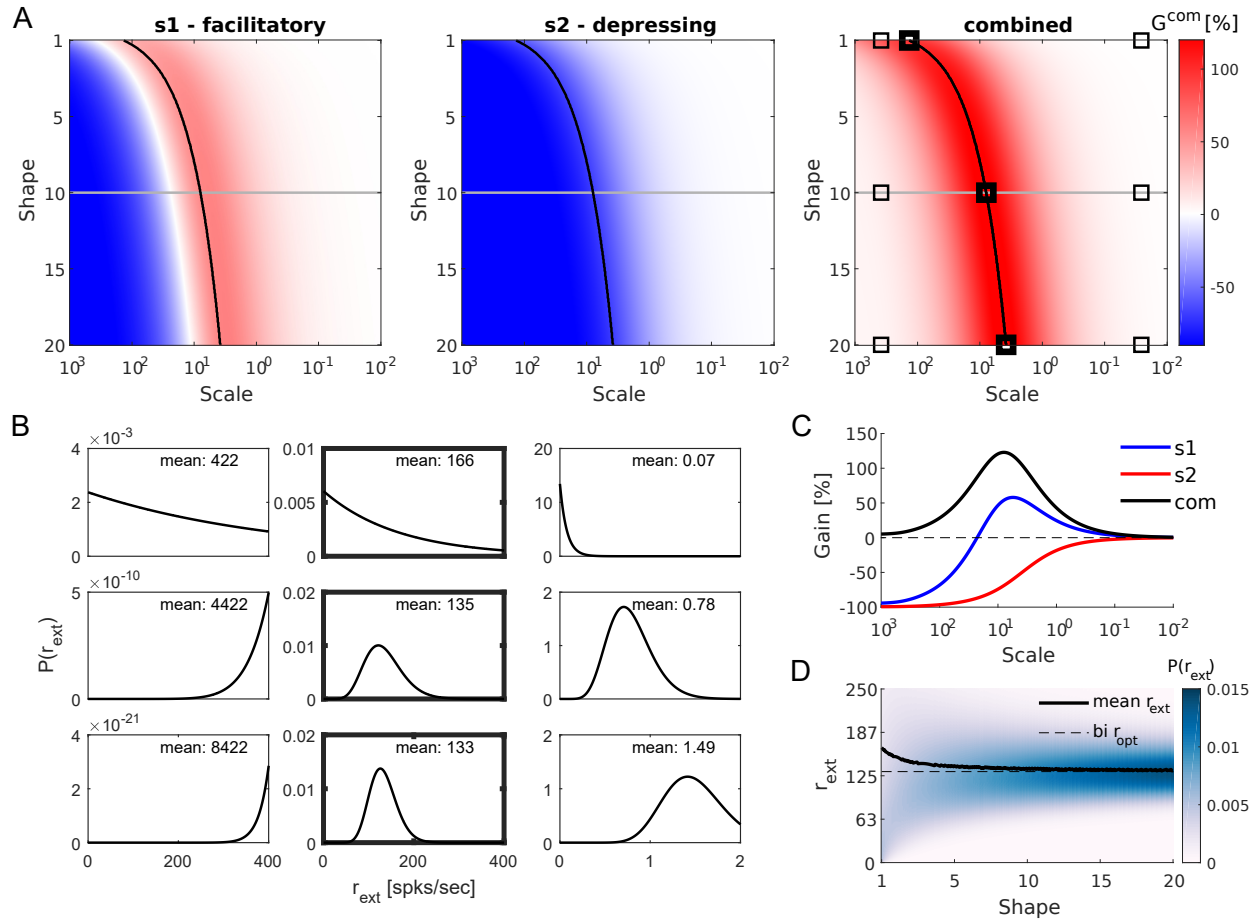


Figure 7: Synaptic gain when firing rates of individual neurons (r_{ext}) were drawn from continuous distributions (Gamma distribution). **A** Gain surfaces for a facilitatory synapse (left), for a depressing synapse (middle) and for the combined effect in a FF-EI (right). Increasing the shape parameter moved the distribution from an exponential to a right-skewed to an approximately Gaussian one. Decreasing the scale parameter moved the distribution from a high mean and high variance (sparse) to a low mean and low variance one (dense). Black lines mark the θ that resulted in maximum combined gain for each value of k . Similar to the binary distribution, the G_{max}^{com} is obtained by putting $s1$ in positive and $s2$ in negative gain regions. Note that the gain values are in the same range as in Figure 5A,B. **B** Representative examples of Gamma functions used to model the distribution r_{ext} . The nine examples correspond to the points marked in panel A. Changing k affects the shape of the distribution: exponential (top row), right-skewed (middle row) and approximately Gaussian (bottom row) shapes. Changing θ affects the scale of the distribution (sparsity of the population code): high mean and variance (left column), optimal mean and variance (middle column) and low mean and variance (right column). **C** Gain curves for a facilitatory synapse (blue), for a depressing synapse (red) and for the combined effect in a FF-EI (black). These curves were obtained for a fixed $k = 10$ (gray lines on panel A) and gradually changing θ . The gains as a function of the Gamma activity distribution follow a profile similar to the binary distribution (compare with Figure 5C). **D** Gamma-distributed r_{ext} (color plot) as a function of shape parameter. Mean r_{ext} from the Gamma distributions obtained with the optimal θ for each value of k (black lines on panel A). As the Gamma shape moves from an exponential to a Gaussian one (increasing k), the mean of the optimal distribution approaches the r_{opt} for the binary distribution.

the fluctuations in basal rate are as big as the event-related high-frequency bursts. As our analysis shows, if PF synapses were static, the PuC would not be able to discriminate between high frequency bursts and background fluctuations. However, PF synapses show short-term facilitation when targeting PuC and short-term depression when targeting Basket cells (Atluri and Regehr, 1996; Bao et al., 2010; Blackman et al., 2013). Basket cells perform strong, phasic somatic inhibition to PuCs (Jörntell et al., 2010). This circuit motif closely matches with the FFI circuit investigated in the Figure 7. Based on these similarities, we argue that one of the functional implications of the specific properties

of STP is to enable the PuC to discriminate between information encoded in high frequency bursts and background fluctuations .

In the neocortex, the population code in the layer 2/3 of the somatosensory (De Kock and Sakmann, 2008) and visual cortex of rats (Greenberg et al., 2008) and mice (Rocheffort et al., 2009) is believed to be sparse (Petersen and Crochet, 2013), with short-lived bursts (usually $< 20ms$) of high firing rates occurring over low rate spontaneous activity ($< 0.5Hz$). Additionally, it has been recently found that pyramidal cells at layer 2/3 of the mouse somatosensory cortex show short-term facilitation when targeting cells at layers 2/3 and 5 (Lefort and Petersen, 2017). These characteristics suggest that the mechanism to discriminate between weak signals and background fluctuations may also be present in the neocortex. It is believed that such sparse representation at superficial cortical layers indicates strong stimulus selectivity (Petersen and Crochet, 2013), in which case the transient gain, provided by the target-dependent STP configuration of local pyramidal neurons, would be a suitable property for inter-layer communication.

In the hippocampus, the Schaffer collaterals bringing signals from CA3 to CA1 operate under low basal firing rates with evoked bursts of high frequency activity during short periods of time (Schultz and Rolls, 1999). The synapses from pyramidal cells in CA3 to pyramidal cells in CA1 are facilitatory and provide this pathway with extra gain control (Klyachko and Stevens, 2006). Simultaneously, Schaffer collaterals synapses to CA1 stratum radiatum interneurons show larger release probability than to pyramidal neurons (Sun et al., 2005). It is, therefore, likely that this STP-based stimulus/noise discrimination mechanism is also used to improve, transmission of sequential activity from CA3 to CA1.

As we have pointed above, STP configuration in the neocortex, hippocampus and cerebellum are consistent with the configuration that enables the neural networks to take advantage of sparse coding. It is however, not a global feature of STP in the brain, and it is important to notice that facilitatory excitatory inputs to other inhibitory cells also exist in the aforementioned circuits. These facilitatory inputs mostly target interneurons that form synapses on distal dendrites (see the next paragraph). The presence of facilitatory excitatory drive to these classes of inhibitory neurons is, however, unlikely to counteract the distribution-dependent transient gains, because they produce weaker, slower and persistent dendritic inhibition. Consistent with this idea, only parvalbumin-expressing neurons (that synapse on the soma), but not somatostatin-expressing neurons (that synapse on distal dendrites), modulate stimulus response gain (Wilson et al., 2012).

The initial release probability is the most distinguishable STP parameter between Schaffer collaterals synapses onto CA1 pyramidal cells versus CA1 interneurons (Sun et al., 2005). In line with that, our approach predicts that facilitatory mechanisms that steadily increase a low initial release probability during a fast sequence of spikes (low U) will have a greater impact on the optimal OD and gain amplitude than mechanisms for fast replenishment of resources (low τ_{rec}). However, the speed of recovery has been shown to be itself an activity-dependent feature (Fuhrmann et al., 2004; Crowley et al., 2007; Valera et al., 2012; Doussau et al., 2017) and this could in principle increase the relevance of τ_{rec} .

The facilitatory or depressing nature of STP depends on the postsynaptic neuron type (Reyes et al., 1998; Markram et al., 1998; Rozov et al., 2001; Sun et al., 2005; Pelkey and McBain, 2007; Bao et al., 2010; Blackman et al., 2013; Larsen and Sjöström, 2015; Éltés et al., 2017). Target-dependent STP is a strong indication that such short living dynamics are relevant for specific types of information processing in the brain (Naud and Sprekeler, 2018). Here we predict that, when accompanied by specific arrangements of target-dependent STP found experimentally in different brain regions, disynaptic inhibition could further increase the gain of sparse over dense distributions and make it robust even at higher basal activity, when most of the gain possible at facilitatory excitation vanishes.

Disynaptic inhibition following excitation is a common motif throughout the brain, and different classes of inhibitory neurons are believed to serve distinct computations within their local circuits (Wilson et al., 2012; Jiang et al., 2015). Despite a wide diversity of inhibitory cell types, a classification of FF-I into two main types, perisomatic- and dendritic-targeting, seems to be coherent with findings throughout the central nervous system. A remarkable attribute of this configuration is the consistency of the short-term dynamics of excitatory synapses across local circuits: depressing to perisomatic- and facilitating to dendritic interneurons (Sun et al., 2005; Bao et al., 2010; Blackman et al., 2013; Éltés et al., 2017).

Disynaptic inhibition has been implicated in controlling the precision of a postsynaptic neuron's response to brief stimulation in the cerebellum (Mittmann et al., 2005; Ito, 2014) and hippocampus (Pouille and Scanziani, 2001). Additionally, the combination of disynaptic inhibition with target-dependent STP has been recently associated with the ability of networks to decode multiplexed neural signals in the cortex (Naud and Sprekeler, 2018). In line with these, our results show a bimodal profile of the readout neuron response to sparse or dense input code. We also demonstrate that, coexisting with the sustained gain during sparse code transmission, in a dense coding scenario, the system produces shorter periods ($\sim 10ms$) of increased (decreased) spike probability right after stimulus onset (offset) (Figure 6B, gray line). This results from inhibitory conductances (GABA) which are slower than the excitatory conductances (AMPA).

This very short period of firing rate modulation might work as an indication of a widespread baseline rate change in the presynaptic population.

We showed how the activity distribution of an input can exploit the nonlinearities of short-term synaptic plasticity and, with that, the theoretical potential of synaptic dynamics to endow a postsynaptic target with activity discrimination capabilities. Such mechanisms have the advantage of being in-built in synapses, not requiring further recurrent computation or any sort of supervised learning to take place. This feature is likely to be present in different brain regions, from the cerebellum to the hippocampus, and might have critical implications for general information processing in the brain.

Experimentally, these results can be tested by measuring the distribution of evoked firing rates of the neurons and STP properties of the synapses in the same brain area. Recent technological advances in stimulation systems, allowing for sub-millisecond manipulation of single and multiple cells spike activity, might soon provide means for fine control of population spike codes in intact tissues (Shemesh et al., 2017). These, together with refined methods for single cell resolution imaging of entire populations (Xu et al., 2017; Weisenburger and Vaziri, 2018) might soon allow for scrutinizing the extent of which the proposed synaptic mechanisms for distribution-dependent gain are present in neural networks.

Materials and Methods

Model of short-term plasticity

One parsimonious and yet powerful mathematical description of short-term synaptic dynamics was proposed already 20 years ago (Tsodyks and Markram, 1997). The Tsodyks-Markram (TM) model could first account for activity-dependent synaptic depression observed in pairs of neocortical pyramidal neurons and was soon extended to cover for facilitation (increase in probability) of vesicle release (Tsodyks et al., 1998). With a small set of parameters, the TM model is able to explain the opposed effects of depletion of available synaptic vesicles and of the increase in release probability caused by accumulation of residual calcium in the presynaptic terminal, making it suitable as a framework to conjecture general impact of STP in neural information processing.

Here, we use the TM model (equation 4) to describe the short-term synaptic dynamics. The effect of STD is modeled by depletion of the proportion of available resources, represented by the variable x ($0 \leq x \leq 1$), which instantaneously decreases after each spike and returns to 1 with recovery time τ_{rec} . The gain effect of short-term facilitation is modeled by the facilitation factor U ($0 \leq U \leq 1$), which accounts for the accumulation of calcium at the presynaptic terminal after the arrival of an action potential. U transiently increases the release probability u ($0 \leq u \leq 1$), which returns to 0 with time constant τ_f :

$$\begin{aligned} \frac{du^-}{dt} &= -\frac{u^-}{\tau_f} + U(1 - u^-)\delta(t - t_{sp}) \\ u^+ &= u^- + U(1 - u^-) \\ \frac{dx}{dt} &= \frac{1 - x}{\tau_{rec}} - u^+x^-\delta(t - t_{sp}) \\ \frac{dg^s}{dt} &\propto B^s u^+(t_{sp}) x^-(t_{sp}) \end{aligned} \quad (4)$$

where t_{sp} is the last spike time.

Proportion of released resources (PRR)

The change in the postsynaptic conductance g^s after a presynaptic spike is proportional to the instantaneous proportion of released resources ($PRR(t_{sp}) \propto u^+(t_{sp})x^-(t_{sp})$) and to the absolute synaptic strength B^s . The average instantaneous

437 PRR of a presynaptic unit can also be described as a function of a time-dependent Poissonian firing rate $r(t)$ (Tsodyks
438 et al., 1998) as:

$$\begin{aligned}\frac{d\langle u \rangle}{dt} &= -\frac{\langle u \rangle}{\tau_f} + U(1 - \langle u^- \rangle)r(t) \\ \langle u^+ \rangle &= \langle u^- \rangle + U(1 - \langle u^- \rangle) \\ \frac{d\langle x \rangle}{dt} &= \frac{1 - \langle x \rangle}{\tau_{rec}} - \langle u^+ \rangle \langle x^- \rangle r(t) \\ PRR^s(t) &= \langle u^+ \rangle \langle x^- \rangle r(t),\end{aligned}\tag{5}$$

439 where the brackets denote the average over many realizations.

440 Total effective input to a postsynaptic neuron

441 The total PRR contribution of a single synapse for a time window of duration T_s is then obtained by integrating
442 equation 5 over this period:

$$Q^s = \int_0^{T_s} PRR^s(t) dt\tag{6}$$

443 In the event of arrival of stimulus encoded as extra firing rate R_{ext} in a homogeneous presynaptic population of size
444 N , a number N_{ext} of units will carry the stimulus encoded as a firing rate increase by $r_{ext} = R_{ext}/N_{ext}$ while the
445 remaining units will keep their basal activity r_{bas} . The average proportion of resources released to a target neuron by
446 the entire population, during T_s , will then be

$$Q_{ext}^p = (N - N_{ext}) \cdot Q_{bas}^s + N_{ext} \cdot Q_{ext}^s\tag{7}$$

447 where Q_{bas}^s and Q_{ext}^s are the total PRR delivered by a stationary unit (firing at r_{bas}) and a stimulus encoding unit
448 (firing at $r_{bas} + r_{ext}$), respectively.

449 Gain in the effective input

450 We quantify the gain in Q_{ext}^p for a given N_{ext} relative to the Q_{δ}^p caused by an input of the same intensity but with dense
451 distribution (when $N_{ext} = N$) as

$$G = 100 \cdot \left(\frac{Q_{ext}^p - Q_{bas}^p}{Q_{\delta}^p - Q_{bas}^p} - 1 \right).\tag{8}$$

452 G is a non-monotonic function of N_{ext} with a single maximum value G_{max} at N_{opt} (see Figure 2D).

453 Optimal distribution

454 The optimal distribution of the activity (OD) is defined as the fraction of the optimal number of encoding units N_{opt} in
455 a given population of size N , that is, $OD = N_{opt}/N$. Because the optimal code, $N_{opt} = R_{ext}/r_{opt}$, is the distribution
456 that maximizes the gain over the dense distribution with the same input magnitude, $N = R_{ext}/r_{\delta}$, OD can be written
457 as

$$\begin{aligned}OD &= \frac{N_{opt}}{N} \\ &= \frac{R_{ext}}{r_{opt}} \cdot \frac{r_{\delta}}{R_{ext}} \\ &= \frac{r_{\delta}}{r_{opt}}.\end{aligned}\tag{9}$$

458 Because r_{δ} is defined as a fraction of r_{bas} and r_{opt} is fixed given the STP parameters and r_{bas} , an interesting consequence
459 is that OD becomes independent of any particular choices of N and N_{opt} . Since the optimal encoding rate is constrained

by $r_\delta < r_{opt} < \infty$, the optimal distribution will be constrained to $0 < OD < 1$ (see Figure 3D), with values close to zero or one characterizing, respectively, sparse or dense distributions.

Optimum rate (r_{opt}) and maximum gain (G_{max}) estimation

Equation 8 describes the gain G obtained by encoding a stimulus R_{ext} into N_{ext} units (with rates increased by $r_{ext} = R_{ext}/N_{ext}$) as opposed to N units (with rates increased by $r_\delta = R_{ext}/N$). The peak of this function (G_{max}) is achieved by an optimum number of encoding units N_{opt} with their rate increased by $r_{opt} = R_{ext}/N_{opt}$. This maximum point can be found by taking the derivative of the gain function with respect to r_{ext} and setting it equal to zero,

$$\frac{dG}{dr_{ext}} = \frac{d}{dr_{ext}} \left(\frac{(N - N_{ext}) \cdot Q_{bas}^s + N_{ext} \cdot Q_{ext}^s - N \cdot Q_{bas}^s}{N \cdot Q_\delta^s - N \cdot Q_{bas}^s} - 1 \right) = 0, \quad (10)$$

this can be further simplified into

$$\frac{Q_{ext}^s - Q_{bas}^s}{Q_{ext}^{s'}} - r_{ext} = 0, \quad (11)$$

where $Q_{ext}^{s'} = dQ_{ext}^s/dr_{ext}$ and r_{ext} that solves the equation is denominated r_{opt} . This solution is independent of the stimulus intensity R_{ext} and population size N , resulting in the iso-frequency line in Figure 2F.

For the optimum rate r_{opt} , the gain (equation 8) can be written as

$$G_{max} = 100 \cdot \left(\frac{r_\delta \cdot (Q_{opt}^s - Q_{bas}^s)}{r_{opt} \cdot (Q_\delta^s - Q_{bas}^s)} - 1 \right). \quad (12)$$

Assuming that Q_δ^s is linear with slope S^s for small r_δ , that is, $Q_\delta^s = Q_{bas}^s + S^s \cdot r_\delta$ (See *Linear approximation of Q_δ^s* , below), then G_{max} can be further simplified into

$$G_{max} = 100 \cdot \left(\frac{Q_{opt}^s - Q_{bas}^s}{r_{opt} \cdot S^s} - 1 \right), \quad (13)$$

which makes G_{max} independent of the stimulus intensity R_{ext} and population size N .

Combined optimum rate (r_{opt}^{com}) and maximum gain (G_{max}^{com}) estimation

When an axon branches to connect to different targets, STP properties might be target dependent. In the case of excitatory fibers driving FF-EI motifs, with $s1$ directly exciting a readout neuron and $s2$ driving the local FF-I circuit, the gain is

$$\frac{G^{com}}{100} = \frac{r_\delta}{r_{ext}} \cdot \left(\frac{Q_{ext}^{s1} - Q_{bas}^{s1}}{Q_\delta^{s1} - Q_{bas}^{s1}} - \frac{Q_{ext}^{s2} - Q_{bas}^{s2}}{Q_\delta^{s2} - Q_{bas}^{s2}} \right), \quad (14)$$

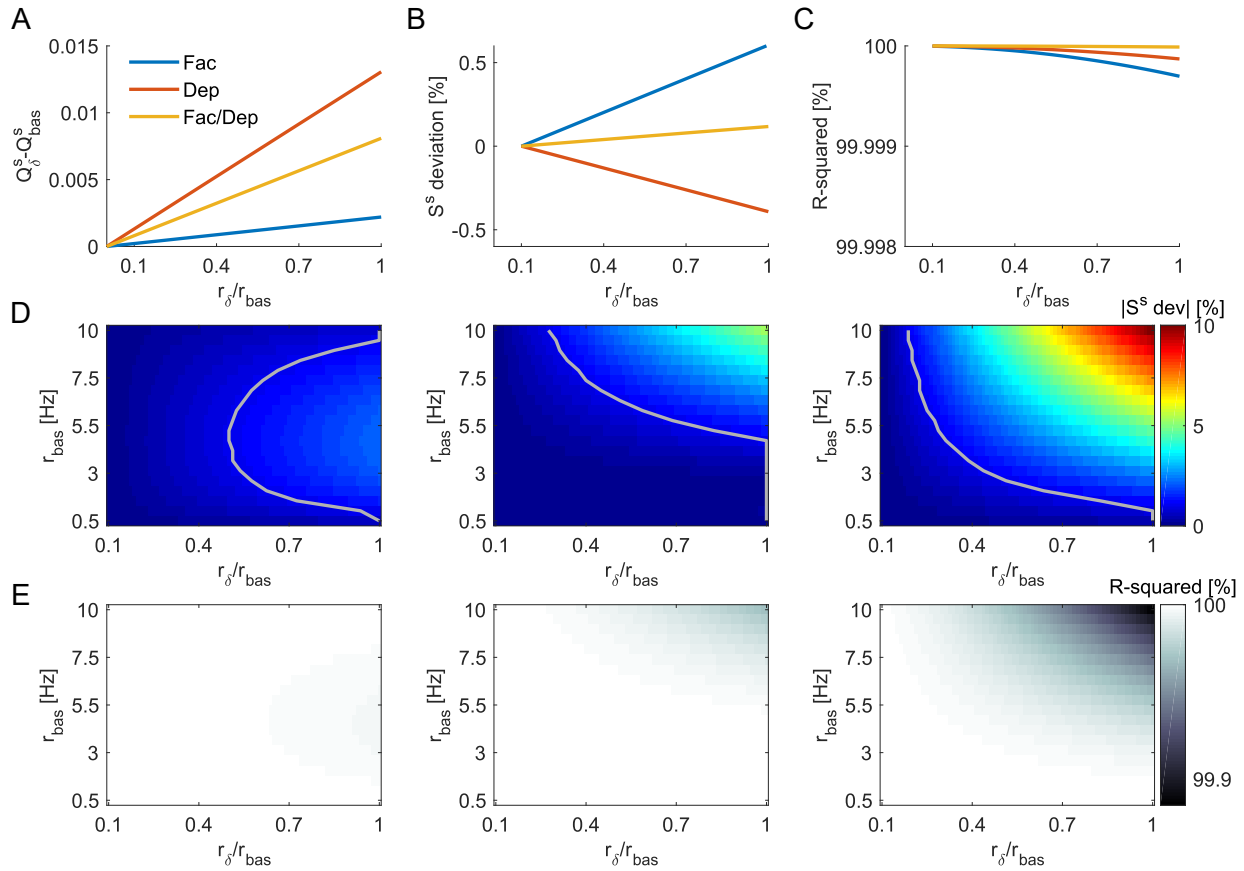
Taking the derivative of G^{com} with respect to r_{ext} , setting it equal to zero and assuming again that Q_δ^s is linear with slope S^s for both synapses, it yields

$$\frac{Q_{bas}^{s1} - Q_{ext}^{s1} + r_{ext} \cdot Q_{ext}^{s1'}}{S^{s1}} = \frac{Q_{bas}^{s2} - Q_{ext}^{s2} + r_{ext} \cdot Q_{ext}^{s2'}}{S^{s2}}, \quad (15)$$

for which the solution, $r_{ext} = r_{opt}^{com}$, is independent of the stimulus intensity R_{ext} and population size N . The optimum combined gain is then

$$G_{opt}^{com} = 100 \cdot \left(\frac{Q_{opt}^{s1} - Q_{bas}^{s1}}{r_{opt}^{com} \cdot S^{s1}} - \frac{Q_{opt}^{s2} - Q_{bas}^{s2}}{r_{opt}^{com} \cdot S^{s2}} \right), \quad (16)$$

which is also independent of the stimulus intensity R_{ext} and population size N .



Linear approximation of Q_{δ}^s

We solve Q^s numerically (equation 6) and show that it behaves linearly for a moderate range of rates in different STP regimes (Figure 8). The approximation by a linear function, $Q_{\delta}^s = Q_{bas}^s + S^s \cdot r_{\delta}$, allows G_{max} to be independent of the stimulus intensity and population size (equation 13).

To which extent is the linear approximation valid? To investigate this, we solve Q_{δ}^s for gradually increasing r_{δ} departing from a range of different basal levels $r_{bas} = 0.5 \dots 10\text{Hz}$. We then compare the slopes for each r_{δ} to the slope for $r_{\delta} = 0.1 \cdot r_{bas}$ and see how much they deviate from it (Figure 8B,D). If, for a given r_{bas} , increasing r_{δ} would result in significant change in the regressed S^s , then G_{max} would be dependent on the stimulus intensity R_{ext} . We also show the R-squared statistics to confirm the accuracy of the linear approximation (Figure 8C,E).

As we observe, for low signal-to-basal ratios ($r_{\delta}/r_{bas} < 1$), there is a wide range of rates for which the approximation is good enough, with $|S^s \text{ dev}| < 1\%$ and $R\text{-squared} > 99.9\%$. Specially for low r_{bas} , the approximation is valid for the whole length of r_{δ} .

Readout neuron model

We make our proof of concept with a simulated conductance-based integrate-and-fire (IF) neuron with its membrane voltage V_m described by

$$C_m \frac{dV_m}{dt} = g_e(V_e - V_m) + g_i(V_i - V_m) \quad (17)$$

where $C_m = 250pF$ is the membrane capacitance, g_e and g_i are respectively the excitatory and inhibitory input conductances and $V_e = 0mV$ and $V_i = -75mV$ are the excitatory and inhibitory synaptic reverse potentials. When a spike occurs, the membrane voltage is reset at $V_{reset} = -60mV$ and held at this value for a refractory period of $2ms$. The synapses were modeled by α -functions (Kuhn et al., 2004) with time constants $\tau_e = .5ms$ for excitatory and $\tau_i = 2ms$ for inhibitory synapses.

Feedforward inhibitory circuit and input

The presynaptic population consisted of $N = 160000$ units that connected to the postsynaptic neuron in a feedforward excitation-inhibition arrangement (Figure 1C). The population stationary basal rate was $R_{bas} = 80kHz$, with the individual basal rate $r_{bas} = 0.5Hz$.

At the stationary basal rate, the synaptic states are described by

$$\begin{aligned} u_{bas} &= U \frac{1 + \tau_f r_{bas}}{1 + U \tau_f r_{bas}} \\ x_{bas} &= \frac{1}{1 + u_{bas} \tau_{rec} r_{bas}} \\ PRR_{bas}^s &= u_{bas} x_{bas} r_{bas} \end{aligned} \quad (18)$$

where PRR_{bas}^s is the expected rate of proportion of released resources by each synapse with STP parameters $\{U, \tau_{rec}, \tau_f\}$.

We simulate a neuron that, during stationary basal activity, is kept in the fluctuation-driven regime through excitation-inhibition input balance (Kuhn et al., 2004). While excitation is provided directly by $s1$, disinhibitory inhibition is modulated by $s2$ in a linear fashion,

$$\lambda_i = a PRR^{s2}. \quad (19)$$

The inhibitory firing rate that keeps the target neuron membrane potential fluctuating around the mean value of $\mu(V_m)$ during stationary basal activity can be approximated by a linear function of the excitation (adapted from Kuhn et al. (2004)),

$$\begin{aligned} \lambda_i &\approx -\frac{(V_e - \mu(V_m))B_e \tau_e}{(V_i - \mu(V_m))B_i \tau_i} PRR_{bas}^{s1} \\ a &\approx -\frac{(V_e - \mu(V_m))B_e \tau_e}{(V_i - \mu(V_m))B_i \tau_i} \frac{PRR_{bas}^{s1}}{PRR_{bas}^{s2}} \end{aligned} \quad (20)$$

where B_e and B_i are the maximum amplitudes for the excitatory and inhibitory synaptic conductances. Equation 20 allows to find the linear scale of equation 19 that fulfill the condition $\mu(V_m) = -53mV$. The inhibitory synapses are kept static (no STP). The extra presynaptic activity happens in blocks of $T_s = 40ms$ and is defined as *sparse* ($N_{ext} = N_{opt}$) or *dense* ($N_{ext} = N$).

All the analyses, simulations and figures were performed in Matlab and Python. The IF model simulations were performed using Euler's method with time step of $0.1ms$ implemented in the neural simulator Brian2 (Stimberg et al., 2014).

Limitations and possible extensions

The transient enhancement or depression of synaptic efficacy by presynaptic mechanisms consists of many independent processes (Zucker and Regehr, 2002). The TM model is a tractable and intuitive way to account for these two phenomena

of interest, but this parsimony comes at the cost of biophysical simplifications. For example, it assumes the space of available resources is a continuum ($0 < x < 1$) as opposed to the known discrete nature of transmitter-carrying vesicles. However, we argue that when modeling a large number of simultaneously active synapses, the variable of interest (population *PRR*) can be approximated by a continuous variable. The nonuniform amount of transmitters per vesicle might further contribute to the validity of this assumption.

Detailed STP models that try to account for specific intracellular mechanisms were already proposed in the literature (Dittman et al., 2000), also accounting for the stochasticity of the release process (Sun et al., 2005; Kandaswamy et al., 2010). We argue that, although the quantitative results could possibly differ with more complex models, the qualitative outcome of our analysis would remain: that presynaptic short-term facilitation (depression) yields a substantial positive (negative) gain to sparse over dense population codes. Nevertheless, it would be interesting to see how the gain and optimum rate predictions would be shaped by more detailed models.

Our analyses do not account for use-dependent recovery time, changes in the readily releasable pool size (Kaesler and Regehr, 2017) or vesicles properties heterogeneity. The effects of postsynaptic receptor desensitization and neurotransmitter release inhibition by retrograde messengers (Brown et al., 2003) are likely to decrease the estimated gains by counteracting facilitation. Another interesting extension could be done with multi-compartment neuron models, which could be used to further investigate the effects of input STP heterogeneity at compartment-dependent input (Vetter et al., 2001; Grillo et al., 2018).

If the same patterns of bursts tend to happen repeatedly (e.g. parallel fibers in cerebellum during continuously repetitive movement), there might be an optimum inter-burst interval IBI^{opt} for which, if bursts happen faster than IBI^{opt} , the signal would be compromised (by a not fast enough vesicles recovery time) and if bursts happen separated by intervals longer than IBI^{opt} no extra gain will happen. Experimental evidence points to resonance at the theta-frequency oscillations range ($5.0 \sim 10.0 Hz$, $100 \sim 200 ms$ IBI) as important for cortical-cerebellar drive (Gandolfi et al., 2013; Chen et al., 2016) and hippocampus (Buzsáki, 2002). In these cases, the slower interaction between different pools of vesicles (Rizzoli and Betz, 2005) are likely to play a role in information transfer. Augmentation, a form of transient synaptic enhancement that can last for seconds, is also likely to play a role in these cases (Kandaswamy et al., 2010; Deng and Klyachko, 2011).

Acknowledgments

We would like to thank Dr. Gilad Silberberg, Dr. Erik Fransén, Dr. Philippe Isope and Martino Sindaci for helpful suggestions and feedback. This work was funded in part by: the EU Erasmus Mundus Joint Doctorate Program EUROSPIN, The International Graduate Academy (IGA) of the Freiburg Research Services (to Luiz Tauffer), Swedish Research Council (Research Project Grant, StratNeuro, India-Sweden collaboration grants to Arvind Kumar).

References

- Abbott, L. F. and Regehr, W. G. (2004). Synaptic computation. *Nature*, 431(7010):796–803.
- Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science*, 273(5283):1868–1871.
- Atluri, P. P. and Regehr, W. G. (1996). Determinants of the Time Course of Facilitation at the Granule Cell to Purkinje Cell Synapse. *The Journal of Neuroscience*, 16(18):5661–5671.
- Bao, J., Reim, K., and Sakaba, T. (2010). Target-Dependent Feedforward Inhibition Mediated by Short-Term Synaptic Plasticity in the Cerebellum. *Journal of Neuroscience*, 30(24):8171–8179.
- Billings, G., Piasini, E., Lorincz, A., Nusser, Z., and Silver, R. A. (2014). Network Structure within the Cerebellar Input Layer Enables Lossless Sparse Encoding. *Neuron*, 83(4):960–974.
- Blackman, A. V., Abrahamsson, T., Costa, R. P., Lallanne, T., and Sjöström, P. J. (2013). Target-cell-specific short-term plasticity in local circuits. *Frontiers in Synaptic Neuroscience*, 5(DEC):1–13.
- Brown, S. P., Brenowitz, S. D., and Regehr, W. G. (2003). Brief presynaptic bursts evoke synapse-specific retrograde inhibition mediated by endogenous cannabinoids. *Nature Neuroscience*, 6(10):1048–1057.
- Brunel, N. (2000). Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. *Journal of Computational Neuroscience*, 8(3):183–208.
- Buonomano, D. V. (2000). Decoding temporal information: A model based on short-term synaptic plasticity. *The Journal of neuroscience*, 20(3):1129–41.
- Buzsáki, G. (2002). Theta oscillations in the hippocampus. *Neuron*, 33(3):325–340.
- Chadderton, P., Margrie, T. W., and Häusser, M. (2004). Integration of quanta in cerebellar granule cells during sensory processing. *Nature*, 428(6985):856–860.
- Chen, H., Wang, Y. J., Yang, L., Sui, J. F., Hu, Z. A., and Hu, B. (2016). Theta synchronization between medial prefrontal cortex and cerebellum is associated with adaptive performance of associative learning behavior. *Scientific Reports*, 6:20960.
- Crowley, J. J., Carter, A. G., and Regehr, W. G. (2007). Fast Vesicle Replenishment and Rapid Recovery from Desensitization at a Single Synaptic Release Site. *Journal of Neuroscience*, 27(20):5448–5460.
- De Kock, C. P. J. and Sakmann, B. (2008). High frequency action potential bursts (≥ 100 Hz) in L2/3 and L5B thick tufted neurons in anaesthetized and awake rat primary somatosensory cortex. *The Journal of Physiology*, 586(14):3353–3364.
- Dean, P., Porrill, J., Ekerot, C.-F., and Jörntell, H. (2009). The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nature Reviews Neuroscience*, 11 VN - re(1):30–43.
- Deng, P. Y. and Klyachko, V. A. (2011). The diverse functions of short-term plasticity components in synaptic computations. *Communicative and Integrative Biology*, 4(5):543–548.
- Dittman, J. S., Kreitzer, A. C., and Regehr, W. G. (2000). Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 20(4):1374–1385.
- Doussau, F., Schmidt, H., Dorgans, K., Valera, A. M., Poulain, B., and Isope, P. (2017). Frequency-dependent mobilization of heterogeneous pools of synaptic vesicles shapes presynaptic plasticity. *eLife*, 6:e28935.
- Éltes, T., Kirizs, T., Nusser, Z., and Holderith, N. (2017). Target Cell Type-Dependent Differences in Ca(2+) Channel Function Underlie Distinct Release Probabilities at Hippocampal Glutamatergic Terminals. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 37(7):1910–1924.
- Fuhrmann, G., Cowan, A., Segev, I., Tsodyks, M., and Stricker, C. (2004). Multiple mechanisms govern the dynamics of depression at neocortical synapses of young rats. *The Journal of Physiology*, 557(2):415–438.
- Fuhrmann, G., Segev, I., Markram, H., and Tsodyks, M. (2002). Coding of Temporal Information by Activity-Dependent Synapses. *Journal of Neurophysiology*, 87(1):140–148.

- Gandolfi, D., Lombardo, P., Mapelli, J., Solinas, S., and D'Angelo, E. (2013). Theta-Frequency Resonance at the Cerebellum Input Stage Improves Spike Timing on the Millisecond Time-Scale. *Frontiers in Neural Circuits*, 7:64.
- Grangeray-Vilmin, A., Valera, A. M., Kumar, A., and Isope, P. (2018). Short-Term Plasticity Combines with Excitation-Inhibition Balance to Expand Cerebellar Purkinje Cell Dynamic Range. *The Journal of Neuroscience*, 38(22):5153–5167.
- Greenberg, D. S., Houweling, A. R., and Kerr, J. N. D. (2008). Population imaging of ongoing neuronal activity in the visual cortex of awake rats. *Nature Neuroscience*, 11(7):749–751.
- Grillo, F. W., Neves, G., Walker, A., Vizcay-Barrena, G., Fleck, R. A., Branco, T., and Burrone, J. (2018). A Distance-Dependent Distribution of Presynaptic Boutons Tunes Frequency-Dependent Dendritic Integration. *Neuron*, 99(2):275–282.
- Gütig, R. and Sompolinsky, H. (2006). The tempotron: A neuron that learns spike timing-based decisions. *Nature Neuroscience*, 9(3):420–428.
- Hawkins, J. and Ahmad, S. (2016). Why Neurons Have Thousands of Synapses, a Theory of Sequence Memory in Neocortex. *Frontiers in Neural Circuits*, 10:23.
- Isaacson, J. and Scanziani, M. (2011). How Inhibition Shapes Cortical Activity. *Neuron*, 72(2):231–243.
- Ito, M. (2006). Cerebellar circuitry as a neuronal machine. *Progress in Neurobiology*, 78(3-5):272–303.
- Ito, M. (2014). Cerebellar Microcircuitry. *Reference Module in Biomedical Sciences*.
- Jackman, S. L. and Regehr, W. G. (2017). The Mechanisms and Functions of Synaptic Facilitation. *Neuron*, 94(3):447–464.
- Jackman, S. L., Turecek, J., Belinsky, J. E., and Regehr, W. G. (2016). The calcium sensor synaptotagmin 7 is required for synaptic facilitation. *Nature*, 529(7584):88–91.
- Jiang, X., Shen, S., Cadwell, C. R., Berens, P., Sinz, F., Ecker, A. S., Patel, S., and Tolias, A. S. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex. *Science (New York, N.Y.)*, 350(6264):aac9462.
- Jörntell, H., Bengtsson, F., Schonewille, M., and De Zeeuw, C. I. (2010). Cerebellar molecular layer interneurons - computational properties and roles in learning. *Trends in Neurosciences*, 33(11):524–535.
- Jörntell, H. and Ekerot, C.-F. F. (2006). Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. *J Neurosci*, 26(45):11786–11797.
- Kaesler, P. S. and Regehr, W. G. (2017). The readily releasable pool of synaptic vesicles. *Current Opinion in Neurobiology*, 43:63–70.
- Kandaswamy, U., Deng, P.-Y., Stevens, C. F., and Klyachko, V. A. (2010). The Role of Presynaptic Dynamics in Processing of Natural Spike Trains in Hippocampal Synapses. *Journal of Neuroscience*, 30(47):15904–15914.
- Klyachko, V. A. and Stevens, C. F. (2006). Excitatory and feed-forward inhibitory hippocampal synapses work synergistically as an adaptive filter of natural spike trains. *PLoS Biology*, 4(7):1187–1200.
- Kuhn, A., Aertsen, A., and Rotter, S. (2004). Neuronal integration of synaptic input in the fluctuation-driven regime. *The Journal of neuroscience*, 24(10):2345–56.
- Kumar, A., Schrader, S., Aertsen, A., and Rotter, S. (2008). The High-Conductance State of Cortical Networks. *Neural Computation*, 20(1):1–43.
- Larsen, R. S. and Sjöström, P. J. (2015). Synapse-type-specific plasticity in local circuits. *Current Opinion in Neurobiology*, 35:127–135.
- Lefort, S. and Petersen, C. C. (2017). Layer-Dependent Short-Term Synaptic Plasticity between Excitatory Neurons in the C2 Barrel Column of Mouse Primary Somatosensory Cortex. *Cerebral Cortex*, 27(7):3869–3878.
- Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc Natl Acad Sci U S A*, 95(9):5323–8.

- Mittmann, W., Koch, U., and Häusser, M. (2005). Feed-forward inhibition shapes the spike output of cerebellar Purkinje cells. *Journal of Physiology*, 563(2):369–378.
- Naud, R. and Sprekeler, H. (2018). Sparse bursts optimize information transmission in a multiplexed neural code. *Proceedings of the National Academy of Sciences*, 115(27):E6329–E6338.
- Pelkey, K. A. and McBain, C. J. (2007). Differential regulation at functionally divergent release sites along a common axon. *Current Opinion in Neurobiology*, 17(3):366–373.
- Petersen, C. C. H. and Crochet, S. (2013). Synaptic Computation and Sensory Processing in Neocortical Layer 2/3. *Neuron*, 78(1):28–48.
- Pouille, F. and Scanziani, M. (2001). Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. *Science*, 293(5532):1159–1163.
- Reyes, A., Lujan, R., Rozov, A., Burnashev, N., Somogyi, P., and Sakmann, B. (1998). Target-cell-specific facilitation and depression in neocortical circuits. *Nature neuroscience*, 1(4):279–285.
- Rizzoli, S. O. and Betz, W. J. (2005). Synaptic vesicle pools. *Nature Reviews Neuroscience*, 6(1):57–69.
- Rocheffort, N. L., Garaschuk, O., Milos, R.-I., Narushima, M., Marandi, N., Pichler, B., Kovalchuk, Y., and Konnerth, A. (2009). Sparsification of neuronal activity in the visual cortex at eye-opening. *Proceedings of the National Academy of Sciences of the United States of America*, 106(35):15049–54.
- Rotman, Z., Deng, P.-Y. P.-Y., and Klyachko, V. a. (2011). Short-Term Plasticity Optimizes Synaptic Information Transmission. *Journal of Neuroscience*, 31(41):14800–14809.
- Rotman, Z. and Klyachko, V. a. (2013). Role of synaptic dynamics and heterogeneity in neuronal learning of temporal code. *Journal of neurophysiology*, 110(10):2275–86.
- Rozov, A., Burnashev, N., Sakmann, B., and Neher, E. (2001). Transmitter release modulation by intracellular Ca²⁺ buffers in facilitating and depressing nerve terminals of pyramidal cells in layer 2/3 of the rat neocortex indicates a target cell-specific difference in presynaptic calcium dynamics. *The Journal of physiology*, 531(Pt 3):807–26.
- Schultz, S. R. and Rolls, E. T. (1999). Analysis of information transmission in the Schaffer collaterals. *Hippocampus*, 9(5):582–598.
- Scott, P., Cowan, A. I., and Stricker, C. (2012). Quantifying impacts of short-term plasticity on neuronal information transfer. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 85(4):1–15.
- Shemesh, O. A., Tanese, D., Zampini, V., Linghu, C., Piatkevich, K., Ronzitti, E., Papagiakoumou, E., Boyden, E. S., and Emiliani, V. (2017). Temporally precise single-cell-resolution optogenetics. *Nature Neuroscience*, 20(12):1796–1806.
- Stevens, C. F. and Wang, Y. (1995). Facilitation and depression at single central synapses. *Neuron*, 14(4):795–802.
- Stimberg, M., Goodman, D. F. M., Benichoux, V., and Brette, R. (2014). Equation-oriented specification of neural models for simulations. *Frontiers in Neuroinformatics*, 8:6.
- Sun, H. Y., Lyons, S. A., and Dobrunz, L. E. (2005). Mechanisms of target-cell specific short-term plasticity at Schaffer collateral synapses onto interneurons versus pyramidal cells in juvenile rats. *Journal of Physiology*, 568(3):815–840.
- Thanawala, M. S. and Regehr, W. G. (2013). Presynaptic Calcium Influx Controls Neurotransmitter Release in Part by Regulating the Effective Size of the Readily Releasable Pool. *Journal of Neuroscience*, 33(11):4625–4633.
- Tsodyks, M., Markram, H., Pawelzik, K., and Markram, H. (1998). Neural Networks with Dynamic Synapses. *Neural Computation*, 8(1996):821–835.
- Tsodyks, M. V. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences of the United States of America*, 94(2):719–23.
- Tyrrell, T. and Willshaw, D. (1992). Cerebellar cortex: its simulation and the relevance of Marr’s theory. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 336(1277):239–57.
- Valera, A. M., Doussau, F., Poulain, B., Barbour, B., and Isope, P. (2012). Adaptation of Granule Cell to Purkinje Cell Synapses to High-Frequency Transmission. *Journal of Neuroscience*, 32(9):3267–3280.

- 691 van Beugen, B. J., Gao, Z., Boele, H.-J., Hoebeek, F., and De Zeeuw, C. I. (2013). High frequency burst firing of granule
692 cells ensures transmission at the parallel fiber to purkinje cell synapse at the cost of temporal coding. *Frontiers in*
693 *neural circuits*, 7(May):95.
- 694 Vetter, P., Roth, A., and Häusser, M. (2001). Propagation of action potentials in dendrites depends on dendritic
695 morphology. *Journal of neurophysiology*, 85(2):926–937.
- 696 Weisenburger, S. and Vaziri, A. (2018). A guide to emerging technologies for large-scale and whole brain optical
697 imaging of neuronal activity. *Annual Review of Neuroscience*, 41(1):431–452.
- 698 Wilson, N. R., Runyan, C. A., Wang, F. L., and Sur, M. (2012). Division and subtraction by distinct cortical inhibitory
699 networks in vivo. *Nature*, 488(7411):343–348.
- 700 Xu, Y., Zou, P., and Cohen, A. E. (2017). Voltage imaging with genetically encoded indicators. *Current Opinion in*
701 *Chemical Biology*, 39:1–10.
- 702 Zucker, R. S. and Regehr, W. G. (2002). Short-Term Synaptic Plasticity. *Annual Review of Physiology*, 64(1):355–405.