eNeuro

Short-term synaptic plasticity makes neurons sensitive to the distribution of presynaptic population firing rates

eNeuro Accepted Manuscript

Manuscript Title

Short-term synaptic plasticity makes neurons sensitive to the distribution of presynaptic population firing rates

Abbreviated Title

Short-term synaptic plasticity and neural code

List of Authors

Luiz Tauffer[1,2] Arvind Kumar[1,2†]
   [1] Dept. of Computational Science and Technology, School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden
[2] Bernstein Center Freiburg, University of Freiburg, Germany

Author Contributions

AK : Designed Research, Contributed unpublished reagents/ analytic tools, Analyzed data, Wrote the paper. TF : Designed Research, Performed research, Contributed unpublished reagents/ analytic tools, Analyzed data, Wrote the paper.

† Corresponding author

Arvind Kumar [arvkumar@kth.se]
Computational Science and Technology,
School of Electrical Engineering and Computer Science,
KTH Royal Institute of Technology, Stockholm
Lindstedtsvägen 5, Stockholm   11428, Sweden

Manuscript info

Number of pages: 31
Number of figures: 10
Number of extended figures: 2
Number of tables: 1
Number of words for Abstract: 164
Number of words for Introduction: 629
Number of words for Discussion: 3043
Conflict of Interest: The authors declare no competing financial interests.

Acknowledgement

Funding sources

Title

Short-term synaptic plasticity makes neurons sensitive to the distribution of presynaptic population firing rates

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 this 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.

Significance statement

What is the optimal way to distribute a fixed number of spikes over a set of neurons so the we get a maximal response in the downstream neuron? This question is at the core of neural coding. Here we show that when synapses show short-term facilitation, sparse code (when a few neurons increase their firing rate in a task-dependent manner) is more effective than dense code (when many neurons increase their firing rate in a task-dependent manner). By contrast, when synapses show short-term depression a dense code is more effective than a sparse code. Thus, for the first time, we show that the dynamics of synapses itself has an effect in deciding the most effective neural code.

1 Introduction

2 The brain is a highly noisy system. At the cellular level, the neurons are unreliable in eliciting
3 spikes and synapses are unreliable in transmitting the spikes to the postsynaptic neurons. At the
4 network level, the connectivity and balance of excitation and inhibition gives rise to fluctuations
5 in the background activity (Brunel, 2000; Kumar et al., 2008) which can be as high as the mean
6 stimulus response (Arieli et al., 1996; Kenet et al., 2003). In such a noisy environment, a neuron is
7 faced with a crucial task: how to discriminate stimulus-induced firing rate changes from fluctuations
8 in the firing rate of the background activity of the same magnitude?
9   If synapses were static (that is, when the postsynaptic conductances do not depend on the
10 immediate spike history) this task could not be accomplished, unless synapses are specifically tuned
11 to do so. For instance, the identification of specific spiking patterns, filtering out presumed noise
12 sequences, can be accomplished by precise tuning of synaptic weights (Gütig and Sompolinsky,
13 2006). This solution, however, relies on training synaptic weights using a certain supervised learn-
14 ing rule and even then it could only work for a specific set of spike timing sequences. Active
15 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 aforementioned task. 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), that is, STP arises because neurotransmitter release dynamics is history dependent. Thus, STP becomes a crucial part of neural hardware when information is encoded as firing rate. Indeed, STP has been suggested to play several important roles in neural information processing (Buonomano, 2000; Fuhrmann et al., 2002; Izhikevich et al., 2003; Abbott and Regehr, 2004; Rotman et al., 2011; Middleton et al., 2011; Scott et al., 2012; Rotman and Klyachko, 2013; Jackman and Regehr, 2017; Naud and Sprekeler, 2018; Grangeray-Vilmint et al., 2018).

An immediate 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 among different input firing rate 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 firing 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 highly negative gains in relation to dense background changes of the same magnitude. We also investigated feedforward networks with excitation and disynaptic inhibition, with target-dependent STP, and found that this arrangement allows for extra robustness for the output gain.

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 firing rate-based population code.

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 1) to describe the short-term synaptic dynamics. The effect of depression 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

62 ($0 \leq U \leq 1$), which accounts for the accumulation of calcium at the presynaptic terminal after the
63 arrival of an action potential. U transiently increases the release probability u ($0 \leq u \leq 1$), which
64 returns to 0 with time constant τ_f:

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

65 where t_{sp} is the last spike time.

Proportion of released resources (PRR)

67 The change in the postsynaptic conductance g^s after a presynaptic spike is proportional to the
68 instantaneous proportion of released resources ($PRR(t_{sp}) \propto u^+(t_{sp})x^-(t_{sp})$) and to the absolute
69 synaptic strength B^s. The average instantaneous PRR of a presynaptic unit can also be described
70 as a function of a time-dependent Poissonian firing rate $r(t)$ (Tsodyks et al., 1998) as:

$$
\begin{aligned}
\frac{d\langle u \rangle}{dt} &= -\frac{\langle u \rangle}{\tau_f} + U(1 - \langle u^- \rangle)r(t) \\
\langle u^+ \rangle &= \langle u^- \rangle + U(1 - \langle u^- \rangle) \\
\frac{d\langle x \rangle}{dt} &= \frac{1 - \langle x \rangle}{\tau_{rec}} - \langle u^+ \rangle \langle x^- \rangle r(t) \\
PRR^s(t) &= \langle u^+ \rangle \langle x^- \rangle r(t),
\end{aligned}
\tag{2}
$$

71 where the brackets denote the average over many realizations. The total PRR contribution of a
72 single synapse, for a time window of duration T_s, can then be obtained by integrating equation 2
73 over this period:

$$
Q^s = \int_0^{T_s} PRR^s(t)dt
\tag{3}
$$

Total effective input to a postsynaptic neuron

75 For a homogeneous presynaptic population with same STP parameters and individual basal firing
76 rate r_{bas}, the population basal activity is $R_{bas} = N \cdot r_{bas}$, where N is the population size. We quantify
77 R_{ext} as a multiple of R_{bas}. Our analysis is restricted to the case of low signal-to-noise ratio i.e.
78 $R_{ext} < 0.1R_{bas}$. We consider a simplified scenario where R_{ext} is distributed homogeneously through
79 a number N_{ext} of selected presynaptic units, which will increase their firing rate by $r_{ext} = R_{ext}/N_{ext}$
80 while the remaining presynaptic units will keep their activity unchanged.
81   The total PRR released to a target neuron by the entire population, during T_s, will then be

$$
Q^p_{ext} = (N - N_{ext}) \cdot Q^s_{bas} + N_{ext} \cdot Q^s_{ext}
\tag{4}
$$

82 where Q^s_{bas} and Q^s_{ext} are the total PRR (Equation 3) delivered by a stationary unit (firing at r_{bas})
83 and a stimulus encoding unit (firing at $r_{bas} + r_{ext}$), respectively.

Table 1: List of recurrent symbols.

Symbol	Description	Equation
U	Facilitation factor	1
τ_{rec}	Recovery time constant	1
τ_f	Facilitation time constant	1
r	Firing rate of a single synapse	2
PRR	Proportion of released resources	2
Q	Total PRR (integrated over a time interval)	3
T_s	Duration of the stimulus interval	3
N	Number of presynaptic neurons	4
G	Gain of any given distribution of presynaptic activity over the maximally distributed presynaptic activity ($N_{ext} = N$)	6
OD	Optimal distribution of the presynaptic activity	7
$s1$, $s2$	Synaptic type 1, synaptic type 2	12
r_{bas}	Individual basal firing rate	
r_{ext}	Individual increase in firing rate	
r_δ	Individual increase in firing rate when $N_{ext} = N$	
r_{opt}	Optimal individual increase in firing rate when $N_{ext} = N_{opt}$	
R_{bas}	Total population basal firing rate	
R_{ext}	Total population increase in firing rate	
X^s	A s superscript indicates a quantity for a single synapse	
X^p	A p superscript indicates a quantity for a population of synapses	
X_{bas}	A bas subscript indicates a quantity for basal level activity	
X_{ext}	A ext subscript indicates a quantity for extra activity	
X_δ	A $delta$ subscript indicates a quantity for the maximally distributed presynaptic activity ($N_{ext} = N$)	
X_{opt}	A opt subscript indicates a quantity for the optimally distributed presynaptic activity ($N_{ext} = N_{opt}$)	

Gain in the effective input

We are interested in the effects of varying the presynaptic distribution (over N_{ext} inputs) of this total extra rate (R_{ext}) on the effective input to postsynaptic targets. To estimate the change in the gain due to STP we used the maximally dense distribution, when $N_{ext} = N$ as the reference point:

$$Q_\delta^p = N \cdot Q_\delta^s \tag{5}$$

where the δ subscript denotes the smallest possible increase in individual firing rates, r_δ (maximally distributed R_{ext}). We refer to this as the dense distribution case and it ideally represents a situation of homogeneous increase in the basal activity of the system, against which a stimulus would need to be distinguished from. $N_{ext} = N$ also implies smallest increase in individual operating rates ($r_{ext} = r_\delta$), therefore in the dense distribution case STP nonlinearities will be minimal. In other words, $N_{ext} = N$ is the point where dynamic synapses will operate as close to static as possible.

We then quantify the gain in Q_{ext}^p for a given N_{ext} always relative to Q_δ^p caused by an input of the same intensity but with dense distribution, as

$$G = 100 \cdot \left(\frac{Q_{ext}^p - Q_{bas}^p}{Q_\delta^p - Q_{bas}^p} - 1 \right). \tag{6}$$

We calculate the curves of G as a function of N_{ext} for different sets of STP parameters and

5

97 basal rates and search for points where it is maximized ($N_{ext}=N_{opt}$), which we call the optimal
98 distribution (see example in Figure 2D).

Optimal distribution

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

$$
\begin{aligned}
OD &= \frac{N_{opt}}{N} \\
&= \frac{R_{ext}}{r_{opt}} \cdot \frac{r_\delta}{R_{ext}} \\
&= \frac{r_\delta}{r_{opt}}.
\end{aligned}
\tag{7}
$$

104 We define R_{ext} as a fraction of R_{bas} in order to keep the same signal-to-noise ratio (R_{ext}/R_{bas})
105 for populations of different sizes N. We find that r_{opt} is fixed given the STP parameters and r_{bas} (see
106 Results), therefore by defining r_δ (R_{ext}/N) as a fraction of r_{bas} (R_{bas}/N) we reach the interesting
107 consequence of OD being independent of any particular choices of population size (equation 7).
108 That is, given the same STP parameters and value of r_{bas}, populations of different sizes will
109 optimally encode the same stimulus intensity (relative to their basal activity) with the same OD.
110 Because the optimal encoding rate is constrained by $r_\delta < r_{opt} < \infty$, the optimal distribution will
111 also be constrained to $0 < OD < 1$ (see Figure 3D), with values close to zero or one characterizing
112 sparse or dense distributions, respectively.

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

114 Equation 6 describes the gain G obtained by encoding a stimulus R_{ext} into N_{ext} units (with rates
115 increased by $r_{ext} = R_{ext}/N_{ext}$) as opposed to N units (with rates increased by $r_\delta = R_{ext}/N$). The
116 peak of this function (G_{max}) is achieved by an optimal number of encoding units $N_{ext} = N_{opt}$ with
117 their rate increased by $r_{opt} = R_{ext}/N_{opt}$. This maximum point can be found by taking the derivative
118 of the gain function with respect to r_{ext} and setting it equal to zero,

$$
\frac{dG}{dr_{ext}} = \frac{d}{dr_{ext}}\left(\frac{(N - N_{ext}) \cdot Q_{bas}^s + N_{ext} \cdot Q_{ext}^s - N \cdot Q_{bas}^s}{N \cdot Q_\delta^s - N \cdot Q_{bas}^s} - 1 \right) = 0,
\tag{8}
$$

119 Given that $\frac{dQ_\delta}{dr_{ext}} = 0$, this can be further simplified into

$$
\frac{Q_{ext}^s - Q_{bas}^s}{Q_{ext}^{s\prime}} - r_{ext} = 0,
\tag{9}
$$

120 where $Q_{ext}^{s\prime} = dQ_{ext}^s/dr_{ext}$ and the value of r_{ext} that solves the equation is denominated r_{opt}. This
121 solution is independent of the stimulus intensity R_{ext} and population size N (see results in Figure
122 2F).
123 For the optimal rate r_{opt}, the gain (equation 6) can be written as

$$
G_{max} = 100 \cdot \left(\frac{r_\delta \cdot (Q_{opt}^s - Q_{bas}^s)}{r_{opt} \cdot (Q_\delta^s - Q_{bas}^s)} - 1 \right).
\tag{10}
$$

124 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*
125 *approximation of Q_δ^s*, below), then G_{max} can be further simplified into

6

$$G_{max} = 100 \cdot \left(\frac{Q_{opt}^s - Q_{bas}^s}{r_{opt} \cdot S^s} - 1 \right), \tag{11}$$

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

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

128 When an axon branches to connect to different targets, STP properties might be target dependent.
129 In the case of excitatory fibers driving feedforward excitation-inhibition (FF-EI) motifs, with synapses
130 type 1 ($s1$) directly exciting a readout neuron and synapses type 2 ($s2$) driving the local inhibitory
131 circuit (Figure 1C), the combined gain is given by

$$\frac{G^{com}}{100} = \frac{r_\delta}{r_{ext}} \cdot \left(\frac{Q_{ext}^{s1} - Q_{bas}^{s1}}{Q_\delta^{s1} - Q_{bas}^{s1}} - \frac{Q_{ext}^{s2} - Q_{bas}^{s2}}{Q_\delta^{s2} - Q_{bas}^{s2}} \right), \tag{12}$$

132 To find the activity distribution that maximizes the combined gain, we take the derivative of G^{com}
133 with respect to r_{ext}, set it equal to zero and, assuming again that Q_δ^s is linear with slope S^s for
134 both synapses, find the equivalence

$$\frac{Q_{bas}^{s1} - Q_{ext}^{s1} + r_{ext} \cdot Q_{ext}^{s1\prime}}{S^{s1}} = \frac{Q_{bas}^{s2} - Q_{ext}^{s2} + r_{ext} \cdot Q_{ext}^{s2\prime}}{S^{s2}}, \tag{13}$$

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

$$G_{opt}^{com} = 100 \cdot \left(\frac{Q_{opt}^{s1} - Q_{bas}^{s1}}{r_{opt}^{com} \cdot S^{s1}} - \frac{Q_{opt}^{s2} - Q_{bas}^{s2}}{r_{opt}^{com} \cdot S^{s2}} \right), \tag{14}$$

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

Numerical simulations

139 As a prof of concept of the potential relevance that the estimated presynaptic gains could have on
140 postsynaptic targets, we performed numerical simulations of a conductance-based integrate-and-fire
141 (I&F) neuron model acting as the readout device for a FF-EI circuit (Figure 6A). The I&F model's
142 membrane voltage V_m is described by

$$C_m \frac{dV_m}{dt} = g_e(V_e - V_m) + g_i(V_i - V_m) \tag{15}$$

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

149 The presynaptic population consisted of $N = 160000$ units that connected to the I&F neuron in
150 a feedforward excitation-inhibition arrangement (Figure 6A). The population stationary basal rate
151 was $R_{bas} = 80kHz$, with the individual basal rate $r_{bas} = 0.5Hz$. At the stationary basal rate, the
152 synaptic states are described by

$$\begin{aligned}
u_{bas} &= U\frac{1+\tau_f r_{bas}}{1+U\tau_f r_{bas}}\\
x_{bas} &= \frac{1}{1+u_{bas}\tau_{rec}r_{bas}}\\
PRR^s_{bas} &= u_{bas}x_{bas}r_{bas}
\end{aligned} \tag{16}$$

where PRR^s_{bas} 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 $s1$, disynaptic inhibition is modulated by $s2$ in a linear fashion,

$$\lambda_i = aPRR^{s2}. \tag{17}$$

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

$$\begin{aligned}
\lambda_i &\approx -\frac{(V_e-\overline{V_m})B_e\tau_e}{(V_i-\overline{V_m})B_i\tau_i}PRR^{s1}_{bas}\\
a &\approx -\frac{(V_e-\overline{V_m})B_e\tau_e}{(V_i-\overline{V_m})B_i\tau_i}\frac{PRR^{s1}_{bas}}{PRR^{s2}_{bas}}
\end{aligned} \tag{18}$$

where B_e and B_i are the maximum amplitudes for the excitatory and inhibitory synaptic conductances. Equation 18 allows to find the linear scale of equation 17 that fulfills the condition $\overline{V_m} = -53mV$. The inhibitory synapses are kept static (no STP). The extra presynaptic activity happens in blocks of $T_s = 40ms$ and is defined as *sparse* (when $N_{ext} = N_{opt}$) or *dense* (when $N_{ext} = N$).

Continuous rate distribution

Although some bursting networks (e.g. cerebellar parallel fibers) 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):

$$r_{ext} \sim Gamma(k, \theta), \tag{19}$$

where k is the shape parameter and θ is the scale parameter. When $k = 1$, this is equivalent to an exponential distribution and, for increasing values of k, this 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

$$E[r_{ext}] = k\theta. \tag{20}$$

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

$$E[Q^p_{ext}] = \int_{r_\delta}^\infty \frac{Gamma(r_{ext};k,\theta)}{r_{ext}}Q^s(r_{ext})dr_{ext}, \tag{21}$$

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

All the analyses and simulations were performed in Matlab and Python. The model simulations were performed using Euler's method with time step of $0.1ms$ implemented in the neural simulator Brian2 (Stimberg et al., 2014). The simulation and analysis code is available on GitHub: https://github.com/luiztauffer/stp-activity-distribution

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 spiking distributions with the same intensity (e.g. the same number of spikes). This problem is schematically illustrated in Figure 1. Consider two scenarios: In the first scenario 7 spikes arrive from a single presynaptic neuron while others 6 remain silent (Figure 1A, sparse distribution). In the second scenario each of the 7 presynaptic neurons spikes once. In both trial the postsynaptic neuron receives 7 spikes (Figure 1B, dense distribution). Here we test the hypothesis that when synapses exhibit short term plasticity (facilitation or depression) the two scenarios can be differentiated without any specific tuning of synaptic weights.

Static synapses evoke exactly the same postsynaptic conductance (PSC) sequence for both sparse and dense distributions (black lines), making them indistinguishable for a readout neuron. However, when synapses are dynamic, short-term facilitation (blue line) enhances the PSC amplitudes compared to the static synapses (cf. Figure 1A,B, bottom traces). Short-term depression (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 2). In this model, the instantaneous proportion of released resources (PRR) depends on the resource release probability (u^+) and the proportion of available resources (x^-), which have their dynamics guided by the choice of STP model parameters $\{U, \tau_f, \tau_{rec}\}$. For a transient increase in firing rate, a facilitatory synapse produces an average profile of sustained PRR, while a depressing synapse produces an average profile of rapid decaying PRR (Figure 2A). Throughout this work, the two reference sets of values for STP types are: $U = 0.1$, $\tau_f = 200ms$ and $\tau_{rec} = 50ms$ (facilitatory) and $U = 0.7$, $\tau_f = 50ms$ and $\tau_{rec} = 200ms$ (depressing).

To quantify the effects that different profiles will have on the presynaptic output, for varying transient increases in firing rate (r_{ext}), we calculate the total amount of extra resources (Q_{ext}^s) a synapse releases over a time period of T_s (equation 3; Figure 2B). We found that Q_{ext}^s varied in a nonlinear fashion as a function of r_{ext}, with depressing dynamics approaching Q_{ext}^s saturation much faster than facilitatory dynamics. The slope of Q_{ext}^s (Figure 2B *inset*) for depressing synapses is

225 monotonically decreasing, indicating that any increase in the firing rate in those synapses will produce
226 sublinear increase in Q_{ext}^s, whereas for facilitatory synapses the slope initially grows, indicating that
227 increases in the firing rate of those synapses, up to some point, will produce supralinear increase in
228 Q_{ext}^s.

229   In the brain, neurons typically receive inputs from a large ensemble of presynaptic neurons.
230 In the ongoing activity state, these neurons spike at a low-basal firing rate (r_{bas}) with the total
231 synaptic output of $Q_{bas}^p = N \cdot Q_{bas}^s$. In the event-related activity state, the firing rate of a subset
232 of presynaptic neurons (N_{ext}) is transiently increased and the total synaptic output (equation 4)
233 changes accordingly. We distribute a fixed event-related population rate increase R_{ext} into varied
234 numbers of chosen synapses N_{ext}, each of these chosen synapses increasing it's firing rate by r_{ext},
235 that is, $R_{ext} = N_{ext} \times r_{ext}$, and report the changes in Q_{ext}^p.

236   We found that, for a population of facilitatory synapses, Q_{ext}^p varied in a non-monotonic fashion
237 as a function of N_{ext}, initially increasing up to a peak point, then decreasing (Figure 2C left).
238 By contrast, for depressing synapses (Figure 2C right), Q_{ext}^p varied in a monotonically increasing
239 fashion. For both facilitatory and depressing synapses, Q_{ext}^p converged to their respective Q_δ^p when
240 the total extra input rate R_{ext} was distributed over all the neurons such that $N_{ext} = N$ and
241 $r_{ext} = r_\delta = R_{ext}/N$.

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

Activity distribution-dependent Gain

253 To further quantify the effect of distribution of event-related activity over the input ensemble (that
254 is, how neurons increase their rate in the event-related phase), we defined the distribution gain G as
255 the proportional change in Q_{ext}^p in relation to Q_δ^p (Equation 5 and Equation 6). We found that Q_δ^p
256 is approximately a linear function of r_δ for a wide range of scenarios (see Materials and Methods)
257 and, because of that, with the dense distribution of the activity (when all the presynaptic neurons
258 change their firing rate by a small amount r_δ in the event-related activity state), even dynamic
259 synapses behave approximately as static synapses. Therefore, G can be understood either as a gain
260 over a dense distribution or as a gain over static synapses. For facilitatory synapses, just as for Q_{ext}^p,
261 G follows a non-monotonic curve as a function of N_{ext}, with a single peak at N_{opt} (Figure 2D, blue
262 line). By contrast, depressing synapses resulted in negative gains to every distribution, except for
263 $N_{ext} = N$ where $G = 0\%$ (Figure 2D, red line).

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

271   Because the presynaptic neurons are assumed to be Poisson processes, an advantage of parametrize
272 R_{ext} in terms of fraction of R_{bas} is that it directly translates to signal-to-noise ratio. For the exam-

²⁷³ ple shown in Figure 2G, we found that STP could amplify the presynaptic output for weak signals
²⁷⁴ (which were less than 10% of the basal activity) by up to 60% if the extra rate was distributed
²⁷⁵ over N_{opt} synapses as opposed to N synapses. For low signal-to-noise ratios ($r_\delta < r_{bas}$), the gain
²⁷⁶ at the optimal distribution (G_{max}) was approximately constant and always positive for facilitatory
²⁷⁷ synapses, while depressing synapses keep $G_{max} = 0$ at $N_{opt} = N$ (Figure 2G). Finally, we ana-
²⁷⁸ lytically show that the independence of r_{opt} and G_{max} from the extra activity intensity is a good
²⁷⁹ approximation for a wide range of basal rates and STP types (see Materials and Methods).

²⁸⁰ These results suggest that when synapses are facilitatory, the input should be distributed sparsely
²⁸¹ (or sparse code, that is, only a small set of neurons change their firing rate in the event-related state)
²⁸² to maximize the total amount of synaptic resources released at the downstream neuron. By contrast,
²⁸³ when synapses are depressing, the input should be distributed densely (or dense code, that is, all
²⁸⁴ the neurons change their firing rate in the event-related state) to maximize the synaptic resources
²⁸⁵ released at the downstream neuron. Thus, for sparse population activity, facilitatory synapses get
²⁸⁶ optimally utilized and depressing synapses get subutilized.

²⁸⁷ Effects of STP parameters on optimal rate and gain

²⁸⁸ Next, we investigated how N_{opt}, r_{opt} and G_{max} vary with STP parameters. To this end we system-
²⁸⁹ atically changed synapses from facilitatory to depressing by jointly varying the set of parameters:
²⁹⁰ $U = \{0.05, ..., 0.9\}$, $\tau_{rec} = \{0.02, ..., 0.5\}$ and $\tau_f = \{0.5, ..., 0.02\}$. We found that r_{opt} decayed
²⁹¹ exponentially as the synapses became more depressing (Figure 3A). This follows from the fact that
²⁹² facilitatory synapses profit from high firing rates and depressing synapses avoid negative gains at
²⁹³ lower rates.

²⁹⁴ The maximum gain G_{max} also decreased exponentially as synapses were systematically changed
²⁹⁵ from facilitatory to depressing (Figure 3B). We found that the relationship between gain and optimal
²⁹⁶ rate was linear from mildly to strongly facilitatory synapses (Figure 3C), with larger basal rates
²⁹⁷ constraining the optimal conditions to lower rates with lower gains.

²⁹⁸ Interestingly, increasing the basal firing rate r_{bas} substantially reduced r_{opt} and G_{max}. This
²⁹⁹ is surprising because, at such low values of spiking rates, STP effects are hardly perceivable in
³⁰⁰ traditional paired-pulse ratio analyses. The high value of G_{max}, when the system operates at
³⁰¹ low r_{bas}, happens because of synapses taking advantage of the nonlinearities in their individual
³⁰² Q_{ext}^s (Figure 2B). Increased basal activity attenuates these nonlinearities, therefore impairing the
³⁰³ distribution-dependent gain.

³⁰⁴ Relationship between facilitatory synapses and sparse coding

³⁰⁵ We quantified the optimal distribution of an evoked neural signal by OD (see Materials and
³⁰⁶ Methods). High OD ($OD \to 1$) indicates a dense distribution in which many neurons spike to
³⁰⁷ encode the extra activity, whereas low OD ($OD \to 0$) indicates a sparse distribution. We found
³⁰⁸ that OD changed abruptly from sparse to dense as synapses were changed from facilitatory to
³⁰⁹ depressing (Figure 3D). Facilitatory synapses yielded maximum response for sparse while depressing
³¹⁰ synapses yielded maximum response (avoid negative gains) for dense distributions. The transition
³¹¹ point from sparse to dense OD did not depend on the stimulus duration. However, the basal rate
³¹² strongly modified the transition point, with higher r_{bas} allowing only strongly facilitatory synapses
³¹³ to take advantage of sparse distributions. This configuration remained independent of the stimulus
³¹⁴ intensity as long as the circuit operates at low signal-to-noise conditions ($r_\delta/r_{bas} < 1$, Figure 2G).

³¹⁵ In the above, we changed the synapses from facilitatory to depressing by linearly modifying
³¹⁶ the whole set of parameter together. Next, we systematically varied each of the STP parameters
³¹⁷ independently and measured the OD for maximum gain. We found that the transition region was
³¹⁸ primarily governed by the facilitation factor U ($U \sim 0.45$), with a weak dependence on τ_{rec} and τ_f

₃₁₉ (Figure 3E). The relative contribution of τ_{rec} and τ_f became more relevant at higher T_s (3-1 Supp.
₃₂₀ Fig).

₃₂₁ These results clearly highlight the importance of the stationary basal rate in how well the synaptic
₃₂₂ gain modulation operates, as only low r_{bas} allows for significant gains. Importantly, the switch-like
₃₂₃ behavior of the optimal distribution indicates that, for a given population code, there is a robust
₃₂₄ range of STP attributes that could produce positive gains. This transition point seems to be
₃₂₅ relatively independent of the signal duration but is strongly affected by r_{bas}. Finally, having a low
₃₂₆ initial release probability (defined in the model by a low U) seems to be the preeminent feature in
₃₂₇ defining the optimal OD.

₃₂₈ The equation 7 suggests that OD is independent of the population size (N). However, there is
₃₂₉ a lower limit of N below which the sparsity argument does not hold. We have shown that given
₃₃₀ the STP parameters, there is an optimum firing rate r_{opt} at which signal carrying neurons should
₃₃₁ operate to maximize the gain (Figure 2F). For a given rate it is optimal to distribute spikes over N_{opt}
₃₃₂ input channels (e.g. Figure 2F). However, when $N = N_{opt}$, then clearly the optimal distribution
₃₃₃ will not be sparse. The argument for sparseness arises when $N >> N_{opt}$. When $N >> N_{opt}$,
₃₃₄ and we increase N while keeping all other parameters constant, OD will decrease. However, if we
₃₃₅ change N while keeping all other parameters constant, the signal-to-noise ratio will change. The
₃₃₆ signal-to-noise ratio is defined as R_{ext}/R_{bas}, where $R_{bas} = N \times r_{bas}$ and if we change N, R_{bas}
₃₃₇ will also change. In order to maintain the signal-to-noise ratios comparable for low and high N
₃₃₈ scenarios, we need to scale R_{ext} accordingly. Therefore, here we defined R_{ext} in proportion to R_{bas}
₃₃₉ so that it can accommodate the changes in N. With this choice OD is indeed independent of N
₃₄₀ (see equation 7).

₃₄₁ Effects of different sources of enhancement on G_{max}

₃₄₂ The enhancement of the output at facilitatory synapses could, in principle, have many causes
₃₄₃ (Valera et al., 2012; Thanawala and Regehr, 2013; Jackman and Regehr, 2017). Using the TM
₃₄₄ model (equation 1), we phenomenologically accounted for two important sources: a low initial
₃₄₅ release probability which sequentially increases with each incoming spike (Jackman et al., 2016)
₃₄₆ and fast replenishment of readily available resources (Crowley et al., 2007). The first characteristic
₃₄₇ is mimicked by a low facilitation factor U, which determines the initial release probability after a
₃₄₈ long quiescent period and the proportional increase in it after each spike. The second mechanism
₃₄₉ is captured by a fast recovery time constant τ_{rec}.

₃₅₀ We systematically varied U and τ_{rec} and measured G_{max} and r_{opt}. We found that several different
₃₅₁ combinations of U and τ_{rec} resulted in the same optimal distribution gain and rate. However, when
₃₅₂ we changed U and τ_{rec} while keeping the r_{opt} fixed, G_{max} could no longer be kept constant and vice
₃₅₃ versa. For instance, the two parameter sets $\{U = 0.05, \tau_{rec} = 90ms\}$ and $\{U = 0.1, \tau_{rec} = 15ms\}$
₃₅₄ gave $r_{opt} = 150Hz$ (Figure 4A), but the first parameter set gave $G_{max} = 109\%$ and the second
₃₅₅ parameter set gave $G_{max} = 92\%$ (Figure 4B). Holding U fixed and choosing τ_{rec} to match with
₃₅₆ different r_{opt} showed that G_{max} consistently dropped for higher U (Figure 4C,D).

₃₅₇ These results indicate that, in terms of maximum gain G_{max}, the fine tuning of intracellular
₃₅₈ mechanisms that work to steadily increase a low initial release probability might be more important
₃₅₉ than fast vesicle replenishment mechanisms. This remains true for larger T_s (4-1 Supp. Fig).

₃₆₀ In summary, our results show that a set of presynaptic STP parameters generates a gain surface
₃₆₁ G that, in principle, could be tuned to match presynaptic population activity characteristics. The
₃₆₂ optimal rate and the maximum gain are independent of the stimulus intensity for a low signal-to-
₃₆₃ noise ratio, with facilitatory synapses yielding high gains for sparse distributions while depressing
₃₆₄ synapses avoid negative gains only with dense distributions. For low basal activity ($r_{bas} = 0.5Hz$)
₃₆₅ and short duration integration window ($T_s = 40ms$) conditions, the parameter U is the principal
₃₆₆ determinant of the optimal distribution. Furthermore, lower U yields a higher gains than lower τ_{rec}

367   when the optimal encoding rate is kept constant.

Feedforward inhibition and heterogeneous STP

369   In the above we ignored the fact that presynaptic STP can be target-dependent (Reyes et al.,
370   1998; Markram et al., 1998; Rozov et al., 2001; Sun et al., 2005; Pelkey and McBain, 2007; Bao
371   et al., 2010; Blackman et al., 2013; Larsen and Sjöström, 2015; Éltes et al., 2017) and the spike
372   trains coming from the same axon can be modulated by different short-term dynamics at different
373   synapses. In the following, we describe the effects of such heterogeneity in a feedforward excitation
374   and inhibition (FF-EI) motif (Figure 1C), an ubiquitous circuit motif across the brain (Klyachko and
375   Stevens, 2006; Dean et al., 2009; Isaacson and Scanziani, 2011; Wilson et al., 2012; Jiang et al.,
376   2015; Grangeray-Vilmint et al., 2018).
377   We extend our previous analysis to a scenario in which the presynaptic population makes synaptic
378   contacts not only with a readout neuron, but also with the local inhibitory population which projects
379   to the readout neuron creating the FF-EI motif. Both, the readout neuron and the inhibitory group
380   receive the same spike trains via two different types of synapses, $s1$ and $s2$ (Figure 1C). Because
381   the presynaptic population activity is the same for both synapses ($r_{bas} = 0.5Hz$, $T_s = 40ms$), the
382   differences in gain (G) are governed by the STP properties of the two synapses. Figure 5A shows
383   G for a facilitatory ($s1$, $U = 0.1$, $\tau_f = 200ms$, $\tau_{rec} = 50ms$) and a depressing ($s2$, $U = 0.7$,
384    $\tau_f = 50ms$, $\tau_{rec} = 200ms$) synapse.
385   In the case of a FF-EI network, those two synapse types may be associated with the two branches,
386   for example $s1$ to the feedforward excitation (FFE) branch (targeting a principal neuron) and $s2$
387   to the feedforward inhibition (FFI) branch (targeting local interneurons which eventually project to
388   principal neurons) (Figure 5B inset). In this arrangement, the combined gain is determined by the
389   two branches $G^{com} = G^{s1} - G^{s2}$. We found that the combined gain of the FF-EI circuit also varied
390   non-monotonically as a function of N_{ext} and peaked at N_{opt}^{com} which corresponded to the combined
391   optimal encoding rate r_{opt}^{com} (Figure 5B,C). Note that the combined maximum gain of the FF-EI
392   circuit is larger than the gain obtained via the FFE branch with facilitatory synapses alone (Figure
393   2C). This substantial increase is a consequence of the strictly negative profile of G^{s2}. When the
394   extra input is distributed in N_{opt}^{com} units (sparse coding), the depressing branch of the FF-EI drove
395   the local inhibitory group with weaker strength than a scenario in which $N_{ext} = N$ (dense coding).
396   Therefore, with sparse distribution of the input, the readout neuron experienced stronger excitation
397   from the FFE branch and weaker inhibition from FFI branch.
398   Similar to the behavior of facilitatory synapses, in the FF-EI network N_{opt}^{com} increased linearly as
399   a function of r_δ, maintaining a constant optimal encoding rate r_{opt}^{comb} (Figure 2D, *top*). We also
400   observed that r_{opt}^{comb} was larger than r_{opt}^{s1} ($N_{opt}^{com} < N_{opt}^{s1}$), making the isolated gain of $s1$ suboptimal.
401   However, this can be compensated by putting $s2$ into a very negative gain region (Figure 5D *bottom*,
402   red dashed line), with a sparse distribution of the inputs. We show analytically that r_{opt}^{com} and G_{max}^{com}
403   are independent of the extra rate for a wide range of conditions (see Materials and Methods).
404   We extended this analysis to a large range of $\{s1, s2\}$ STP combinations by gradually changing
405   the set of parameters $\{U, \tau_f, \tau_{rec}\}$ (Figure 5E). We found that G_{max}^{com} increased monotonically when
406   we made the synapse $s1$ more facilitatory or when we made the synapse $s2$ more depressing. The
407   anti-diagonal (where $s1 = s2$) marked the region of zero gain and any point above it ($s2$ more
408   facilitatory than $s1$) resulted in $G_{max}^{com} = 0$, whereas any point below it ($s1$ more facilitatory than $s2$)
409   resulted in $G_{max}^{com} > 0$. As expected, if $s1$ is highly facilitatory and $s2$ highly depressing the combined
410   effect will be of very high gains, given that the presynaptic activity is optimally distributed.

13

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 5F). 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 lose most of its activity distribution-dependent gain as the basal 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 activity more robust than what could be achieved by the feedforward excitation alone. This can be achieved when the excitatory 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 5E).

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 (I&F) neuron model (equation 15) 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 ($R_{ext} = 1 \ldots 10\%$ of R_{bas}) during a brief time window ($T_s = 40ms$) to mimic an 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 9) 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 $\tau_f : mean = 50, s.d. = 10ms$). Maximum weights of each excitatory and inhibitory were drawn from Gaussian distributions ($Be : mean = 25, s.d. = 2.5\ nS$, $Bi : mean = 2., s.d. = .2\ nS$).

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

458 stimulus was sparse (Figure 6D, *leftmost*), the average excitation was larger (because of synaptic
459 facilitation) and the average inhibition was lower (because of synaptic depression) than the average
460 changes caused by a stimulus of the same intensity but with dense distribution (Figure 6D, *right-*
461 *most*). Note how, with dynamic synapses and dense distribution of the stimulus, the conductance
462 changes matched the expected change for static synapses (dashed line). When we kept the stimulus
463 distribution sparse, but interchangeably set $s1$ and $s2$ to static, the conductance trace related to
464 the static branch reached the same value as for the dense distribution and the system was left with
465 the gain produced at the dynamic branch. Dense distributions, therefore, do not exploit the STP
466 nonlinearities and the synapses behave approximately as static, as predicted.

467   Next, we systematically changed N_{ext} as percentages of N_{opt} ($N_{ext} =$1, 10, 25, 50, 100, 200,
468 400, 1000% of N_{opt}, black circles in Figure 6E) and found that both the mean membrane potential
469 and the average spike count during the stimulus period followed profiles that closely matched the
470 predicted G^{com} curve (Figure 6E). This result confirms that the modulation of the proportion of
471 released resources from the presynaptic population is faithfully translated into postsynaptic variables
472 (gain estimated at the presynaptic side and membrane potential and spike rate measured on the
473 postsynaptic neuron side). Furthermore, this result also highlights the robustness of this mechanism
474 – even with considerable deviations from the optimal encoding distribution ($N_{ext} = 50\%$ or $N_{ext} =$
475 200% of N_{opt}, marked as the first black points at left and right from $N_{ext} = N_{opt}$), the evoked
476 responses remained reasonably close to the optimal.

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

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

Continuous extra rate distribution

494 Thus far, we have considered a binary distribution of the extra rate: a fraction of presynaptic cells
495 increased their rate by r_{ext} or not at all. Although some neural networks might roughly operate in this
496 binary fashion, it is important to ask how would such STP-driven gains operate under continuous
497 distributions, a perhaps more comprehensive way of describing the activity distribution of many
498 neural populations. We therefore estimate the optimal conditions for when the extra presynaptic
499 activity follows a Gamma distribution (Equation 19).

500   The variation of the shape parameter ($1 < k < 20$) changes the distribution from an exponential
501 to a quasi-Gaussian. For each fixed shape, we control the mean (therefore the sparsity, Equation
502 20) of the distribution with the scale parameter ($10^{-2} < \theta < 10^3$). For each set $\{k, \theta\}$ we calculate
503 the expected gain (Equation 21) yielded by a population of facilitatory synapses ($s1$, 7A left), of
504 depressing synapses ($s2$, Figure 7A center) and the combined gain (Figure 7A right). Nine particular
505 parameters choices are demonstrated in Figure 7B, where the central panels follow the choices that

506 maximize the combined gain.

507   We found that, similar to the binary distribution case, the gain for facilitatory synapses followed
508 a non-monotonic curve as a function of θ (for a fixed k), with negative values at high θ (overly
509 sparse distribution), a single peak at the optimal θ choice and convergence to 0 at low θ (dense
510 distribution). By contrast, depressing synapses showed negative gains, monotonically converging to
511 zero at low θ. The combined gain reached high values when $s1$ synapses were in very positive and
512 $s2$ synapses were in very negative operating regions (Figure 7C, gray line at Figure 7A).

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

Continuous basal rate distribution

522 In the preceding analysis we assumption that r_{bas} is fixed and the same for all presynaptic units. A
523 more natural scenario, however, would be to consider a continuous distribution of basal firing rates.
524 We extend our analysis to account for this continuous r_{bas} scenario in a similar way to what we did
525 for r_{ext}: we modelled the distribution of basal rates with a Gamma distribution, with varying shape
526 and scale parameters. The variation of the shape parameter ($1 < k < 20$) changed the distribution
527 from an exponential to a quasi-Gaussian (Figure 8A). For each fixed shape, we controlled the mean
528 of the distribution (Equation 20) with the scale parameter.

529   We calculated the expected values for Q_{bas}^s, Q_δ^s and Q_{opt}^s for each distribution shape (Figure
530 8B) and found that these values converged to the values estimated with a fixed r_{bas} for higher
531 k (quasi-Gaussian) and diverged for lower k (exponential). Using these traces, we calculated the
532 gains (Figure 8C) and again found that the results converged to the estimated values for fixed r_{bas}.
533 For higher r_{bas}, however, the differences between fixed and continuously distributed r_{bas} were more
534 pronounced.

535   The divergence we observed can be explained by the probabilities that any chosen unit will
536 have a rate below or above the mean r_{bas} value. As discussed previously, higher r_{bas} will hinder
537 the exploitation of STP nonlinearities and, therefore, reduce the possibility of higher gains. For
538 exponential-like distributions, a higher proportion of the population has $r_{bas} < \overline{r_{bas}}$, which diminishes
539 this hindering effect, even if a smaller part of the population gets extra impaired by being at
540 $r_{bas} > \overline{r_{bas}}$. As the distribution gets closer to a Gaussian (increase in shape parameter), the
541 proportions of the population with r_{bas} below or above the $\overline{r_{bas}}$ become almost equal. In the limit
542 of $k \to \infty$, the variance of the Gamma distribution will approximate zero (for our fixed mean) and
543 the gains will converge to the values estimated with fixed r_{bas}.

544   It is worth noting that, for higher $\overline{r_{bas}}$, the spike rate variances are also higher and the estimates
545 of gains with fixed r_{bas} become less accurate. This means that our simplified predictions of the
546 hindering impact that higher r_{bas} have on the distribution-dependent gains will likely be an overes-
547 timate of the actual effects in real neuron populations. In other words, the distribution-dependent
548 gains in facilitatory populations can be more resilient to higher r_{bas} than what is predicted by a
549 fixed r_{bas} models.

From presynaptic gains to postsynaptic rate changes

The readout neuron in our simulations operates in a regime where the presynaptic gains are reliably translated into readout firing rate gains, which is equivalent to saying that the postsynaptic transfer function is independent of the input distribution. However, both synapses and readout neuron dendrites/soma can operate in a non-linear regime and further transform the presynaptic gain described above. These non-linearities reflect in the transfer function of the neuron i.e. the probability of an output spike given a certain input.

To identify in which circumstances changes in post-synaptic transfer function may affect the transfer of presynaptic gains into output firing rate, let us consider a neuron with two possible transfer functions (Figure 9, blue and red curves). The transfer function TF-1 is similar to the one we have considered previously in (Figure 6). The TF-2 shows a sharp change. Such a sharp change in the transfer function may arise, for example, because of NMDA receptors. When input is strong, postsynaptic depolarization can remove the Mg^{2+} block and create a larger EPSP and increase the spike probability (Du et al., 2017). Similarly, non-linear local dendritic integration (Polsky et al., 2004), input correlations (de la Rocha and Parga, 2005) and voltage dependent ion channels may also create input-dependent changes in the neuron transfer function. When the neuron transfer function can change between TF-1 and TF-2, the output firing rate is not only determined by the effective input (sparse > dense, for s1 - facilitatory) but also by the qualitative differences in transfer functions.

Sparse input distributions will allocate extra incoming spikes as bursts, which could potentially cause extra accumulation of neurotransmitters (for s1 - facilitatory) in specific dendritic sites, triggering supralinear integration (TF-2, red curve). If a dense input distribution does not attain the triggering of TF-2 and instead keeps operating under TF-1, the difference between presynaptic gains of sparse and dense distributions will be further increased (see the difference between points 1-4 Figure 9).

In cases where both input distributions operate under the same TF the presynaptic gains will be reliably transferred into output rates (compare points 1 and 2 for TF-1 and points 3 and 4 for TF-2 in Figure 9). Finally, when a dense distribution of inputs makes the output neuron operate under TF-2 and a sparse distribution brings the neuron to operate under TF-1, the presynaptic gains could potentially be overcome (compare points 2 and 3 Figure 9).

Linear approximation of Q_δ^s

We solve Q^s numerically (equation 3) and show that it behaves linearly for a moderate range of rates in different STP regimes (Figure 10). The approximation by a linear function, $Q_\delta^s = Q_{bas}^s + S^s \cdot r_\delta$, allows G_{max} and r_{opt} to be independent of the stimulus intensity and population size (Equation 13 and Equation 14).

To which extent is the linear approximation valid? To investigate this, we solve Q_δ^s for gradually increasing r_δ (Figure 10A) departing from a range of different basal levels $r_{bas} = 0.5...10$ 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 10B,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_δ. We also show the R-squared statistics to confirm the accuracy of the linear approximation (Figure 10C,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 $|S^s$ dev$| < 1\%$ and R-squared$> 99.9\%$. Specially for low r_{bas}, the approximation is valid for the whole length of r_δ.

Discussion

Our results suggest how the activity distribution of a presynaptic population can exploit the non-linearities of short-term synaptic plasticity and, with that, the theoretical potential of synaptic dynamics to endow a postsynaptic target with the ability to discriminate between weak signals and background activity fluctuations of the same amplitude. 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.

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. Sparse coding provides many advantages for neural representations (Babadi and Sompolinsky, 2014) and associative learning (Litwin-Kumar et al., 2017). As discussed in the following, a number of experimental studies provide support for sparse coding in several brain regions such as the neocortex, cerebellum and hippocampus.

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 ($\sim 40\,ms$), high frequency (usually $> 100Hz$) 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 to a PuC, 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; Grangeray-Vilmint et al., 2018). Basket cells perform strong, phasic somatic inhibition to PuCs (Jörntell et al., 2010). This circuit motif closely matches the FF-EI circuit investigated in this work (e.g. Figure 6). 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 activity 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 (Rochefort 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). The receptive field properties in the visual cortex are also consistent with the sprase code (Olshausen and Field, 1996). 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). Therefore, it is likely that this STP-based stimulus/noise discrimination mechanism is also used to improve the 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. However, 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. 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; Éltes 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 (Middleton et al., 2011; 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 the gain at facilitatory excitation decreases substantially.

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; Éltes 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)

691 which are slower than the excitatory conductances (AMPA). This very short period of firing rate
692 modulation might work as an indication of a widespread basal rate change in the presynaptic
693 population.

Relationship with previous work

695 Historically, STP has been prominently explored as a frequency filter which renders an individual
696 neuron as a low-pass filter (when synapses are depressing) or high pass filter (when synapses are
697 facilitatory) (Abbott and Regehr, 2004; Markram et al., 1998; Dittman et al., 2000). It has been
698 suggested that under some conditions STD can also interact with subthreshold oscillation to mod-
699 ulate the gain of the neurons (Latorre et al., 2016). With STP the synaptic strength depends on
700 recent history of the incoming spikes in a particular synapse. This automatically makes the down-
701 stream neurons more sensitive to transient fluctuations in input spike trains. In most of the previous
702 work this specific property has been exploited for neural coding.

703 For instance, history dependence of STP means that the effect of serial correlations (that can
704 be seen in the autocorrelogram of spike trains) and spike bursts in the presynaptic activity depends
705 on whether the synapses express STF or STD. Synapses with STD reduce redundancy in the input
706 spike train by reducing the PSPs of spikes that appear with a certain serial correlation or periodicity
707 (Goldman et al., 2002). By contrast, when synapses express STF, they enhance the effect of
708 serial correlations or spike bursts and the readout neuron can function as a burst detector (Lisman,
709 1997). In fact, both STF and STD can be combined to de-multiplex spike bursts from single spikes
710 (Izhikevich et al., 2003; Middleton et al., 2011; Naud and Sprekeler, 2018). Thus, much emphasis
711 has been put on understanding how STP can be used to extract information encoded in the pattern
712 of spikes of a single input neuron.

713 Here we extend this line of work and show how STP may affect the impact of a neuron ensemble
714 on downstream neurons. Previous work has suggested that STP makes a neuron sensitive to
715 transient rate changes. Given this property, when synapses show STD, input correlations can still
716 modulate the neuron output for a wide range of firing rates (de la Rocha and Parga, 2005). Our
717 work reveals a new consequence of the same effect as we show that STP renders the neurons in the
718 brain with input distribution dependent gain, through which sparse-bursty codes could have stronger
719 downstream impact than dense codes with same intensity. Furthermore, we investigate the relative
720 importance of different STP parameters and baseline firing rates for these gains. This novel feature
721 could be a highly valuable asset in low signal-to-noise ratio conditions. Moreover, Our results also
722 show how synapses can impose further constrains on the neural code

Experimental verification of model predictions

724 Experimentally, these results can be tested by measuring the distribution of evoked firing rates
725 of the neurons and STP properties of the synapses in the same brain area. Recent technological
726 advances in stimulation systems, allowing for sub-millisecond manipulation of single and multiple
727 cells spike activity, might soon provide means for fine control of population spike codes in intact
728 tissues (Shemesh et al., 2017). These, together with refined methods for single cell resolution
729 imaging of entire populations (Xu et al., 2017; Weisenburger and Vaziri, 2018) might soon allow for
730 scrutinizing the extent of which the proposed synaptic mechanisms for distribution-dependent gain
731 are present in neural networks. Our prediction about the role of background activity in determining
732 the gain of the sparse or dense codes can be tested by changing the overall background activity
733 using chemogenetic techniques.

Limitations and possible extensions

Here we made several simplifications and assumptions to reveal that STP of synapses has important consequences for neural coding. Relaxing each of these simplifications and assumptions may affect our conclusions in certain conditions and should be investigated in further studies. In the following we briefly discuss a few crucial simplifications and how they might affect our results.

Our analyses consider the presynaptic activity to be comprised of independent Poisson processes, that is, whenever we choose a set of N_{ext} presynaptic units to increase their firing rates, we choose them randomly. Because STP is a synapse specific property, input cross-correlation will not affect PRR and Q and therefore, the presynaptic gain. However, it is well known that input correlation can change the gain of a neuron (Kuhn et al., 2003; de la Rocha and Parga, 2005).

It is conceivable that in some conditions, input correlations can potentially neutralize the advantage of sparse codes over dense codes. The readout neuron fluctuations (and therefore, their output firing rate) are dependent on the input correlation. For the same amount of pairwise correlation, the size of fluctuations in the readout neurons is directly proportional to the number of signal carrying units (N_{ext}). A larger N_{ext} (dense distribution) will elicit larger fluctuations than a smaller N_{ext} (sparse distribution). This is because for a larger N_{ext} more input spikes can occur together in the same time bin. Thus, input correlations may amplify the downstream impact of dense input distributions more than the sparse input distributions. The size of this effect of correlations depends on the number of inputs (N_{ext}) and the amount of correlations (both pairwise and higher-order). However, because cortical activity is weakly correlated (Ecker et al., 2010) such an effect of correlation may not be enough to completely neutralize the advantage of sparse distribution over dense distribution.

We also did not study the effect of spatial location of synapses in transferring the effect of sparse codes over dense codes to the readout neurons. There are at least two possibilities in which dendritic locations of the synapses may weaken the advantage of a sparse input distributions over a dense input distributions. First, when synaptic strength decreases as a function of distance from the soma: it is possible that in the sparse input distribution case, the synapses bringing the information are located far away from the soma while for dense input distribution at least some inputs will be closer to the soma. Therefore, even if on the presynaptic side, a sparse input distribution generates stronger outputs than a dense input distribution, its effect on the postsynaptic neurons may be weakened because of weaker synapses. However, even in this case, because of dendritic nonlinearities and Na+/Ca2+ spikes (Larkum et al., 2009), distally located sparse distribution may still have a stronger response than proximally located dense input distributions. Second, the effect of synapses on certain dendrites is cancelled by strategically placed inhibitory synapses (Gidon and Segev, 2012): It is possible that a sparse distribution (because of fewer synapses) may get cancelled or weakened by strategically placed inhibition. The effect of such inhibition will indeed be weaker on dense input distributions as many more synapses will carry the input information. Thus, for sparse input distributions, their location on the neuron may be an important factor. A proper treatment of this question requires the knowledge of e.g. neuron morphology, distribution of inhibition and dendritic non-linearities, and should addressed in a separate study.

We also assumed that all the synaptic weights are sampled from the same Gaussian distribution as our goal was to consider a naive situation in which weights have not been 'trained' for any specific task. Having different synaptic weight distributions may affect the value of the gains, specially when synaptic weights and input are associated (stimulus-specific tuning). Such different distributions may arise because of supervise/unsupervised learning. A systematic study of a network with stimulus-specific tuning of synaptic weights raises several pertinent questions and should be investigated in a separate study.

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 optimal 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 (Kaeser 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 optimal 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 then IBI^{opt} no extra gain will happen. Experimental evidence points to resonance at the theta-frequency oscillations range ($5.0 \sim 10.0Hz$, $100 \sim 200ms$ IBI) as being important for cortical-cerebellar drive (Gandolfi et al., 2013; Chen et al., 2016) and for 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).

22

References

Abbott LF, Regehr WG (2004) Synaptic computation. *Nature* 431:796–803.

Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science* 273:1868–1871.

Atluri PP, Regehr WG (1996) Determinants of the Time Course of Facilitation at the Granule Cell to Purkinje Cell Synapse. *Journal of Neuroscience* 16:5661–5671.

Babadi B, Sompolinsky H (2014) Sparseness and expansion in sensory representations. *Neuron* 83:1213–1226.

Bao J, Reim K, Sakaba T (2010) Target-Dependent Feedforward Inhibition Mediated by Short-Term Synaptic Plasticity in the Cerebellum. *Journal of Neuroscience* 30:8171–8179.

Billings G, Piasini E, Lorincz A, Nusser Z, Silver RA (2014) Network Structure within the Cerebellar Input Layer Enables Lossless Sparse Encoding. *Neuron* 83:960–974.

Blackman AV, Abrahamsson T, Costa RP, Lalanne T, Sjöström PJ (2013) Target-cell-specific short-term plasticity in local circuits. *Frontiers in Synaptic Neuroscience* 5:1–13.

Brown SP, Brenowitz SD, Regehr WG (2003) Brief presynaptic bursts evoke synapse-specific retrograde inhibition mediated by endogenous cannabinoids. *Nature Neuroscience* 6:1048–1057.

Brunel N (2000) Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience* 8:183–208.

Buonomano DV (2000) Decoding temporal information: A model based on short-term synaptic plasticity. *Journal of Neuroscience* 20:1129–41.

Buzsáki G (2002) Theta oscillations in the hippocampus. *Neuron* 33:325–340.

Chadderton P, Margrie TW, Häusser M (2004) Integration of quanta in cerebellar granule cells during sensory processing. *Nature* 428:856–860.

Chen H, Wang YJ, Yang L, Sui JF, Hu ZA, 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 JJ, Carter AG, Regehr WG (2007) Fast Vesicle Replenishment and Rapid Recovery from Desensitization at a Single Synaptic Release Site. *Journal of Neuroscience* 27:5448–5460.

De Kock CPJ, 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. *Journal of Physiology* 586:3353–3364.

de la Rocha J, Parga N (2005) Short-term synaptic depression causes a non-monotonic response to correlated stimuli. *Journal of Neuroscience* 25:8416–8431.

Dean P, Porrill J, Ekerot CF, Jörntell H (2009) The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nature Reviews Neuroscience* 11 VN - re:30–43.

Deng PY, Klyachko VA (2011) The diverse functions of short-term plasticity components in synaptic computations. *Communicative and Integrative Biology* 4:543–548.

Dittman JS, Kreitzer AC, Regehr WG (2000) Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *Journal of Neuroscience* 20:1374–1385.

Doussau F, Schmidt H, Dorgans K, Valera AM, Poulain B, Isope P (2017) Frequency-dependent mobilization of heterogeneous pools of synaptic vesicles shapes presynaptic plasticity. *eLife* 6:e28935.

Du K, Wu YW, Lindroos R, Liu Y, Rózsa B, Katona G, Ding JB, Kotaleski JH (2017) Cell-type–specific inhibition of the dendritic plateau potential in striatal spiny projection neurons. *Proceedings of the National Academy of Sciences* 114:E7612–E7621.

Ecker AS, Berens P, Keliris GA, Bethge M, Logothetis NK, Tolias AS (2010) Decorrelated neuronal firing in cortical microcircuits. *science* 327:584–587.

Éltes T, Kirizs T, Nusser Z, Holderith N (2017) Target Cell Type-Dependent Differences in Ca(2+) Channel Function Underlie Distinct Release Probabilities at Hippocampal Glutamatergic Terminals. *Journal of Neuroscience* 37:1910–1924.

Fuhrmann G, Cowan A, Segev I, Tsodyks M, Stricker C (2004) Multiple mechanisms govern the dynamics of depression at neocortical synapses of young rats. *Journal of Physiology* 557:415–438.

Fuhrmann G, Segev I, Markram H, Tsodyks M (2002) Coding of temporal information by activity-dependent synapses. *Journal of Neurophysiology* 87:140–148.

Gandolfi D, Lombardo P, Mapelli J, Solinas S, DeAngelo E (2013) Theta-Frequency Resonance at the Cerebellum Input Stage Improves Spike Timing on the Millisecond Time-Scale. *Frontiers in Neural Circuits* 7:64.

Gidon A, Segev I (2012) Principles governing the operation of synaptic inhibition in dendrites. *Neuron* 75:330–341.

Goldman MS, Maldonado P, Abbott LF (2002) Redundancy reduction and sustained firing with stochastic depressing synapses. *Journal of Neuroscience* 22:584–91.

Grangeray-Vilmint A, Valera AM, Kumar A, Isope P (2018) Short-term plasticity combines with excitation/iInhibition balance to expand cerebellar Purkinje cell dynamic range. *Journal of Neuroscience* 38:5153–5167.

Greenberg DS, Houweling AR, Kerr JND (2008) Population imaging of ongoing neuronal activity in the visual cortex of awake rats. *Nature Neuroscience* 11:749–751.

Grillo FW, Neves G, Walker A, Vizcay-Barrena G, Fleck RA, Branco T, Burrone J (2018) A Distance-Dependent Distribution of Presynaptic Boutons Tunes Frequency-Dependent Dendritic Integration. *Neuron* 99:275–282.

Gütig R, Sompolinsky H (2006) The tempotron: A neuron that learns spike timing-based decisions. *Nature Neuroscience* 9:420–428.

Hawkins J, Ahmad S (2016) Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Frontiers in Neural Circuits* 10:23.

Isaacson J, Scanziani M (2011) How Inhibition Shapes Cortical Activity. *Neuron* 72:231–243.

Ito M (2014) Cerebellar microcircuitry. *Reference Module in Biomedical Sciences* .

Ito M (2006) Cerebellar circuitry as a neuronal machine. *Progress in Neurobiology* 78:272–303.

Izhikevich EM, Desai NS, Walcott EC, Hoppensteadt FC (2003) Bursts as a unit of neural information: selective communication via resonance. *Trends in neurosciences* 26:161–167.

Jackman SL, Regehr WG (2017) The mechanisms and functions of synaptic facilitation. *Neuron* 94:447–464.

Jackman SL, Turecek J, Belinsky JE, Regehr WG (2016) The calcium sensor synaptotagmin 7 is required for synaptic facilitation. *Nature* 529:88–91.

Jiang X, Shen S, Cadwell CR, Berens P, Sinz F, Ecker AS, Patel S, Tolias AS (2015) Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* 350:aac9462.

Jorntell H, Ekerot CFF (2006) Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. *J Neurosci* 26:11786–11797.

Jörntell H, Bengtsson F, Schonewille M, De Zeeuw CI (2010) Cerebellar molecular layer interneurons - computational properties and roles in learning. *Trends in Neurosciences* 33:524–535.

Kaeser PS, Regehr WG (2017) The readily releasable pool of synaptic vesicles. *Current Opinion in Neurobiology* 43:63–70.

Kandaswamy U, Deng PY, Stevens CF, Klyachko VA (2010) The Role of Presynaptic Dynamics in Processing of Natural Spike Trains in Hippocampal Synapses. *Journal of Neuroscience* 30:15904–15914.

Kenet T, Bibitchkov D, Tsodyks M, Grinvald A, Arieli A (2003) Spontaneously emerging cortical representations of visual attributes. *Nature* 425:954–956.

Klyachko VA, Stevens CF (2006) Excitatory and feed-forward inhibitory hippocampal synapses work synergistically as an adaptive filter of natural spike trains. *PLoS Biology* 4:1187–1200.

Kuhn A, Aertsen A, Rotter S (2003) Higher-Order Statistics of Input Ensembles and the Response of Simple Model Neurons. *Neural Computation* 15:67–101.

Kuhn A, Aertsen A, Rotter S (2004) Neuronal integration of synaptic input in the fluctuation-driven regime. *Journal of Neuroscience* 24:2345–56.

Kumar A, Schrader S, Aertsen A, Rotter S (2008) The high-conductance state of cortical networks. *Neural Computation* 20:1–43.

Larkum ME, Nevian T, Sandler M, Polsky A, Schiller J (2009) Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. *Science* 325:756–760.

Larsen RS, Sjöström PJ (2015) Synapse-type-specific plasticity in local circuits. *Current Opinion in Neurobiology* 35:127–135.

Latorre R, Torres JJ, Varona P (2016) Interplay between subthreshold oscillations and depressing synapses in single neurons. *PLOS ONE* 11:1–31.

Lefort S, Petersen CC (2017) Layer-Dependent Short-Term Synaptic Plasticity between Excitatory Neurons in the C2 Barrel Column of Mouse Primary Somatosensory Cortex. *Cerebral Cortex* 27:3869–3878.

Lisman J (1997) Bursts as a unit of neural information: making unreliable synapses reliable. *Trends in Neurosciences* 20:38–43.

927 Litwin-Kumar A, Harris KD, Axel R, Sompolinsky H, Abbott L (2017) Optimal degrees of synaptic
928   connectivity. *Neuron* 93:1153–1164.

929 Markram H, Wang Y, Tsodyks M (1998) Differential signaling via the same axon of neocortical
930   pyramidal neurons. *Proceedings of the National Academy of Sciences* 95:5323–8.

931 Middleton JW, Yu N, Longtin A, Maler L (2011) Routing the flow of sensory signals using
932   plastic responses to bursts and isolated spikes: experiment and theory. *Journal of Neuro-*
933   *science* 31:2461–2473.

934 Mittmann W, Koch U, Häusser M (2005) Feed-forward inhibition shapes the spike output of
935   cerebellar Purkinje cells. *Journal of Physiology* 563:369–378.

936 Naud R, Sprekeler H (2018) Sparse bursts optimize information transmission in a multiplexed neural
937   code. *Proceedings of the National Academy of Sciences* 115:E6329–E6338.

938 Olshausen BA, Field DJ (1996) Emergence of simple-cell receptive field properties by learning a
939   sparse code for natural images. *Nature* 381:607–609.

940 Pelkey KA, McBain CJ (2007) Differential regulation at functionally divergent release sites along a
941   common axon. *Current Opinion in Neurobiology* 17:366–373.

942 Petersen CCH, Crochet S (2013) Synaptic Computation and Sensory Processing in Neocortical
943   Layer 2/3. *Neuron* 78:28–48.

944 Polsky A, Mel B, Schiller J (2004) Computational subunits in thin dendrites of pyramidal cells.
945   *Nature neuroscience* 7:621–627.

946 Pouille F, Scanziani M (2001) Enforcement of temporal fidelity in pyramidal cells by somatic feed-
947   forward inhibition. *Science* 293:1159–1163.

948 Reyes A, Lujan R, Rozov A, Burnashev N, Somogyi P, Sakmann B (1998) Target-cell-specific
949   facilitation and depression in neocortical circuits. *Nature Neuroscience* 1:279–285.

950 Rizzoli SO, Betz WJ (2005) Synaptic vesicle pools. *Nature Reviews Neuroscience* 6:57–69.

951 Rochefort NL, Garaschuk O, Milos RI, Narushima M, Marandi N, Pichler B, Kovalchuk Y, Konnerth
952   A (2009) Sparsification of neuronal activity in the visual cortex at eye-opening. *Proceedings of*
953   *the National Academy of Sciences* 106:15049–54.

954 Rotman Z, Deng PYPY, Klyachko Va (2011) Short-term plasticity optimizes synaptic information
955   transmission. *Journal of Neuroscience* 31:14800–14809.

956 Rotman Z, Klyachko VA (2013) Role of synaptic dynamics and heterogeneity in neuronal learning
957   of temporal code. *Journal of Neurophysiology* 110:2275–86.

958 Rozov A, Burnashev N, Sakmann B, Neher E (2001) Transmitter release modulation by intracellular
959   Ca2+ buffers in facilitating and depressing nerve terminals of pyramidal cells in layer 2/3 of the
960   rat neocortex indicates a target cell-specific difference in presynaptic calcium dynamics. *Journal*
961   *of Physiology* 531:807–26.

962 Schultz SR, Rolls ET (1999) Analysis of information transmission in the Schaffer collaterals. *Hip-*
963   *pocampus* 9:582–598.

964 Scott P, Cowan AI, Stricker C (2012) Quantifying impacts of short-term plasticity on neuronal
965   information transfer. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 85:1–15.

Shemesh OA, Tanese D, Zampini V, Linghu C, Piatkevich K, Ronzitti E, Papagiakoumou E, Boyden ES, Emiliani V (2017) Temporally precise single-cell-resolution optogenetics. *Nature Neuroscience* 20:1796–1806.

Stevens CF, Wang Y (1995) Facilitation and depression at single central synapses. *Neuron* 14:795–802.

Stimberg M, Goodman DFM, Benichoux V, Brette R (2014) Equation-oriented specification of neural models for simulations. *Frontiers in Neuroinformatics* 8:6.

Sun HY, Lyons SA, Dobrunz LE (2005) Mechanisms of target-cell specific short-term plasticity at Schaffer collateral synapses onto interneurones versus pyramidal cells in juvenile rats. *Journal of Physiology* 568:815–840.

Thanawala MS, Regehr WG (2013) Presynaptic Calcium Influx Controls Neurotransmitter Release in Part by Regulating the Effective Size of the Readily Releasable Pool. *Journal of Neuroscience* 33:4625–4633.

Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Sciences* 94:719–23.

Tsodyks M, Markram H, Pawelzik K, Markram H (1998) Neural networks with dynamic synapses. *Neural Computation* 835:821–835.

Tyrrell T, 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* 336:239–57.

Valera AM, Doussau F, Poulain B, Barbour B, Isope P (2012) Adaptation of granule Cell to Purkinje cell synapses to high-frequency transmission. *Journal of Neuroscience* 32:3267–3280.

van Beugen BJ, Gao Z, Boele HJ, Hoebeek F, De Zeeuw CI (2013) High frequency burst firing of granule cells ensures transmission at the parallel fiber to purkinje cell synapse at the cost of temporal coding. *Frontiers in neural circuits* 7:95.

Vetter P, Roth A, Häusser M (2001) Propagation of action potentials in dendrites depends on dendritic morphology. *Journal of Neurophysiology* 85:926–937.

Weisenburger S, Vaziri A (2018) A guide to emerging technologies for large-scale and whole brain optical imaging of neuronal activity. *Annual Review of Neuroscience* 41:431–452.

Wilson NR, Runyan CA, Wang FL, Sur M (2012) Division and subtraction by distinct cortical inhibitory networks in vivo. *Nature* 488:343–348.

Xu Y, Zou P, Cohen AE (2017) Voltage imaging with genetically encoded indicators. *Current Opinion in Chemical Biology* 39:1–10.

Zucker RS, Regehr WG (2002) Short-Term Synaptic Plasticity. *Annual Review of Physiology* 64:355–405.

Figure 1: Distribution of the spiking activity over presynaptic neurons and short-term-plasticity. **A** Top: A neuron receives input from a presynaptic population, with only one of the neurons eliciting 7 spikes. Bottom: The PSC generated by the 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: Similar to panel A, a neuron receives inputs from a presynaptic population, but in this scenario the spikes were distributed among all presynaptic neurons. 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 $r_\delta = R_{ext}/N$.

Figure legends

Figure 2: A Temporal profile of the PRR for a facilitatory synapse (*top*, $U = 0.1$, $\tau_f = 200ms$ and $\tau_{rec} = 50ms$) and a depressing synapse (*bottom*, $U = 0.7$, $\tau_f = 50ms$ and $\tau_{rec} = 200ms$), 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 curve in *A*, equation 3). Q^s_{ext} for depressing synapses saturates at lower firing rates than facilitatory synapses. The inset shows the derivative of Q^s_{ext} and highlights the nonlinearities in Q^s_{ext}, 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 ($Q^p_{ext} - Q^p_{bas}$) 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 ($Q^p_\delta - Q^p_{bas}$). 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 5) 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 maximum 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.

Figure 3: Effects of STP attributes on maximum gain of the neural population. Here the STP parameters U, τ_{rec} and τ_f were varied in the following range, from more facilitatory to more depressing: $U : 0.05 \rightarrow 0.9$, $\tau_{rec} : 20 \rightarrow 500ms$ and $\tau_f : 500 \rightarrow 20ms$. **A** The optimal 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 circles mark the parameters position for the facilitatory ($s1$, blue) and a depressing ($s2$, red) synapses used as reference in this work. **B** G_{max} as a function of STP properties that gradually and monotonically change the synapse from facilitatory to depressing. Similar to r_{opt}, G_{max} also decays rapidly for more depressing types and for higher basal rates. **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}. This summarizes our prediction that higher basal firing rates will constrain the amplitudes of activity distribution-dependent gains and optimal encoding rates. **D** Optimal distribution of presynaptic activity, that maximizes the gain G. The change from sparse ($OD \sim 0$) to dense ($OD \sim 1$) optimal distribution is abrupt and occurs approximately at the same STP region for different stimulus durations (T_s). However, the transition point where OD changes from sparse to dense code is strongly modulated by r_{bas} – higher basal rates allow for sparse code only for much more facilitatory synapses. **E** Optimal distribution of rate as a function of the three key model parameters (U, τ_{rec} and τ_f). 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 $OD \sim 0$. Effects of STP attributes on maximum gain of the neural population can be seen in the Figure 3-1

Figure 4: Effects of resources recovery time constant τ_{rec} and facilitation factor U on G_{max} and r_{opt} for $T_s = 40ms$, for facilitatory synapses. **A** The r_{opt} surface, as a function of U and τ_{rec}, shows that a given optimal encoding rate can be matched by different combinations of synaptic parameters. For example, the iso-frequency curve of $150Hz$ (black line) can be achieved with either $\{U = 0.05, \tau_{rec} = 90ms\}$ (○) or $\{U = 0.1, \tau_{rec} = 15ms\}$ (□). **B** The G_{max} surface, as a function of U and τ_{rec}, shows that the same maximum gain can 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 (○ = 109\%, □ = 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 resources recovery time constant τ_{rec} and facilitation factor U on G_{max} and r_{opt} for facilitatory synapses can be seen in the Figure 4-1.

Figure 5: Combined optimal distribution of activity 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** 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} (*top*). 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 (*bottom*). **E** G_{max}^{com} surface for different combinations of STP characteristics of $s1$ and $s2$. Notice that G_{max}^{com} steadily increases for $s1 \to Fac$ or $s2 \to Dep$, and that $G_{max}^{com} = 0$ whenever $s1$ is more depressing than $s2$. The circle marks the specific $\{s1, s2\}$ combination used in the other panels and throughout the work. **F** Effects of ongoing basal activity r_{bas} on optimal conditions for $T_s = 40ms$. Increasing basal activity decreases the combined optimal rate (*top*) and combined maximum gain (*center*). Results for 3 different U at the facilitatory synapse. Increasing basal activity consistently decreases G_{max}. 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 (*bottom*).

Figure 6: Transfer of STP gain from presynaptic population to postsynaptic neuron. **A** Schematic of a readout neuron receiving feedfoward excitation and feedforward inhibitory input. The inhibitory group was driven by the PRR of synapses of type s_2 (Equation 17). Sparse and dense activity patterns are schematically shown. The population temporally ($T_s = 40ms$ 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 (feedforward excitation branch) and depression (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. **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*).

Figure 7: Synaptic gain when firing rates of individual neurons (r_{ext}) were draw 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 (k) moved the distribution from exponential to approximately Gaussian. Decreasing the scale parameter (θ) 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.

Figure 8: Continuous distribution of r_{bas}. **A** Two configurations of a Gamma distribution with the same mean value $\overline{r_{bas}} = 0.5$ Hz. When $k = 1$, the distribution is exponential and there is a higher probability that any chosen neuron will have $r_{bas} < 0.5$ Hz as opposed to $r_{bas} > 0.5$ Hz. A higher shape value $k = 20$ brings the distribution closer to a Gaussian and decreases variance. **B** Expected values of Q_{bas}^s, Q_δ^s and Q_{opt}^s for varied Gamma shape values. The higher k is, the smaller is the variance and the results converge to the estimated values with fixed r_{bas}. **C** Similar to the other variables, the optimal gains start with higher values (when k is low) and converge to the estimated values with fixed r_{bas} for higher k. For lower rates the estimates do not diverge much (see overlapping solid and dashed lines when $r_{bas} = 0.5$ Hz), but for higher r_{bas} the increased variance of the continuous distribution has non trivial effects. This indicates that distribution-dependent gains from facilitatory populations might be more resilient to higher basal levels than initially predicted.

Figure 9: Transfer of presynaptic gains onto postsynaptic rate changes. Distribution-dependent postsynaptic transfer function. For facilitatory synapses, sparse input distributions will yield higher effective input than dense distributions. If both distributions keep the neuron operating under the same transfer function, the presynaptic gain will be efficiently transferred to the postsynaptic rate (1-2 or 3-4). If sparse distributions put the neuron operating under TF-2 while dense distributions keep the neuron under TF-1, the presynaptic gains will be magnified (1-4). Otherwise, if dense distributions put the neuron under TF-2 while sparse distributions keep it under TF-1, the presynaptic gains might be overtaken by the postsynaptic integration effects (2-3).

Figure 10: Linear approximation of Q_δ^s. **A** For basal rate $r_{bas} = 0.5 Hz$ and $T_s = 40ms$, Q_δ^s is approximately linear for all STP regimes, facilitatory ($U = .1, \tau_{rec} = 50ms, \tau_f = 2000ms$), facilitatory/depressing ($U = .4, \tau_{rec} = 100ms, \tau_f = 1000ms$) or depressing ($U = .7, \tau_{rec} = 2000ms, \tau_f = 500ms$). **B** Slope deviation for increasing r_δ in comparison to the slope for $r_\delta = 0.1 \cdot r_{bas}$ is always smaller than 0.6% for the three synapse types. **C** R-squared is always close to 100% for the three synapse types. **D** Absolute value of slope deviation, similar to panel B, but for r_δ departing from several different values of r_{bas}. The gray line marks $|S^s \text{ dev}| = 1\%$. We observe that the linear approximation will work well throughout a large space (left from the gray line) for facilitation (*left*) and facilitation/depression (*middle*) regimes, and gets a bit more constrained for depressing (*right*) synapses. **E** Similar to panel C, but for r_δ departing from several different values of r_{bas}.

Extended Figure 3-1: Effects of STP attributes on maximum gain of the neural population. **A** Similar to Figure 3 but with $T_s = 100ms$. **B** Similar to Figure 3 but with $T_s = 300ms$. Both panels reproduce the main findings of 3, showing that the distribution-dependent gain is high for facilitatory synapses and it is strongly affected by the basal rate even if we consider longer integration time windows. The relative importance of the recovery time constant and facilitation time constant in defining the optimal distribution OD grow with larger T_s (large circles for $OD = 1$ and small circles for $OD \sim 0$), but the facilitation factor U keeps being the most relevant attribute in defining the OD.

Extended Figure 4-1: Effects of resources recovery time constant τ_{rec} and facilitation factor U on G_{max} and r_{opt} for facilitatory synapses. **A** Similar to Figure 4 but with $T_s = 100$ ms. **B** Similar to Figure 4 but with $T_s = 300$ ms. Both panels reproduce the main findings of 4, showing that a given optimal encoding rate can be matched by different combinations of synaptic parameters, but resulting in different gains.

A

B

C