frontiers

RESEARCH TOPICS
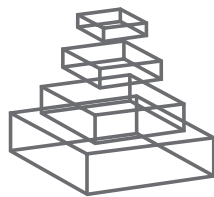
NEURAL INFORMATION PROCESSING WITH DYNAMICAL SYNAPSES

Topic Editors
Si Wu, K. Y. Michael Wong and Misha Tsodyks

frontiers in
COMPUTATIONAL NEUROSCIENCE

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

NEURAL INFORMATION PROCESSING WITH DYNAMICAL SYNAPSES

Topic Editors:
Si Wu, Beijing Normal University, China
K. Y. Michael Wong, Hong Kong University of Science & Technology, China
Misha Tsodyks, Weizmann Institute of Science, Israel

Table of Contents

Neural information processing with dynamical synapses

Si Wu[1,2]*, K. Y. Michael Wong[3]* and Misha Tsodyks[4]*

[1] State Key Laboratory of Cognitive Neuroscience & Learning and IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China
[2] Center for Innovation and Collaboration in Brain and Learning Sciences, Beijing Normal University, Beijing, China
[3] Department of Physics, Hong Kong University of Science & Technology, Hong Kong, China
[4] Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel
*Correspondence: wusi@bnu.edu.cn; phkywong@ust.hk; misha@weizmann.ac.il

Experimental data have consistently revealed that the neuronal connection weight, which models the efficacy of firing of a pre-synaptic neuron in modulating the state of the post-synaptic neuron, varies on short time scales, ranging from tens to thousands of milliseconds (Markram and Tsodyks, 1996; Zucker and Regehr, 2002). This is called short-term plasticity (STP). Two types of STP, with opposite effects on the connection efficacy, have been observed in experiments, which are known as short-term depression (STD) and short-term facilitation (STF).

Computational studies have explored the impact of STP on single neuron and network dynamics, and found that STP can generate very rich intrinsic dynamical behaviors, including adaptation, temporal filtering, damped oscillation, state hopping with transient population spike, traveling front and pulse, spiral wave, rotating bump state, robust self-organized critical activity and so on. These studies also strongly suggest that STP may play many important roles in neural computation. For instances, STD may generate a dynamic control mechanism that allows equal fractional changes on rapidly and slowly firing afferents to produce post-synaptic responses, realizing Weber's law (Abbott et al., 1997); STD may generate a mechanism to close down network activity naturally, achieving iconic sensory memory (Fung et al., 2012); STD may provide a mechanism for memory searching by destabilizing attractor states (Torres et al., 2007); and STF may provide a mechanism for implementing work memory without recruiting neural firing (Mongillo et al., 2008).

From the computational point of view, the time scale of STP resides between fast neural signaling (on the order of milliseconds) and slow experience-induced learning (on the order of minutes or above), and it is on the time order of many important temporal processes occurring in our daily lives, such as motion control, speech recognition and working memory. Thus, STP may serve as a substrate for neural systems manipulating temporal information on the relevant time scales.

This *Research Topic* presents new results in the study of STP and summarizes some recent progress in the field. It includes the works on analyzing the phenomenological models of STP, the effects of STP on single neuron and network dynamics, and the roles of STP in a number of neural information processes.

REFERENCES

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 221–224. doi: 10.1126/science.275.5297.221

Fung, C. C. A., Wong, K. Y. M., Wang, H., and Wu, S. (2012). Dynamical synapses enhance neural information processing: gracefulness, accuracy, and mobility. *Neural Comput.* 24, 1147–1185. doi: 10.1162/NECO_a_00269

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382, 807–810.

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science* 319, 1543–1546. doi: 10.1126/science.1150769

Torres, J., Cortes, J., Marro, J., and Kappen, H. J. (2007). Competition between synaptic depression and facilitation in attractor neural networks. *Neural Comput.* 19, 2739–2755. doi: 10.1162/neco.2007.19.10.2739

Zucker, R., and Regehr, W. (2002). Short-term synaptic plasticity. *Annu. Rev. Physiol.* 64, 355–405. doi: 10.1146/annurev.physiol.64.092501.114547

Theoretical models of synaptic short term plasticity

Matthias H. Hennig*

School of Informatics, Institute for Adaptive and Neural Computation, University of Edinburgh, Edinburgh, UK

Short term plasticity is a highly abundant form of rapid, activity-dependent modulation of synaptic efficacy. A shared set of mechanisms can cause both depression and enhancement of the postsynaptic response at different synapses, with important consequences for information processing. Mathematical models have been extensively used to study the mechanisms and roles of short term plasticity. This review provides an overview of existing models and their biological basis, and of their main properties. Special attention will be given to slow processes such as calcium channel inactivation and the effect of activation of presynaptic autoreceptors.

Keywords: short term plasticity, synaptic transmission, mathematical model, synaptic depression, synaptic facilitation

INTRODUCTION

Chemical synapses are highly specialized structures that enable neurons to exchange signals, or to send signals to non-neural cells such as muscle fibers. Even though there is a staggering diversity of synapse morphologies and types in the brain, the fundamental process of synaptic transmission is always the same. A presynaptic membrane potential depolarization, typically caused by the arrival of an action potential, triggers the release of neurotransmitter, which then binds to receptors that, in turn, generate a response in the postsynaptic neuron.

A key quantity in neural circuits is the synaptic efficacy or strength, which varies over time. Cellular processes such as long-term potentiation and depression contribute to the patterning of the nervous system during development, and are thought to constitute the basis of learning and memory (Morris, 2003). Slow and long-lasting homeostatic processes adjust synaptic strength to maintain circuit activity within functional regimes (Turrigiano and Nelson, 2004). In addition, a whole range of activity-dependent processes exist that modulate synaptic efficacy continuously on very short time scales ranging from milliseconds to minutes (for reviews, see Zucker and Regehr, 2002; Fioravante and Regehr, 2011). Unlike long-term and homeostatic plasticity, *short term plasticity*, the topic of this review, has a direct influence on the computation performed by neural circuits as these dynamics take place on the time scale of stimulus-driven activity, neural computations and behavior.

Broadly, short term plasticity can be classified as synaptic depression and facilitation. Depression refers to the progressive reduction of the postsynaptic response during repetitive presynaptic activity, while facilitation is an increase synaptic efficacy. Each of these may be caused by a range of different mechanisms with different time constants, and the two forms are not mutually exclusive. For instance, a particularly well-studied example of a strongly depressing synapse is the calyx of Held, a giant synaptic terminal in the mammalian auditory brainstem (Schneggenburger and Forsythe, 2006). A closer look at the underlying mechanisms, however, reveals that the response is also modulated by facilitation, which is however, partially masked by depression. In fact, most synapses express some combination of these two mechanisms, but with considerable variability between different neuron types (Wang et al., 2006).

The purpose of this review is to summaries models of short term plasticity, to discuss their biological background and plausibility, and to provide a guide for selecting an appropriate model and level of detail. The focus here is on the mechanistic aspects of these models, for a review of functional implications see Abbott and Regehr (2004). The review begins with a reminder of the main processes involved in synaptic transmission. Next, the vesicle depletion model and its variants will be introduced as a canonical model for short term plasticity. Finally, several additions to this class of models will be discussed that were required to explain more recent experimental findings.

PRINCIPLES OF SYNAPTIC TRANSMISSION

Almost all factors contributing to short term plasticity are located in the presynaptic terminal. To identify the relevant variables required in models, we begin with a brief review of the main events following the arrival of a presynaptic action potential at a synapse, as illustrated in **Figure 1**. The site where synaptic transmission of neural activity is initiated is called the *active zone* (AZ), a presynaptic morphological specialization where vesicles containing neurotransmitter and proteins required for the release process are clustered. The AZ is opposed by the *postsynaptic density* (PSD), an area that contains a large number of different proteins implicated in synapse maintenance and plasticity. In addition to a whole variety of structural and signaling complexes, the PSD contains the bulk of the neurotransmitter receptors mediating the postsynaptic response.

Neurotransmitter release from vesicles located at the AZ is initiated by an elevation of the intracellular calcium concentration $[Ca^{2+}]_i$ due to opening of voltage gated calcium channels

FIGURE 1 | Schematic illustration of the main steps involved in synaptic transmission, and of variables subject to use-dependent modification. Symbols refer to quantities used in the model equations in this review.
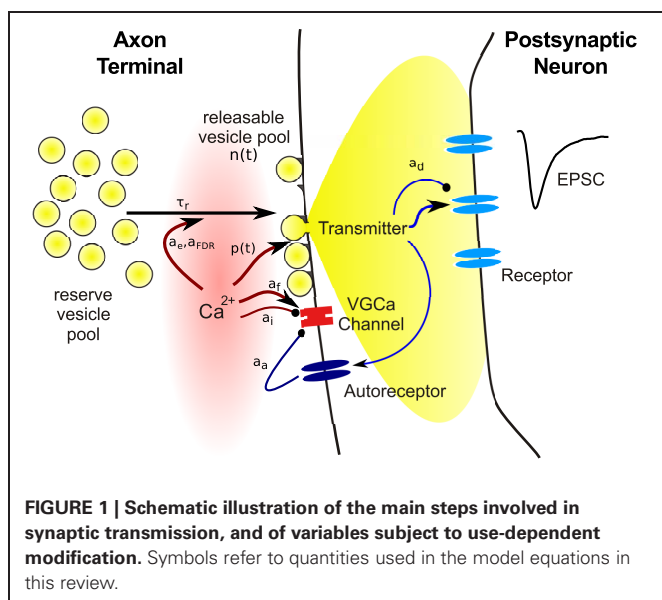
(VGCC). VGCCs are thought to be tightly co-localized with AZs, such that the arrival of the presynaptic action potential causes an increase of $[Ca^{2+}]_i$ within a localized nanodomain from around 30 nM at rest to about 10–30 μM. This brief elevation of $[Ca^{2+}]_i$ increases the probability of vesicle fusion with the cell membrane and subsequent release of transmitter into the synaptic cleft. Hence the release probability $p(t)$ is the first variable required in a model of short term plasticity. Importantly, the relationship between $[Ca^{2+}]_i$ and release probability p is not linear, but follows a steep power function relationship with an exponent between three and four (Bollmann et al., 2000; Schneggenburger and Neher, 2000; Lou et al., 2005). The release probability is often modulated in an activity-dependent manner, hence it is expressed as a function of time.

Electronmicrographs show that presynaptic terminals contain vesicles filled with neurotransmitter. The release of a single vesicle then constitutes the smallest signal (or *quantum*) that can be transmitted to the postsynaptic neuron, which can be seen as spontaneous miniature postsynaptic current at an unstimulated synapse. Usually only a small fraction of the vesicles in the terminal are located in close vicinity of the cell membrane at the AZ. These vesicles are assumed to be release-ready or "primed," while the remaining are assumed to be on hold to replace empty vesicles following transmitter release. The existence of anatomically distinguishable vesicle populations has led to the concept of *vesicle pools*: docked vesicles form the *releasable pool* and those in waiting the *reserve pool*. The release process is termed *excocytosis*, which is followed by the retrieval of empty vesicles through *endocytosis*, and *replenishment* of vesicles on available release sites from the reserve pool. There is evidence that more than two vesicle pools may exist, which differ in release probability and retrieval rate (Trommershäuser et al., 2003; Wölfel et al., 2007), which may be due to their distance from VGCCs (Wadel et al., 2007). However, the details of this matter are still debated and will not further discussed here (for reviews, see Sudhof, 2004; Rizzoli and Betz, 2005).

Hence the second variable required in a synapse model is the number of vesicles $N(t)$ available for release. Again, as will be discussed in more detail below, the number of release-ready vesicles changes over time since the occupancy of the pool changes during neural activity. Vesicle number and release probability are the key ingredients for a model of presynaptic transmitter release:

$$T(t) = p(t) \cdot N(t) \tag{1}$$

Here $T(t)$ is the amount of transmitter released into the synaptic cleft at time t. Simulating a highly realistic synapse model using this expression would require a precise, time continuous model of calcium influx and vesicle cycling. However, since the release probability dramatically increases upon the arrival of a presynaptic action potential from a resting value of almost zero, it is usually sufficient to update these quantities only once every time a presynaptic action potential arrives.

Finally, the released transmitter diffuses through the synaptic cleft and binds to receptors to generate a postsynaptic response, the main quantity of interest in synapse models. Here, we focus on the action of ionotropic receptors, which contain an ion channel that opens when transmitter is bound. The kinetics of such a response is determined by the rates of transmitter binding and unbinding and opening and closing of the channel, as well as transitions to and from desensitized states. The simplest model of this process is when the postsynaptic conductance is proportional to the amount of transmitter released:

$$g(t) = g_m T(t) \tag{2}$$

The peak conductance is denoted by g_m. If the time course of the response is relevant, for instance to distinguish between fast AMPA receptor and slow NMDA receptor mediated transmission, alpha functions, double exponential models, or simple kinetic models are useful to model this process (Destexhe et al., 1994b; Roth and Rossum, 2009).

Numerous studies have been devoted to assessing the release probability and quantal content of synapses in various brain areas and neurons types. As will be shown below, this is generally achieved through model-based analysis, which is possible because the synapse models provide a good mapping between experimental observables, usually the postsynaptic current and its variance, and the underlying synaptic parameters. A comprehensive overview of parameters of a range of neuron types assessed in this way can be found in a review by Branco and Staras (2009).

THE VESICLE DEPLETION MODEL AND EXTENSIONS
VESICLE DEPLETION AS MAIN CAUSE OF SYNAPTIC DEPRESSION
The outline in the preceding section hints that presynaptic vesicles are a limited resource, and that their depletion during ongoing activity can lead to a suppression of the postsynaptic response. The first formal model of such a process was published by Liley and North (1953), even before synaptic vesicles were discovered by De Robertis and Bennett (1955). It sought to explain synaptic depression during brief tetanic stimulation of the rat neuromuscular junction, and was based on

the assumption that releasable neurotransmitter is produced at a limited rate. Tetanic stimulation was assumed to cause transmitter depletion and a concomitant reduction in postsynaptic response. This process is described by a simple first order kinetic model:

$$\frac{dn(t)}{dt} = \underbrace{\frac{1 - n(t)}{\tau_r}}_{\text{replenishment}} \underbrace{- \sum_j \delta(t - t_j) \cdot p \cdot n(t)}_{\text{release}} \qquad (3)$$

where $n(t)$ is the occupancy of the release pool, bounded between zero and one, τ_r the time constant of the vesicle replenishment, and t_j the presynaptic spike times. Note that in this and all following equations, the dynamic quantity, here $n(t)$, is evaluated before the delta function [as in $n(t - \epsilon)$, here the ϵ is omitted for clarity]. The release term reduces the vesicle pool occupancy by $T(t) = p \cdot n(t)$, which is proportional to the postsynaptic response (see Equation 2). Experiments suggest that the recovery time constant is typically in the order of seconds. Equation (3) describes a continuous form of the model, which may be inappropriate for synapses with a small number of releasable vesicles, as it is often the case. Then a discrete form should be used where the release pool occupancy $n(t)$ is replaced by the vesicle number $N(t)$. In this case, a discrete form is also required to accurately model the stochasticity of synapses.

This model predicts an exponential decay of the postsynaptic response during stimulation at a constant rate, and an inverse relation between input frequency ν and steady state level of depression $n_\infty = 1/(p\nu\tau_r + 1)$ (**Figures 2A,E**). It was found to fit responses recorded from some depressing synapses very well (Liley and North, 1953; Tsodyks and Markram, 1997), including EPSCs during stimulation of the calyx of Held with *in vivo*-like activity patterns (Hermann et al., 2009). However, often synapses show substantial deviations. In particular, the steady state values decrease more slowly with increasing frequency than the inverse behavior predicted here.

SYNAPTIC FACILITATION

To explain such deviations from the deletion model, it was first suggested by Betz (1970) to extend it by release probability facilitation that counteracts depression. Potential underlying mechanism of facilitation are an accumulation of residual calcium in the synaptic terminal (Atluri and Regehr, 1996; Blatow et al., 2003; Felmy et al., 2003), which causes rapid VGCC facilitation (Katz and Miledi, 1968; Borst and Sakmann, 1998; Cuttle et al., 1998; Mochida et al., 2008). A simple phenomenological model of such processes is to increase the release probability after each presynaptic spike (Betz, 1970; Varela et al., 1997; Markram et al., 1998):

$$\frac{dp(t)}{dt} = \frac{p_0 - p(t)}{\tau_f} + \sum_j \delta(t - t_j) \cdot a_f \cdot (1 - p(t)) \qquad (4)$$

Here p_0 is the baseline release probability, a_f the amount of facilitation per action potential and τ_f the recovery time constant. The time constant is typically in the range of tens of milliseconds, much faster than vesicle replenishment. Therefore, facilitation is

usually observed during more intense periods of activity. Steady-state facilitation approaches $p_\infty = (p_0 + \nu a_f \tau_f)/(1 + \nu a_f \tau_f)$ for a stimulus with constant frequency ν (**Figure 2E**).

The net effect of the combined model of facilitation and vesicle depletion depends strongly on the basal release probability: for a small p_0, facilitation can have a substantial effect since it is not masked by rapid vesicle pool depletion, and for large values depression will dominate over depletion (**Figure 2B**). As a general rule, it appears that synapses with a larger vesicle pool also tend to have a higher release probability (Dobrunz and Stevens, 1997). Hence facilitation is expected to be more dominant at "smaller" synapses.

This extension of the depletion model can account quite well for data where the simpler depletion model fails, in particular the relationship between stimulus frequency and steady-state response amplitude (Varela et al., 1997; Markram et al., 1998). For instance, a comprehensive survey of cells in the medial prefrontal cortex has shown that this model can fit a wide range of different behaviors encountered in such data sets, despite large variability in the relative contribution of depression and facilitation (Wang et al., 2006).

This depletion model with facilitation has become very popular as a canonical model for short term plasticity. It has, either in the form given here (Equations 3, 4) or using a slightly different set of equations as introduced by Tsodyks et al. (1998), been used in many studies investigating the functional importance of short term plasticity (see e.g., Abbott et al., 1997; Tsodyks et al., 1998; Fuhrmann et al., 2002; Mongillo et al., 2008; Pfister et al., 2010). As usual, however, a closer experimental investigation of synapses has shown that this relatively simple and intuitive model lacks potentially important detail, as will be discussed in the following sections.

USE-DEPENDENT VESICLE REPLENISHMENT

An important observation at odds with the depletion model is that vesicle replenishment can accelerate after intensive stimulation. This effect was found to depend on an increase in intracellular calcium concentration, and to occur in a physiological range of input firing rates (Dittman and Regehr, 1998; Stevens and Wesseling, 1998; Wang and Kaczmarek, 1998; Sakaba and Neher, 2001; Fuhrmann et al., 2004; Hosoi et al., 2007). Enhanced vesicle replenishment can be included in the depletion model by adding some form of activity-dependent component to Equation (3). Two slightly different approaches have been proposed, both capable of explaining the slow reduction in steady state depression for strong stimuli that the simple depletion model fails to replicate.

The first model, introduced by Fuhrmann et al. (2004) to reproduce depression at cortical synapses, was based on the idea that presynaptic activity directly modulates the time constant τ_r of vesicle replenishment in Equation (3) above:

$$\frac{d\tau_r(t)}{dt} = \frac{\tau_{r0} - \tau_r(t)}{\tau_{\text{FDR}}} - a_{\text{FDR}}\tau_r(t) \cdot \sum_j \delta(t - t_j) \qquad (5)$$

Here each presynaptic action potential reduces the time constant by $a_{\text{FDR}}\tau_r(t)$, which recovers to its resting value τ_{r0} with
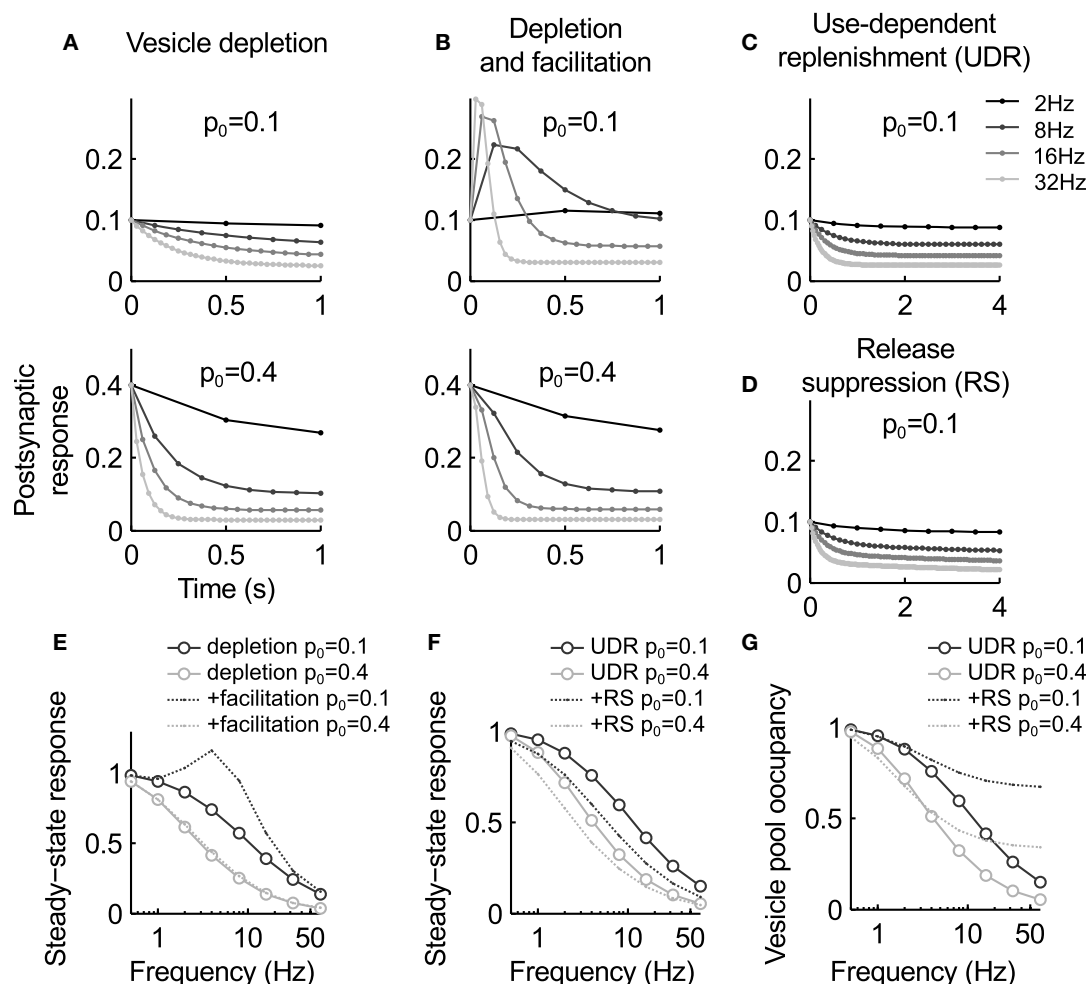
FIGURE 2 | Summary of the key characteristics of the models discussed in this review. (A–D) Postsynaptic response for the different models during stimulation at different frequencies. **(A)** The vesicle depletion model (Equation 3) predicts exponential decay of the response and an inverse relation between stimulus frequency and steady-state amplitude. A higher release probability causes faster and stronger depression [compare upper and lower graph, see also panel **(E)**]. **(B)** The depletion model with facilitation (Equations 3, 4) predict a transient response increase during high-frequency stimulation. For a low basal release probability p_0 the response remains elevated (top graph), while for higher p_0 vesicle depletion masks facilitation [bottom graph, see also panel **(E)**]. **(C)** Use-dependent vesicle replenishment (Equation 6) increases the steady-state response. **(D)** As panel **(C)**, but with

added slow use-dependent suppression of release probability. Here the postsynaptic response continues to slowly decay when the depletion model reaches steady-state [compare **(C)** and **(D)**]. **(E)** Steady-state response magnitude as a function of input frequency for the depletion model (circles) and the depletion model with facilitation (dashed lines). **(F)** Same as **(E)**, but for the depletion model with use-dependent replenishment (UDE, circles) and the UDE model with slow suppression of release probability (RS, dashed). Note that the latter increases depression in particular at low frequencies. **(G)** Occupancy of the releasable vesicle pool for the models in panel **(F)**. It is less depleted for the RS model as steady-state depression is mediated by the reduction in release probability. Parameters: $\tau_r = 1$ s, $a_f = 0.3$, $\tau_f = 0.2$ s [no facilitation in **(C,D)**], $a_e = 0.4$, $\tau_e = 0.1$ s, $a_i = 0.01$, $\tau_i = 10$ s.

a time constant τ_{FDR} in the order of hundreds of milliseconds. A very similar model with a non-linear relation between intracellular calcium concentration and recovery rate was proposed to explain the different kinetics observed at hippocampal and cerebellar synapses (Dittman and Regehr, 1998; Dittman et al., 2000).

Alternatively, it may be assumed activity leads to a temporary enhancement of vesicle replenishment. This is based on the observation that high-frequency stimulation causes a fast but short-lived component of recovery from depression, which is absent after weaker stimulation (Wang and Kaczmarek, 1998).

In these experiments, the recovery time course was fit by two exponential functions, suggesting the combined action of at least two processes. This can be modeled by augmenting a constant background replenishment with a low rate ($k_r = \frac{1}{\tau_r}$) with an activity-dependent component:

$$\frac{dk_e(t)}{dt} = -\frac{k_e(t)}{\tau_e} + a_e \cdot \sum_j \delta(t - t_j) \cdot (1 - k_e(t)) \qquad (6)$$

This process is activated by presynaptic activity, leads to an increment a_e of the replenishment rate for each action potential,

and decays with a time constant τ_e in the range of 10–100 ms. Equation (3) then becomes:

$$\frac{dn(t)}{dt} = \underbrace{(k_r + \tilde{k}_e k_e(t))(1 - n(t))}_{\text{replenishment}} - \underbrace{\sum_j \delta(t - t_j) \cdot p(t) \cdot n(t)}_{\text{release}}$$

(7)

where $\tilde{k}_e$ is the peak rate of activity-dependent vesicle replenishment. This model predicts weaker steady-state depression at high frequencies (**Figures 2C,F**), and has been shown to rather accurately reproduce the vesicle pool kinetics (Hosoi et al., 2007) and steady-state behavior at the calyx of Held (Wong et al., 2003; Hennig et al., 2008).

The biophysical mechanism behind use-dependent vesicle replenishment is still not well understood. It appears clear that it depends on calcium influx (Wang and Kaczmarek, 1998; Sakaba and Neher, 2001; Hosoi et al., 2007), but it has been difficult to experimentally disentangle the role of calcium-dependent vesicle recruitment and calcium-dependent endocytosis, perhaps because most studies so far used extremely strong and unphysiological stimuli to deplete the vesicle pool. A recent study suggests that these two processes may in fact be linked, and that perhaps the speed at which release sites are made available by endocytosis is an important rate limiting step during high frequency transmission (Yao and Sakaba, 2012). Use-dependent replenishment may then reflect faster recruitment due to more efficient endocytosis.

A main function of this mechanism appears to maintain the ability of a synapse to transmit during sustained periods of high activity (Wong et al., 2003; Hosoi et al., 2007). It is as such an important, and often overlooked component of short term plasticity that has implications for transmission of varying firing rates. In addition, it has been suggested to improve transmission by broadening the range over which information about rate and rate changes are reliably transmitted (Fuhrmann et al., 2004; Yang et al., 2009). Which of the two models discussed here is more appropriate is unclear. The difference between the two models is that enhanced replenishment is unbounded in Equation (5), but bounded in Equation (6). Hence the former predicts a faster decrease of the steady state response amplitude with increasing frequency, which more quickly settles to a constant value. It is therefore possible that it underestimates the amount of depression at some synapses, but this would require a more exhaustive comparison with data.

SLOW MODULATION OF RELEASE PROBABILITY

A further omission of the depletion model is that activity-dependent release probability suppression may also contribute to synaptic depression (Xu and Wu, 2005; Mochida et al., 2008). Potential mechanisms include VGCC inactivation (Forsythe et al., 1998; Patil et al., 1998) or activation of presynaptic autoreceptors such as mGluRs or AMPARs, which in turn can cause a reduction of the release probability (Takahashi et al., 1996; Takago et al., 2005). A possible molecular route of such effects is calcium/calmodulin (Lee et al., 1999). Postsynaptic release of endocannabinoids has also been shown to suppress synaptic strength

over short time scales, but the mechanisms are currently not well understood (Brenowitz and Regehr, 2005). Overall, the degree to which these mechanisms are relevant under physiological conditions is still not fully understood. For instance, release probability suppression has been reported to strongly contribute to synaptic depression during weak activity at the calyx of Held (Xu and Wu, 2005), but this effect may be more pronounced at immature synapses were morphological development renders synaptic transmission is less effective (Renden et al., 2005; Nakamura et al., 2008).

A generic model incorporating both release probability facilitation and depression can be constructed by extending Equation (4) by an activity-dependent modulation of the baseline release probability p_0 (Billups et al., 2005; Hennig et al., 2008):

$$\frac{dp_0(t)}{dt} = -\frac{\tilde{p}_0 - p_0(t)}{\tau_i} - \sum_j \delta(t - t_j) \cdot a_i \cdot p_0(t) \qquad (8)$$

Here the baseline release probability $p_0(t)$ is reduced by a constant fraction a_i after each spike, and recovers back to $\tilde{p}_0$ with a time constant τ_i in the order of several seconds. Then depression of release probability is proportional to the incoming spike rate. An alternative form, which models the activation of autoreceptors, is to replace the term on the right-hand side with $\sum_j \delta(t - t_j) \cdot a_a \cdot p_0(t) \cdot p(t) \cdot n(t)$. In this case, depression of release probability is release-dependent. Combinations of both mechanisms are also possible, as shown by Hennig et al. (2008). In combination with the depletion model and facilitation (Equations 3 or 6, and Equation 4), this model can account for a slow form of depression that follows initial rapid vesicle depletion (**Figures 2D,F**), as observed at GABAergic synapses (Kraushaar and Jonas, 2000) or the calyx of Held (Hennig et al., 2008) during prolonged stimulation.

The analysis of the steady-state behavior the model reveals an interesting further property (Hennig et al., 2007). If the release probability is assumed to vary slowly compared to the effective vesicle replenishment rate $\tilde{k}_e$, the quasi-stationary solution of Equation (3) with use-dependent vesicle replenishment (Equation 6) is $n_\infty p_c = \tilde{k}_e(1 - n_\infty)$, where the index c indicates that p_c is constant over the time interval considered, and we obtain $n_\infty = k_e/(p_c + k_e)$. This solution is valid when all fast processes (e.g., facilitation) have settled to their stationary values. If the release probability is now changed by a small amount to $p_c' = \alpha p_c$, then the vesicle pool occupancy settles to a value that differs by a factor of $n_\infty'/n_\infty = (p_c + k_e)/(\alpha p_c + k_e)$.

Hence a slow reduction in release probability will not only slowly depress the postsynaptic response, but also increase the size of the releasable vesicle pool (**Figure 2G**). This corresponds to a transfer of depression from vesicle depletion to a reduction of release probability. The net effect is a decrease in postsynaptic response that is slower than the change in release probability, and a concomitant refilling of the vesicle pool. Analysis of synaptic depression at the calyx of Held during prolonged stimulation support this conclusion, and suggest that it is, in part, mediated by mGluR autoreceptor activation (Billups et al., 2005; Hennig et al., 2008).

A CLOSER LOOK AT RELEASE PROBABILITY

A central variable in the models discussed is the release probability, and so far the effect of activity was assumed to be linear. This is however, incompatible with the steep non-linearity that couples presynaptic calcium influx to release rate (Bollmann et al., 2000; Schneggenburger and Neher, 2000; Lou et al., 2005). If we assume that the effects of facilitation and depression discussed above such as accumulation of residual calcium, channel facilitation or inactivation, have a linear effect on the calcium concentration, this non-linearity would predict a far more drastic effect on the release rate. In fact, early studies already found that a third to fourth-power relationship is a better model for facilitation than a linear model (Zengel and Magleby, 1982).

An analysis of synaptic depression at the calyx of Held by Xu and Wu (2005) further confirms this intuition. This study suggested that depression during slow stimulation (in the range between 1 and 10 Hz) is primarily mediated by a reduction in calcium influx, while vesicle depletion is only effective at higher frequencies. Interestingly the model presented in the preceding section qualitatively reproduces this effect. As shown in **Figure 2F**, slow depression of release probability has a significant effect at low frequencies when compared to an equivalent depletion model, which becomes weaker with increasing frequency. However, as shown above this model also predicts that the depression at higher frequencies is still due to reduced release probability, which replaces vesicle depletion during sustained activity. There is some experimental evidence based on fluctuation analysis in support of this hypothesis (Hennig et al., 2008), but it will be interesting to see if alternative vesicle depletion models can also account for these findings.

AUGMENTATION AND POST-TETANIC POTENTIATION

Augmentation and post-tetanic potentiation are two slowly developing and long-lasting forms of synaptic enhancement (Fisher, 1997; Zucker and Regehr, 2002). They are induced by prolonged stimulation of the synapse, and vary in their activation and relaxation kinetics. The faster form, with time constants of seconds, is typically referred to as augmentation, whereas post-tetanic potentiation operates on time scales of tens of seconds. It appears that these processes are caused by an increase in release probability, which can be occluded by depression due to vesicle depletion during ongoing stimulation (Habets and Borst, 2007). While early studies proposed accumulation of residual calcium at the synaptic terminal as a primary mechanism (Zengel and Magleby, 1982; Zucker and Lara-Estrella, 1983; Habets and Borst, 2007), more recent work also implicated PKC activation (Korogod et al., 2007) or calmodulin/CaM kinase II activity (Fiumara et al., 2007).

A model which could account for a range of findings in data from the frog and toad neuromuscular junction was proposed by Zengel and Magleby (1982). They proposed that facilitation (F), augmentation (A) and post-tetanic potentiation (P) affect release probability in a multiplicative manner:

$$P(t) \propto F(t)A(t)P(t) \qquad (9)$$

Each process follows first order kinetics, and facilitation was best captured by including a fast and a slow component (see also

Zucker and Lara-Estrella, 1983). While facilitation required a fourth-power relationship between the corresponding state variables and release rate, it was sufficient to assume a linear dependence for augmentation and potentiation. This points to different potential sites of action of these mechanisms as outlined above. In addition, it was found that augmentation increases with longer stimuli. This was modeled by including a time-dependent increase in activation rate $a = a_0 z^{\nu T}$ (where ν is the stimulus frequency, T is stimulus time and z a constant), but could also indicate the presence of multiple first-order processes acting on different time scales (Drew and Abbott, 2006; Hennig et al., 2008). For instance, activation of presynaptic NMDA receptors has also been shown to enhance release probability, with a time course in the order of minutes (Duguid and Smart, 2004).

So far, few theoretical studies have investigated the implications of slow enhancement of release using detailed models. A simple, phenomenological model based on Equation (4) above, where time constants were chosen in the range of augmentation, suggests a potential role in short term memory (Mongillo et al., 2008).

THE OTHER SIDE: RECEPTOR DESENSITIZATION

The time course of the postsynaptic response depends not only on the amount of released transmitter and its time course, but also on the kinetics of the receptors. The interplay of these factors with synapse morphology has been investigated in great detail with Monte Carlo simulations (Stiles et al., 1996; Franks et al., 2003; Coggan et al., 2005; Postlethwaite et al., 2007), which are in particular useful to understand the sources of variability at synapses. The semi-quantitative models discussed in this review cannot easily accommodate this level of detail, but can still be extended to include salient aspects of the postsynaptic response (Destexhe et al., 1994a; Roth and Rossum, 2009).

Apart from the response latency and duration, desensitization is an important property of receptors which has been shown to contribute to synaptic depression during physiological activity levels (Trussell et al., 1988, 1993; Jones and Westbrook, 1996; Neher and Sakaba, 2001). A simple but effective approximation of the state of the population of receptors $D(t)$, can be modeled using first order kinetics:

$$\frac{dD(t)}{dt} = \frac{1 - D(t)}{\tau_D} - \sum_j \delta(t - t_j) \cdot a_D \cdot p(t) \cdot n(t) \cdot D(t) \qquad (10)$$

The quantity $D(t)$ represents the fraction of non-desensitized receptors. Recovery from desensitization τ_D is typically in the order of tens of milliseconds, such that it is only effective during intense episodes of activity. The postsynaptic response is then expressed as $R(t) = g_m D(t) \cdot n(t) \cdot p(t)$, where g_m is the peak conductance.

This basic model captures synaptic depression due to desensitization well. In particular, simulations have shown that a main effect is the masking of presynaptic facilitation at high stimulus frequencies (Jones and Westbrook, 1996; Wong et al., 2003). Yet in this form the model obviously neglects the time course of the postsynaptic potential, which can also be affected by

desensitization. To model this, it is possible to extend it by adding more states, such as closed, open and desensitized, and to model state transitions in a transmitter concentration-dependent manner as a Markov process. Such models have been proposed to better account for the kinetics of the postsynaptic response, in particular for kinetics of different receptor subunit composition (Destexhe et al., 1994a; Robert et al., 2005; Postlethwaite et al., 2007). A drawback of this approach is that this also requires an appropriate model of the time course of neurotransmitter seen by the receptors, which has to be obtained by more detailed diffusion models (see e.g., Franks et al., 2003; Postlethwaite et al., 2007). Finally it is also worth mentioning that potentially other postsynaptic mechanisms exist that contribute to short term plasticity, which have not yet been investigated in models. For instance, AMPA receptors can show an increased paired-pulse facilitation during activity-dependent relief of polyamide block (Rozov and Burnashev, 1999). This effect is potentially important at immature synapses lacking the GluR2 AMPA receptor subunit.

STOCHASTICITY OF SYNAPSES

Transmitter release is a stochastic process, and as a consequence the magnitude of the postsynaptic current evoked by each presynaptic action potential fluctuates from time to time. Due to the quantal nature of synaptic transmission, the variance of the postsynaptic response is described by binomial statistics, with a predicted variance of $\text{Var}(g(t)) = g_m \cdot N(t) \cdot p(t) \cdot (1 - p(t))$ (Del Castillo and Katz, 1954). This shows that changes in the synaptic parameters due to short term plasticity will not only cause changes in the average postsynaptic response, but also in the magnitude of the fluctuations, as measured by the coefficient of variation:

$$CV(g(t)) = \sqrt{\frac{1 - p(t)}{N(t)p(t)g_m}} \qquad (11)$$

This value is high when the release probability or the number of release-ready vesicles is small, as, for instance, often found for cortical neurons (Wang et al., 2006; Brémaud et al., 2007). The expression also shows that stochastic effects are bound to be more important when synaptic depression is dominated by vesicle depletion. In addition, the entire vesicle cycle, which includes vesicle replenishment, consists of stochastic events. In contrast, the influx of calcium during an action potential, which triggers transmitter release, is considered a much more salient event, and is therefore expected to contribute much less to postsynaptic response variability. To model the main sources of stochasticity of synaptic, the models discussed above can be directly cast into a stochastic form by simulating vesicle release and replenishment as random events (see e.g., de la Rocha and Parga, 2005; Yang et al., 2009; Rosenbaum et al., 2012).

Stochastic models are extremely useful for quantitative evaluation of models of synaptic transmission and plasticity, since postulated changes in N and p have predictable effects on variability that can be directly tested experimentally (see e.g., Quastel, 1997; Scheuss and Neher, 2001; Brémaud et al., 2007). This type of analysis requires a careful dissection of synaptic function, since

for instance conductance changes through receptor desensitization may be mistaken for presynaptic effects if not properly controlled for. Recent work exploiting this approach provided evidence for substantial variability of synaptic parameters for synapses between cortical pyramidal neurons, the presence of multi-quantal release (Loebel et al., 2009), and the coordination of pre-synaptic release probability and postsynaptic synaptic strength (Hardingham et al., 2010).

Studies investigating stochastic synapse models have reported several effects indicating that this also has important implications for neural computations and network function. For instance, shot noise due to stochastic release can increase the output firing rate of a neuron operating in a fluctuation driven regime when compared to deterministic dynamics (de la Rocha and Parga, 2005). The same study also showed that short term depression in stochastic synapses causes a further, non-monotonic modulation of output firing in presence of input correlations (see also Rosenbaum et al., 2012). Such effects were further analyzed by Rosenbaum et al. (2013), who showed that, unlike for a deterministic model, a stochastic synapse with short term depression can significantly de-correlate neural activity. Finally, an analysis of stochastic models including slow release probability modulation and activity-dependent vesicle replenishment suggested that multiple mechanisms of short term plasticity may act synergistically to maintain stable information transmission over a broad range of input frequencies (Yang et al., 2009). Overall, however, the models used so far to analyze stochastic effects were mostly rather simple, typically only the depletion model was considered, and assumed constant random inputs to the neuron (see Merkel and Lindner, 2010, for an extension).

OUTLOOK

Theoretical models have contributed much to our understanding of synaptic transmission and short term plasticity by providing a framework to express conceptual models in rigorous terms, and to derive quantitatively testable hypotheses. The models discussed here capture the central biophysical processes involved in synaptic transmission in relatively simple mathematical form, such that an exact or at least approximate analytical treatment is possible. Moreover, key variables in these models have direct measurable correlates. This supports analysis and comparison with data, as often exploited for deriving synaptic parameters from experimentally recorded synaptic currents. It is however not straight forward to experimentally interfere with short term plasticity in intact neural circuits in a targeted manner, for instance to assess functional implications and consequences. Therefore, these models are also a valuable tool that enables analysis beyond the experimentally feasible.

The basic depletion model with facilitation has passed the test of time, which nicely illustrates the success of simple, mathematically tractable phenomenological models in biology. However, as shown here, short term plasticity can be more complicated. In particular slow forms of synaptic depression and facilitation merit more thorough investigation, both in terms of mechanisms and their relevance for neural computations. While the depletion model can very successfully replicate even synaptic responses during *in vivo*-like activity patterns (Hermann et al., 2009),

slow synaptic modulation may have important effects during firing rate modulation on time scales of tens of seconds (see e.g., Mongillo et al., 2008). A combination of slow facilitation and depression has also been shown to support differential responses to time varying stimuli (Barak and Tsodyks, 2007). These studies show that these mechanisms certainly warrant further investigation.

As shown in this review, even the extended and more complete models of short term plasticity have a relatively simple mathematical form, which will greatly facilitate the understanding of their effects in networks. Perhaps a central question in this context is in how far the different mechanisms discussed here have direct functional implications, or rather reflect the biophysical properties and limitations of chemical signaling between neurons. Some of the studies touched upon above and in the previous section suggest the former may be the case (for a more detailed discussion, see e.g., Abbott and Regehr, 2004). On the other hand it is

equally plausible some aspects of short term plasticity may related to homeostatic effects or metabolic efficacy of synapses, issues that have received little attention so far and are now easily testable in models. Addressing such questions may require the analysis of the models under more physiologically relevant conditions. For example, recent experiments indicate that unreliable synapses with short term plasticity are particularly suited to transmit information contained in brief bursts of activity typically observed in hippocampus (Rotman et al., 2011). Therefore, modeling studies specifically investigating synapses in their "natural habitat" of recurrent networks should allow us to refine and consolidate such hypotheses, and to establish more of the much sought-after links between neural biophysics and brain function and dysfunction.

REFERENCES

Abbott, L. F., and Regehr, W. G. (2004). Synaptic computation. *Nature* 431, 796–803.

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224.

Atluri, P. P., and Regehr, W. G. (1996). Determinants of the time course of facilitation at the granule cell to Purkinje cell synapse. *J. Neurosci.* 16, 5661–5671.

Barak, O., and Tsodyks, M. (2007). Persistent activity in neural networks with dynamic synapses. *PLoS Comput. Biol.* 3:e35. doi: 10.1371/journal.pcbi.0030035

Betz, W. J. (1970). Depression of transmitter release at the neuromuscular junction of the frog. *J. Physiol.* 206, 629–644.

Billups, B., Graham, B. P., Wong, A. Y. C., and Forsythe, I. D. (2005). Unmasking group III metabotropic glutamate autoreceptor function at excitatory synapses in the rat CNS. *J. Physiol.* 565, 885–896.

Blatow, M., Caputi, A., Burnashev, N., Monyer, H., and Rozov, A. (2003). Ca2+ buffer saturation underlies paired pulse facilitation in calbindin-D28k-containing terminals. *Neuron* 38, 79–88.

Bollmann, J. H., Sakmann, B., and Borst, J. G. (2000). Calcium sensitivity of glutamate release in a calyx-type terminal. *Science* 289, 953–957.

Borst, J. G., and Sakmann B. (1998). Calcium current during a single action potential in a large presynaptic terminal of the rat brainstem. *J. Physiol.* 506, 143–157.

Branco, T., and Staras K. (2009). The probability of neurotransmitter release: variability and feedback control at single synapses. *Nat. Rev. Neurosci.* 10, 373–383.

Brémaud, A., West, D. C., and Thomson, A. M. (2007). Binomial parameters differ across neocortical layers and with different classes of connections in adult rat and cat neocortex. *PNAS* 104, 14134–14139.

Brenowitz, S. D., and Regehr, W. G. (2005). Associative short-term synaptic plasticity mediated by endocannabinoids. *Neuron* 45, 419–431.

Coggan, J. S., Bartol, T. M., Esquenazi, E., Stiles, J. R., Lamont, S., Martone, M. E., et al. (2005). Evidence for ectopic neurotransmission at a neuronal synapse. *Science* 309, 446–451.

Cuttle, M. F., Tsujimoto, T., Forsythe, I. D., and Takahashi, T. (1998). Facilitation of the presynaptic calcium current at an auditory synapse in rat brainstem. *J. Physiol.* 512, 723–729.

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

De Robertis, E. D., and Bennett, H. S. (1955). Some features of the submicroscopic morphology of synapses in frog and earthworm. *J. Biochem. Cytol.* 1, 47–58.

Del Castillo, J., and Katz, B. (1954). Quantal components of the end-plate potential. *J. Physiol.* 124, 560–573.

Destexhe, A., Mainen, Z. F., and Sejnowski, T. J. (1994a). An efficient method for computing synaptic conductances based on a kinetic model of receptor binding. *Neural Comput.* 6, 14–18.

Destexhe, A., Mainen, Z. F., and Sejnowski, T. J. (1994b). Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *J. Comput. Neurosci.* 1, 195–231.

Dittman, J. S., Kreitzer, A. C., and Regehr, W. G. (2000). Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *J. Neurosci.* 20, 1374–1385.

Dittman, J. S., and Regehr, W. G. (1998). Calcium dependence and recovery kinetics of presynaptic depression at the climbing fiber to Purkinje cell synapse. *J. Neurosci.* 18, 6147–6162.

Dobrunz, L. E., and Stevens, C. F. (1997). Heterogeneity of release probability, facilitation, and depletion at central synapses. *Neuron* 18, 995–1008.

Drew, P. J., and Abbott, L. F. (2006). Models and properties of power-law adaptation in neural systems. *J. Neurophysiol.* 96, 826–833.

Duguid, I. C., and Smart, T. G. (2004). Retrograde activation of presynaptic NMDA receptors enhances GABA release at cerebellar interneuron-Purkinje cell synapses. *Nat. Neurosci.* 7, 525–533.

Felmy, F., Neher, E., and Schneggenburger, R. (2003). Probing the intracellular calcium sensitivity of transmitter release during synaptic facilitation. *Neuron* 37, 801–811.

Fioravante, D., and Regehr, W. G. (2011). Short-term forms of presynaptic plasticity. *Curr. Opin. Neurobiol.* 21, 269–274.

Fisher, S. (1997). Multiple overlapping processes underlying short-term synaptic enhancement. *Trends Neurosci.* 20, 170–177.

Fiumara, F., Milanese, C., Corradi, A., Giovedi, S., Leitinger, G., Menegon, A., et al. (2007). Phosphorylation of synapsin domain A is required for post-tetanic potentiation. *J. Cell. Sci.* 120, 3321.

Forsythe, I. D., Tsujimoto, T., Barnes-Davies, M., Cuttle, M. F., and Takahashi, T. (1998). Inactivation of presynaptic calcium current contributes to synaptic depression at a fast central synapse. *Neuron* 20, 797–807.

Franks, K. M., Stevens, C. F., and Sejnowski, T. J. (2003). Independent sources of quantal variability at single glutamatergic synapses. *J. Neurosci.* 23, 3186–3195.

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

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

Habets, R. L. P., and Borst, J. G. G. (2007). Dynamics of the readily releasable pool during post-tetanic potentiation in the rat calyx of Held synapse. *J. Physiol.* 581, 467–478.

Hardingham, N. R., Read, J. C. A., Trevelyan, A. J., Nelson, J. C., Jack, J. J. B., and Bannister, N. J. (2010). Quantal analysis reveals a functional correlation between presynaptic and postsynaptic efficacy in excitatory

connections from rat neocortex. *J. Neurosci.* 30, 1441–1451.

Hennig, M. H., Postlethwaite, M., Forsythe, I. D., and Graham, B. P. (2007). A biophysical model of short-term plasticity at the calyx of Held. *Neurocomputing* 70, 1626–1629.

Hennig, M. H., Postlethwaite, M., Forsythe, I. D., and Graham, B. P. (2008). Interactions between multiple sources of short-term plasticity during evoked and spontaneous activity at the rat calyx of Held. *J. Physiol.* 586, 3129–3146.

Hermann, J., Grothe, B., and Klug, A. (2009). Modeling short-term synaptic plasticity at the calyx of held using *in vivo*-like stimulation patterns. *J. Neurophysiol.* 101, 20–30.

Hosoi, N., Sakaba, T., and Neher, E. (2007). Quantitative analysis of calcium-dependent vesicle recruitment and its functional role at the calyx of Held synapse. *J. Neurosci.* 27, 14286–14298.

Jones, M. V., and Westbrook, G. L. (1996). The impact of receptor desensitization on fast synaptic transmission. *Trends Neurosci.* 19, 96–101.

Katz, B., and Miledi, R. (1968). The role of calcium in neuromuscular facilitation. *J. Physiol.* 195, 481–492.

Korogod, N., Lou, X., and Schneggenburger, R. (2007). Posttetanic potentiation critically depends on an enhanced Ca(2+) sensitivity of vesicle fusion mediated by presynaptic PKC. *PNAS* 104, 15923–15928.

Kraushaar, U., and Jonas, P. (2000). Efficacy and stability of quantal GABA release at a hippocampal interneuron-principal neuron synapse. *J. Neurosci.* 20, 5594–5607.

Lee, A., Wong, S. T., Gallagher, D., Li, B., Storm, D. R., Scheuer, T., and Catterall, W. A. (1999). Ca2+/calmodulin binds to and modulates P/Q-type calcium channels. *Nature* 399, 155–159.

Liley, A. W., and North, K. A. (1953). An electrical investigation of effects of repetitive stimulation on mammalian neuromuscular junction. *J. Neurophysiol.* 16, 509–527.

Loebel, A., Silberberg, G., Helbig, D., Markram, H., Tsodyks, M., and Richardson, M. J. E. (2009). Multiquantal release underlies the distribution of synaptic efficacies in the neocortex. *Front. Comput. Neurosci.* 3:27. doi: 10.3389/neuro.10.027.2009

Lou, X., Scheuss, V., and Schneggenburger, R. (2005).

Allosteric modulation of the presynaptic Ca2+ sensor for vesicle fusion. *Nature* 435, 497–501.

Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *PNAS* 95, 5323–5328.

Merkel, M., and Lindner, B. (2010). Synaptic filtering of rate-coded information. *Phys. Rev. E* 81:041921. doi: 10.1103/PhysRevE.81.041921

Mochida, S., Few, A. P., Scheuer, T., and Catterall, W. A. (2008). Regulation of presynaptic Ca(V)2.1 channels by Ca2+ sensor proteins mediates short-term synaptic plasticity. *Neuron* 57, 210–216.

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science* 319, 1543–1546.

Morris, R. G. M. (2003). Long-term potentiation and memory. *Philos. Trans. R. Soc. Lond. Ser. B* 358, 643–647.

Nakamura, T., Yamashita, T., Saitoh, N., and Takahashi, T. (2008). Developmental changes in calcium/calmodulin-dependent inactivation of calcium currents at the rat calyx of Held. *J. Physiol.* 586, 2253–2261.

Neher, E., and Sakaba, T. (2001). Combining deconvolution and noise analysis for the estimation of transmitter release rates at the calyx of held. *J. Neurosci.* 21, 444–461.

Patil, P. G., Brody, D. L., and Yue, D. T. (1998). Preferential closed-state inactivation of neuronal calcium channels. *Neuron* 20, 1027–1038.

Pfister, J.-P., Dayan, P., and Lengyel, M. (2010). Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials. *Nat. Neurosci.* 13, 1271–1275.

Postlethwaite, M., Hennig, M. H., Steinert, J. R., Graham, B. P., and Forsythe, I. D. (2007). Acceleration of AMPA receptor kinetics underlies temperature-dependent changes in synaptic strength at the rat calyx of Held. *J. Physiol.* 579, 69–84.

Quastel, D. M. (1997). The binomial model in fluctuation analysis of quantal neurotransmitter release. *Biophys. J.* 72, 728–753.

Renden, R., Taschenberger, H., Puente, N., Rusakov, D. A., Duvoisin, R., Wang, L. Y., et al. (2005). Glutamate transporter studies reveal the pruning of metabotropic glutamate receptors and absence of AMPA receptor desensitization

at mature calyx of held synapses. *J. Neurosci.* 25, 8482–8497.

Rizzoli, S. O., and Betz, W. J. (2005). Synaptic vesicle pools. *Nat. Rev. Neurosci.* 6, 57–69.

Robert, A., Armstrong, N., Gouaux, J. E., and Howe, J. R. (2005). AMPA receptor binding cleft mutations that alter affinity, efficacy, and recovery from desensitization. *J. Neurosci.* 25, 3752–3762.

Rosenbaum, R., Rubin, J., and Doiron, B. (2012). Short term synaptic depression imposes a frequency dependent filter on synaptic information transfer. *PLoS Comput. Biol.* 8:e1002557. doi: 10.1371/journal.pcbi.1002557

Rosenbaum, R., Rubin, J. E., and Doiron, B. (2013). Short-term synaptic depression and stochastic vesicle dynamics reduce and shape neuronal correlations. *J. Neurophysiol.* 109, 475–484.

Roth, A., and Rossum, M. C. W. V. (2009). "Modeling synapses," in *Computational Modeling Methods for Neuroscientists*. Vol. 06, Chapter Modelings, ed E. De Schutter (Cambridge, MA: MIT Press), 139–160.

Rotman, Z., Deng, P.-Y., and Klyachko, V. A. (2011). Short-term plasticity optimizes synaptic information transmission. *J. Neurosci.* 31, 14800–14809.

Rozov, A., and Burnashev, N. (1999). Polyamine-dependent facilitation of postsynaptic AMPA receptors counteracts paired-pulse depression. *Nature* 401, 594–598.

Sakaba, T., and Neher, E. (2001). Quantitative relationship between transmitter release and calcium current at the calyx of held synapse. *J. Neurosci.* 21, 462–476.

Scheuss, V., and Neher, E. (2001). Estimating synaptic parameters from mean, variance, and covariance in trains of synaptic responses. *Biophys. J.* 81, 1970–1989.

Schneggenburger, R., and Forsythe, I. D. (2006). The calyx of Held. *Cell Tissue Res.* 326, 311–337.

Schneggenburger, R., and Neher, E. (2000). Intracellular calcium dependence of transmitter release rates at a fast central synapse. *Nature* 406, 889–893.

Stevens, C. F., and Wesseling, J. F. (1998). Activity-dependent modulation of the rate at which synaptic vesicles become available to undergo exocytosis. *Neuron* 21, 415–424.

Stiles, J. R., Van Helden, D., Bartol, T. M., Salpeter, E. E., and Salpeter, M. M. (1996). Miniature endplate current rise times less than 100 microseconds from improved

dual recordings can be modeled with passive acetylcholine diffusion from a synaptic vesicle. *PNAS* 93, 5747–5752.

Sudhof, T. C. (2004). The synaptic vesicle cycle. *Annu. Rev. Neurosci.* 27, 509–547.

Takago, H., Nakamura, Y., and Takahashi, T. (2005). G protein-dependent presynaptic inhibition mediated by AMPA receptors at the calyx of Held. *PNAS* 102, 7368–7373.

Takahashi, T., Forsythe, I. D., Tsujimoto, T., Barnes-Davies, M., and Onodera, K. (1996). Presynaptic calcium current modulation by a metabotropic glutamate receptor. *Science* 274, 594–597.

Trommershäuser, J., Schneggenburger, R., Zippelius, A., and Neher, E. (2003). Heterogeneous presynaptic release probabilities: functional relevance for short-term plasticity. *Biophys. J.* 84, 1563–1579.

Trussell, L. O., Thio, L. L., Zorumski, C. F., and Fischbach, G. D. (1988). Rapid desensitization of glutamate receptors in vertebrate central neurons. *PNAS* 85, 2834–2838.

Trussell, L. O., Zhang, S., and Raman, I. M. (1993). Desensitization of AMPA receptors upon multiquantal neurotransmitter release. *Neuron* 10, 1185–1196.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835.

Tsodyks, M. V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *PNAS* 94, 719–723.

Turrigiano, G. G., and Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nat. Rev. Neurosci.* 5, 97–107.

Varela, J. A., Sen, K., Gibson, J., Fost, J., Abbott, L. F., and Nelson, S. B. (1997). A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J. Neurosci.* 17, 7926–7940.

Wadel, K., Neher, E., and Sakaba, T. (2007). The coupling between synaptic vesicles and Ca2+ channels determines fast neurotransmitter release. *Neuron* 53, 563–575.

Wang, L. Y., and Kaczmarek, L. K. (1998). High-frequency firing helps replenish the readily releasable pool of synaptic vesicles. *Nature* 394, 384–388.

Wang, Y., Markram, H., Goodman, P. H., Berger, T. K., Ma, J., and

Goldman-Rakic, P. S. (2006). Heterogeneity in the pyramidal network of the medial pre-frontal cortex. *Nat. Neurosci.* 9, 534–542.

Wölfel, M., Lou, X., and Schneggenburger, R. (2007). A mechanism intrinsic to the vesicle fusion machinery determines fast and slow transmitter release at a large CNS synapse. *J. Neurosci.* 27, 3198–3210.

Wong, A. Y. C., Graham, B. P., Billups, B., and Forsythe, I. D. (2003). Distinguishing between presynaptic and postsynaptic mechanisms of short-term depression during action potential trains. *J. Neurosci.* 23, 4868–4877.

Xu, J., and Wu, L.-G. (2005). The decrease in the presynaptic calcium current is a major cause of short-term depression at a calyx-type synapse. *Neuron* 46, 633–645.

Yang, Z., Hennig, M. H., Postlethwaite, M., Forsythe, I. D., and Graham, B. P. (2009). Wide-band information transmission at the calyx of Held. *Neural Comput.* 21, 991–1017.

Yao, L., and Sakaba, T. (2012). Activity-dependent modulation of endocytosis by calmodulin at a large central synapse. *PNAS* 109, 291–296.

Zengel, J. E., and Magleby, K. L. (1982). Augmentation and facilitation of transmitter release. A quantitative description at the frog neuromuscular junction. *J. Gen. Physiol.* 80, 583–611.

Zucker, R. S., and Lara-Estrella, L. O. (1983). Post-tetanic decay of evoked and spontaneous transmitter release and a residual-calcium model of synaptic facilitation at crayfish neuromuscular junctions. *J. Gen. Physiol.* 81, 355–372.

Zucker, R. S., and Regehr, W. G. (2002). Short-term synaptic plasticity. *Annu. Rev. Neurosci.* 64, 355–405.

Emerging phenomena in neural networks with dynamic synapses and their computational implications

Joaquin J. Torres[1] and Hilbert J. Kappen[2]*

[1] Granada Neurophysics Group at Institute "Carlos I" for Theoretical and Computational Physics, University of Granada, Granada, Spain
[2] Donders Institute for Brain Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands

In this paper we review our research on the effect and computational role of dynamical synapses on feed-forward and recurrent neural networks. Among others, we report on the appearance of a new class of dynamical memories which result from the destabilization of learned memory attractors. This has important consequences for dynamic information processing allowing the system to sequentially access the information stored in the memories under changing stimuli. Although storage capacity of stable memories also decreases, our study demonstrated the positive effect of synaptic facilitation to recover maximum storage capacity and to enlarge the capacity of the system for memory recall in noisy conditions. Possibly, the new dynamical behavior can be associated with the voltage transitions between up and down states observed in cortical areas in the brain. We investigated the conditions for which the permanence times in the up state are power-law distributed, which is a sign for criticality, and concluded that the experimentally observed large variability of permanence times could be explained as the result of noisy dynamic synapses with large recovery times. Finally, we report how short-term synaptic processes can transmit weak signals throughout more than one frequency range in noisy neural networks, displaying a kind of *stochastic multi-resonance.* This effect is due to competition between activity-dependent synaptic fluctuations (due to dynamic synapses) and the existence of neuron firing threshold which adapts to the incoming mean synaptic input.

Keywords: short-term synaptic plasticity, emergence of dynamic memories, memory storage capacity, criticality in up–down cortical transitions, neural stochastic multiresonances

1. INTRODUCTION

In the last decades many experimental studies have reported that transmission of information through the synapses is strongly influenced by the recent presynaptic activity in such a way that the postsynaptic response can decrease (that is called *synaptic depression*) or increase (or *synaptic facilitation*) at short time scales under repeated stimulation (Abbott et al., 1997; Tsodyks and Markram, 1997). In cortical synapses it was found that after induction of long-term potentiation (LTP), the temporal synaptic response was not uniformly increased. Instead, the amplitude of the initial postsynaptic potential was potentiated whereas the steady-state synaptic response was unaffected by LTP (Markram and Tsodyks, 1996).

From a biophysical point of view it is well accepted that short-term synaptic plasticity including synaptic depression and facilitation has its origin in the complex dynamics of release, transmission and recycling of neurotransmitter vesicles at the synaptic buttons (Pieribone et al., 1995). In fact, synaptic depression occurs when the arrival of presynaptic action potentials (APs) at high frequency does not allow an efficient recovering at short time scales of the available neurotransmitter vesicles to be released near the cell membrane (Zucker, 1989; Pieribone et al., 1995). This causes a decrease of the postsynaptic response for successive APs. Other possible mechanisms responsible for

synaptic depression have been described including feedback activation of presynaptic receptors and from postsynaptic processes such as receptor desensitization (Zucker and Regehr, 2002). On the other hand, synaptic facilitation is a consequence of residual cytosolic calcium—that remains inside the synaptic buttons after the arrival of the firsts APs—which favors the release of more neurotransmitter vesicles for the next arriving AP (Bertram et al., 1996). This increase in neurotransmitters causes a potentiation of the postsynaptic response or synaptic facilitation. It is clear that strong facilitation causes a fast depletion of available vesicles so at the end it also induces a strong depressing effect. Other possible mechanisms responsible for short-term synaptic plasticity include, for instance, glial-neuronal interactions (Zucker and Regehr, 2002).

In the two seminal papers (Tsodyks and Markram, 1997) and (Abbott et al., 1997) a simple phenomenological model has been proposed based in these biophysical principles which nicely fits the evoked postsynaptic responses observed in cortical neurons. The model is characterized by three variables $x_j(t)$, $y_j(t)$, $z_j(t)$ that follow the dynamics

$$\frac{dx_j(t)}{dt} = \frac{z_j(t)}{\tau_{rec}} - U_j \cdot x_j(t) \cdot \delta(t - t_{sp}^j)$$

$$\frac{dy_j(t)}{dt} = \frac{-y_j(t)}{\tau_{\text{in}}} + U_j \cdot x_j(t) \cdot \delta(t - t_{\text{sp}}^j)$$

$$\frac{dz_j(t)}{dt} = \frac{y_j(t)}{\tau_{\text{in}}} - \frac{z_j(t)}{\tau_{\text{rec}}} \qquad (1)$$

where $y_j(t)$ is the fraction of neurotransmitters which is released into the synaptic cleft after the arrival of an AP at time t_{sp}^j, $x_j(t)$ is the fraction of neurotransmitters which is recovered after previous arrival of an AP near the cell membrane and $z_j(t)$ is the fraction of inactive neurotransmitters. The model assumes conservation of the total number of neurotransmitter resources in time so one has $x_j(t) + y_j(t) + z_j(t) = 1$. The released neurotransmitter inactivates with time constant τ_{in} and the inactive neurotransmitter recovers with time constant τ_{rec}. The synaptic current received by a postsynaptic neuron from its neighbors is then defined as $I_i(t) = \sum_j A_{ij} y_j(t)$ where A_{ij} represents the maximum synaptic current evoked in the postsynaptic neuron i by an AP from presynaptic neuron j which in cortical neurons is around $40\,pA$ (Tsodyks et al., 1998).

For constant release probability U_j, the model describes the basic mechanism of synaptic depression. The model is completed to account for synaptic facilitation by considering that U_j increases in time to its maximum value U as the consequence of the residual cytosolic calcium that remains after the arrival of very consecutive APs, and follows the dynamics

$$\frac{dU_j(t)}{dt} = \frac{[U - U_j(t)]}{\tau_{\text{fac}}} + U \cdot [1 - U_j(t)] \cdot \delta(t - t_{\text{sp}}^j). \qquad (2)$$

Short term synaptic plasticity has profound consequences on information transmission by individual neurons as well as on network functioning and behavior. Previous works have shown this fact on both feed-forward and recurrent networks. For instance, in feed-forward networks activity-dependent synapses act as nonlinear filters in supervised learning paradigms (Natschläger et al., 2001), being able to extract statistically significant features from noisy and variable temporal patterns (Liaw and Berger, 1996).

For recurrent networks, several studies revealed that populations of excitatory neurons with depressing synapses exhibit complex regimes of activity (Senn et al., 1996; Tsodyks et al., 1998, 2000; Bressloff, 1999; Kistler and van Hemmen, 1999), such as short intervals of highly synchronous activity (population bursts) intermittent with long periods of asynchronous activity, as is observed in neurons throughout the cortex (Tsodyks et al., 2000). Related with this, it was proposed (Senn et al., 1996, 1998) that synaptic depression may serve as a mechanism for rhythmic activity and central pattern generation. Also, recent studies on rate models have reported the importance of dynamic synapses in the emergence of persistent activity after removal of stimulus which is the base of the so called working memories (Barak and Tsodyks, 2007), and in particular it has been also reported the relevant role of synaptic facilitation, mediated by residual calcium, as the main responsible for appearance of working memories (Mongillo et al., 2008).

All these phenomena have stimulated much research to elucidate the effect and possible functional role of short term synaptic plasticity. In this paper we review our own efforts over the last decade in this research field. In particular, we have demonstrated both theoretically and numerically the appearance of different non-equilibrium phases in attractor networks as the consequence of the underlying noisy activity in the network and of the existence of synaptic plasticity (see section 2). The emergent phenomenology in such networks includes a high sensitivity of the network to changing stimuli and a new phase in which dynamical attractors or dynamical memories appear with the possibility of regular and chaotic behavior and rapid "switching" between different memories (Pantic et al., 2002; Cortes et al., 2004, 2006; Torres et al., 2005, 2008; Marro et al., 2007). The origin of such new phases and the extraordinary sensibility of the system to varying inputs—even in the memory phase—is precisely the "fatigue" of synapses due to heavy presynaptic activity competing with different sources of noise which induces a destabilization of the regular stable memory attractors. One of the main consequences of this behavior is the strong influence of short-term synaptic plasticity on storage capacity of such networks (Torres et al., 2002; Mejias and Torres, 2009) as we will explain in section 3.

The switching behavior is characterized by a characteristic time scale during which the memory is retained. The distribution of time scale depends in a complex way on the parameters of the dynamical synapse model and is the result of a phase transition. We have investigated the conditions for the appearance of power-law behavior in the probability distribution of the permanence times in the Up state, which is a sign for criticality (see section 4). This dynamical behavior has been associated (Holcman and Tsodyks, 2006) to the empirically observed transitions between states of high activity (Up states) and low activity (Down states) in the mammalian cortex (Steriade et al., 1993a,b).

The enhanced sensibility of neural networks with dynamic synapses to external stimuli could provide a mechanism to detect relevant information in weak noisy external signals. This can be viewed as a form of *stochastic resonance* (SR), which is the general phenomenon that enhances the detection by a non-linear dynamical system of weak signals in the presence of noise. Recent experiments in auditory cortex have shown that synaptic depression improves the detection of weak signals through SR for a larger noise range (Yasuda et al., 2008). In a feed-forward network model of spiking neurons, we have modeled these experimental findings (Mejias and Torres, 2011; Torres et al., 2011). We demonstrated theoretically and numerically that, in fact, short-term synaptic plasticity together with non-linear neuron excitability induce a new type of SR where there are multiple noise levels at which weak signals can be detected by the neuron. We denoted this novel phenomenon by *bimodal stochastic resonances* or *stochastic multiresonances* (see section 5) and, very recently, we have proved that this intriguing phenomenon not only occurs in feed-forward neural networks but also in recurrent attractor networks (Pinamonti et al., 2012).

2. APPEARANCE OF DYNAMICAL MEMORIES

In this section we review our work on the appearance of dynamical memories in attractor neural networks with dynamical synapses as originally reported in (Pantic et al., 2002; Torres et al.,

2002, 2008; Mejias and Torres, 2009). For simplicity and in order to obtain straightforward mean-field derivations we have considered the case of a network of N binary neurons (Hopfield, 1982; Amit, 1989). However, we emphasize that the same qualitative behavior emerges in networks of integrate and fire (IF) neurons (Pantic et al., 2002).

Each neuron in the network, whose state is $s_i = 1, 0$ depending if the neuron is firing or not an AP, receives at time t from its neighbor neurons a total synaptic current, or local field, given by

$$h_i(t) = \sum_j \omega_{ij}(t) s_j(t) \qquad (3)$$

where $\omega_{ij}(t)$ is the synaptic current received by the postsynaptic neuron i from the presynaptic neuron j when this fires an AP ($s_j(t) = 1$). If the synaptic current to neuron i, $h_i(t)$, is larger than some neuron threshold value θ_i, neuron i fires an AP with a probability that depends on the intrinsic noise present in the network. The noise is commonly modeled as a thermal bath at temperature T. We assume parallel dynamics (Little dynamics) using the probabilistic rule

$$\text{Prob}(s_i(t+1) = \sigma) = \frac{1}{2} + \left(\sigma - \frac{1}{2}\right) \tanh[2T^{-1}(h_i(t) - \theta_i)] \qquad (4)$$

with $\sigma = 1, 0$.

To account for short-term synaptic plasticity in the network we consider

$$\omega_{ij}(t) = \overline{\omega}_{ij} D_j(t) F_j(t) \qquad (5)$$

where $D_j(t)$ and $F_j(t)$ are dynamical variables representing synaptic depression and synaptic facilitation mechanisms. The constants $\overline{\omega}_{ij}$ denote static maximal synaptic conductances, that contain information concerning a number P of random patterns of neural activity, or *memories*, $\boldsymbol{\xi}^\mu \equiv \{\xi_i^\mu = 1, 0; i = 1, \ldots, N, \mu = 1, \ldots, P\}$ previously learned and stored in the network. Such static memories can be achieved in actual neural systems by LTP or depression of the synapses due to network stimulation with these memories. For concreteness, we assume here that these weights are the result of a Hebbian-like learning process that takes place on a time scale that is long compared to the dynamical time scales of the neurons and the dynamical synapses. The Hebbian learning takes the form

$$\overline{\omega}_{ij} = \frac{1}{Na(1-a)} \sum_{\mu=1}^{P} (\xi_i^\mu - a)(\xi_j^\mu - a) \quad \overline{\omega}_{ii} = 0, \qquad (6)$$

also known as the *covariance learning rule*, with $a = \langle \xi_i^\mu \rangle$ representing the mean level of activity in the patterns. It is well-known that a recurrent neural network with synapses (Equation 6) acts as an *associative memory* (Amit, 1989). That is, the stored patterns $\boldsymbol{\xi}^\mu$ become local minima of the free-energy and within the basin of attraction of each memory, the neural dynamics (Equation 4) drives the network activity toward this memory. Thus, appropriate stimulation of (a subset of) neurons that are active in the

stored pattern initiates a memory recall process in which the network converges to the memory state.

To model the dynamics of the synaptic depression $D_j(t)$ and facilitation $F_j(t)$, we simplify the phenomenological model of dynamic synapses described by Equations (1, 2), taking into account that in actual neural systems such as the cortex $\tau_{in} \ll \tau_{rec}$, which implies that $y_i(t) = 0$ for most of the time and only at the exact point at which the AP arrives has a non-zero value $y_j(t_{sp}) = x_j(t_{sp}) U_j(t_{sp})$. Thus, the synaptic current evoked in the postsynaptic neuron i by a presynaptic neuron j every time it fires is approximatively $I_{ij}(t) = A_{ij} x_j(t_{sp}^j) U_j(t_{sp}^j)$ which has the form given by Equation (5) with $\overline{\omega}_{ij} = A_{ij}$, $D_j(t) \equiv x_j(t)$ and $F_j(t) \equiv U_j(t)$. We set $U = 1$ without loss of generality in order to have $D_j(t) = F_j(t) = 1 \forall j, t$ for $\tau_{rec}, \tau_{fac} \ll 1$, that corresponds to the well know limit of *static synapses* without depressing and facilitating mechanism. In this limit, in fact, one recover the classical Amari–Hopfield model of associative memory (Amari, 1972; Hopfield, 1982) when one chooses the neuron thresholds as

$$\theta_i = \frac{1}{2} \sum_j \overline{\omega}_{ij}. \qquad (7)$$

It is important to point out that due to the discrete nature of the probabilistic neuron dynamics (Equation 4) together with the approach $\tau_{in} \ll \tau_{rec}$, only discrete versions of the dynamics for $x_i(t)$ and $U_i(t)$ [see for instance (Tsodyks et al., 1998)] are needed here, namely

$$x_j(t+1) = x_j(t) + \frac{1 - x_j(t)}{\tau_{rec}} - U_j(t) \cdot x_j(t) \cdot s_j(t)$$

$$U_j(t+1) = U_j(t) + \frac{[U - U_j(t)]}{\tau_{fac}} + U \cdot [1 - U_j(t)] \cdot s_j(t). \qquad (8)$$

Equations (4–8) completely define the dynamics of the network. Note, that in the limit of $\tau_{rec, fac} \to 0$ the model reduces to the standard Amari–Hopfield model with static synapses.

To numerically and analytically study the emergent behavior of this attractor neural network with dynamical synapses, it is useful to measure the degree of correlation between the current network state $\mathbf{s} \equiv \{s_i; i = 1, \ldots, N\}$ and each one of the stored patterns $\boldsymbol{\xi}^\mu$ by mean of the overlap function

$$m^\mu(\mathbf{s}) = \frac{1}{N a(1-a)} \sum_i (\xi_i^\mu - a) s_i. \qquad (9)$$

Monte Carlo simulations of the network storing a small number of random patterns (loading parameter $\alpha \equiv P/N \to 0$), each pattern having 50% active neurons ($a = 0.5$), no facilitation ($U_j(t) = 1$) and an intermediate value of τ_{rec} is shown in **Figures 1A,B**. It shows a new phase where dynamical memories characterized by quasi-periodic switching of the network activity between pattern ($\boldsymbol{\xi}^\mu$) and anti-pattern ($1 - \boldsymbol{\xi}^\mu$) configurations appear. For lower values of τ_{rec} the network reduces to the attractor network with static synapses and shows the emergence of the traditional *ferromagnetic* or associative memory phase at relatively low T, where network activity reaches a steady state
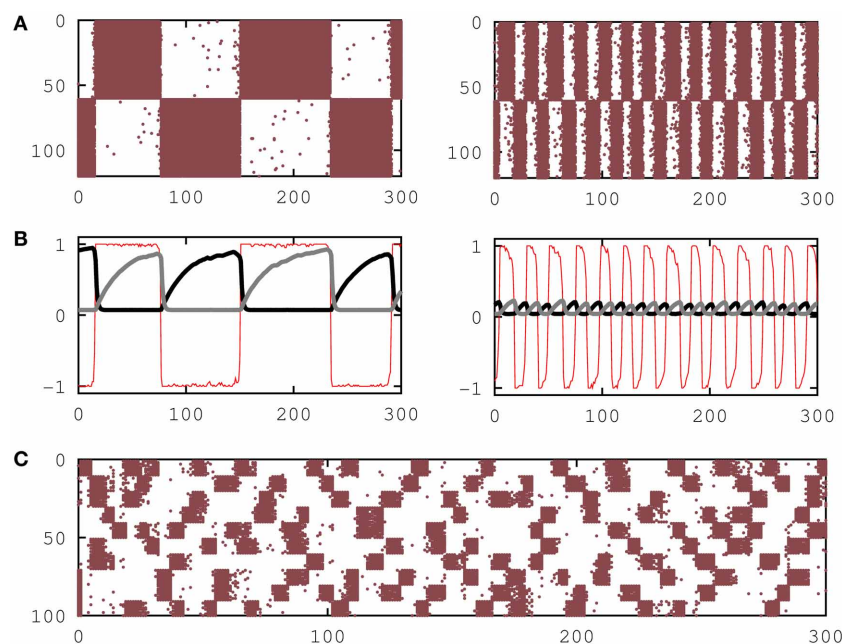
FIGURE 1 | Emergence of dynamic memories in attractor neural networks. (A) Raster plots showing the switching behavior of the network neural activity between one activity pattern and its anti-pattern for $\tau_{rec} = 26$ and $T = 0.025$ (left) and for $\tau_{rec} = 50$ and $T = 0.025$ (right), respectively. **(B)** The behavior of the overlap function $m^\mu(\mathbf{s})$ (thin red line), the mean recovering variable x_+^μ of active neurons in the pattern (thick black line) and the mean recovering variable x_-^μ of not active neurons in the pattern (thick gray line) are plotted for the two cases depicted in **(A)**. **(C)** Raster plot that shows the emergence of dynamic memories when 10 activity patterns are stored in the synapses for $\tau_{rec} = 50$. In all panels the firing threshold was set to $\theta_i = 0$, and the network size was $N = 120$ in **(A)** and **(B)** and $N = 100$ in **(C)**.

that is highly correlated with one of the stored patterns, and a *paramagnetic* or no-memory phase at high T where the network activity reaches a highly fluctuating disordered steady state.

The **Figure 1C** shows simulation results of a network with $P = 10$ patterns and $a = 0.1$, demonstrating that switching behavior is also obtained for relatively large number of patterns and sparse network activity. **Figure 2B** shows that the switching behavior is not an artifact of the binary neuron dynamics and is also obtained in a network of more realistic networks of spiking integrate-and-fire neurons.

All time constants, such τ_{rec} or τ_{fac} are given in units of Monte Carlo steps (MCS) a temporal unit that in actual systems can be associated, for instance, with the duration of the refractory period and therefore of order of 5 ms.

In the limit of $N \to \infty$ (thermodynamic limit) and $\alpha \to 0$ (finite number of patterns) the emergent behavior of the model can be analytically studied within a standard mean field approach [see for details (Pantic et al., 2002; Torres et al., 2008)]. The dynamics of the system then is described by a $6P$-dimensional discrete map

$$\mathbf{v}_{t+1} = \mathcal{F}(\mathbf{v}_t) \tag{10}$$

where $\mathcal{F}$ is a $6P$-dimensional non-linear function of the order parameters

$$\mathbf{v}_t \equiv \{m_+^\mu(t), m_-^\mu(t), x_+^\mu(t), x_-^\mu(t), U_+^\mu(t), U_-^\mu(t);$$
$$\mu = 1, \ldots, P\} \tag{11}$$

that are averages of the microscopic dynamical variables over the sites that are active and quiescent, respectively, in a given pattern μ, that is

$$c_+^\mu(t) \equiv \frac{1}{Na} \sum_{i \in Act(\mu)} c_i(t),$$

$$c_-^\mu(t) \equiv \frac{1}{N(1-a)} \sum_{i \notin Act(\mu)} c_i(t), \tag{12}$$

with $c_i(t)$ being $m_i(t)$, $x_i(t)$, and $U_i(t)$, respectively.

Local stability analysis of the fixed point solutions of the dynamics (Equation 10) shows that, similarly to the Amari–Hopfield standard model and in agreement with Monte Carlo simulations described above, the stored memories ξ^μ are stable attractors in some regions of the space of relevant parameters, such as T, U, τ_{rec}, and τ_{fac}. Varying these parameters, there are, however, some critical values for which the memories destabilize and an oscillatory regime, in which the network visits different memories, can emerge. These critical values are depicted in **Figures 2A,C,D** in the form of transition lines between phases or dynamical behaviors in the system. For instance, for only depressing synapses ($\tau_{fac} = 0$, $U_j(t) = 1$), there is a critical monotonic line $\tau_{rec}^*(T^{-1})$, as in a second order phase transition, separating the no-memory phase and the oscillatory phase (solid line in **Figure 2A**) where oscillations start to appears with small amplitude as in a supercritical Hopf bifurcation. Also there is a transition line $\tau_{rec}^{**}(T^{-1})$, also monotonic, between the oscillatory phase
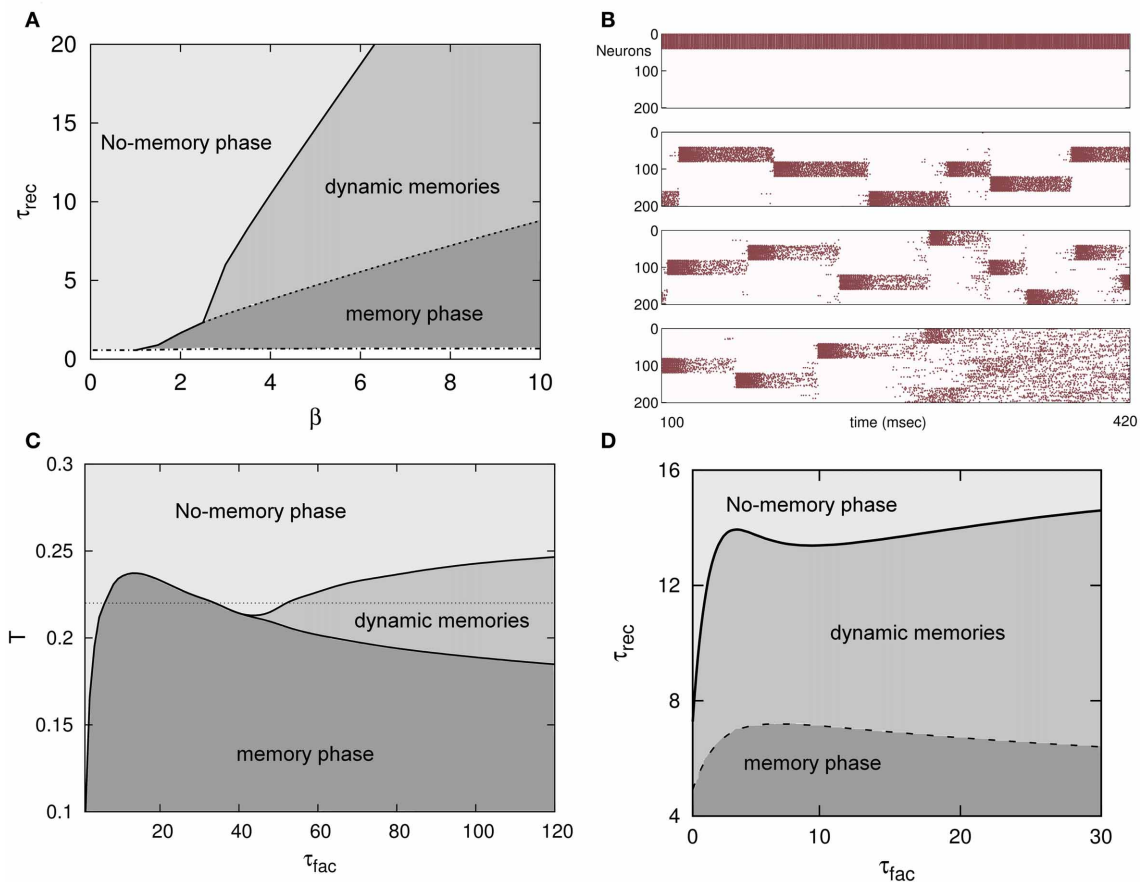
FIGURE 2 | (A) Phase diagram ($\tau_{rec}, \beta \equiv T^{-1}$) of an attractor binary neural network with depressing synapses for $\alpha = 0$. A new phase in which dynamical memories appear—with the network activity switching between the different memory attractors—emerges between the traditional memory and no-memory phases that characterize the behavior of attractors neural networks with static synapses. **(B)** The emergent behavior depicted in **(A)** is robust when a more realistic attractor network of IF neurons and more stored patterns are considered (5 in this simulation). From top to bottom, the behavior of the network activity for $\tau_{rec} = 0$, 300, 800 and 3000 ms is depicted, respectively. For some level of noise the network activity pass from the memory phase to the dynamical phase and from this to the no-memory phase when τ_{rec} is increased. **(C)** Phase diagram (T, τ_{fac}) for $\tau_{rec} = 3$ and $U = 0.1$ of an attractor binary neural network with short-term depression and facilitation mechanisms in the synapses and $\alpha = 0$. **(D)** Phase diagram (τ_{rec}, τ_{fac}) for $T = 0.1$ and $U = 0.1$ in the same system than in **(C)**. In both, **(C,D)**, the diagrams depict the appearance of the same memory, oscillatory and no-memory phases than in the case of depressing synapses. The transition lines between different phases, however, show here a clear non-linear and non-monotonic dependence with relevant parameters consequence of the non-trivial competition between depression and facilitation mechanisms. This is very remarkable in **(C)** where for a given level of noise, namely $T = 0.22$ (horizontal dotted line), the increase of facilitation time constant τ_{fac} induces the transition of the activity of the network from a no-memory state to a memory state, from this one to a no-memory state again, and finally from this last to an oscillatory regime.

and the memory phase which occurs sharply as in a first order phase transition (dashed line in **Figure 2A**). When facilitation is included, the picture is more complex, although similar critical and sharp transitions lines appear separating the same phases. Now, however, the lines separating different phases are non-monotonic and highly non-linear which shows the competition between *a priori* opposite mechanisms, depressing and facilitating, as is depicted in **Figures 2C,D**. In fact, among other features, synaptic depression induces fatigue at the synapses which destabilizes the attractors, and synaptic facilitation allows a fast access to the memory attractors and to stay there during a shorter period of time (Torres et al., 2008). As in **Figure 1**, in all phase diagrams appearing in **Figure 2**, τ_{rec} and τ_{fac} are given in MCS units

(see above) with a value for that temporal unit of around the typical duration of the refractory period in actual neurons (~5ms).

The attractor behavior of the recurrent neural network has the important property to complete a memory based on partial or noisy stimulus information. In this section we have seen that memories that are stable with static synapses become metastable with dynamical synapses, inducing a switching behavior among memory patterns in the presence of noise. In this manner, dynamic synapses provide the associative memory with a natural mechanism to dissociate from a memory in order to associate with a new memory pattern. In contrast, with static synapses the network would stay in the stable memory state forever, preventing recall of new memories. Thus, dynamic synapses change stable

3. STORAGE CAPACITY

It is important to analyze how short-term synaptic plasticity affects the maximum number of patterns of neural activity the system is able to store and efficiently recall, that is, the so called *maximum storage capacity*. In a recent paper we have addressed this important issue using a standard mean field approach in the model described by Equations (3–8) when it stored $P = \alpha N$ activity random patterns with $\alpha > 0$ and $N \to \infty$, $a = 1/2$ and in the absence of noise ($T = 0$). In fact, for very low temperature ($T \ll 1$), redefining the overlaps as $M^\nu \equiv m^\nu - \frac{1}{N}\sum_i (2\xi_i^\nu - 1) \equiv m^\nu - B^\nu$ and assuming steady-state conditions in which there is only one pattern (condensed pattern) with overlap $M \equiv M^1 \sim \mathcal{O}(1)$ and the remaining patters $M^\nu \sim \mathcal{O}(1/\sqrt{N})$, $\nu = 2, \ldots, P$, it is straightforward (Hertz et al., 1991) to see that the steady state of the system is described by the set of mean field equations

$$M = \frac{1}{N}\sum_i \tanh\left[\beta\left(\frac{\gamma'}{1+\gamma\gamma'}M + \zeta_i\right)\right]$$

$$q = \frac{1}{N}\sum_i \tanh^2\left[\beta\left(\frac{\gamma'}{1+\gamma\gamma'}M + \zeta_i\right)\right]$$

$$r = \frac{q}{\left(1 - \beta\frac{\gamma'}{1+\gamma\gamma'}(1-q)\right)^2} \tag{13}$$

where $\gamma \equiv U\tau_{\text{rec}}$, $\gamma' \equiv \frac{1+\tau_{\text{fac}}}{1+U\tau_{\text{fac}}}$, $q \equiv \frac{1}{N}\sum_i \tanh^2[2\beta(h_i(t) - \theta_i)]$ is the spin-glass order parameter, $r = \frac{1}{\alpha}\sum_{\nu \neq 1}(M^\nu)^2$ is the pattern interference parameter and

$$\zeta_i \equiv \sum_{\nu \neq 1}(2\xi_i^1 - 1)(2\xi_i^\nu - 1)$$

$$\left[\frac{\gamma'}{1+\gamma\gamma'}M^\nu + \left(1 - \frac{\gamma'}{1+\gamma\gamma'}\right)B^\nu\right]$$

which in the limit of $N \to \infty$ becomes a Gaussian variable

$$\zeta \approx \frac{\gamma'}{1+\gamma\gamma'}\left(\alpha r + \alpha\left(\frac{1+\gamma\gamma'-\gamma'}{\gamma'}\right)^2\right)^{1/2} z$$

where z is a random normal-distributed variable $N[0, 1]$—see details in (Mejias and Torres, 2009). Then, the $\frac{1}{N}\sum_i$ appearing in Equation (13) becomes an average over $\mathcal{P}(\zeta)$. Using standard techniques in the limit $T = 0$ (Hertz et al., 1991), the set of the resulting three mean-field equations reduces to a single mean-field equation which gives the maximum number of patterns that the system is able to store and retrieve, namely (see mathematical details in Mejias and Torres, 2009)

$$y\left[\sqrt{2\alpha\left(\frac{1+\gamma\gamma'-\gamma'}{\gamma'}\right)^2 + \frac{2}{\sqrt{\pi}}\exp(-y^2)}\right] = \text{erf}(y) \tag{14}$$

where $y \equiv M/\sqrt{\left(2\alpha r + 2\alpha\left(\frac{1+\gamma\gamma'-\gamma'}{\gamma'}\right)^2\right)}$ with M being the overlap of the current state of the network activity with the pattern that is being retrieved. The Equation (14) has a trivial solution $y = 0$ ($M = 0$). Non-zero solutions (with non-zero overlap M) exist for α less than some critical α, which defines the maximum storage capacity of the system α_c.

A complete study of the system by means of Monte Carlo simulations (in a network with $N = 3000$ neurons) has demonstrated the validity of this mean field result and is depicted in **Figure 3A**. The figure shows the behavior of α_c obtained from Equation (14) (different solid lines), when some relevant parameters of the synapse dynamics are varied, and it is compared with the maximum storage capacity obtained from the Monte Carlo simulations (different symbols). The most remarkable feature is that in the absence of facilitation the storage capacity decreases when the level of depression increases (that is, large release probability U, or large recovering time τ_{rec}; see black curves in the top and middle panels of **Figure 3A**. This decrease is caused by the loss of stability of the memory fixed points of the network due to depression. Facilitation (see dark and light gray curves) allows to recover the maximal storage capacity of static synapses, which is the well know limit $\alpha_c \approx 0.14$ (dotted horizontal line), in the presence of some degree of synaptic depression. In general the competition between synaptic depression and facilitation induces a complex non-linear and non-monotonic behavior of α_c for different synaptic dynamics parameters as is shown in different panels of **Figure 3B**. In general, large values of α_c appear for moderate values of U and τ_{rec}, and large values of τ_{fac}. These values qualitatively agree with those described in facilitating synapses in some cortical areas, where U is lower than in the case of depressing synapses and τ_{rec} is several times lower than τ_{fac} (Markram et al., 1998). Note that facilitation or depression never increases the storage capacity of the network above the maximum value $\alpha_c \approx 0.14$.

4. CRITICALITY IN UP–DOWN TRANSITIONS

In a recent paper (Holcman and Tsodyks, 2006), the emergent dynamic memories described in section 2 that result from short-term plasticity have been related to the voltage transitions observed in cortex between a high-activity state (the Up state) and a low-activity state (the Down state). These transitions have been observed in simultaneous individual single neuron recordings as well as in local field measurements.

Using a simple but biologically plausible neuron and synapse model similar to the models described in sections 1 and 2, we have theoretically studied the conditions for the emergence of this intriguing behavior, as well as their temporal features (Mejias et al., 2010). The model consists of a simple stochastic bistable rate model which mimics the average dynamics of a population of interconnected excitatory neurons. The neural activity is summarized by a single activity $\nu(t)$, whose dynamics follows a stochastic mean field equation

$$\tau_\nu \frac{d\nu(t)}{dt} = -\nu(t) + \nu_m \mathcal{S}[J\nu(t)x(t) - \theta] + \zeta(t) \tag{15}$$
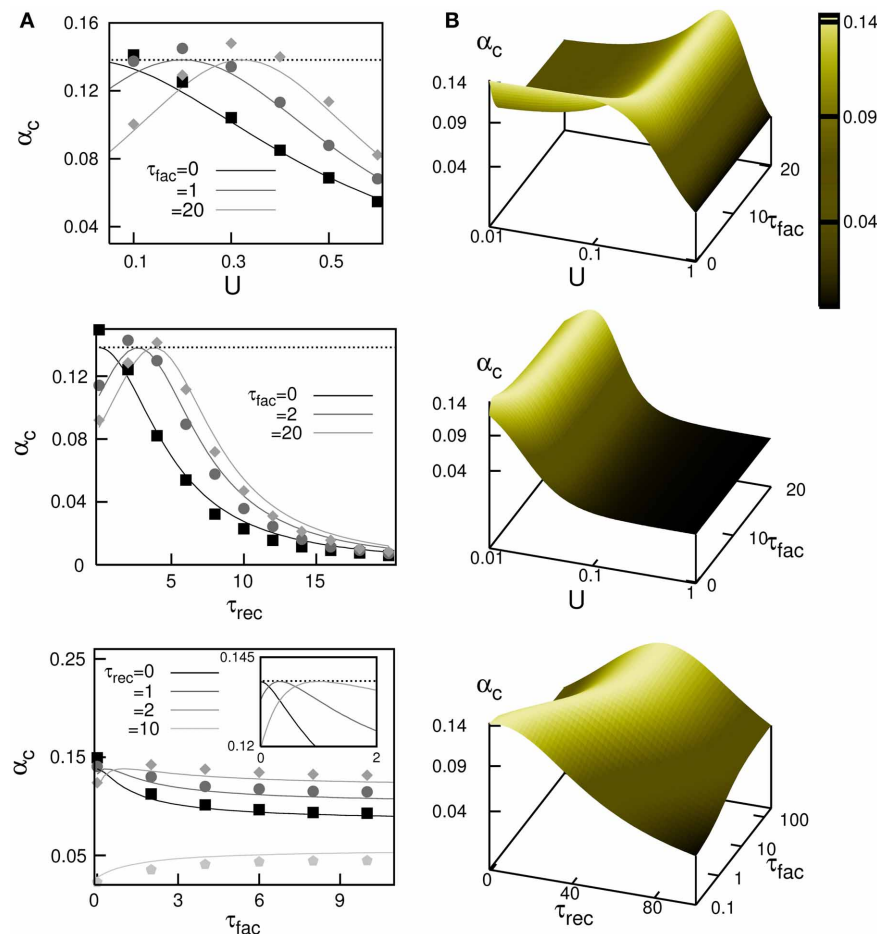
FIGURE 3 | Maximum storage capacity obtained in attractor neural networks with dynamic synapses with both depressing and facilitating mechanisms. (A) Behaviour of α_c as a function of U, τ_{rec} and τ_{fac}. The solid lines correspond the theoretical prediction from the mean field Equation (14) and symbols are obtained from Monte Carlo simulations. From top to bottom, it is shown $\alpha_c(U)$ for $\tau_{rec} = 2$ and different values of τ_{fac}, $\alpha_c(\tau_{rec})$ for $U = 0.2$ and different values of τ_{fac} and $\alpha_c(\tau_{fac})$ for $U = 0.2$ and different values of τ_{rec}, respectively. The

horizontal dotted lines correspond to the static synapses limit $\alpha_c \approx 0.138$. **(B)** Mean-field results from Equation (14) for the dependence of α_c for different combinations of relevant parameters. This corresponds—from top to bottom—to the surfaces $\alpha_c(U, \tau_{fac})$ for $\tau_{rec} = 2$, $\alpha_c(U, \tau_{fac})$ for $\tau_{rec} = 50$ and $\alpha_c(\tau_{rec}, \tau_{fac})$ for $U = 0.02$. In all panels, τ_{rec} and τ_{fac} are given in MCS units that can be associated to a value of 5 ms if one assumes that a MCS corresponds to the duration of the refractory period in actual neurons.

where τ_v is the time constant for the neuron dynamics, v_m is the maximum synaptic input to the neuron population, J is the (static) synaptic strength and θ is the neuron threshold. The function $\mathcal{S}[X]$ is a sigmoidal function which models the excitability of neurons in the population.

The synaptic input from other neurons is modulated by a short-term dynamic synaptic process $x(t)$ which satisfies the stochastic mean field equation

$$\frac{dx(t)}{dt} = \frac{1 - x(t)}{\tau_r} - U\,x(t)v(t) + \frac{D}{\tau_r}\xi(t). \quad (16)$$

The parameters τ_r, U and D are, respectively, the recovery time constant for the stochastic short-term synaptic plasticity mechanism, a parameter related with the reliability of the synaptic transmission (the average release probability in the population)

and the amplitude of this synaptic noise. The explanation of each term appearing in the rhs of Equation (16) is the following: the first term accounts for the slow recovery of neurotransmitter resources, the second term represents a decrease of the available neurotransmitter due to the level of activity in the population and the third term is a noise term that accounts for all possible sources of noise affecting transmission of information at the synapses of the population and that remains at the mesoscopic level.

A complete analysis of this model, both theoretically and by numerical simulations, shows the appearance of complex transitions between high (up) and low (down) neural activity states driven by the synaptic noise $x(t)$, with permanence times in the up state distributed according to a power-law for some range of the synaptic dynamic parameters. The main results of this study are summarized in **Figure 4**. On **Figure 4A**, a typical time series of the temporal behavior of the mean neural activity $v(t)$ of the
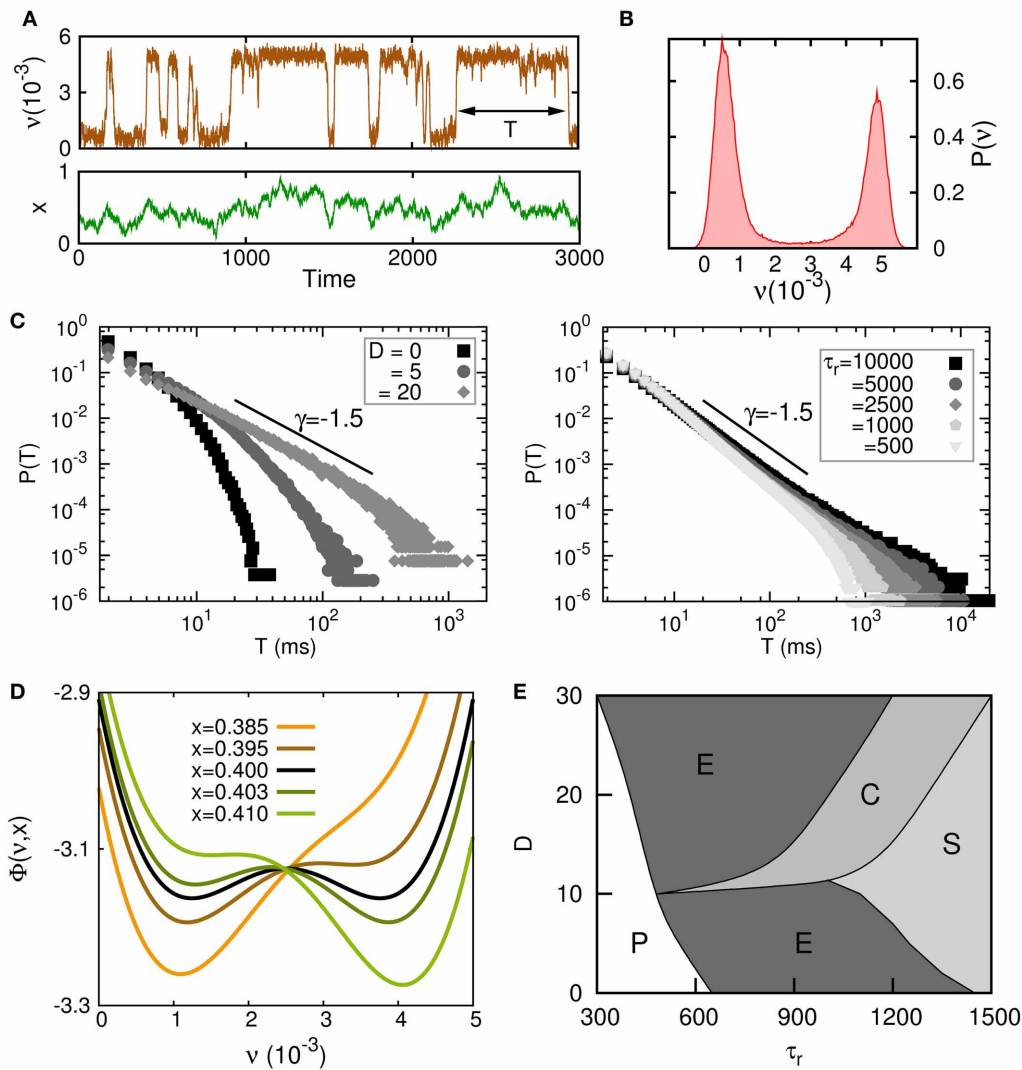
FIGURE 4 | Criticality in up–down transitions. (A) Typical times series for the neuron population rate variable $\nu(t)$ and the mean depression variable $x(t)$ in the neuron population when irregular up–down transitions emerge. Parameter values were $J = 1.2$ V, $\tau_r = 1000\,\tau_\nu$ $U = 0.6$, $D = 0$, $\delta = 0.3$, and $\nu_m = 5 \cdot 10^{-3}$. **(B)** Histogram of the same time series for $\nu(t)$ which presents bimodal features corresponding to two different levels of activity. **(C)** Transitions from exponential to power law behavior for the probability distribution for the permanence time in the up or down state $P(T)$ when parameters D (left panel)

and τ_r (right panel) are varied. Model parameters were the same than in panel **(A)** except that $J = 1.1$ V in the left panel and $U = 0.04$ and $D/\tau_r = 0.02/\tau_\nu$ in the right panel. **(D)** A variation of $x(t)$ induces a change in the shape of the potential function Φ—driving the dynamics of the rate variable $\nu(t)$—which causes transitions between the up and down states. Parameters were the same than in panel **(A)** except that $J = 1.1$ V. **(E)** Complete phase diagram (D, τ_r), for the same set of parameters than in panel **(D)**, where different phases characterize different dynamics of $\nu(t)$, $x(t)$ (see main text for the explanation).

system in the regime in which irregular up–down transitions occur is depicted. In **Figure 4B**, the histogram of $\nu(t)$ for this time series shows a clear bimodal shape corresponding to the two only possible states for $\nu(t)$. **Figure 4C** shows how the parameters τ_r and D, that control the stochastic dynamics of $x(t)$, also are relevant for the appearance of power law distributions $P(T)$ for the permanence time in the up or down state T. As is outlined in (Mejias et al., 2010), the dynamics can be approximately described in an adiabatic approximation, in which the neuron dynamics is subject to an effective potential Φ. **Figure 4D** shows how Φ changes for different values of the mean synaptic depression x.

For relatively small x (orange and brown lines) all synapses in the population have a strong degree of depression and the population has a small level of activity, that is, the global minimum of the potential function is the low-activity state (the down state). On the other hand, when synapses are poorly depressed and x takes relatively large values (dark and light green lines) the neuron activity level is high and the potential function has its global minimum in a high-activity state (up state). For intermediate values of x (black line) the potential becomes bistable. **Figure 4E** shows the complete phase diagram of the system and illustrates the regions in the parameter space (D, τ_r) where different behaviors emerge.

In the phase (P) no transition between a high-activity state and low-activity state occurs. In phase (E) transitions between up and down states are exponentially distributed. The phase (C) is characterized by the emergence of power-law distributions $P(T)$, and therefore is the most intriguing phase since it could be associated to a critical state. Finally, phase (S) is characterized by a highly fluctuating behavior of both $v(t)$ and $x(t)$. In fact, $v(t)$ is behaving as a slave variable of $x(t)$ and, therefore, it presents the dynamical features of the dynamics (Equation 16), which has some similarities with those of colored noise for U small. In fact for $U = 0$, and making the change $z(t) = x(t) - 1$ the dynamics (Equation 16) transforms in that for an Ornstein–Uhlenbeck (OU) process (van Kampen, 1990).

From these studies, we can conclude that the experimentally observed large fluctuations in up and down permanence times in the cortex can be explained as the result of sufficiently noisy dynamical synapses (large D) with sufficiently large recovery times (large τ_r). Static synapses ($\tau_r = 0$) or dynamical synapses in the absence of noise ($D = 0$) cannot account for this behavior, and only exponential distributions for $P(T)$ emerge in this case.

5. STOCHASTIC MULTIRESONANCE

In section 2 we mentioned that short-term synaptic plasticity induces the appearance of dynamic memories as the consequence of the destabilization of memory attractors due to synapse fatigue. The synaptic fatigue in turn is due to strong neurotransmitter vesicle depletion as the consequence of high frequency presynaptic activity and large neurotransmitter recovering times. Also, we concluded that this fact induces a high sensitivity of the system to respond to external stimuli, even if the stimulus is very weak and in the presence of noise. The source of the noise can be due to the neural dynamics as well as the synaptic transmission. It is the combination of non-linear dynamics and noise that causes the enhanced sensitivity to external stimuli. This general phenomenon is the so called *stochastic resonance* (SR) (Benzi et al., 1981; Longtin et al., 1991).

In a set of recent papers we have studied the emergence of SR in feed-forward neural networks with dynamic synapses (Mejias and Torres, 2011; Torres et al., 2011). We considered a post-synaptic neuron which receives signals from a population of N presynaptic neurons through dynamic synapses modeled by Equations (1, 2). Each one of these presynaptic neurons fires a train of Poisson distributed APs with a given frequency f_n. In addition the postsynaptic neuron receives a weak signal $S(t)$ which we can assume sinusoidal. In addition, we assume a stationary regime, where the dynamic synapses have reached their asymptotic values $u_\infty = \frac{U + U\tau_{fac}f_n}{1 + U\tau_{fac}f_n}$ and $x_\infty = \frac{1}{1 + u_\infty \tau_{rec}f_n}$. If all presynaptic neurons fire independently the total synaptic current is a noisy quantity with mean $\bar{I}_N$ and variance σ_N^2 given by

$$\bar{I}_N = N f_n \tau_{in} I_p$$
$$\sigma_N^2 = \frac{1}{2} N f_n \tau_{in} (I_p)^2 \tag{17}$$

with $I_p = A u_\infty x_\infty$ and A the synaptic strength. To explore the possibility of SR, we vary the firing frequency of the presynaptic population f_n. The reason for this choice is that varying f_n

changes the output variance σ_N^2 and f_n can also be relatively easily controlled in an experiment.

To quantify the amount of signal that is present in the output rate we use the standard input–output cross-correlation or *power norm* (Collins et al., 1995) during a time interval Δt and defined as:

$$C_0 = \langle S(t)v(t) \rangle = \frac{1}{\Delta t} \int_t^{t+\Delta t} S(t)v(t)dt, \tag{18}$$

where $v(t)$ is the firing rate of the post-synaptic neuron. The behavior of C_0 as a function of f_n for static synapses is depicted in **Figure 5A** which clearly shows a resonance peak at certain non-zero input frequency f_n. The output of the postsynaptic neuron at the positions in the frequency domain labeled with "a," "b," and "c" is illustrated in **Figure 5B** and compared with the weak input signal. This shows how stochastic resonance emerges in this system. For low firing frequency (case labeled with "a") in the presynaptic population the generated current is so small that the postsynaptic neuron only has sub-threshold behavior weakly correlated with $S(t)$. For very large f_n (case labeled with "c") both $\bar{I}_N$ and σ_N^2 are large and the postsynaptic neuron is firing all the time, so it can not detect the temporal features of $S(t)$. However, there is an optimal value of f_n at which the postsynaptic neuron fires strongly correlated with $S(t)$; in fact it fires several APs each time a maximum in $S(t)$ occurs (case labeled with "b").

This behavior dramatically changes when dynamic synapses are considered, as is depicted in **Figures 5C,D**. In fact, for dynamic synapses there are two frequencies at which resonance occurs. That is, short-term synaptic plasticity induces the appearance of stochastic multi-resonances (SMR). Interestingly, the position of the peaks is controlled by the parameters that control the synapse dynamics. For instance, in **Figure 5C** it is shown how for a fixed value of facilitation and increasing depression (increasing τ_{rec}) the second resonance peak moves toward low values of f_n while the position of the first resonance peak remains unchanged. On the other hand, for a given value of depression, the increase of facilitation time constant τ_{fac} moves the first resonance peak while the position of the second resonance peak is unaltered (see **Figure 5D**). This clearly demonstrates that in actual neural systems synapses with different levels of depression and facilitation can control the signal processing at different frequencies.

The appearance of SMR in neural media with dynamic synapses is quite robust: SMR also appears when the post-synaptic neuron is model with different types of spiking mechanisms, such as the FitzHugh–Nagumo (FHN) model or the integrate and fire model (IF) with an adaptive threshold dynamics (Mejias and Torres, 2011). SMR also appears with more realistic stochastic dynamic synapses and more realistic weak signals such as a train of inputs with small amplitude and short durations distributed in time according to a rate modulated Poisson process (Mejias and Torres, 2011).

The physical mechanism behind the appearance of SMR is the existence of a non-monotonic dependence of the synaptic current fluctuations with f_n—due to the dynamic synapses—together with the existence of an adaptive threshold mechanism in the postsynaptic neuron to the incoming synaptic current. In this

FIGURE 5 | Appearance of stochastic multiresonances in feed forward neural networks of spiking neurons with dynamic synapses. (A) Behaviour of C_0—defined in Equation (18)—as a function of f_n for static synapses showing the phenomenon of stochastic resonance. **(B)** Temporal behavior for the response of the postsynaptic neuron at each labeled position of the resonance curve in panel **(A)**. **(C)** Resonance curve for C_0 when dynamic synapses are included. The most remarkable feature is the appearance of a two-peak resonance in the frequency domain, with the position of high frequency peak controlled by the particular value of τ_{rec}. **(D)** The panel shows another interesting feature of the two-peak resonance curve for C_0, that is, the control of the position of the low frequency peak by τ_{fac}.

way, the distance in voltage between the mean post-synaptic sub-threshold voltage and the threshold for firing remains constant or decreases very slowly for increasing presynaptic frequencies. This implies the existence of two values of f_n at which current fluctuations are enough to induce firing in the post-synaptic neuron [see Mejias and Torres (2011) for more details].

In light of these findings, we have reinterpreted recent SR experimental data from psycho-physical experiments on human blink reflex (Yasuda et al., 2008). In these experiments the neurons responsible for the blink reflex receive inputs from neurons in the auditory cortex, which are assumed to be uncorrelated due to the action of some external source of white noise. The subject received in addition a weak signal in the form of a periodic small air puff into the eyes. The authors measured the correlation between the air puff signal and the blink reflex and their results are plotted in **Figure 6A** (dark gray square error-bar symbols). They used a feed-forward neural network with a postsynaptic neuron with IF dynamics with fixed threshold to interpret their findings (light-gray dashed line). With this model, only the high-frequency correlation points can be fitted. Using instead a FHN model or an IF with adaptive threshold dynamics, we were able

to fit all experimental data points (black solid line). The SMR is also observed with more realistic rate-modulated weak Poisson pulses (light-gray filled circles) instead of the sinusoidal input (black solid line). Both model predictions are consistent with the SMR that is observed in this experiments. In **Figure 6B** we summarize the conditions that neurons and synapses must satisfy for the emergence of SMR in a feed forward neural network.

6. RELATION WITH OTHER WORKS

The occurrence of non-fixed point behavior in recurrent neural networks due to dynamic synapses has also been reported by others (Senn et al., 1996; Tsodyks et al., 1998; Dror and Tsodyks, 2000). These studies differ from our work because one assumes continuous deterministic neuron dynamics (instead of binary and stochastic, as in our work). The oscillations observed in these networks do not have the rapid switching behavior as we observe and seem unrelated to the metastability that we have found in our work.

In addition, it has been reported that oscillations in the firing rate can be chaotic (Senn et al., 1996; Dror and Tsodyks, 2000) and present some intermittent behavior that resembles observed

FIGURE 6 | (A) Appearance of stochastic multi-resonance in experiments in the brain. Dark gray square symbols represent the values of C_0 obtained in the experiments performed in the human auditory cortex. Dashed light gray line corresponds to best model prediction using a neuron with fixed threshold (Yasuda et al., 2008). Solid black line correspond to our model consisting of a FHN neuron and depressing synapses. Gray filled circle symbols shows C_0 when the weak signal is a train of (uncorrelated) Poisson pulses instead of the sinusoidal input (solid line). **(B)** Schematic overview showing the neuron and synapse mechanisms needed for the appearance of stochastic multi-resonances in feed-forward neural networks. (see (Torres et al., 2011) for more details).

patterns of EEG. The chaotic regime in these continuous models seems unrelated to the existence of fixed point behavior and most likely understood as a generic feature of non-linear dynamical systems.

It is worth noting that for each neuron, the effect of dynamic synapses is modeled through a single variable x_i that multiplies the synaptic strength w_{ij} for all synapses that connect to i. There is one depression variable per *neuron* and not per connection. As a result, one can obtain the same behavior of the network by interpreting x_i as implementing a dynamic firing thresholds (Horn and Usher, 1989) instead of a dynamic synapse.

The switching behavior that we described in this paper, is somewhat similar to the neural network with chaotic neurons that displays a self-organized chaotic transition between memories (Tsuda et al., 1987; Tsuda, 1992).

The possible interpretation of the switching behavior as up/down cortical transitions is controversial, because similar

cortical oscillations can be generated without synaptic dynamics, where the up state is terminated because of hyperpolarizing potassium ionic currents (Compte et al., 2003). However, a very recent study has focused on the interplay between synaptic depression and these inhibitory currents and concludes that synaptic depression is relevant for maintaining the up state (Benita et al., 2012). The reason for that counterintuitive behavior is that synaptic depression decreases the firing rate in the up state which also decreases the effect of the hyper-polarizing potassium currents and, as a consequence, the prolongation of the up state.

Related also is a recent study on the effect of dynamic synapses on the emergence of a coherent periodic rhythm within the Up state which results in the phenomenon of *stochastic amplification* (Hidalgo et al., 2012). It has been shown that this rhythm is an emergent or collective phenomenon given that individual neurons in the up state are unlocked to such a rhythm.

The relation between dynamic synapses and storage capacity has also been studied by others. For very sparse stored patterns ($a \ll 1$) it has been shown that storage capacity decreases with synaptic depression (Bibitchkov et al., 2002), in agreement with our findings. On the other hand, it has been reported that the basin of attraction of the memories are enlarged by synaptic depression (Matsumoto et al., 2007) and these are even enlarged more when synaptic facilitation is taken into account (Mejias and Torres, 2009).

(Otsubo et al., 2011) reported a theoretical and numerical study on the role of short-term depression on memory storage capacity in the presence of noise, showing that noise reduces the storage capacity (as is also the case for static synapses). (Mejias et al., 2012) shows the important role of facilitation to enlarge the regions for memory retrieval even in the presence of high noise.

In the last decade there has been some discussion whether neural systems, or even the brain as a whole, can work in a critical state using the notion of self-organized criticality (Beggs and Plenz, 2003; Tagliazucchi et al., 2012). As we stated in section 4, the combination of colored synaptic noise and short-term depression can cause power-low distributed permanence times in the Up and Down states, which is a signature of criticality. The emergence of critical phenomena as a consequence of dynamic synapses has also been explored by others (Levina et al., 2007, 2009; Bonachela et al., 2010; Millman et al., 2010).

Finally, it is worth mentioning a recent work that has investigated the formation of spatio-temporal structures in an excitatory neural network with depressing synapses (Kilpatrick and Bressloff, 2010). As a result of dynamic synapses, robust complex spatio-temporal structures, including different types of travelling waves, appear in such a system.

7. CONCLUSIONS

It is well-known that during transmission of information, synapses show a high variability with a diverse origin, such as the stochastic release and transmission of neurotransmitter vesicles, variations in the Glutamate concentration through synapses and the spatial heterogeneity of the synaptic response in the dendrite tree (Franks et al., 2003). The cooperative effect of all these mechanisms is a noisy post-synaptic response which depends on past pre-synaptic activity. The strength of the postsynaptic response

can decrease or increase and can be modeled as dynamical synapses.

In a large number of papers, we have studied the effect of dynamical synapses in recurrent an feed-forward networks, the result of which we have summarized in this paper. The main findings are the following:

Dynamic memories: Classical neural networks of the Hopfield type, with symmetric connectivity, display attractor dynamics. This means that these networks act as memories. A specific set of memories can be stored as attractors by Hebbian learning. The attractors are asymptotically stable states. The effect of synaptic depression in these networks is to make the attractors lose stability. Oscillatory modes appear where the network rapidly switches between memories. Instead, the permanence time to stay in a memory can have any positive value and becomes infinite in the regime where memories are stable. Thus, the recurrent network with dynamical synapses implements a form of dynamical memory.

Input sensitivity: The classical Hopfield network is relatively insensitive to external stimuli, once it has converged into one of its stable memories. Synaptic depression improves the sensitivity to external stimuli, because it destabilizes the memories. In addition, synaptic facilitation further improves the sensitivity of the attractor network to external stimuli.

Storage capacity: The storage capacity of the attractor neural network, i.e., the maximum number of memories that can be stored in a network, is proportional to the number of neurons N and scales as $P_{max} = \alpha N$ with $\alpha = 0.138$. Synaptic depression causes a decrease of the maximum storage capacity but facilitation allows to recover the capacity of the network with static synapses under some conditions.

Up and down states: The emergence of dynamic memories has been related to the well-known up–down transitions observed in local-field recording in the cortex. We demonstrated that the observed distributions of permanence times can be explained by a stochastic synaptic dynamics. Scale free permanence time distributions could signal a critical state in the brain.

Stochastic multiresonance: Whereas static synapses in a stochastic network give rise to a single stochastic resonance peak, dynamical synapses produce a double resonance. This phenomenon is robust for different types of neurons and input signals. Thus, dynamic synapses may explain recently observed SMR in psychophysical experiments. SMR also seems to occur in recurrent neural networks with dynamic synapses as it has been recently reported (Pinamonti et al., 2012). This work demonstrates the relevant role of short-term synaptic plasticity for the appearance of the SMR phenomenon in recurrent networks, although the exact underlying mechanism behind it is slightly different than in the case described here, namely feed-forward neural networks.

It is important to point out that although the phenomenology reported in this review has been obtained using different models, *all* the reported phenomena can be also derived in a single model consisting in a network of binary neurons with dynamic synapses as described in section 1. The phenomena reported in sections 2 and 3 have in fact been obtained using this model and the phenomenon of stochastic multiresonance (section 5) has been reported recently in such a model by Pinamonti et al. (2012). The results on critical up and down states that are reported in section 4 have been obtained in a mean-field model that can be derived from the same binary model and by assuming in addition sparse neural activity and sparse connectivity, which increases the stochasticity in the synaptic transmission through the whole network.

In addition, our studies show that the reported phenomena are robust to detailed changes in the model, such as replacing the binary neurons by graded response neurone or integrate-and-fire neurone.

REFERENCES

Abbott, L. F., Valera, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224.

Amari, S. (1972). Learning patterns and pattern sequences by self-organizing nets of threshold elements. *IEEE Trans. Comput.* 21, 1197–1206.

Amit, D. J. (1989). *Modeling Brain Function: The World of Attractor Neural Network.* Cambridge, UK: Cambridge University Press.

Barak, O., and Tsodyks, M. (2007). Persistent activity in neural networks with dynamic synapses. *PLoS Comput. Biol* 3:e35. doi: 10.1371/journal.pcbi.0030035

Beggs, J. M., and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *J. Neurosci.* 23, 11167–11177.

Benita, J. M., Guillamon, A., Deco, G., and Sanchez-Vives, M. V. (2012). Synaptic depression and slow oscillatory activity in a biophysical network model of the cerebral cortex. *Front. Comp. Neurosci.* 6:64. doi: 10.3389/fncom.2012.00064

Benzi, R., Sutera, A., and Vulpiani, A. (1981). The machanims of stochastic resonance. *J. Phys. A Math. Gen.* 14, L453.

Bertram, R., Sherman, A., and Stanley, E. F. (1996). Single-domain/bound calcium hypothesis of transmitter release and facilitation. *J. Neurophysiol.* 75, 1919–1931.

Bibitchkov, D., Herrmann, J. M., and Geisel, T. (2002). Pattern storage and processing in attractor networks with short-time synaptic dynamics. *Netw. Comput. Neural Syst.* 13, 115–129.

Bonachela, J. A., de Franciscis, S., Torres, J. J., and Muñoz, M. A. (2010). Self-organization without conservation: are neuronal avalanches generically critical? *J. Stat. Mech. Teor Exp.* 2010:P02015. doi: 10.1088/1742-5468/2010/02/P02015

Bressloff, P. C. (1999). Mean-field theory of globally coupled integrate-and-fire neural oscillators with dynamic synapses. *Phys. Rev. E* 60, 2160–2170.

Collins, J. J., Carson Chow, C. C., and Imhoff, T. T. (1995). Aperiodic stochastic resonance in excitable systems. *Phys. Rev. E* 52, R3321–R3324.

Compte, A., Sanchez-Vives, M. V., McCormick, D. A., and Wang. X.-J. (2003). Cellular and network mechanisms of slow oscillatory activity (<1 hz) and wave propagations in a cortical network model. *J. Neurophysiol.* 89, 2707–2725.

Cortes, J. M., Garrido, P. L., Marro, J., and Torres, J. J. (2004). Switching between memories in neural

automata with synaptic noise. *Neurocomputing* 58–60. 67–71.

Cortes, J. M., Torres, J. J., Marro, J., Garrido, P. L., and Kappen, H. J. (2006). Effects of fast presynaptic noise in attractor neural networks. *Neural Comput.* 18, 614–633.

Dror, G., and Tsodyks, M. (2000). Chaos in neural networks with dynamic synapses. *Neurocomputing* 32–33, 365–370.

Franks, K. M., Stevens, C. F., and Sejnowski, T. J. (2003). Independent sources of quantal variability at single glutamatergic synapses. *J. Neurosci.* 23, 3186–3195.

Hertz, J., Krogh, A., and Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation*. Redwood City, CA: Addison-Wesley.

Hidalgo, J., Seoane, L. F., Cortés, J. M., and Muñoz, M. A. (2012). Stochastic amplification of fluctuations in cortical up-states. *PLoS ONE* 7:e40710. doi: 10.1371/journal.pone.0040710

Holcman, D., and Tsodyks, M. (2006). The emergence of up and down states in cortical networks. *PLoS Comput. Biol.* 2, 174–181. doi: 10.1371/journal.pcbi.0020023

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.

Horn, D., and Usher, M. (1989). Neural networks with dynamical thresholds. *Phys. Rev. A* 40, 1036–1044.

Kilpatrick, Z. P., and Bressloff, P. C. (2010). Spatially structured oscillations in a two-dimensional excitatory neuronal network with synaptic depression. *J. Comput. Neurosci.* 28, 193–209.

Kistler, W. M., and van Hemmen, J. L. (1999). Short-term synaptic plasticity and network behavior. *Neural Comput.* 11, 1579–1594.

Levina, A., Herrmann, J. M., and Geisel, T. (2007). Dynamical synapses causing self-organized criticality in neural networks. *Nat. Phys.* 3, 857–860.

Levina, A., Herrmann, J. M., and Geisel, T. (2009). Phase transitions towards criticality in a neural system with adaptive interactions. *Phys. Rev. Lett.* 102, 118110.

Liaw, J. S., and Berger, T. W. (1996). Dynamic synapse: a new concept of neural representation and computation. *Hippocampus* 6, 591–600.

Longtin, A., Bulsara, A., and Moss, F. (1991). Time-interval sequences in bistable systems and the noise-induced transmission of information by sensory neurons. *Phys. Rev. Lett.* 67, 656–659.

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382, 759–760.

Markram, H., Wang, Y., and Tsodyks. M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5323–5328.

Marro, J., Torres, J. J., and Cortes, J. M. (2007). Chaotic hopping between attractors in neural networks. *Neural Netw.* 20, 230–235.

Matsumoto, N., Ide, D., Watanabe, M., and Okada, M. (2007). Retrieval property of attractor network with synaptic depression. *J. Phys. Soc. Jpn.* 76, 084005.

Mejias, J. F., Hernandez-Gomez, B., and Torres, J. J. (2012). Short-term synaptic facilitation improves information retrieval in noisy neural networks. *EPL (Europhys. Lett.)* 97, 48008.

Mejias, J. F., Kappen, H. J., and Torres, J. J. (2010). Irregular dynamics in up and down cortical states. *PLoS ONE* 5:e13651. doi: 10.1371/journal.pone.0013651

Mejias, J. F., and Torres, J. J. (2009). Maximum memory capacity on neural networks with short-term depression and facilitation. *Neural Comput.* 21, 851–871.

Mejias, J. F., and Torres, J. J. (2011). Emergence of resonances in neural systems: the interplay between adaptive threshold and short-term synaptic plasticity. *PLoS ONE* 6:e17255. doi: 10.1371/journal.pone.0017255

Millman, D., Mihalas, S., Kirkwood, A., and Niebur, E. (2010). Self-organized criticality occurs in non-conservative neuronal networks during up states. *Nat. Phys.* 6, 801–805.

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science* 319, 1543–1546.

Natschläger, T., Maass, W., and Zador, A. (2001). Efficient temporal processing with biologically realistic dynamic synapses. *Netw. Comput. Neural Syst.* 12, 75–87.

Otsubo, Y., Nagata, K., Oizumi, M., and Okada, M. (2011). Influence of synaptic depression on memory storage capacity. *J. Phys. Soc. Jpn.* 80, 084004.

Pantic, L., Torres, J. J., Kappen, H. J., and Gielen, S. C. A. M. (2002).

Associative memory with dynamic synapses. *Neural Comput.* 14, 2903–2923.

Pieribone, V. A., Shupliakov, O., Brodin, L., Hilfiker-Rothenfluh, S., Zernik, A. J., and Greengard, P. (1995). Distinct pools of synaptic vesicles in neurotransmitter release. *Nature* 375, 493–497.

Pinamonti, G., Marro, J., and Torres, J. J. (2012). Stochastic resonance crossovers in complex networks. *PLoS ONE* 7:e51170. doi: 10.1371/journal.pone.0051170

Senn, W., Wyler, K., Streit, J., Larkum, M., Lüscher, H. r., Merz, F. et al. (1996). Dynamics of random neural network with synaptic depression. neural networks. *Neural Netw.* 9, 575–588.

Senn, W., Segev, I., and Tsodyks, M. (1998). Reading neuronal synchrony with depressing synapses. *Neural Comput.* 10, 815–819.

Steriade, M., McCormick, D. A., and Sejnowski. T. J., (1993a). Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262, 679–685.

Steriade, M., Nunez, A., and Amzica, F. (1993b). A novel slow (<1hz) oscillation of neocortical neurons *in vivo*: depolarizing and hyperpolarizing components. *J. Neurosci.* 13, 3252–3265.

Tagliazucchi, E., Balenzuela, P., Fraiman, D., and Chialvo, D. R. (2012). Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Front. Physiol.* 3:15. doi: 10.3389/fphys.2012.00015

Torres, J. J., Cortes, J. M., and Marro, J. (2005). Instability of attractors in auto-associative networks with bioinspired fast synaptic noise. *LNCS* 3512, 161–167.

Torres, J. J., Cortes, J. M., Marro, J., and Kappen, H. J. (2008). Competition between synaptic depression and facilitation in attractor neural networks. *Neural Comput.* 19, 2739–2755.

Torres, J. J., Marro, J., and Mejias, J. F. (2011). Can intrinsic noise induce various resonant peaks? *New J. Phys.* 13:053014. doi: 10.1088/1367-2630/13/5/053014

Torres, J. J., Pantic, L., and Kappen, H. J. (2002). Storage capacity of attractor neural networks with depressing synapses. *Phys. Rev. E.* 66:061910. doi: 10.1103/PhysRevE.66.061910

Tsodyks, M., Uziel, A., and Markram, H. (2000). Synchrony generation in recurrent networks with

frequency-dependent synapses. *J. Neurosci.* 20, RC50 (1–5).

Tsodyks, M. V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U.S.A.* 94, 719–723.

Tsodyks, M. V., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835.

Tsuda, I. (1992). Dynamic link of memory–chaotic memory map in nonequilibrium neural networks. *Neural Netw.* 5, 313–326.

Tsuda, I., Koerner, E., and Shimizu, H. (1987). Memory dynamics in asynchronous neural networks. *Prog. Theor. Phys.* 78, 51–71.

van Kampen, N. G. (1990). *Stochastic Processes in Physics and Chemistry*. Amsterdam: North-Holland Personal Library (Elsevier).

Yasuda, H., Miyaoka, T., Horiguchi, J., Yasuda, A., Hanggi, P., and Yamamoto, Y. (2008). Novel class of neural stochastic resonance and error-free information transfer. *Phys. Rev. Lett.* 100:118103. doi: 10.1103/PhysRevLett.100.118103

Zucker, R. S. (1989). Short-term synaptic plasticity. *Annu. Rev. Neurosci.* 12, 13–31.

Zucker, R. S., and Regehr, W. G. (2002). Short-term synaptic plasticity. *Annu. Rev. Physiol.* 64, 355–405.

Mathematical analysis and algorithms for efficiently and accurately implementing stochastic simulations of short-term synaptic depression and facilitation

Mark D. McDonnell[1]*, Ashutosh Mohan[2] and Christian Stricker[2]

[1] Computational and Theoretical Neuroscience Laboratory, Institute for Telecommunications Research, University of South Australia, Mawson Lakes, SA, Australia
[2] John Curtin School of Medical Research, Australian National University, Canberra, ACT, Australia

The release of neurotransmitter vesicles after arrival of a pre-synaptic action potential (AP) at cortical synapses is known to be a stochastic process, as is the availability of vesicles for release. These processes are known to also depend on the recent history of AP arrivals, and this can be described in terms of time-varying probabilities of vesicle release. Mathematical models of such synaptic dynamics frequently are based only on the mean number of vesicles released by each pre-synaptic AP, since if it is assumed there are sufficiently many vesicle sites, then variance is small. However, it has been shown recently that variance across sites can be significant for neuron and network dynamics, and this suggests the potential importance of studying short-term plasticity using simulations that do generate trial-to-trial variability. Therefore, in this paper we study several well-known conceptual models for stochastic availability and release. We state explicitly the random variables that these models describe and propose efficient algorithms for accurately implementing stochastic simulations of these random variables in software or hardware. Our results are complemented by mathematical analysis and statement of pseudo-code algorithms.

Keywords: short term synaptic dynamics, short term depression, facilitation, stochastic simulation, stochastic synapse, vesicle site model, synaptic plasticity models, short term plasticity

1. INTRODUCTION

The release of vesicles following arrival of a pre-synaptic action potential (AP) at a synapse is inherently probabilistic (Vere-Jones, 1966; Melkonian and Kostopoulos, 1996; Branco and Staras, 2009). The amount of neurotransmitter released by each AP can also vary stochastically over time, in a manner dependent on the timing of previously arriving APs (Dobrunz and Stevens, 1997). These effects result in what is called short-term synaptic plasticity (Zucker and Regehr, 2002; Klug et al., 2012; Regehr, 2012). It has been suggested that the short term dynamics such plasticity introduces may play an important role in information processing in the cortex (Abbott and Regehr, 2004; Branco and Staras, 2009). This has been demonstrated in studies of the influence of short-term plasticity on: gain control (Abbott et al., 1997); coding and detection mechanisms (Tsodyks and Markram, 1997; Maass and Zador, 1999); filtering effects (Matveev and Wang, 2000a; Merkel and Lindner, 2010; Rosenbaum et al., 2012); redundancy reduction (Goldman et al., 2002); information transmission (Goldman, 2004); membrane potential estimation (Pfister et al., 2010); attractor networks (Fung et al., 2012); and correlations in neural activity (Rosenbaum et al., 2013).

Popular mathematical models of short term synaptic plasticity effects, such as depression and facilitation, typically are expressed in term of differential equations that describe how the mean number of available and/or released vesicles changes with time in response to pre-synaptic spiking (Tsodyks and Markram, 1997; Tsodyks et al., 1998). The mean is an ensemble-average over multiple repeats of the same pre-synaptic spike train, and is often the focus of study because if the number of vesicles in the model is large, the variance across trials is small and assumed to be negligible in its impact. The consequence of this assumption is that simulations of this type of model of short term plasticity provides deterministic outcomes, in the sense that they do not produce varying outcomes if repeated trials with identical initial conditions are simulated.

However, variability in the number of vesicles available/released has also been studied mathematically (Vere-Jones, 1966), as has the covariance in the response to consecutive pre-synaptic APs (Quastel, 1997). Recently, it has been shown mathematically that explicit inclusion of the variance in models of short-term plasticity leads to significant differences in terms of frequency-dependent information transmission, in comparison with models that study only the mean (Rosenbaum et al., 2012). This mathematical finding that variance can be influential is consistent with previous simulation results (discussed in following paragraphs) that found that the mean-model underestimates post-synaptic firing rate (de la Rocha and Parga, 2005).

As well as mathematical analysis, the conceptual models of stochastic vesicle availability and release that these models are based on can also be studied by implementing stochastic simulations. We use the term "stochastic simulation" to mean a software (or, potentially, hardware) implementation that explicitly generates random or pseudo-random numbers for the purposes of simulating outcomes of a model's random variables (Gillesple, 1977).

By doing this, repeated runs with identical initial conditions and identical external input to the model results in randomly varying outcomes, i.e., trial-to-trial variability. Such simulations have, for example, been used to study ion-channel noise and its impact on AP generation (Faisal and Laughlin, 2007).

Although the mean model described above has been used frequently, results based on stochastic simulations of short term plasticity models have also been described previously (Melkonian and Kostopoulos, 1996; Quastel, 1997; Matveev and Wang, 2000b,a; Fuhrmann et al., 2002; de la Rocha and Parga, 2005; Loebel et al., 2009; Rosenbaum et al., 2012, 2013; Scott et al., 2012; Reich and Rosenbaum, 2013) and comparisons between simulations of the deterministic and stochastic models have been shown to give rise to different outcomes in neural activity (de la Rocha and Parga, 2005; Rosenbaum et al., 2012; Scott et al., 2012).

In general, it may be important to implement stochastic simulations for synaptic connections where only a very small number of vesicles are available for release, which is often the case (Branco and Staras, 2009). In this case the mean model might be very inaccurate in scenarios where ensemble averaging across multiple repeated trials is not possible, such as in large network simulations.

As noted above, previous work has published results from stochastic simulations as a complement to mathematical analysis. However, as far as we are aware, the implementation details have not been discussed at a level of detail that will enable researchers whose primary expertise and experience is not in implementing stochastic simulations, or who have little mathematical training, to introduce trial-to-trial variation in simulations.

The primary aim of this paper is, therefore, to articulate precisely how to efficiently implement stochastic simulations that accurately reflect several of the most well-known conceptual models of vesicle availability and release processes. In our discussion, and associated pseudo-code algorithms, we assume that the algorithms would be applied under conditions where the number of vesicles available may be small, and that therefore stochastic simulation of all random variables in the conceptual models may be important. We also aim to present mathematical descriptions of key random variables that must be simulated in stochastic models, as well as relating these descriptions to existing equations describing mean numbers of vesicles. A secondary aim is to show how existing algorithms may be made more efficient and general.

As well as the usual models of release dependent depression and facilitation, the content of this paper is equally applicable to the case of release-independent depression and associated frequency dependent recovery (Fuhrmann et al., 2004; Scott et al., 2012; Mohan et al., 2013).

The paper is organized as follows. In section 2, we review conceptual models that we will use in this paper and in section 3 we mathematically introduce notation to describe the random variables implied by each conceptual model. Next, section 4 contains descriptions of correct and incorrect implementations of stochastic simulations of the conceptual models, and relates these to the random variables we described. Section 5 describes example simulation results, and shows that incorrect implementations can significantly miscalculate the number of vesicles that should be released in response to sequences of pre-synaptic AP arrivals.

Finally, the conclusions drawn from our paper are summarized in section 6.

2. CONCEPTUAL MODELS OF SHORT TERM PLASTICITY

The first step in computational modeling is to state a conceptual model; once stated, a primary goal of computational modeling is to faithfully implement simulations of the conceptual model (Carnevale and Hines, 2005). We therefore first clearly articulate conceptual stochastic models in this section, and discuss algorithms for faithfully implementing stochastic simulations of them in the following sections. Other conceptual models exist, but the ones we consider serve to illustrate important principles that should be reflected in stochastic simulations.

2.1. AVAILABILITY OF A SINGLE VESICLE FOLLOWING RELASE

In this paper we consider two conceptual "release-site" models (Sterratt et al., 2011) for short term synaptic depression, due to stochastic unavailability of a vesicle:

- **Availability Model 1:** In this model it is assumed that once a specific vesicle is released, the time at which it is next available for release is a random variable that depends only on the time since it was released. This random variable is not affected by subsequent pre-synaptic spikes.
- **Availability Model 2:** Like Availability Model 1, it is also assumed that after the vesicle is released, the time that passes before it is next available for release is a random variable. However, now if a pre-synaptic spike arrives before the vesicle becomes available, the time before the unavailable vesicle then becomes available is recalculated in a manner dependent only on the time of the latest pre-synaptic spike.

Note that these models treat a single vesicle as if it is a conserved object that switches between two states. Of course in reality the vesicle is not conserved, and a more accurate description is to say that a vesicle release site that can contain at most a single vesicle either (1) does contain a vesicle, or (2) does not contain one.

Below we show that Availability Model 1 and Availability Model 2 are mathematically equivalent, given an assumption that the random variable describing availability times is exponential. This is a standard assumption, because it provides good fits to experimental data, and therefore underpins models developed in conjunction with experimental data on short term depression (for example, Tsodyks and Markram, 1997). However, it is feasible that better fits to data might discard the exponential assumption, and in that case it would be necessary to consider how stochastic simulations need to differ for each model. As we show below for a non-exponential example (**Figure 6**), the two models provide significantly different outcomes.

2.2. RELEASE OF A SINGLE VESICLE UPON ARRIVAL OF A PRE-SYNAPTIC SPIKE

In this paper we consider two conceptual models for the stochastic release of a single vesicle upon arrival of a pre-synaptic spike:

- **Release Model 1:** In this model it is assumed that if the single vesicle is available, then it is released with a constant probability

upon arrival of a pre-synaptic spike, and this probability does not change over time.

- **Release Model 2:** In this model it is assumed that if the single vesicle is available, then it is released upon arrival of a pre-synaptic spike with a certain time-varying probability.

Release Model 1 is a classical model of probabilistic release (Vere-Jones, 1966). Release Model 2 is appropriate when a synapse is known to exhibit facilitation. Usually, based on experimental evidence (see, for example, Markram et al., 1998), the change in release probability (given availability) over time is modeled as increasing by a percentage of the current probability of non-release, whenever a pre-synaptic spike arrives (usually independently of whether a vesicle is available or released) and then decaying exponentially over time to a constant rest probability, for as long as no more pre-synaptic APs arrive (Tsodyks et al., 1998).

Release Model 2 is also appropriate when a synapse is known to exhibit a different form of short term depression to that modeled by the lack of vesicle availability. In this type of depression, known as "release-independent depression," the probability of vesicle release (given its availability) is reduced by arriving pre-synaptic spikes independently of whether the vesicle is released, due to different mechanisms from those that cause facilitation (Fuhrmann et al., 2004). In some models, facilitation and release-independent depression are assumed to be present simultaneously (Graham and Stricker, 2008; Scott et al., 2012).

2.3. COMBINING AVAILABILITY AND RELEASE
A single vesicle obviously cannot be released if it is not available, but it is assumed that an available vesicle remains available until released. This is the key feature of the conceptual models we study where both availability and release are modeled as stochastic.

2.4. AVAILABILITY AND RELEASE FOR A POOL OF N VESICLES
In this paper, when we consider a conceptual model where there is a pool consisting of at most N vesicle release sites, each containing at most a single vesicle, we use the typical assumption that the release and availability of each single vesicle occurs independently of that in the other vesicle release sites. Note that although this model is typical, it may not always be accurate (Quastel, 1997).

2.5. MULTIPLE TRIALS OF AVAILABILITY AND RELEASE
In this paper, when we consider a conceptual model where there are N repeated trials for the same sequence of pre-synaptic APs, and a single vesicle in a single release site, we assume that the availability or release of the vesicle is independent for each trial.

Note that the outcome for a model where there are N such repeated trials is equivalent mathematically to a conceptual model where there is a pool of N vesicle release sites with at most a single vesicle available, for a single trial of the sequence of APs.

Since an experimental protocol is more amenable to studying repeated trials for a single release site and the same sequence of APs, we will refer to the case of N trials rather than N release sites.

2.6. VESICLE RELEASE SITES CONTAINING MULTIPLE VESICLES
The content of this paper regarding stochastic simulations can be extended to a scenario where multiple vesicles are available in a release site, and also where multiple such sites are available, potentially each with different numbers of vesicles. However, we do not discuss this further, as the most important observations are relevant to sites containing single vesicles. Further discussion of evidence for multiple release sites can be found in Loebel et al. (2009).

3. RANDOM VARIABLES IMPLIED BY STOCHASTIC CONCEPTUAL MODELS
The purpose of this section is to explicitly describe all random variables inherent in the conceptual models we study, since correct stochastic simulations of the models relies on correct simulation of outcomes from these random variables.

3.1. AVAILABILITY MODELS
There is a specific random variable that exists in both conceptual availability models: the time taken for vesicle to become available following a successful release at time $t = t_s$. We label this random variable as T_{a1} for Availability Model 1, and as T_{a2} Availability Model 2.

3.1.1. Availability model 1
In standard existing models, the random variable describing the time until a release site contains an available vesicle, following release of its vesicle, is exponentially distributed with a known mean, τ_a. In this section we generalize to arbitrary positive and continuously valued distributions for the availability time. We write the probability density function describing the random variable T_{a1} as $f_{T_{a1}}(T_{a1} = x)$, and its cumulative distribution function [describing $\text{Prob}(T_{a1} \leq y)$] as $F_{T_{a1}}(y)$. We introduce $P_{a,1}(t|t_s)$ to describe the probability of availability at time t, given that the most recent successful release was at time t_s. We can write

$$P_{a,1}(t|t_s) = F_{T_{a1}}(t - t_s), \, t \geq t_s. \tag{1}$$

Below, we note how this probability describes a distribution of the potential times, immediately following a successful release at $t = t_s$, at which the released vesicle will next become available. However, it is crucial to note that for a stochastic simulation to be faithful to Availability Model 1, the released vesicle must always be in one of two states (available or not available) and that once it switches from not-available to available, it must stay available, until released again. Ignoring this fact can lead to incorrect implementations of the conceptual model.

We now derive an expression for a conditional probability that is potentially useful in some stochastic simulation implementations. Suppose a vesicle was released at the i–th AP. We introduce notation for the time interval between APs, i and j as $\theta_{i,j} = t_{AP,j} - t_{AP,i} > 0$, where j may be any AP after the i–th one. The probability that the vesicle becomes available by the j-th AP is $\text{Prob}(T_{a1} \leq \theta_{i,j}) = F_{T_{a1}}(\theta_{i,j})$, $j = i+1, i+2, \ldots$. However, we also are interested in $\text{Prob}(T_{a1} \leq \theta_{i,j}|T_{a1} > \theta_{i,j-1})$ $j = i+2, i+3, \ldots$, which is the probability that the vesicle does not become available before

the $j-1$–th AP, but does becomes available before the j–th AP. By Bayes' rule, this probability can be written as

$$\text{Prob}(T_{a1} \leq \theta_{i,j}|T_{a1} > \theta_{i,j-1}) = \frac{\text{Prob}(T_{a1} \in [\theta_{i,j-1}, \theta_{i,j}])}{\text{Prob}(T_{a1} > \theta_{i,j-1})}$$

$$= \frac{F_{T_{a1}}(\theta_{i,j}) - F_{T_{a1}}(\theta_{i,j-1})}{1 - F_{T_{a1}}(\theta_{i,j-1})} \quad (2)$$

Special Case: For the case where $F_{T_{a1}}(y) = 1 - \exp(-y/\tau_a)$, i.e., T_{a1} is exponentially distributed with mean equal to τ_a, it is simple to derive $\text{Prob}(T_{a1} \leq \theta_{i,j}|T_{a1} > \theta_{i,j-1}) = F_{T_{a1}}(t_{AP,j} - t_{AP,j-1})$. So, in this special case, the probability of a vesicle becoming available after the j–th spike, given it wasn't available at the time of the $j-1$–th spike, is independent of the time at which the vesicle actually became unavailable in the first place. This observation is actually a well known property of Poisson point processes: events in every increment of time are independent of the past history of the process. These processes have exponentially distributed inter-event distributions, as we assumed in this discussion.

3.1.2. Availability model 2

A direct translation of this conceptual model implies that a random variable must be evaluated for every pre-synaptic AP that arrives while a vesicle remains unavailable. We write the time of the k–th pre-synaptic AP after the most recent release as $t_{AP,k}$, where $k = 0, 1, 2, \ldots, K$, $t_{AP,0} = t_s$ is the time at which the vesicle was previously released, and K is the number of AP arrivals before the vesicle actually becomes available. We write the random variable evaluated at the k–th AP as $T_{a2,k}$.

Under Availability Model 2, we can write that if the vesicle did not become available by the k–th AP, then the conditional probability that a vesicle is available by time t is

$$P_{a,2}(t|t > t_{AP,k}) = \text{Prob}(t_{AP,k} + T_{a2,k} \leq t)$$

$$= F_{T_{a2}}(t - t_{AP,k}), \quad t \in (t_{AP,k}, t_{AP,k+1}], \quad (3)$$

where it is assumed that each $T_{a2,k}$ is drawn independently from the same distribution with cumulative distribution function $F_{T_{a2}}(y)$.

For this model, the probability of availability by time t, given only the most recent release time, t_s is given by

$$P_{a,2}(t|t_s) = 1 - (1 - F_{T_{a2}}(t - t_{AP,K})) \prod_{k=0}^{K-1} \times$$

$$(1 - F_{T_{a2}}(t_{AP,k+1} - t_{AP,k})), \quad (4)$$

which clearly in general is different from $P_{a,1}(t|t_s)$ for Availability Model 1.

This direct translation of the conceptual model to obtain $P_{a,2}(t|t > t_{AP,k})$ suggests a stochastic simulation implementation where a new random number is drawn for an unavailable vesicle, upon every AP arrival. However, if we can derive the cumulative distribution function of the total time to availability under this release model, T_{a2}, a stochastic simulation that only draws a single random number upon every vesicle release is feasible. Such a random variable would have to produce $P_{a,2}(t|t_s)$ according to the above expression, and in general such a random variable is not readily obtainable. The following describes a special case where it is.

Special case: For an exponential distribution of T_a we can easily derive from Equation (4) that

$$P_{a,2}(t|t_s) = 1 - \exp(-(t - t_s)/\tau_a), \quad t \geq t_s. \quad (5)$$

Consequently, by inspection of Equation (1), Availability Model 1 is equivalent to Availability Model 2, for exponential availability times. This equivalence can also be seen by considering Equation (2).

There are, of course, other possible models for the distribution of the release time, such as a Rayleigh or lognormal model, and it is feasible that such models may be a better fit to data than the assumed exponential model. For example, more complex models exist that describe the biophysics of vesicle generation, and how release probability depends on calcium concentration (Meddis, 1986; Sumner et al., 2002; McDonnell et al., 2008). Discussing the accuracy of simplifying such models to the phenomenological model used here is beyond the scope of this paper. In general, however, any non-exponentially distributed T_a will not lead to equivalence between Availability Model 1 and Availability Model 2.

3.2. RELEASE MODELS

There is a specific random variable that exists in both conceptual release models: the event that a vesicle is released, or not released, upon arrival of the i–th pre-synaptic AP at time $t = t_{AP,i}$. We label this random variable as $R(t_{AP,i})$. This random variable is binary, it exists only at each AP time, it depends on the last time at which a vesicle was released, t_s, and we denote its outcomes as α if a vesicle is released and as β if it is not. We denote the probability that the event α occurs at time t, given the vesicle is available, as $P_{r|a}(t)$. The random variable has a probability mass function, and this is given for Availability Model 1 by

$$\text{Prob}(R(t_{AP,i}) = \alpha|t_s) = P_{r|a}(t_{AP,i})F_{T_a}(t_{AP,i} - t_s);$$

$$\text{Prob}(R(t_{AP,i}) = \beta|t_s) = 1 - P_{r|a}(t_{AP,i})F_{T_a}(t_{AP,i} - t_s), \quad (6)$$

where $t_{AP,i} > t_s$, and for Availability Model 2 by

$$\text{Prob}(R(t_{AP,i}) = \alpha|t_s) = P_{r|a}(t_{AP,i})P_{a,2}(t_{AP,i}|t_s);$$

$$\text{Prob}(R(t_{AP,i}) = \beta|t_s) = 1 - P_{r|a}(t_{AP,i})P_{a,2}(t_{AP,i}|t_s). \quad (7)$$

Note that in Release Model 1, $P_{r|a}$ has no time dependence [i.e., $P_{r|a}(t_{AP,i}) = P_{r|a}$], but this is the only difference in comparison with Release Model 2 (see Equation 1). Consequently, provided the release probability has been calculated correctly at each point in time during a simulation, there are no other differences in a stochastic simulation implementation in comparison with Release Model 2.

3.3. COMBINING AVAILABILITY AND RELEASE

A relevant binary-valued stochastic process can be stated based on the random variables described above, namely, the process describing whether a vesicle is available at any point in time. A succinct description of this process is given in Loebel et al. (2009, Equation 4), where it is expressed in terms of a differential equation. Following the notation used in that description, we label the stochastic process as $\sigma(t)$, and let $\sigma(t) = 1$ when the vesicle is available and let $\sigma(t) = 0$ otherwise. The process is fully described by the following equation:

$$\frac{d\sigma(t)}{dt} = -\sigma(t^-)R(t)\delta(t - t_{AP,i})$$
$$+ (1 - \sigma(t^-))\delta(t - t_{AP,i} - T_a), \qquad (8)$$

where the notation $^-$ in t^- is used as shorthand to represent $t^- = t - \epsilon$, where ϵ is a very short time period; thus when $t = t_{AP,i}$ then t^- is the time instant immediately prior to AP i arriving. We have assigned $\alpha = 1$ and $\beta = 0$ as the possible values of the random variable $R(t)$. During intervals of time for which $\sigma(t) = 0$, the right hand side of Equation (8) is just $\delta(t - t_{AP,i} - T_a)$, and mathematically, the remaining terms in Equation (8) describe the fact that $\sigma(t)$ can jump from 0 to 1 only at the time $t = t_{AP,i} + T_a$. Similarly, Equation (8) is such that $\sigma(t)$ can jump from 1 to 0 only when both $R(t) = \alpha = 1$ and $t = t_{AP,i}$, or equivalently, $R(t_{AP,i}) = 1$, which means a vesicle is released when AP i arrives.

Note that in Loebel et al. (2009), the event where $\sigma(t)$ jumps from 0 to 1 is stated to be modeled as a Poisson process. A Poisson process has exponentially distributed times between events, and therefore the conceptual model in Loebel et al. (2009) is in this sense the same as our Availability Model 1 with exponentially distributed T_{a1}, with mean τ_a. However, for an actual Poisson process, events will continue to occur for all time, not just when vesicles are currently unavailable, which is at odds with our stated conceptual model. Nevertheless, it can be inferred that in Loebel et al. (2009) that Poisson events are ignored when $\sigma(t) = 1$.

Does this mean that the distribution of times until a vesicle becomes available is different in each conceptual model, since in the Poisson process, the exponential time to arrival begins at the time of the previous Poisson event, whereas in Availability Model 1 begins at the most recent release time? The answer is no, due to the independence of events in Poisson processes (the same reason that Availability Models 1 and 2 are equivalent for exponentially distributed availability times). Therefore, there will be no difference when a stochastic simulation implementation of the Loebel et al. (2009) conceptual model is carried out, compared with an implementation of our Availability Model 1 with exponentially distributed arrival times. However, if the arrival times are not exponential, and the corresponding non-Poisson process replaces the Poisson process in the Loebel et al. (2009) conceptual model, the results will not be the same.

3.4. DETERMINISTIC MEAN MODELS FOR AVAILABILITY MODELS 1 AND 2

Differential equation notation is often used to express how the mean fraction of available vesicles, $N_a(t)$, changes over time in two ways: either upon a spike arrival, or between spike arrivals (Tsodyks and Markram, 1997; Fuhrmann et al., 2002;

Scott et al., 2012). The typical form of such expressions is

$$\frac{dN_a(t)}{dt} = \frac{1 - N_a(t)}{\tau_a} - N_{r|a}N_a(t^-) \sum_{i=1}^{K} \delta(t - t_{AP,i}),$$

where $t_{AP,i}$ is the arrival time of the i–th AP, out of a total of K, and $N_{r|a}$ is the mean fraction of available vesicles released by the i–th AP. This differential equation can be easily solved in closed form (e.g., Tsodyks and Markram, 1997) to get

$$
\begin{aligned}
N_a(t) &= 1, & t &< t_{AP,1}, \\
N_a(t_{AP,i}) &= (1 - N_{r|a})N_a(t_{AP,i}^-), & t &= t_{AP,i}, \\
N_a(t) &= 1 - (1 - N_a(t_{AP,i})) & t &\in [t_{AP,i}, t_{AP,i+1}), \quad (9) \\
&\quad \exp\left(-(t - t_{AP,i})/\tau_a\right)
\end{aligned}
$$

where $i = 1, \ldots, K$.

Note that the change over time in the fraction of trials in which the vesicle is available clearly has a dependence on both (1) the time since the most recent pre-synaptic AP and (2) on the fraction of vesicles available at the time of the most-recent pre-synaptic AP. Consequently, the deterministic mean model should be interpreted as explicitly solving for the *conditional* mean number of vesicles released at each AP arrival, *given* the number that are available for release.

Remark 1: It is clear that the mean model accurately reflects Availability Model 2 generally, and in the specific case stated above, assumes exponential availability times following each AP arrival. Moreover, we have discussed that Availability Models 1 and 2 are equivalent for exponentially distributed availability times, and hence the stated mean model also accurately reflects Availability Model 1 for this specific case.

Remark 2: The deterministic mean model does not, however, accurately reflect Availability Model 1 for non-exponentially distributed availability times, since under Availability Model 1, the fraction of trials in which a vesicle should be released, given that it is available, should be based on the trial-dependent time since a vesicle was released, not solely on the time since the most recent AP. Therefore, the right hand side of a differential equation describing the mean number of trials in which a vesicle is available should have an additional term for each AP that occurs prior to the current AP. Moreover, if each additional term describes the mean number of trials in which vesicles have not become available since the i–th AP, the results will potentially become increasingly inaccurate with the time elapsed since the i–th AP.

One possibly useful element in any extension of mathematical analysis to this case of multiple trials might be an iterative expression articulated in a different context by McDonnell et al. (2002, 2008), that can be adapted to describe the conditional probability that u vesicles are available across Z trials, even if the time they were released differs. This approach does not suggest a straightforward method for implementing a stochastic simulation, but as described by McDonnell et al. (2002, 2008), there are simple expressions for the conditional mean and variance, and these could potentially be used within a deterministic equation that describes how the mean number of trials in which a vesicle is available changes with time.

In section 5, we compare the results of stochastic simulations with results for the mean obtained from Equation (9). We also use a result for a scenario where pre-synaptic APs arrive at the synapse periodically with frequency f Hz so that the AP times are $t_{\mathrm{AP},i} = i/f$, $i = 1, 2, \ldots$. In this case, it is well known that the mean fraction of vesicles available quickly decays to a constant steady state value, $N_{\mathrm{ss}}^- := N_a(t_{\mathrm{AP},i+1}^-) = N_a(t_{\mathrm{AP},i}^-)$. As shown in (Abbott et al., 1997; Matveev and Wang, 2000b), this can be obtained from Equation (9) (which hold for Availability Model 2 generally, and for Availability Model 1 with exponentially distributed availability times) to get the mean fraction of vesicles available for release just prior to a pre-synaptic spike as

$$N_{\mathrm{ss}}^- = \frac{\left(1 - \exp\left(-1/(f\tau_a)\right)\right)}{1 - \exp\left(-1/(f\tau_a)\right)(1 - N_{\mathrm{r|a}})}. \tag{10}$$

4. CORRECT AND INCORRECT STOCHASTIC SIMULATIONS, IN RESPONSE TO PRE-SYNAPTIC SPIKE TRAINS

We consider how a synaptic vesicle release site, containing at most a single vesicle, responds over time ($t \geq 0$), to a sequence of K arriving pre-synaptic APs. A stochastic simulation implementation that is faithful to the conceptual models is one that accurately produces vesicle releases that reflect the probabilities stated in Equation (6) or in (7).

In order to carry this out, it is necessary at every time step of the simulation to have a determined state of the availability of the vesicle. In other words, the vesicle is either available or not available. It switches from available to not available in the event that it is released, and it switches from not-available to available once, and only once, in the time following its last release. Therefore, once the vesicle becomes available according to the stochastic simulation, after a time T since the previous release, the probability of availability that must be used within the simulation is given by

$$P_{\mathrm{a,f}}(t|t_s) = 1, \quad t \geq T$$
$$0, \quad t < T. \tag{11}$$

This holds for both Availability Model 1 and Availability Model 2.

There are a number of parameter values that are required to be set in order to simulate a stochastic synapse model, as introduced above. These are summarized in **Table 1**.

In implementations of stochastic simulations it is necessary to generate random numbers from particular probability distributions. If a uniform random number generator is available, then its output, $U \in (0, 1)$, represents a number drawn from

Table 1 | List of parameter notation for single-vesicle release models.

Parameter name	Notation
Time-dependent probability of release, given vesicle is available	$P_{\mathrm{r|a}}(t)$
Mean time to restore a released vesicle	τ_a
Simulation duration	T
Number of pre-synaptic spikes	K
Time of arrival of each pre-synaptic action potential (AP)	$t_{\mathrm{AP},i} \in [0, T]$, $i = 1, 2, \ldots, K$

a continuous probability distribution. Random numbers from many other distributions can be generated from uniform random numbers. For example, exponentially distributed random numbers can be obtained by the operation $T_a = -\tau_a \ln(U)$.

In the pseudo-code below, we assume exponentially distributed availability times as our example, but if other distributions for this random variable are desirable, then the only change required is to generate random numbers from that distribution instead.

4.1. SINGLE VESICLE AVAILABILITY AND RELEASE: AVAILABILITY MODEL 1

The following pseudo-code illustrates how simulations of the random variables described above can be implemented in stochastic simulations.

Correct Implementation 1, for AM1

```
Set: NextAvailabilityTime = 0
For each pre-synaptic spike, i=1:K,
          occurring at time t_i
  if t_i >= NextAvailabilityTime
    //vesicle is available
    if Pr_given_a(t_i) > unifrand()
      //Release the vesicle
      Set: LastReleaseTime = t_i
      //reset the next availability
        time, for
      //exponentially distributed
        availability times
      NextAvailabilityTime = LastReleaseTime
      +exprand(tau)
    end
  end
end
//unifrand() generates a uniformly
  distributed random number between
  0 and 1
//exprand() generates an exponentially
  distributed random number with mean tau
```

The pseudo-code variable **LastReleaseTime** represents our mathematical variable, t_s. A direct translation of this pseudo-code into the probability that the vesicle will be released upon the arrival of AP i, *given* t_s, obtains $P_{\mathrm{r|a}}(t_{\mathrm{AP},i})\mathrm{Prob}(t_{\mathrm{AP},i} > t_s + T_a)$. This can be expressed as $P_{\mathrm{r|a}}(t_{\mathrm{AP},i})\mathrm{Prob}(T_a < t_{\mathrm{AP},i} - t_s) = P_{\mathrm{r|a}}(t_{\mathrm{AP},i})F_{T_a}(t_{\mathrm{AP},i} - t_s)$, and thus exactly matches Equation (6), as required.

There are also several ways in which the conceptual model has been, or could be, erroneously translated into a stochastic simulation, and these are described in the following subsections.

4.1.1. First incorrect implementation of availability Model 1

It is stated in Scott et al. (2012) that "Following successful vesicle release, [the availability probability] is set to zero and relaxes back to 1 ..." The exact form of this time changing probability [which we introduced above as $P_a(t|t_s)$] is expressed in Scott et al. (2012, Equation 14) as the solution to a differential equation, which has an exact solution equivalent to stating that

$$P_a(t|t_s) = 1 - \exp\left(-(t - t_s)/\tau_a\right), \quad t > t_s, \tag{12}$$

where t_s was the last successful release time. Clearly Equation (12) is equivalent to Equation (1). However, it is also stated in Scott et al. (2012) that in order to create a stochastic model, "... we allowed vesicle release following comparison of" $P_a(t|t_s)P_{r|a}(t)$ "with a random number between 0 and 1."

The following pseudo-code illustrates how this statement would be correctly implemented:

Incorrect Implementation 1.1, for AM1

```
For each pre-synaptic spike, i=1:K,
    occurring at time t_i
  Set: Pa = 1-exp(-(t_i-LastReleaseTime)/
                  tau_r)
  if Pa*Pr_given_a(t_i) > unifrand()
    //Release the vesicle
    Set: LastReleaseTime = t_i
  end
end
```

A direct translation of this pseudo-code into the probability that the vesicle will be released upon the arrival of the first AP after t_s, at time $t_{AP,j}$, given t_s, obtains $P_{r|a}(t_{AP,j})P_a(t_{AP,j}|t_s)$ which is in agreement with the correct implementation. However, this implementation also imposes a probability that the vesicle will be released upon the arrival of the second AP after t_s, at time $t_{AP,j+1}$, as $(1 - P_{r|a}(t_{AP,j})P_a(t_{AP,j}|t_s)) \times P_{r|a}(t_{AP,j+1})P_a(t_{AP,j+1}|t_s)$, which is the product of the probabilities of non-release at the j–th AP, and the calculated probability of release at the $j + 1$–th AP. This is not in agreement with Equation (6). Similar holds for the case where the vesicle is not released within the simulation after the $j + 2$–th AP, the $j + 3$–th and so forth.

The reason that the implementation is incorrect is that it does not take into account that the non-release at the j–th AP could have been due to release failure for an available vesicle, and this distorts the simulated probability of when the vesicle is released.

This fact might be more readily apparent by considering the following different incorrect implementation that achieves equivalent, but slightly less efficient, results:

Incorrect Implementation 1.2, for AM1

```
For each pre-synaptic spike, i=1:K,
    occurring at time t_i
  Set: Pa = 1-exp(-(t_i-LastReleaseTime)/
                  tau_r)
  if Pa > unifrand1()
    if Pr_given_a(t_i) > unifrand2()
      //Release the vesicle
      Set: LastReleaseTime = t_i
    end
  end
end
```

A direct translation of this pseudo-code into the probability that the vesicle will be released upon the arrival of the first AP after t_s, at time $t_{AP,j}$ is also in agreement with the correct implementation. However, the probability that the vesicle will be released upon the arrival of the second AP after t_s, at time $t_{AP,j+1}$, translates as $[(1 - P_a(t_{AP,j}|t_s)) + P_a(t_{AP,j}|t_s)(1 - P_{r|a}(t_{AP,j}))] \times P_{r|a}(t_{AP,j+1})P_a(t_{AP,j+1}|t_s)$, which is also not in agreement with Equation (6). Rearranging this gives $(1 - P_{r|a}(t_{AP,j})P_a(t_{AP,j}|t_s)) \times P_{r|a}(t_{AP,j+1})P_a(t_{AP,j+1}|t_s)$, which is identical to the result for the first stated incorrect pseudo-code.

Both cases of incorrect pseudo-code are incorrect because, as is clear in the second version, a random number can be drawn that is less than the probability of availability, which represents the vesicle being available. But the code does not take into account that once this happens once, there should never be a failure of availability before the vesicle is released.

The pseudo-code is equivalent to a different conceptual model where the vesicle's availability is reset to zero upon every spike arrival, regardless of whether the vesicle is released or not. Following this reset, the time until availability remains dependent on the time since the last release. This is unlike Availability Model 2, in which the reset causes the time until availability to become dependent on the time since the last spike arrival instead, and only for vesicles that are unavailable.

4.1.2. Second incorrect implementation of availability model 1

A second possible incorrect implementation could result from attempting to address the problem above by implementing the following incorrect pseudo-code:

Incorrect Implementation 2, for AM1

```
Set: IsVesicleAvailable = 1
For each pre-synaptic spike, i=1:K,
    occurring at time t_i
  if IsVesicleAvailable == 0
    //vesicle is not available
    Set: Pa = 1-exp(-(t_i-LastReleaseTime)/
                    tau_r)
    if Pa > unifrand1()
      //vesicle becomes available
      IsVesicleAvailable = 1
    end
  end
  if IsVesicleAvailable == 1
    //vesicle is available
    if Pr_given_a(t_i) > unifrand2()
      //Release the vesicle
      Set: IsVesicleAvailable = 0
      Set: LastReleaseTime = t_i
    end
  end
end
```

A direct translation of this pseudo-code into the probability that the vesicle will be released upon the arrival of the first AP after t_s, at time $t_{AP,j}$ is also in agreement with the correct implementation. However, the probability that the vesicle will be released upon the arrival of the second AP after t_s, at time $t_{AP,j+1}$, translates as $P_{r|a}$, if the vesicle was made available after the first spike, but not released, and as $P_a P_{r|a}$ if it became available after two spikes. When the probability of being in each of these three states is taken into account, the overall probability that the vesicle will be released upon the arrival of the second AP is $P_{r|a}(t_{AP,j+1})[P_a(t_{AP,j})(1 - P_{r|a}(t_{AP,j})) + (1 - P_a(t_{AP,j}))P_a(t_{AP,j+1})]$.

As a concrete example of why this implementation is incorrect, consider an example where immediately after a vesicle release, the next arriving AP did not find a vesicle available. Suppose $P_a = 0.4$ at this time, and increases to $P_a = 0.7$ just before the next arriving spike. We should have a vesicle available after the first spike in 40% of repeated trials, and a vesicle available in 70% of repeated

trials after two spikes. However, in this implementation, when the vesicle is not available after the first spike, we compare $P_a = 0.7$ with a random number, and 70% of the time for this case we then say a vesicle will be available after two spikes. This is incorrect, because we will have 40% of trials finding a vesicle on the first spike arrival spike and therefore by comparing P_a with 0.7 we have $100 \times (1 - 0.4) \times 0.7 = 42\%$ of trials finding a vesicle available on the second spike arrival, but not the first. Thus, there are $40 + 42 = 82\%$ of all trials finding a vesicle available after either the first or second spike arrival. The latter value should, however, be 70%, not 82%, according to the conceptual model. Therefore, this implementation causes too many vesicles to become available by the time of the second spike arrival, if they were not available on the first arrival. The correct number to compare with a random variable upon the second spike arrival is 0.5, which would mean 30% of trials find a vesicle available on the second spike arrival, but not the first.

4.1.3. Second correct implementation of availability model 1

Incorrect Implementation 2 can be corrected by changing the calculation of $P_a(t)$, based on Equation (2). For exponential availability times, the correction is a simple matter of replacing the pseudo-code line

```
Set: Pa = 1-exp(-(t_i-LastReleaseTime)/
                tau_r)
```

with

```
Set: Pa = 1-exp(-(t_i-t_(i-1))/tau_r)
```

For non-exponentially distributed arrival times, the required change is more complex, but readily follows in a similar fashion, from Equation (2).

4.1.4. Third correct implementation of availability model 1, for exponential availability times

We stated above that for the special case of exponentially distributed times for a vesicle to become available, Availability Model 1 is equivalent to a conceptual model where a vesicle becomes available upon generation of the next event within a Poisson process with rate $1/\tau_a$, following release, as in Loebel et al. (2009). An implementation of this conceptual model is illustrated in the following pseudo code, where it is assumed that the Poisson events had previously been calculated, and that `NextPoissonTime(x)` is a function that returns the time of the Poisson event immediately following the time given by its argument.

Correct Implementation 3, for AM1
```
Set: NextAvailabilityTime = 0
For each pre-synaptic spike, i=1:K,
    occurring at time t_i
  if t_i >= NextAvailabilityTime
    //vesicle is available
    if Pr_given_a(t_i) > unifrand()
      //Release the vesicle
      Set: NextAvailabilityTime
          = NextPoissonTime(t_i)
    end
  end
end
```

4.2. SINGLE VESICLE AVAILABILITY AND RELEASE: AVAILABILITY MODEL 2

The following pseudo-code illustrates how simulations of Availability Model 2 can be implemented in stochastic simulations. Note that the only difference in comparison with the pseudo-code for Availability Model 1 is that the time of next availability (for unavailable vesicles only) is dependent only on the last spike arrival time, not the last release time, in order to match the conceptual model.

Correct Implementation 1 for AM2
```
Set: NextAvailabilityTime = 0
For each pre-synaptic spike, i=1:K,
    occurring at time t_i
  if t_i >= NextAvailabilityTime
    //vesicle is available
    if Pr_given_a(t_i) > unifrand()
      //Release the vesicle and reset the
        next availability time, for
      // exponentially distributed
        availability times
      NextAvailabilityTime = t_i
                  +exprand(tau)
    end
  else
    //vesicle is unavailable; reset the
      next availability time, for
    //exponentially distributed
      availability times
    NextAvailabilityTime = t_i +exprand(tau)
  end
end
```

A direct translation of this pseudo-code into the probability that the vesicle will be released upon the arrival of AP i, given that it was not released by the time of AP $i-1$, obtains $P_{r|a}(t_{AP,i})\text{Prob}(t_{AP,i} > t_{AP,i-1} + T_a)$. This can be expressed as $P_{r|a}(t_{AP,i})\text{Prob}(T_a < t_{AP,i} - t_{AP,i-1})$, and thus exactly matches Equation (7), as required, upon substitution of Equation (5).

It is possible to incorrectly implement Availability Model 2 in a manner directly analogous to that in the first incorrect implementation of Availability Model 1. However, an implementation analogous to the second incorrect implementation of Availability Model 1, will actually be correct for Availability Model 2, since now the probability of availability is dependent only on the last AP arrival time.

4.3. MULTIPLE TRIALS OF SINGLE VESICLE AVAILABILITY AND RELEASE

4.3.1. Availability model 1 with exponential availability times

For the special case of exponentially distributed availability times, for each trial in which a vesicle is unavailable at the previous AP, the probability of becoming available by the current one will be identical for each trial (provided the input APs occur at the same times in all trials). As a direct consequence of this, the probability that w trials result in a vesicle becoming available, out of v in which a vesicle was not available at time $t_{AP,i}$, is given by the binomial distribution, as mentioned and studied numerous times, e.g., (Vere-Jones, 1966; Melkonian and Kostopoulos, 1996; Quastel, 1997; Matveev and Wang, 2000b; Pfister et al., 2010;

Reich and Rosenbaum, 2013). We introduce a random variable, W, to describe the number of unavailable vesicles that become available. We have

$$\mathrm{Prob}(W = w | v) = \binom{v}{w} (1 - P_{\mathrm{a}})^{(v-w)} (P_{\mathrm{a}})^{w}.$$

That this expression holds enables a stochastic simulation implementation that is far more efficient than repeating each of Z trials independently, as described in the following pseudo-code.

```
Set: NumUnavailable = 0
For each pre-synaptic spike, i=1:K,
    occurring at time t_i
  //Calculate probability of availability
   at t_i for any unavailable vesicles
  Set: Pa = 1-exp(-(t_i-t_(i-1))/tau_r)
  //calculate the number to become
   available by t_i
  Set: NumUnavailable = NumUnavailable
    - binornd(NumUnavailable,Pa);
  //calculate number to release at t_i
  Set: NumUnavailable = NumUnAvailable
    + binornd(NumTrials-NumUnavailable,
     Pr_given_a(t_i))
end
//binornd(v,w) calculates a binomially
  distributed random number with
// a maximum value of v, and mean vw.
```

This algorithm is an extension of an algorithm presented by Quastel (1997) (see also Pfister et al., 2010) for the case where P_{a} is time-independent.

In the above pseudo-code, we have calculated two independent binomially distributed random numbers for each pre-synaptic AP arrival. The second random number describes the number of trials in which an available vesicle is released. This is accurate with respect to both Release Model 1 and Release Model 2 under the assumptions of this paper, since the simulation calculated how many trials have a vesicle available at each time $t_{\mathrm{AP},i}$, and the probability of release is independent and identical for all trials in both release models. Mathematically, if we denote the random variable describing the number of vesicles released as U, when s are available, we have

$$\mathrm{Prob}(U = u | s) = \binom{s}{u} (1 - P_{\mathrm{r|a}})^{(s-u)} (P_{\mathrm{r|a}})^{s}.$$

The use of binomially distributed random numbers in this way will not be correct for a possible alternative release models where the probability of release, given availability, depends on the history of vesicle release in each trial, because the refill events are not independent in that case [see, e.g., Quastel (1997) for mathematical analysis of this case].

4.3.2. Availability model 2 with exponential availability times

The binomial approach described above for the special exponential case of Availability Model 1, will also correctly simulate Availability Model 2 with exponential availability times, since, as discussed above, the two models are equivalent under this special case.

4.3.3. Availability models 1 and 2 with non-exponential availability times

The algorithm above holds only for exponential availability times, as it relies on the fact that in this case $P_{\mathrm{a},1}(t|t_{\mathrm{s}}) = 1 - \exp\left(-(t - t_{\mathrm{s}})/\tau_{\mathrm{a}}\right)$, $t \geq t_{\mathrm{s}}$ for all vesicles. For non-exponential availability times, the number of vesicles unavailable due to release from all previous spikes needs to be tracked, and consequently many more binomial random numbers need to be generated following each AP. Moreover, $P_{\mathrm{a}}(t)$ needs to be calculated using Equation (2).

4.4. COMPARISON OF ALGORITHM IMPLEMENTATION EFFICIENCIES

We have aimed in the pseudo-code implementations above to describe computationally efficient algorithms that require as few random numbers to be generated as possible.

We note that the implementation suggested, for example, in Loebel et al. (2009) [see also Sterratt et al. (2011, p. 188)] involves an accurate approximation of a true Poisson point process, and this approximation is particularly relevant to any simulation in which time is discretised into uniform intervals of Δt, such as in most simulations that involve numerical solution of differential equations. The well-known approximation states that provided that $\Delta t \ll \tau_{\mathrm{a}}$, a Poisson point process event occurs within any given time interval of duration Δt with probability $\frac{\Delta t}{\tau_{\mathrm{a}}}$.

A stochastic simulation based on this approximation requires comparison of $\frac{\Delta t}{\tau_{\mathrm{a}}}$ with a uniform random number at every time step of the simulation between times 0 and t_K. It is possible to alter the implementation so that the Poisson events are only calculated during the simulation, rather than prior, where a comparison of a uniform random number with $\frac{\Delta t}{\tau_{\mathrm{a}}}$ is carried out for every time step following vesicle release, until a random number is generated that is larger than $\frac{\Delta t}{\tau_{\mathrm{a}}}$.

However, such implementations are potentially very inefficient, because many random numbers must usually be generated for every unavailable vesicle, whereas only one random number need be generated in, for example, Correct Implementation 1 for AM1.

5. EXAMPLES: COMPARING STOCHASTIC SIMULATION IMPLEMENTATIONS

5.1. ERRORS IN SIMULATING PROBABILITY OF RELEASE, AND MEAN NUMBER OF RELEASES AFTER *K* SPIKE ARRIVALS, FOR EXPONENTIAL AVAILABILITY TIMES

We consider a scenario where pre-synaptic APs arrive at a synapse periodically with frequency f Hz. We consider Z repeated trials following an initial condition where a vesicle is assumed to have just been released, in all trials, at the start of our simulations. We calculate the number of trials in which the next vesicle release occurs after the first spike, the second spike and so forth. We obtain results for f between 5 and 150 Hz, for $Z = 100,000$, $\tau_{\mathrm{a}} = 0.5$ s and $K = 50$ pre-synaptic spikes (as a maximum; the simulation stops when a vesicle is first released). Thus, the AP times are $t_{\mathrm{AP},i} = i/f$, $i = 1, 2, \ldots, 50$.

We estimated the probability that the vesicle was next released after $i = 1, 2, \ldots 20$ APs following vesicle release at time $t = 0$, by evaluating the fraction of trials in which the vesicle was first released after the i–th spike. We then calculated the absolute value

of the difference in the estimated probability for several correct and incorrect implementations, and also the relative error, relative to the correct version.

In order to clarify the significance of the values we obtained for absolute and relative error, we also considered a simulation where $H = 100$ spikes per trial periodically arrive with frequency f, and for each implementation counted the total number of vesicles released as a function of f. We then compared the mean number released after H spikes, calculated from $Z = 10,000$ repeats of each implementation, as well as the maximum and minimum numbers released.

Finally, in order to show that our simulations and mathematical analysis is correct for Availability Model 1 for both periodic and non-periodic AP arrivals, we compare correct and incorrect implementations for each case with results predicted by the Equations (9) and (10).

5.1.1. Results for availability models 1 and 2 with release model 1

We set the probability of release, given availability to $P_{r|a} = 0.6$. **Figure 1** shows the absolute error, and **Figure 2** shows the relative error between the correct and incorrect implementations, for Availability Model 1.

The absolute error, as predicted by the theory, is zero after the first pre-synaptic spike, for all f. However, it is clear that the absolute error can be as high as 10% for subsequent spikes, and is highest for low frequencies. It is also clear that the relative error can be very high for high frequencies. In these cases, the probability of release is relatively small for all subsequent spikes, and hence the absolute error is low. Yet the relative error can be higher than 500% at $f = 150$ Hz.

Figure 3 shows the absolute error between the correct and incorrect implementation for Availability Model 2 (recall from above that an implementation analogous to the second incorrect implementation of Availability Model 1, is correct for Availability Model 2). The incorrect implementation clearly shows a smaller error than for Availability Model 1.

Results for the mean number of vesicles released after $H = 100$ spikes are shown in **Figure 4**. It is clear for Availability Model 1 that the mean number of vesicles released per trial of 100 spikes becomes more inaccurate for the incorrect models as $f\tau_r$ increases. For example, at $f\tau_r > 10$, the incorrect models can produce more than twice as many vesicles as the correct one. It is clear for Availability Model 2 that the mean number of vesicles released per trial of 100 spikes is inaccurate for the incorrect model, similar to Availability Model 1. However, now the incorrect model underestimates the number of vesicles released, whereas for Availability Model 1, the incorrect models overestimated this number.

The data in **Figure 4** also shows that all models correctly produce a mean of $P_{r|a} = 0.6$ vesicles released at low frequencies, where the availability always has time to recover to close to 100%.

Figure 5 shows the fraction of 1000 trials in which vesicles are released in response to a sequence of 20 periodically arriving APs, with frequency 10 Hz, and to a sequence of 50 APs arriving at times corresponding to a Poisson point process, with mean frequency 10 Hz. In this figure, the data for

A

Incorrect Implementation 1

B

Incorrect Implementation 2
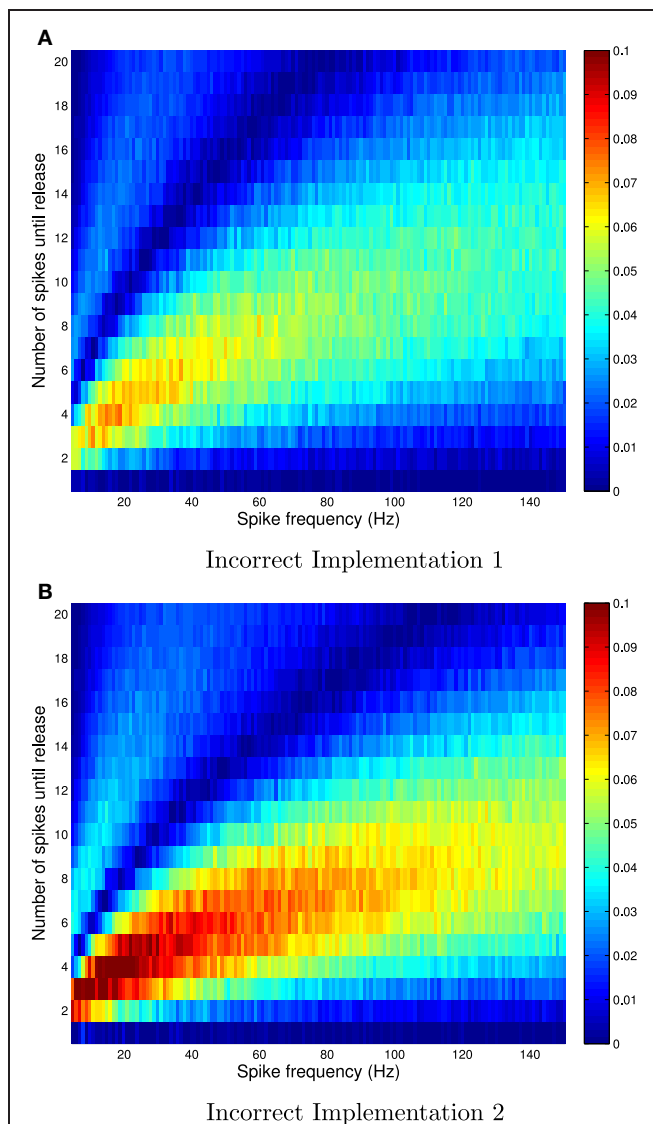
FIGURE 1 | Absolute errors for incorrect implementation 1 (A) and incorrect implementation 2 (B), for Availability Model 1, with exponentially distributed availability times. The data was obtained by empirically estimating the probability of release after i spikes, as a function of the frequency of periodically arriving pre-synaptic action potentials, by stochastically simulating $Z = 100,000$ trials for each condition. The absolute error can be as high as 0.1, and higher errors occur at low frequencies.

the **Deterministic**, and **Steady state** cases were obtained using Equations (9) and (10), respectively (derived previously in the literature, as stated and referenced above) and clearly match the correct stochastic simulations.

5.1.2. Release model 2

In order to demonstrate how to incorporate a time dependent release probability, we consider a standard model of facilitation (see, e.g., Scott et al., 2012). The change in release probability can be expressed as a differential equation, but it is clearer to write a piecewise equation as follows:

Incorrect Implementation 1

Incorrect Implementation 2

FIGURE 2 | Relative errors for incorrect implementation 1 (A) and incorrect implementation 2 (B) for Availability Model 1, with exponentially distributed availability times. The data was obtained by empirically estimating the probability of release after i spikes, as a function of the frequency of periodically arriving pre-synaptic action potentials, by stochastically simulating $Z = 100,000$ trials for each condition. The largest relative errors occur for higher frequencies.

$$P_{r|a}(t) = Q, \qquad\qquad\qquad t < t_{AP,1},$$

$$P_{r|a}(t_{AP,i}) = P_{r|a}(t_{AP,i}^-) + S(1 - P_{r|a}(t_{AP,i}^-)), \quad t = t_{AP,i},$$

$$P_{r|a}(t) = Q - (Q - P_{r|a}(t_{AP,i})) \qquad t \in [t_{AP,i}, t_{AP,i+1}),$$
$$\exp\left(-(t - t_{AP,i})/\tau_f\right)$$

where Q is a parameter that describes the steady-state release probability, when there have been no arriving APs for a long time, and S is a parameter that describes the fractional increase (relative to the maximum possible increase) in release probability that occurs for every arriving pre-synaptic AP. We also have a

FIGURE 3 | Absolute error between correct and incorrect implementation for Availability Model 2, with exponentially distributed availability times. The data was obtained by empirically estimating the probability of release after i spikes, as a function of the frequency of periodically arriving pre-synaptic action potentials, by stochastically simulating $Z = 100,000$ trials for each condition. The largest error occurs for low frequencies, but is much smaller than for Availability Model 1.

time constant of facilitation, τ_f, which determines how quickly the release probability decays back to its resting value, Q. Examples of appropriate parameters might be $Q = 0.4$, and $S = 0.2$, similar to Scott et al. (2012).

Note that this particular function $P_{r|a}(t)$ is determined entirely once the sequence of pre-synaptic spikes is known, and consequently it is easily incorporated into the stochastic simulation algorithms described above, and we do not show example results here.

The same observations hold for release-independent depression with frequency-dependent recovery, in which case τ_f can also change with time (Fuhrmann et al., 2004; Scott et al., 2012).

5.2. COMPARISON OF AVAILABILITY MODELS 1 AND 2 FOR NON-EXPONENTIAL AVAILABILITY TIMES

In order to demonstrate that a non-exponential availability model provides different outcomes for Availability Models 1 and 2, we consider the case of Rayleigh distributed availability times, with mean $\tau_a = 0.5$ s.

Figure 6 shows the fraction of 1000 trials in which vesicles are released, for both availability models, in response to a sequence of 20 periodically arriving APs, with frequency 10 Hz, and to a sequence of 50 APs arriving at times corresponding to a Poisson point process, with mean frequency 10 Hz.

The results shown in **Figure 6** were obtained both by a direct adaptation of the correct stated pseudo-code above to Rayleigh distributed availability times, and also by direct adaption of the binomial approaches described, for Availability Model 2. The binomial results for Availability Model 1 required a more complex algorithm, where the number of vesicles unavailable due to

Availability Model 1

Availability Model 2

FIGURE 4 | Mean (solid traces) number of vesicles released in total after 100 periodic pre-synaptic action potential arrivals for Availability Model 1 (A) and Availability Model 2 (B). The minimum and maximum number of vesicles released are shown with (dashed traces). For each frequency f, all statistics are derived from 10,000 stochastic simulations. Clearly the incorrect implementations can over or under estimate the correct number of vesicles released.

Periodic spiking

Poisson spiking

FIGURE 5 | Fraction of 1000 trials in which vesicles are released, for each of a sequence of 20 periodic spikes (A), and 50 Poisson spikes (B), and vesicles with exponentially distributed availability times. The frequency in both cases is 10 Hz. The traces for **Deterministic**, and **Steady state** were obtained using Equations (9) and (10). This data shows that the incorrect implementations give markedly different outcomes to the correct stochastic simulation implementations, and to the deterministic expression for the mean number of trials in which vesicles are released.

release from all previous spikes needed to be tracked, and consequently many more binomial random numbers generated than for Availability Model 1. The data can be seen to match in either implementation, but to be quite different for each Availability Model.

6. CONCLUSIONS AND EXTENSIONS
6.1. CORRECT AND EFFICIENT STOCHASTIC SIMULATIONS OF SHORT-TERM PLASTICITY
We have shown that various correct implementations of a stochastic simulation of either Availability Model 1 or 2 are possible.
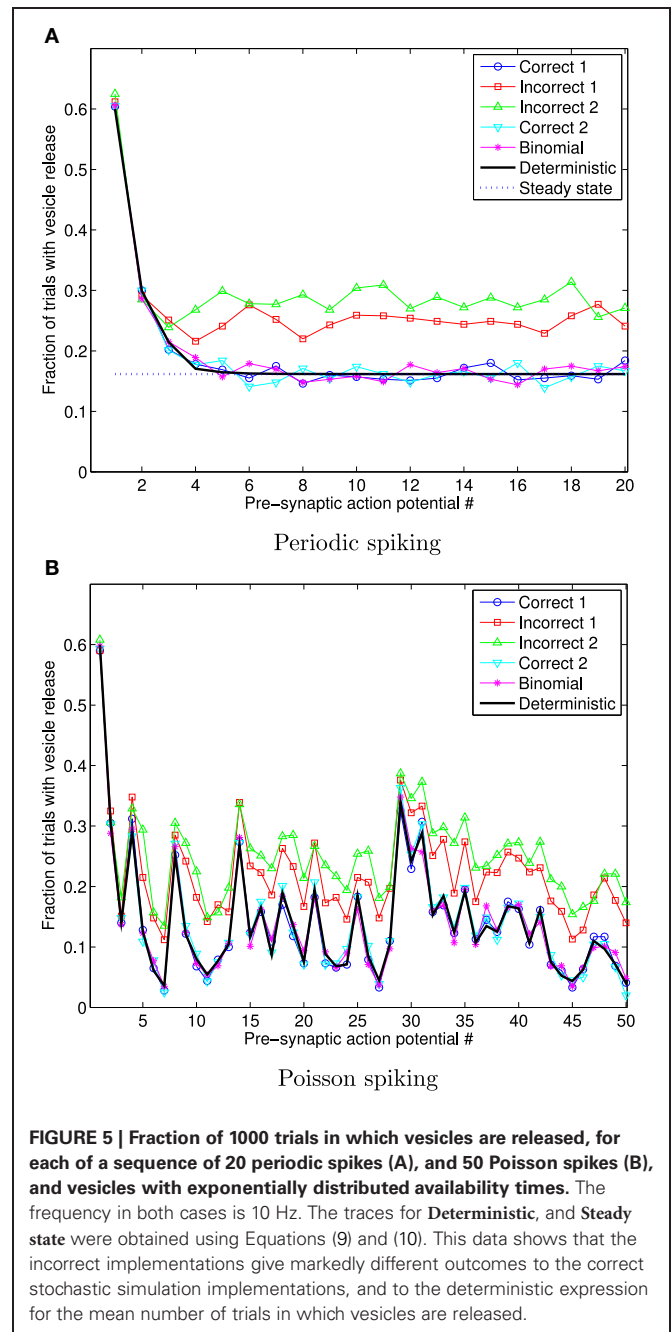
However, it is also possible to incorrectly implement either model. For Availability Model 1, two kinds of incorrect implementation result in more vesicle releases than should be the case. For Availability Model 2, an incorrect implementation results in less vesicle releases than should be the case.

We have also shown that some correct implementations are more efficient than others. In particular, we first stated an implementation that requires only a single random number to be generated each time a vesicle is released. This is more efficient than an implementation based on generation of a Poisson process that determines availability times, and much more efficient than generating a random number for every time step in a simulation.

FIGURE 6 | Fraction of 1000 trials in which vesicles are released, for each of a sequence of 20 periodic spikes (A), and 50 Poisson spikes (B), and vesicles with Rayleigh distributed availability times. The frequency in both cases is 10 Hz. This data shows that Availability Models 1 and 2 provide markedly different outcomes for Rayleigh distributed availability times, unlike the identical outcomes for exponentially distributed times. The data also shows that extension of the binomial approach matches the data where each trial is individually simulated.

We also have shown that when multiple independent vesicle releases are considered, the most efficient stochastic simulation implementation can be achieved by generating binomial random numbers.

6.2. CONSEQUENCES OF EQUIVALENCE OF AVAILABILITY MODELS ONLY FOR EXPONENTIAL AVAILABILITY TIMES

We have discussed that the two availability conceptual models are equivalent when the availability times are exponentially distributed, and this can be derived as a consequence of a well known property of a homogeneous Poisson point process. When the availability times are non-exponential, the two availability conceptual models we consider are generally non-equivalent.

As we have shown, these points are important in terms of their consequences for the implementations that can be used to correctly simulate Availability Model 1. Another consequence is that the popular differential equation approach to describing the mean number of available vesicles could have analogous correct simple forms for Availability Model 2, but not for Availability Model 1.

6.3. BINOMIAL-BASED STOCHASTIC SIMULATIONS FOR NON-EXPONENTIAL AVAILABILITY TIMES

When T_a is not exponentially distributed, no simple adaptation of the binomial approach will work for Availability Model 1, because there is no simple procedure for calculating random numbers corresponding to the random variable W. However, it is possible to keep track of how many vesicles were made unavailable by each AP, and how many are restored by the time of each subsequent AP, and perform a stochastic simulation that calculates as many binomial random numbers as there have been prior APs for which unavailable vesicles still exist. This procedure must make use of Equation (2) to calculate the probability of availability by the time of the next spike, using the time of the previous spike.

Such an algorithm may be more efficient than independently simulating each trial, provided that APs arrive relatively slowly compared with the mean time for a released vesicle to become available. **Figure 6** shows that an implementation of this approach for Availability Model 1 and Rayleigh distributed available times agrees with data from an approach that individually simulates each trial.

For Availability Model 2, the binomial approach will work for arbitrary F_{T_a}, since the probability of an unavailable vesicle becoming available will be the same for all trials, as was the case for the data in **Figure 6**. Moreover, only the cumulative distribution function of the availability times need be known to carry this out, whereas Equation (2) needs to be computed for Availability Model 2. Such an algorithm is a straightforward extension of an algorithm described by Quastel (1997) for constant probabilities for any released vesicle to be available by the next spike.

6.4. OTHER MODELS

We have considered only the simplest models in this paper. Other more complex models of availability and release have been proposed. For example, the mean times to availability for a vesicle may change over time (Wong et al., 2003), vesicles may also be released spontaneously in the absence of pre-synaptic APs (Sterratt et al., 2011), and multiple vesicles may be readily available for release at any site (although in this model only one can be released per pre-synaptic spike) (de la Rocha and Parga, 2005). Stochastic simulations that faithfully reflect these models can be readily devised by extension of the algorithms presented in this paper.

ACKNOWLEDGMENTS

REFERENCES

Abbott, L. F., and Regehr, W. G. (2004). Synaptic computation. *Nature* 431, 796–803.

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224.

Branco, T., and Staras, K. (2009). The probability of neurotransmitter release: variability and feedback control at single synapses. *Nat. Rev. Neurosci.* 10, 373–383.

Carnevale, N. T., and Hines, M. L. (2005). *The NEURON Book*. New York, NY: Cambridge University Press.

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

Dobrunz, L. E., and Stevens, C. F. (1997). Heterogeneity of release probability, facilitation, and depletion at central synapses. *Neuron* 18, 995–1008.

Faisal, A. A., and Laughlin, S. B. (2007). Stochastic simulations on the reliability of action potential propagation in thin axons. *PLOS Comput. Biol.* 3:e79. doi: 10.1371/journal.pcbi.0030079

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

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

Fung, C. C. A., Wong, K. Y. M., Wang, H., and Wu, S. (2012). Dynamical synapses enhance neural information processing: gracefulness, accuracy, and mobility. *Neural Comput.* 24, 1147–1185.

Gillesple, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* 81, 2340–2361.

Goldman, M. S. (2004). Enhancement of information transmission efficiency by synaptic failures. *Neural Comput.* 16, 1137–1162.

Goldman, M. S., Maldonado, P., and Abbott, L. F. (2002). Redundancy reduction and sustained firing with stochastic depressing synapses. *J. Neurosci.* 22, 584–591.

Graham, B. P., and Stricker, C. (2008). "Short term plasticity provides temporal filtering at chemical synapses," in *Proceedings of the 2008 International Conference on Artificial Neural Networks (ICANN 2008)*. Lecture Notes in Computer Science LNCS 5164. eds V. Kůrková, R. Neruda, and J. Koutnik (Berlin; Heidelberg; Germany: Springer-Verlag), 268–276.

Klug, A., Borst, G. G., Carlson, B. A., Kopp-Scheinpflug, C., Klyachko, V. A., and Xu-Friedman, M. A. (2012). How do short-term changes at synapses fine-tune information processing? *J. Neurosci.* 32, 14058–14063.

Loebel, A., Silberberg, G., Helbig, D., Markram, H., Tsodyks, M., and Richardson, M. J. E. (2009). Multiquantal release underlies the distribution of synaptic efficacies in the neocortex. *Front. Comput. Neurosci.* 3:27. doi: 10.3389/neuro.10.027.2009

Maass, W., and Zador, A. M. (1999). Dynamic stochastic synapses as computational units. *Neural Comput.* 11, 903–917.

Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5323–5328.

Matveev, V., and Wang, X.-J. (2000a). Differential short-term plasticity and transmission of complex spike trains: to depress or to facilitate? *Cereb. Cortex* 10, 1143–1153.

Matveev, V., and Wang, X.-J. (2000b). Implications of all-or-none synaptic transmission and short-term depression beyond vesicle depletion: a computational study. *J. Neurosci.* 20, 1575–1588.

McDonnell, M. D., Abbott, D., and Pearce, C. E. M. (2002). An analysis of noise enhanced information transmission in an array of comparators. *Microelectron. J.* 33, 1079–1089.

McDonnell, M. D., Stocks, N. G., Pearce, C. E. M., and Abbott, D. (2008). *Stochastic Resonance: From Suprathreshold Stochastic Resonance to Stochastic Signal Quantisation*. Cambridge: Cambridge University Press.

Meddis, R. (1986). Simulation of mechanical to neural transduction in the auditory receptor. *J. Acoust. Soc. Am.* 79, 702–711.

Melkonian, D. S., and Kostopoulos, G. K. (1996). Stochastic particle formulation of the vesicle hypothesis. relevance to short-term phenomena. *Neuroreport* 7, 937–942.

Merkel, M., and Lindner, B. (2010). Synaptic filtering of rate-coded information. *Phys. Rev. E* 81:041921. doi: 10.1103/PhysRevE.81.041921

Mohan, A., McDonnell, M. D., and Stricker, C. (2013). Interaction of short-term depression and firing dynamics in shaping single neuron encoding. *Front. Comput. Neurosci.* 7:41. doi: 10.3389/fncom.2013.00041

Pfister, J.-P., Dayan, P., and Lengyel, M. (2010). Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials. *Nat. Neurosci.* 13, 1271–1275.

Quastel, D. M. J. (1997). The binomial model in fluctuation analysis of quantal neurotransmitter release. *Biophys. J.* 72, 728–753.

Regehr, W. G. (2012). Short-term presynaptic plasticity. *Cold Spring Harb. Perspect. Biol.* 4:a005702. doi: 10.1101/cshperspect.a005702

Reich, S., and Rosenbaum, R. (2013). The impact of short term synaptic depression and stochastic vesicle dynamics on neuronal variability. *J. Comput. Neurosci.* doi: 10.1007/s10827-012-0438-0. [Epub ahead of print].

Rosenbaum, R., Rubin, J., and Doiron, B. (2012). Short term synaptic depression imposes a frequency dependent filter on synaptic information transfer. *PLoS Comput. Biol.* 8:e1002557. doi:10.1371/journal.pcbi.1002557

Rosenbaum, R., Rubin, J. E., and Doiron, B. (2013). Short-term synaptic depression and stochastic vesicle dynamics reduce and shape neuronal correlations. *J. Neurophysiol.* 109, 475–484.

Scott, P. C., Cowan, A. I., and Stricker, C. (2012). Quantifying impacts of short-term plasticity on neuronal information transfer. *Phys. Rev. E* 85:041921. doi: 10.1103/PhysRevE.85.041921

Sterratt, D., Graham, B., Gillies, A., and Willshaw, D. (2011). *Principles of Computational Modelling in Neuroscience*. New York, NY: Cambridge University Press.

Sumner, C. J., Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R. (2002). A revised model of the inner-hair cell and auditory-nerve complex. *J. Acoust. Soc. Am.* 111, 2178–2188.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835.

Tsodyks, M. V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U.S.A.* 94, 719–723.

Vere-Jones, D. (1966). Simple stochastic models for the release of quanta of transmitter from a nerve terminal. *Aust. J. Stat.* 8, 53–63.

Wong, A. Y. C., Graham, B. P., Billups, B., and Forsythe, I. D. (2003). Distinguishing between presynaptic and postsynaptic mechanisms of short-term depression during action potential trains. *J. Neurosci.* 23, 4868–4877.

Zucker, R. S., and Regehr, W. G. (2002). Short-term synaptic plasticity. *Ann. Rev. Physiol.* 64, 355–405.

Critical dynamics in associative memory networks

Maximilian Uhlig[1,2], **Anna Levina**[1,3], **Theo Geisel**[1,2] **and J. Michael Herrmann**[1,4]*

[1] Bernstein Center for Computational Neuroscience, Göttingen, Germany
[2] Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany
[3] Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany
[4] Institute of Perception, Action and Behaviour, School of Informatics, University of Edinburgh, Edinburgh, UK

Critical behavior in neural networks is characterized by scale-free avalanche size distributions and can be explained by self-regulatory mechanisms. Theoretical and experimental evidence indicates that information storage capacity reaches its maximum in the critical regime. We study the effect of structural connectivity formed by Hebbian learning on the criticality of network dynamics. The network only endowed with Hebbian learning does not allow for simultaneous information storage and criticality. However, the critical regime can be stabilized by short-term synaptic dynamics in the form of synaptic depression and facilitation or, alternatively, by homeostatic adaptation of the synaptic weights. We show that a heterogeneous distribution of maximal synaptic strengths does not preclude criticality if the Hebbian learning is alternated with periods of critical dynamics recovery. We discuss the relevance of these findings for the flexibility of memory in aging and with respect to the recent theory of synaptic plasticity.

Keywords: associative memory, dynamical synapses, SOC, Hebbian learning, homeostatic learning

1. INTRODUCTION

Critical dynamics in neural networks is an experimentally and conceptually established phenomenon which has been shown to be important for information processing in the brain. Critical neural networks have optimal computational capabilities, information transmission and capacity (Beggs and Plenz, 2003; Haldeman and Beggs, 2005; Chialvo, 2006; Shew and Plenz, 2012). At the same time the theoretical understanding of neural avalanches has been developed starting from sandpile-like systems (Herz and Hopfield, 1995) and homogeneous networks (Eurich et al., 2002), but later also including particular structural connectivity (Lin and Chen, 2005; Teramae and Fukai, 2007; Larremore et al., 2011). The network structure in the latter cases was, however, chosen as to support or even to enable criticality, which points obviously to one of the mechanisms criticality is brought about in natural systems. There are, nevertheless, other influences that shape the connectivity structure and weighting. Most prominently, this includes Hebbian learning, but also homeostatic effects or pathological changes. Here we study how such structural changes influence criticality in neural networks.

While homeostatic plasticity may well have a regulatory effect that supports criticality, this cannot be said about Hebbian learning which essentially imprints structure from internally or externally caused activation patterns in the synaptic weighting of the network increasing thus the probability of previous patterns to reoccur. Unless the patterns are carefully chosen to produce critical behavior, these effects have a tendency to counteract critical behavior, e.g., by introducing a particular scale that corrupts the power-law distributions characteristic for critical behavior.

Little is known, in particular, about the influence of criticality on associative memory neural networks. We have chosen this paradigm of long-term memory as a basis for the present model because it is very well understood and because it matches the complexity of models that are typically considered in the study of criticality. Associative memory networks are able to recall stored patterns when a stimulus is presented, that is similar to one of the stored patterns, thus providing a means to implement memory into a neuronal population. If all goes well, the network state follows an attractor dynamics toward the correct memory item when being initialized by a corrupted or incomplete variant of the item as an associative key. Items are stored as activation patterns that are implanted in the network by Hebbian learning. This leads to an effective energy landscape, where the patterns are local minima and as such attractors of the system dynamics (Hopfield, 1982; Herz and Hopfield, 1995). We have studied earlier the effect of dynamical synapses (Markram and Tsodyks, 1996) in associative memory networks (Bibitchkov et al., 2002), now we are interested in the criticalizing role of dynamical synapses.

Other work has shown (Levina et al., 2006, 2007b; Levina, 2008; Levina et al., 2009) that dynamical synapses play an important role in the self-organization of critical neural dynamics. Given the importance of the critical regime for information processing in the brain and the substantial experimental evidence that is available to date, there is a need to consider the compatibility of these two effects and to identify a way to obtain criticality and memory storage simultaneously.

We will discuss here an algorithm to achieve compromise between a critical dynamics that can be seen as exploring the spaces of neural activation patterns, and the attractor dynamics that we assume to underlay the retrieval of content from memory. The present paper continues upon earlier work (Schrobsdorff et al., 2009; Dasgupta and Herrmann, 2011), where the preliminary simulation results were discussed. In our study for the first time conclusive numerical representations are presented, several learning mechanisms are compared and the capacity limit is considered.

2. MATERIALS AND METHODS

2.1. NEURONAL ACTIVITY DYNAMICS IN THE CRITICAL REGIME

We consider a network of N integrate and fire neurons. The membrane potential $h_i \in [0, \theta]$ of a neuron $i \in \{1, \ldots, N\}$ is subject to the dynamics

$$\dot{h}_i = I_{\text{ext}} \delta\left(t - t_{\text{e}}^i\right) + \frac{1}{N} \sum_{j=1}^{N} J_{ij} \delta\left(t - t_{\text{sp}}^j\right). \tag{1}$$

The first term on the right hand side of Equation 1 represents an external excitatory input of size I_{ext} affecting neuron i at time t_{e}^i. The second term describes a recurrent excitatory input, where t_{sp}^j denotes the arrival time of a presynaptic action potential originating from neuron j and J_{ij} is the strength (or weight) of the synapse connecting j to i. Action potentials are generated and delivered to all postsynaptic neurons when neuron i reaches the membrane potential threshold θ. After this depolarization, the potential is reset according to

$$h_i\left(t_{\text{sp}}^+\right) = h_i(t_{\text{sp}}) - \theta. \tag{2}$$

The activity dynamics in this model network depends on the connectivity and the weight matrix $J = \{J_{ij}\}$. For a fully connected network with equal weights the activity forms a series of avalanches that are separated by longer periods of quiescence (Eurich et al., 2002). An avalanche is triggered when external input I_{ext} causes a neuron to fire and consists of a number (the avalanche *size L*) of successive depolarizations. Because some of these activations may occur simultaneously, avalanches are also characterized by their *duration* (*D*), i.e., the time from the start of the avalanche to the firing time of the last neuron that was activated in this way.

For neural networks of this type a critical synaptic weight J_{cr} exists that leads to a scale-free avalanche size distribution $P(L)$ (Eurich et al., 2002). For more complex networks the critical value is usually not explicitly obtainable, except for random (Levina et al., 2007b) or regularly coupled networks (Herz and Hopfield, 1995). This problem can be circumvented by applying an adaptive algorithm that adjusts the weights toward their critical values which do not need to be identical across neurons. Such an adaptation toward criticality can be obtained in form of a homeostatic learning rule (Levina et al., 2007a) which locally regulates the flow of activity from a neuron to its postsynaptic partners. Within the branching process approximation (Otter, 1949; Beggs and Plenz, 2003; Levina et al., 2007a; Levina, 2008) it can be shown that this homeostatic rule causes the network to become critical such that the activity dynamics in the network together with the homeostatic regulation forms a self-organized critical system.

2.2. HOMEOSTATIC REGULATION

Self-organized criticality can be achieved by applying a homeostatic learning rule at the beginning of each avalanche (Levina et al., 2007a) according to

$$J_{ij} = J_{ij}^0 + \varepsilon_{\text{hom}} \left[1 - \ell - N^{-\frac{1}{2}} \right]. \tag{3}$$

Here, J_{ij}^0 denotes the synaptic weights at the time of avalanche initiation, ℓ is the number of active neurons in the second time step of the avalanche, ε_{hom} is a learning rate and $N^{-\frac{1}{2}}$ a finite size correction. According to Equation 3 in the limit $N \to \infty$, the synaptic weights will decrease if $\ell > 1$, and increase if $\ell < 1$. For an infinitely large N a stable configuration is obviously obtained as soon as the triggering neuron causes exactly one other neuron to fire. This corresponds to a mathematical model of a critical branching process, that is known to result in a power-law distribution of the avalanche size. A finite size correction is needed because avalanches cannot spread to infinity but are rather limited to a system of N neurons. Such a learning rule resembles the homeostatic regulation observed in cortical neurons (Abbott and LeMasson, 1993; Turrigiano et al., 1998), with the important difference that the neuron does not keep stable its own firing rate, but that of the postsynaptic population.

2.3. DYNAMICAL SYNAPSES

The empirical observation of criticality in networks of real neurons has initiated a number of alternative explanations by regulatory processes interacting with the neuronal activity dynamics. One mechanism relies on the short-term dynamics of synaptic efficacies (Tsodyks and Markram, 1996, 1997). Given that synaptic resources are limited, high activity of presynaptic neurons will lead to depletion of these resources and thus to a reduced synaptic efficacy. In periods of silence or low activity, the synaptic efficacy will then recover toward its maximum value T_{ij}^{max}. We have modeled this in the following way (Levina et al., 2007b),

$$\dot{T}_{ij} = \frac{1}{\tau_J} \left(T_{ij}^{\text{max}} - T_{ij} \right) - u T_{ij} \delta\left(t - t_{\text{sp}}^j\right), \tag{4}$$

where τ_J sets the time scale of exponential recovery, t_{sp}^j is the presynaptic spike time and u sets the relative amount of resources used upon spike transmission (Markram and Tsodyks, 1996). Note that the T_{ij} are not the synaptic weights J_{ij} used in the sense of the previous section, but are related to Equation 3 by $J_{ij} = u T_{ij}$. The T_{ij}^{max} can be considered to be equal for all the synapses in the network and constant in time, but we will later relax this condition by introducing learning effects on a short time scale. Intuitively, the stabilizing effect of dynamical synapses in this model can be understood in the following way: large avalanches lead to depletion of synaptic resources and thus to series of smaller events, whereas small avalanches lead to an increase of the amount of resources in the synapses resulting in larger avalanches again. Such activity dependent regulation allows for a power-law distribution of avalanche sizes. A mathematical explanation for the success of this model is provided by the fact that $\langle u T_{ij} \rangle \to J_{\text{cr}}$ for a wide range of T^{max}, i.e., the time-averaged synaptic input approaches the critical value J_{cr} of the network with static synaptic weights that was defined in section 2.1 (Eurich et al., 2002; Levina et al., 2007b).

The arrival times t_{e}^i of external inputs of strength I_{ext} in Equation 1 are determined by a random process that selects neurons at a rate τ and increases their membrane potential. The characteristic time scale of synaptic recovery τ_J is related to the

time scale of external input τ via $\tau_J = \tau \nu N$ and $1 < \nu \ll N$. Therefore, the synaptic dynamics of this model is composed of two regimes. In the slow regime, neurons get loaded by external input I_{ext} and synaptic resources slowly recover toward their maximum value T_{ij}^{max}. The activation of a single neuron then marks the transition to the fast "avalanche regime", where the redistribution of neuronal membrane potentials and depression of resources T_{ij} is so fast that we can safely assume external input and synaptic recovery processes to be absent.

Irrespective of the particular recipe used to achieve self-organized criticality in our simulations, we always record A_{ava} avalanches and calculate the mean squared deviation $\Delta \gamma$ between the resulting avalanche size distribution $P(L)$ and the best-matching power law over the range $1 \leq L \leq N/2$. Unless $\Delta \gamma$ is not less than a specified threshold $\Delta \gamma^{\text{max}}$, we keep recording A_{ava} avalanches until the resulting size distribution has converged toward a power law and this sets the end of the critical regime. The resulting synaptic weight configuration $\{J_{ij}\}$ does then represent a neural network operating at the critical point. For small network sizes $\Delta \gamma$ was shown to be as informative about criticality in the network as a Kolmogorov-Smirnov statistic with Monte-Carlo generated p-value (Levina, 2008).

2.4. ASSOCIATIVE MEMORY MODEL

So far we have described the dynamics of the neural network in the critical regime. We now equip the network with the ability to store a set of patterns and to operate as an associative memory of these patterns. The patterns are represented by differences among the synaptic efficacies, and the retrieval of the pattern is understood as an attractor dynamics from a cue toward the pattern. The cue is a stimulus that causes a neuronal activity pattern near one of the memorized patterns and once the stimulus has initiated the attractor dynamics, it is expected that the current activity approaches the memorized pattern even more closely.

Let $\{\xi^\eta\}$, $\eta = 1, \ldots, M$, be a set of binary patterns consisting of pixels $\xi_i^\eta \in \{0, 1\}$. The pattern ξ^η is retrieved if the firing rate of the neurons with $\xi_i^\eta = 1$ is above and of the neurons with $\xi_i^\eta = 0$ is below a certain threshold.

We assume a sparse representation, i.e., only a fraction p of the neurons in a pattern is assumed to be active such that for all η

$$\frac{1}{N} \sum_{i=1}^{N} \xi_i^\eta = p. \tag{5}$$

In order to imprint these M binary patterns on the network connectivity, a matrix $W = \{W_{ij}\}$ in a correlational form is defined as

$$W_{ij} = \frac{1}{p(1-p) N C} \sum_{\eta=1}^{M} \xi_i^\eta \xi_j^\eta \left(1 - \delta_{ij}\right), \tag{6}$$

where C is an additional scaling factor which we choose such that $\sum_{ij} W_{ij} = N$. The structure of this matrix is fixed in time and depends on the specific set of patterns. If we took the synaptic weight matrix in the same way, i.e., $\{J_{ij}\} = \{W_{ij}\}$, the

network would exhibit optimal retrieval quality for the stored patterns η (Tsodyks and Feigel'man, 1988; Tsodyks, 1989). However, this weight configuration cannot be expected to generate critical avalanche size distributions. In order to combine criticality and memory, we therefore start with synaptic weights obtained by homeostatic learning (or dynamical synapses) and then carefully push the $\{J_{ij}\}$ toward the configuration $\{W_{ij}\}$ using the learning rule

$$J_{ij}(t + 1) = J_{ij}(t) + \varepsilon_{\text{hebb}} \left[W_{ij} - J_{ij}(t)\right]. \tag{7}$$

Here, $J_{ij}(t)$ and $J_{ij}(t + 1)$ are respectively the old and the new synaptic strengths and $\varepsilon_{\text{hebb}} \ll 1$ is a learning rate. Note that we do not apply Equation 7 synchronously for all the synapses but rather in a stochastic manner with update probability $p = 1/N$ for each synapse (i, j). This is done until the synaptic weight configuration $\{J_{ij}\}$ allows for associative recall of the stored patterns as specified below. We will refer to the episode during which Equation 7 is applied as *Hebbian learning*.

A modified learning rule is implemented in the case of dynamical synapses, which is given by

$$T_{ij}^{\text{max}}(t + 1) = T_{ij}^{\text{max}}(t) + \varepsilon_{\text{hebb}} \left[u^{-1} W_{ij} - T_{ij}^{\text{max}}(t)\right]. \tag{8}$$

Unlike before, Hebbian learning is not applied to the instantaneous values of synaptic efficacies $\{T_{ij}\}$ but rather to the maximal efficacies $\{T_{ij}^{\text{max}}\}$. Learning of instantaneous efficacies is not reasonable here as the effect of learning would be erased during the critical episode because the $\{T_{ij}\}$ tend to closely recover to their maximum values and these do not contain information about the stored patterns. If, however, the $\{T_{ij}^{\text{max}}\}$ are structured in a way similar to the optimal memory configuration $\{W_{ij}\}$, the instantaneous efficacies $\{T_{ij}\}$ will be affected in favor of the memory configuration too because they recover during episodes of low network activity toward the $\{T_{ij}^{\text{max}}\}$.

Clearly, we need a criterion that sets the end of the Hebbian learning episode. This criterion can only be based on the retrieval quality of the current network state, which is discussed in the following section.

2.5. RETRIEVAL QUALITY

In order to assess the retrieval (or memory) quality in the network with the configuration of synaptic strengths $\{J_{ij}(t)\}$, we construct perturbed versions $\kappa^\eta = Q \xi^\eta$ of the stored patterns. The operator Q selects an active and an inactive neuron at random and swaps their states, thereby keeping the total number of active neurons unchanged. Ideally, the network will be able to reconstruct the original ξ^η from the κ^η using the information that is implicitly stored in the connections. Practically, we can only require that the network produces a state that has less errors than κ^η, i.e., that is closer to the stored pattern than the perturbed version.

In discussing these questions, we will use a simplified model which was chosen mainly in order to be able to relate to results in Levina et al. (2007b, 2009) as well as in Bibitchkov et al. (2002). In addition to the use of binary patterns we will assume a noise-free

dynamics during retrieval and a fixed threshold. The threshold is optimized for achieving maximal overlap in the next state which, however, does not imply an optimal retrieval in the convergent phase (Bibitchkov et al., 2002). More specifically, upon presenting a perturbed pattern κ^η, the network activity will switch to a new configuration S^η given by

$$S_i^\eta \left(\kappa^\eta, \Theta \right) = \text{sign} \left(\sum_j J_{ij} \kappa_j^\eta - \Theta \right), \tag{9}$$

with Θ being some threshold. In what follows, we will refer to S^η as the *retrieved pattern*. To quantify the overlap between two binary patterns ξ and κ we use the correlational measure

$$o(\xi, \kappa) = \frac{1}{N} \frac{\sum_{i=1}^N \left[(\xi_i - m_\xi)(\kappa_i - m_\kappa) \right]}{\sigma_\xi \sigma_\kappa}, \tag{10}$$

where m_ξ and m_κ denote the mean activities and σ_ξ and σ_κ are the standard deviations, respectively. Perfect overlap is obtained for identical patterns where we have $o(\xi, \kappa) = 1$, while we obtain $o(\xi, \kappa) = 0$ for uncorrelated patterns. Thus, the observation that

$$\left\langle o\left(S^\eta, \xi^\eta \right) \right\rangle_\xi - \left\langle o(\kappa^\eta, \xi^\eta) \right\rangle_\xi > \Delta, \tag{11}$$

where Δ is a positive value that sets the minimum required improvement in (average) overlap $\left\langle o\left(S^\eta, \xi^\eta \right) \right\rangle_\xi$ compared to the perturbations, provides evidence that the weight configuration $\{J_{ij}\}$ indeed contains information about the stored pattern ξ^η. If the network realized, in contrast, an identical transformation it would not achieve an improvement of the overlap, but it could "remember" a pattern for a short time in a kind of short-term memory. Typically, we will not only consider a single random perturbation κ^η per pattern but many, so that Equation 11 becomes

$$\left\langle \left\langle o\left(S^\eta, \xi^\eta \right) \right\rangle_\kappa \right\rangle_\xi - \left\langle \left\langle o(\kappa^\eta, \xi^\eta) \right\rangle_\kappa \right\rangle_\xi > \Delta. \tag{12}$$

In a spiking network also temporal averages need to be included in order to obtain a consistent measurement of the retrieval quality. According to the simplifying assumptions above, we will consider only small perturbations which consist in the case of a finite network in single bit swaps. This is done for two reasons. First, such perturbations are used in order to concentrate on the threshold-independent effects of the retrieval dynamics. A perturbed pattern cannot be corrected by the choice of a standard threshold value. Second, near the critical capacity, it is sufficient to study the ability of the network to correct a single error. This is due to the reduction of the size of the basins of attraction of the pattern-related attractor states. As soon as the attractor size has reached zero even an almost correct pattern will typically deteriorate with the dynamics (Equation 9). A persistence of a fixed point state beyond the capacity limit, but without a basin of attraction, is easily achieved, e.g., by avoiding any update of the neurons state, but is not interesting in the present context.

Two comments concerning the threshold Θ in Equation 9 seem to be in order here. First, we choose this threshold such that the average overlap $\left\langle o\left(S^\eta, \xi^\eta \right) \right\rangle_\kappa \rangle_\xi$ is maximal. Second, Θ

may differ from the threshold θ that we use in the critical regime. We assume that both thresholds are the result of a specific action of inhibitory neurons, which we, however, do not model here explicitly.

2.6. OPTIMIZATION TOWARD CONVERGED STATES

In the previous sections we have outlined how the synaptic weights evolve during the critical and Hebbian learning episodes, respectively. The critical episode ends as soon as the sampled avalanche size distribution is close to a power law (see section 2.3), whereas the following phase of Hebbian learning is stopped as soon as the network exhibits good retrieval quality of the stored patterns, when perturbations of the latter are presented. We measure the retrieval quality after each step of Hebbian learning and stop if the improvement in average overlap is at least Δ_{hebb}. For the sake of reduced numerical complexity, we only consider one single perturbation for each stored pattern, i.e., we use Equation 11 instead of Equation 12. After Hebbian learning is over, the network is driven toward the critical regime again, employing either homeostatic regulation of synaptic weights or dynamical synapses, respectively. This alternation between the two episodes may be interpreted as an optimization scheme and the delicate question is if convergence toward a state is obtained in the long run, in which the network is critical and retains an associative memory of the stored patterns at the same time. In what follows, we will refer to these states as *converged states*.

Whether the network is in a converged state is always checked after the critical episode is finished and before the next round of Hebbian learning is started. At this point we already know that the system is operating at the critical point but we still need to make sure, that the critical episode has not erased memory of the stored patterns. We therefore rigorously test the retrieval quality of the network using Equation 12 with a minimum required improvement of Δ_{conv} and take the average over n_p ($n_p \gg 1$) perturbations per pattern. Note that in the case of dynamical synapses, we assess the retrieval quality of $J_{ij} = uT_{ij}^{\text{max}}$ in the Hebbian learning phase, whereas we take $J_{ij} = uT_{ij}$ to check for convergence.

A sketch of the optimization strategy is shown in **Figure 1** and the most important steps are summarized in **Algorithm 1** for the case of homeostatic regulation.

Algorithm 1 | General steps of the optimization strategy for the case of homeostatic regulation (see text for details).

1. Homeostatic learning of synaptic weights (Equation 3) based on T_L avalanches.
2. Recording of A_{ava} avalanches and their sizes L.
3. Calculate mean squared deviation $\Delta \gamma$ of size distribution $P(L)$ from best-fit power law; if $\Delta \gamma < \Delta \gamma^{\text{max}}$ continue, otherwise restart at step 1.
4. Check for convergence using Equation 12 with $\Delta = \Delta_{\text{conv}}$ and averaging over n_p ($n_p \gg 1$) perturbations per pattern; if retrieved states show large enough overlap with stored patterns, network has converged; if not, continue with step 5.
5. Hebbian learning of synaptic weights using Equation 7; after each learning step check retrieval quality according to Equation 11 with $\Delta = \Delta_{\text{hebb}}$; if retrieval quality good enough start at step 1.

3. RESULTS

3.1. SPECIFICATIONS OF THE MODEL USED IN THE EXPERIMENTS

In this study we always simulate networks of $N = 300$ neurons. Other parameters are summarized in **Table 1**.

At the beginning of each critical phase we sample A_0 avalanches without taking them into account in the avalanche size distribution $P(L)$. This is done to ensure that the size distribution is not affected by transient dynamics. The sampling of the distribution is an important contribution to the total simulation time and was the main limitation of the size of the network. Because in larger networks also larger avalanches need to be considered, the sampling time for given $\Delta\gamma^{max}$ increases faster than quadratic which was the main reason for our choice of the network size. Smaller networks, however, are less suitable to store small activity patterns, see section 3.2.

Apart from that, we consider several trials for each number of patterns M stored in the network, where each trial uses a different set of M patterns. Unless otherwise stated, data points are averages over 10 trials for each M and error bars indicate one standard

FIGURE 1 | Schematic representation of the dual optimization algorithm.

deviation from the mean. Instead of the number M of stored patterns in the network, we will typically use the load parameter, defined as $\alpha := M/N$.

3.2. MEMORY NETWORK

Before we study networks that include mechanisms to bring about criticality, we first test pure memory networks. We generate a set of M random binary patterns, calculate the matrix $\{W_{ij}\}$ according to Equation 6 and set the network connectivity to $J_{ij} = W_{ij}$. The memory quality is then assessed by calculating the average overlap $\langle\langle o(S^\eta, \xi^\eta)\rangle_\kappa\rangle_\xi$ between the retrieved patterns S^η and the original patterns ξ^η (see section 2.5 for details). From here on, we always take the average $\langle.\rangle_\kappa$ over $n_p = 1000$ randomly generated perturbations κ^η of each of the M patterns ξ^η. The average overlap is close to 1 up to load parameters $\alpha \approx 0.07$ (**Figure 2A**), indicating perfect retrieval quality of the networks. Around $\alpha \approx 0.11$ it drops below 0.982, which marks the overlap corresponding to an average deviation of one digit from the original pattern. Finally, at $\alpha \approx 0.13$ and beyond the network does not yield retrieved patterns anymore that are closer to the original patterns than the perturbations.

In **Figure 2B** we show results of a criticality test for pure memory networks, where we record A_{ava} avalanches and consider the mean squared deviation $\Delta\gamma$ of the size distribution $P(L)$ from the best-fit power law. Although there is no mechanism in these networks to bring about criticality in a self-organized way, we always choose the normalization C in Equation 6 such that $\langle W_{ij}\rangle$ corresponds to the critical value of the model with fixed synaptic weights [see section 2.1 and Eurich et al. (2002)]. We find that for load parameters $\alpha \gtrsim 0.6$ the network is critical. Below this value it is not critical because the coupling matrix $\{W_{ij}\}$ is too sparse, i.e., many entries are 0. Thus, pure memory networks can become critical but only in a range of load parameters where the quality of retrieval or memory is already poor. In the following sections we show that the critical regime and the memory regime can be brought into agreement by employing the optimization strategy described in section 2.6.

The simulation time depends essentially on the load of the network, see **Figures 3–5**. The numerical complexity of an iteration step depends linearly on M as all patterns are relearned, while the other parameters on which it depends are kept fixed here.

3.3. COMBINED HOMEOSTATIC AND HEBBIAN LEARNING

We now consider simulations that include homeostatic regulation as a means to bring about criticality in a self-organized way.
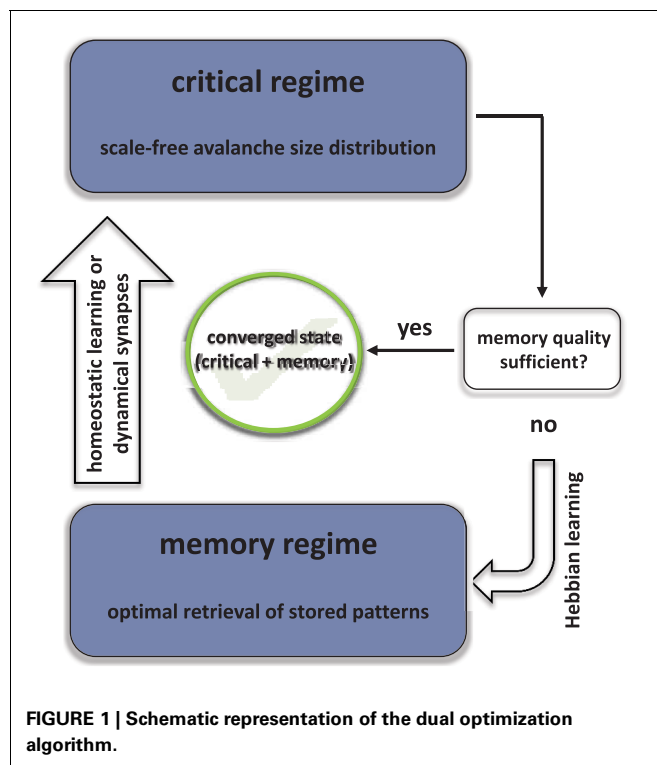
Table 1 | Parameters used in the numerical simulations.

General parameters										
Parameter	N	θ	p	A_0	A_{ava}	ε_{hebb}	$\Delta\gamma^{max}$	Δ_{hebb}	Δ_{conv}	n_p
Value	300	1.0	0.1	10^4	10^6	0.01	0.005	0.035	0.03	1000

Homeostatic plasticity				Dynamical synapses		
Parameter	T_L	ε_{hom}	l_e	Parameter	ν	u	l_e
Value	10^3	0.001	0.0067	Value	10	0.2	0.025

At the beginning, the $\{J_{ij}\}$ are initialized by $\{W_{ij}\}$ but will be modified in the course of the alternating episodes of homeostatic and Hebbian learning, respectively. The most important finding we arrive at is the existence of converged states in which the networks are critical and associative memories of the stored patterns at the same time. The total number of Hebbian learning steps needed to arrive at these states significantly increases with load parameter α (**Figure 3A**), spanning about two orders

of magnitude. In contrast to the pure memory networks studied before, criticality is already achieved for small values of α (**Figure 3B**).

The retrieval quality of the networks in the converged state is again assessed by considering the average overlap $\langle\langle o(S^\eta, \xi^\eta)\rangle_\kappa\rangle_\xi$ of the retrieved solutions S^η and the original patterns ξ^η (**Figure 4A**). For small values of α the networks are able to reconstruct the original patterns almost perfectly. However, the overlap
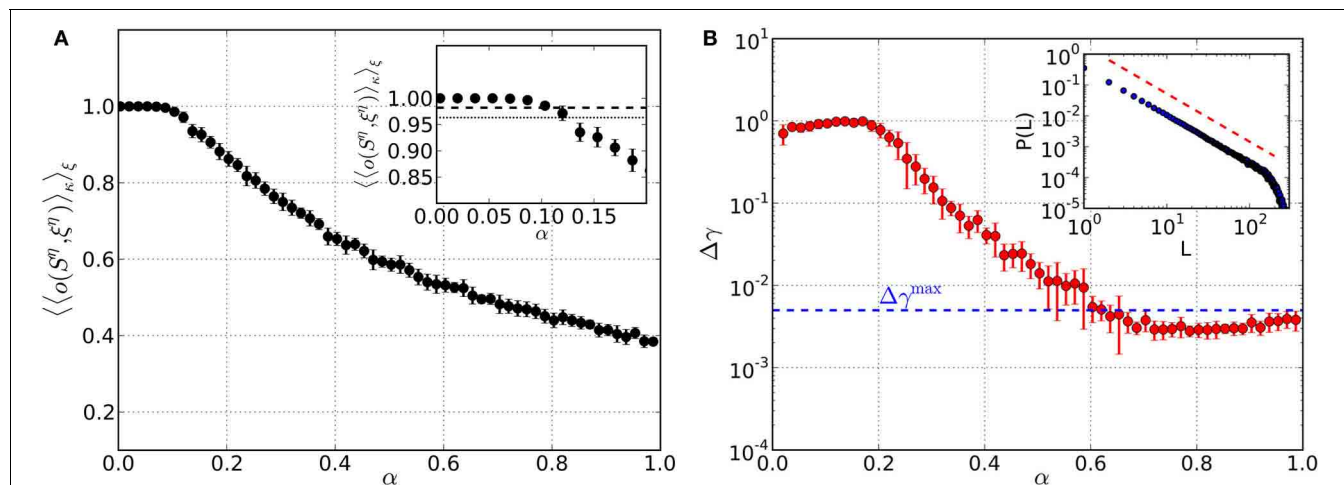
FIGURE 2 | Performance of pure memory networks in terms of retrieval quality (left) and criticality (right) as a function of the load parameter α. Synaptic weights are fixed and defined by $\{W_{ij}\}$. There is neither homeostatic learning nor activity dependent synapses adaptation. **(A)** Average overlap between initially stored patterns and corresponding retrieved patterns. Averages are taken over 10 trials for each α and error bars indicate

one standard deviation from the mean. **(B)** Mean squared deviation of obtained avalanche size distributions from the best-fit power law. The blue dashed line denotes the threshold $\Delta\gamma^{max} = 0.005$, below which avalanche size distributions can be considered critical. The inset shows an example avalanche size distribution $P(L)$ obtained in the range of critical load parameters, along with the slope of the best-fit power law (red dashed line).
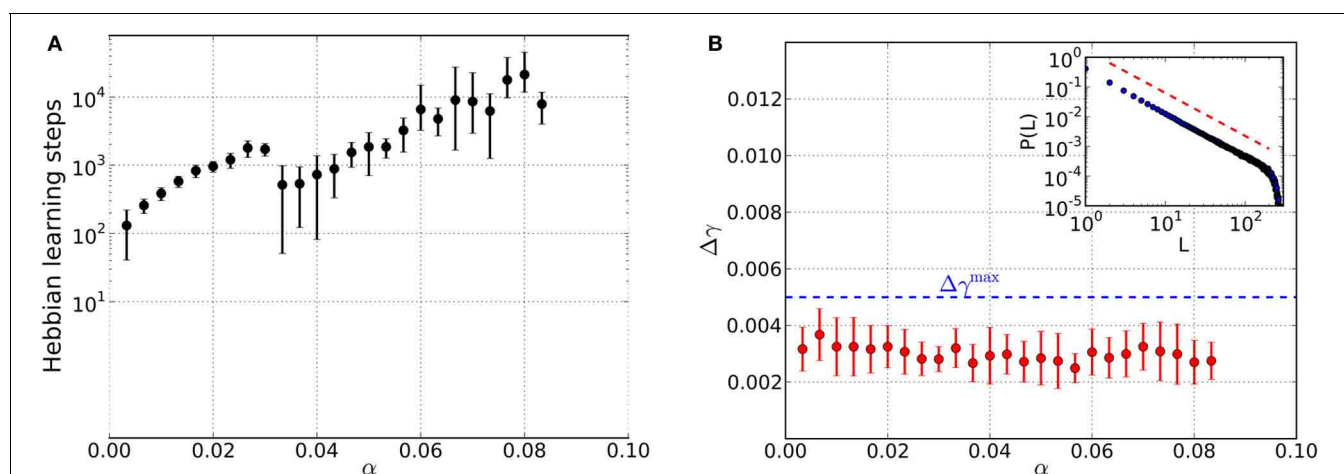
FIGURE 3 | Results from networks including Hebbian and homeostatic learning of synaptic weights, respectively, for different values of the load parameter α. For each value of α data is taken from 10 trials and error bars mark one standard deviation from the mean. **(A)** Total number of steps in Hebbian learning needed to converge to a state that is both critical and an associative memory of the stored patterns. The discontinuity near $\alpha \approx 0.03$ appears to be due to the finite size of the basins of attraction: while for low loading ratios α a basin of attraction of

several bits can be achieved, now only a single bit is corrected in the course of the learning which is faster achievable than before. **(B)** Average mean squared deviation $\Delta\gamma$ from the best-fit power law. Since all data points lie below the threshold of $\Delta\gamma^{max} = 0.005$ (blue dashed line), avalanche size distributions are critical over the whole range of α. An example avalanche size distribution $P(L)$ in the converged state is illustrated in the inset (red dashed line indicates slope of the best-fit power law).

FIGURE 4 | Retrieval performance of networks including Hebbian and homeostatic learning of synaptic weights, respectively, for different load parameters α. For each α data is taken from 10 trials and error bars mark one standard deviation from the mean. **(A)** Average overlap between initially stored patterns and corresponding retrieved patterns. Filled circles include results from converged simulations only, whereas most simulations in the range of open circles did not converge. For comparison, the overlaps corresponding to an average deviation of two digits (dotted line) and one digit (dashed line) from the original patterns are indicated. **(B)** Average fraction of patterns for which the networks yield retrieved patterns with deviation less than one and two digits, respectively. Filled markers again include converged simulations only and open markers mainly have contributions from simulations that did not converge.

is less than in case of the pure memory networks. Compared to the latter, the decrease in retrieval quality also occurs for smaller α and is more abrupt. The open circles mark the range of load parameters, where the majority of simulations does not converge anymore, because the required increase Δ_{hebb} in overlap during the Hebbian learning episode is not reached. Instead, the overlap saturates so that we stop Hebbian learning, add one last excursion toward the critical regime and finally finish the simulations after measuring the retrieval quality.

Since $\left\langle \langle o(S^\eta, \xi^\eta) \rangle_\kappa \right\rangle_\xi$ only measures the overlap averaged over all the M patterns stored in a network, we also assess the overlap on the level of single patterns. For this reason we consider the fraction of patterns stored in a network, that can be reconstructed from perturbed states with a deviation less than a specified number of digits (**Figure 4B**). For the range of load parameters α where the majority of simulations converge, more than 90% of the retrieved patterns can be reconstructed with an average deviation less than one digit. (**Figure 4A**). Also, there are practically no patterns for which the retrieved states deviate more from the original patterns, than the perturbations themselves. Even in the range of load parameters where the average overlap strongly decreases, there is still a small fraction of patterns which is well "remembered" by the networks.

3.4. SYNAPTIC DEPRESSION
We now turn to the second synaptic regulatory mechanism that brings about criticality in our networks (see section 2.3). All the analysis in this part is done along the lines of the previous section, so the only essential difference here is that we substitute homeostatic learning as the mechanism that drives the network into the critical regime by dynamical synapses. At the beginning of the simulations, maximal synaptic resources $\left\{ T_{ij}^{\max} \right\}$ are set

equal to $1.4\,(uN)^{-1}$. Due to Hebbian learning however, structure in the maximal resources will develop in the course of the simulations.

Also the model networks considered here evolve toward converged states that are critical and an associative memory at the same time. While the total number of Hebbian learning steps needed to converge (**Figure 5A**) is comparable to homeostatic learning, the agreement of the avalanche size distributions with scale-free distributions is better for dynamical synapses (**Figure 5B**).

Figure 6 addresses the retrieval quality of the converged network states for the stored patterns. The average overlap $\left\langle \langle o(S^\eta, \xi^\eta) \rangle_\kappa \right\rangle_\xi$ is again close to optimal for small values of α and drops below the overlap of perturbed patterns $\langle o(\kappa^\eta, \xi^\eta) \rangle_\xi$ at $\alpha \approx 0.13$ (**Figure 6A**). This is comparable to what was observed for the pure memory networks in Bibitchkov et al. (2002) so that the overlap decreases less quickly than for homeostatic regulation. This might be attributed to the fact that the structure that was learned into the maximal efficacies $\left\{ T_{ij}^{\max} \right\}$ during episodes of Hebbian learning is not affected during the critical phase, where only the $\{ T_{ij} \}$ are changed. Thus, memory is safely stored within maximal synaptic efficacies.

4. DISCUSSION
In this study we investigated the interplay between criticality and memory in neural networks. We showed that Hebbian learning alone destroys criticality even when the synaptic strength is properly scaled. Applying an optimization procedure that drives the synaptic couplings either toward the critical regime or toward the memory state in an alternating fashion, we finally arrive at a configuration that is both critical and retains an associative

FIGURE 5 | Results from networks that are influenced by Hebbian learning and dynamical synapses, for different values of the load parameter α. For each α data is taken from 10 trials and error bars mark one standard deviation from the mean. **(A)** Total number of steps in Hebbian learning needed to converge to a state that is both critical and an associative memory of the stored patterns. **(B)** Average mean squared deviation $\Delta\gamma$ from the best-fit power law. Since all data points lie below the threshold of $\Delta\gamma^{\max} = 0.005$ (blue dashed line), avalanche size distributions are critical over the whole range of α. The inset shows an example avalanche size distribution $P(L)$ in the converged state and the red dashed line marks the slope of the best-fit power law.

FIGURE 6 | Retrieval performance of networks including Hebbian learning and dynamical synapses for different load parameters α. For each α data is taken from 10 trials and error bars mark one standard deviation from the mean. **(A)** Average overlap between initially stored patterns and corresponding retrieved patterns. Filled circles include results from converged simulations only, whereas most simulations in the range of open circles did not converge. For comparison, the overlaps corresponding to an average deviation of two digits (dotted line) and one digit (dashed line) from the original patterns are indicated. **(B)** Average fraction of patterns for which the networks yield retrieved patterns with deviation less than one and two digits, respectively. Filled markers again include converged simulations only and open markers mainly have contributions from simulations that did not converge.

memory. In the following, we will discuss our findings and possible implications in more detail.

4.1. QUALITY OF CRITICALITY

The mean squared deviation of the avalanche size distributions obtained in the converged network states from their best-fit power law was always smaller than the threshold $\Delta\gamma^{\max} = 0.005$, providing evidence that the networks indeed operate at the finite-size equivalent of a critical point (Levina et al., 2007b). Furthermore, we did not observe oscillatory features in the avalanche size distributions. This suggests that the network dynamics circumvents the attractors (the stored patterns) that were learned into the network structure. The explanation of this observation is probably twofold. First, the majority of activity in the networks consists of small avalanches whose size is smaller than the total activity in the stored patterns, so that there is not enough overlap to be attracted toward the stored patterns. Second, even though larger avalanches have generations of firing neurons with total activity close to that of the stored patterns, the likelihood that they come close enough is very small given the many possible

configurations. Indeed, we found no evidence that avalanches or their sub-generations come close to any of the stored patterns at all during our simulations. There are nevertheless traces of the pattern structures in the avalanches in the sense that a pair of neurons that is active in the same pattern is also correlated in the critical spontaneously active network. Likewise, we consider elevated correlations also between subsequent avalanches. Although these correlations are not unexpected it is interesting that they do not interfere with the criticalization of the network.

4.2. QUALITY OF MEMORY AND CAPACITY

To assess the (associative) memory quality in the converged state, we presented perturbed versions of the initially stored patterns to the networks which resulted in retrieved patterns. The latter almost never deviate from the original patterns more than the perturbed states themselves. More importantly, more than 90% of the patterns are reconstructed with on average less than one digit deviation from the original patterns. We may therefore conclude that the memory quality of the critical networks is very good. However, a pure memory network which has couplings equal to $\{W_{ij}\}$ and does not operate at the critical point, still performs better in terms of reconstruction from perturbed patterns. This might be attributed to the fact that the coupling matrices $\{J_{ij}\}$ obtained in the converged states are not symmetric anymore, as opposed to $\{W_{ij}\}$. But symmetry of the coupling matrix is a major prerequisite for good retrieval quality of traditional Hopfield networks.

4.3. COMPARING HOMEOSTATIC LEARNING AND DYNAMICAL SYNAPSES

We have considered two mechanisms that can regulate a neural network toward criticality (and thus making it truly self-organized critical). Homeostatic learning regulates synaptic weights until the branching ratio approaches the critical value. Dynamical synapses on the other hand represent a biologically more justified regulatory mechanism, where the critical branching ratio is reached through the interplay of synaptic depression and recovery. In the latter model, we found that agreement between criticality and memory can only be achieved if the maximal synaptic weights are structured by Hebbian learning. Homogeneous maximal weights in contrast lead to memory loss during critical episodes because synaptic resources may recover to their maximal values which carry no information of the stored patterns anymore.

4.4. MEMORY STORAGE BY DYNAMICAL SYNAPSES

The apparent contradiction between the classical concept of memory storage by fixed synaptic efficacies which are modifiable only by persistent high-rate stimulation on the one hand, and the realization of adaptive filters based on the short-term dynamics of synaptic resources has been studied already in Tsodyks et al. (1998), Bibitchkov et al. (2002), and in more realistic models in Giudice and Mattia (2001) and Romani et al. (2006). While initially the contradiction between the two modes of operation has been studied in Bibitchkov et al. (2002), later the benefits arising from the combination were uncovered. It is interesting

that the formation of memories which imply a strong structural modification in the context of attractor networks, can even be enhanced in accuracy if short-term synaptic plasticity is used in the model. The finding of critical dynamics in such networks both supports this view and expands it in the sense that a coexistence of a retrieval state and a critical exploratory state becomes possible by dynamical synapses. This suggests a solution of one of the main problems with attractor networks, namely the conditions for the escape from attractors. While this can be achieved by an additional dynamics (Horn and Usher, 1989; Treves, 2005), we have here a form of dynamics that is purely input-driven when an input is available, while it is exploratory if this is not the case. It might be interesting to consider networks with correlated patterns (see Herrmann et al., 1993) where the two effects can become intertwined.

4.5. PATTERN-RELATED CORRELATIONS IN THE CRITICAL REGIME

One of the main points here is that the critical state serves as a ground state of the system which is assumed in the absence of specific external input. But it is, since the completely inactive state is absorbing in our model, constantly fed by spatially and temporally homogeneously distributed external noise. A specific external input has a large overlap with one of the patterns and a small overlap with all the other stored patterns. This is a necessary condition of the model, which in order to be relaxed requires a specific modification. We have dealt with such problems in our previous papers, but assume here uncorrelated patterns. In this way the overlap between any two patterns is negligible in theory. In a finite network this is not necessarily the case, but for a limited number of patterns the overlaps are smaller than the threshold for the spill-over into any of the other patterns. In a memory network below the capacity limit, the activity will therefore be confined to the pattern which is indicated by the input. The dynamics will thus not be critical. In a critical network on the other hand, we conjecture that on short time scales avalanches will be correlated to the existing patterns. Since, however, all neurons and thus all patterns receive constantly external input now, the avalanches are not confined to a pattern but will jump into other patterns on medium time scales.

4.6. IMPLICATIONS FOR AGING AND MEMORY CONSOLIDATION

Although we have not made this explicit here, the memory test can be formulated as an informational criterion (Rieke et al., 1997). The evolution of a network state toward a pattern decreases the distance and thus also the relative entropy as it is not known which bits are wrong or missing. Interestingly, the amount of information (however specified) does not improve when approaching the critical regime. This is contrary to what could be expected by considering the informational optimality of critical neurodynamics (Shew and Plenz, 2012), although the situation there is not comparable to the present attractor dynamics. The optimality concerns information capacity (Haldeman and Beggs, 2005) in the sense that at criticality the entropy of the state space is maximal (Ramo et al., 2007) with respect to a control parameter, i.e., the state of the network returns only rarely to states visited earlier. Obviously, this property is detrimental to the attractor dynamics of a Hopfield model, which pins the

state near or at a certain memory state. In large networks, the number of states far away from any memory state is very large such that for moderate load a critical dynamics is possible. For low memory load the dynamics stays preferentially inside the patterns (Dasgupta and Herrmann, 2011) but is similarly expected to have an entropy maximum near criticality in the absence of a bias toward one of the patterns.

The network considered here can be characterized by the interplay between the attractor dynamics in memory retrieval and critical dynamics that provides optimal exploration of the state space. In a system like the human brain, where the number of memories increases for a large part of the personal history, it seems that there must be eventual a consequence for the flexibility and the ability to explore new patterns (Schrobsdorff et al., 2009). Considering, however, that the breakdown of memory at the critical capacity is not likely to be realistic, the conclusion that cognitive effects of aging can be explained by the effect of memory-dependent structure in the network on critical dynamics in the network does not immediately follow from the present model.

Although we could show here that memory storage and criticality are not irreconcilable, our results support a view that has been adopted by an increasing number of researchers in the last decade, namely that memory traces are not necessarily point attractors but more general dynamics objects (Herrmann et al., 1995; Natschlaeger et al., 2002; Rabinovich et al., 2008). In these approaches stability of the memories leads to a reduction of the capacity, but there may be the possibility of an active stabilization of the memories not necessarily different from the

regulatory mechanism involved in criticalization. The fact that the mechanism for criticalization in the network needs to be counterbalanced by a mechanism for consolidation of the memories, should thus not be surprising, but it would be interesting to identify mechanisms that achieve both goals at the same time.

5. CONCLUDING REMARKS

We have demonstrated that criticality can be preserved in an attractor network if both a memory consolidation process and a mechanism for regulation toward criticality are present. Among the different mechanisms for the maintenance of a critical regime, we found that the dynamics of synaptic resources is both biologically realistic and effective for criticalization, while their effect on memory capacity is moderate. Other mechanisms are possible, but are less easily biologically justifiable and have eventually a disastrous effect on the memory content unless actively counteracted. It is necessary to consider the present results in more realistic network models and under more general learning paradigms in order to understand better their significance for the natural and pathological development of biological neural systems.

REFERENCES

Abbott, L., and LeMasson, G. (1993). Analysis of neuron models with dynamically regulated conductances. *Neural Comput.* 5, 823–842. doi: 10.1162/neco.1993.5.6.823

Beggs, J., and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *J. Neurosci.* 23, 11167–11177.

Bibitchkov, D., Herrmann, J. M., and Geisel, T. (2002). Pattern storage and processing in attractor networks with short-time synaptic dynamics. *Netw. Comput. Neural Syst.* 13, 115–129. doi: 10.1080/net.13.1.115.129

Chialvo, D. (2006). Are our senses critical? *Nat. Phys.* 2, 301–302. doi: 10.1038/nphys300

Dasgupta, S., and Herrmann, J. M. (2011). Critical dynamics in homeostatic memory networks. *Nat. Prec.* doi: 10.1038/npre.2011.5829.1

Eurich, C. W., Herrmann, M., and Ernst, U. (2002). Finite-size effects of avalanche dynamics. *Phys. Rev. E* 66:066137. doi: 10.1103/PhysRevE.66.066137

Giudice, P. D., and Mattia, M. (2001). Long and short-term synaptic

plasticity and the formation of working memory: a case study. *Neurocomputing* 38, 1175–1180. doi: 10.1016/S0925-2312(01)00557-4

Haldeman, C., and Beggs, J. (2005). Critical branching captures activity in living neural networks and maximizes the number of metastable states. *Phys. Rev. Lett.* 94:058101. doi: 10.1103/PhysRevLett.94.058101

Herrmann, J. M., Hertz, J. A., and Prügel-Bennett, A. (1995). Analysis of synfire chains. *Netw. Comput. Neural Syst.* 6, 403–414. doi: 10.1088/0954-898X/6/3/006

Herrmann, M., Ruppin, E., and Usher, M. (1993). A neural model of the dynamic activation of memory. *Biol. Cybern.* 68, 455–463. doi: 10.1007/BF00198778

Herz, A. V. M., and Hopfield, J. J. (1995). Earthquake cycles and neural reverberations: collective oscillations in systems with pulse-coupled threshold elements. *Phys. Rev. Lett.* 75, 1222–1225. doi: 10.1103/PhysRevLett.75.1222

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational

abilities. *PNAS* 79, 2554–2558. doi: 10.1073/pnas.79.8.2554

Horn, D., and Usher, M. (1989). Neural networks with dynamical thresholds. *Phys. Rev. A* 40:1036. doi: 10.1103/PhysRevA.40.1036

Larremore, D. B., Shew, W. L., and Restrepo, J. G. (2011). Predicting criticality and dynamic range in complex networks: effects of topology. *Phys. Rev. Lett.* 106:58101. doi: 10.1103/PhysRevLett.106.058101

Levina, A. (2008). *A Mathematical Approach to Self-organized Criticality in Neural Networks.* Dissertation (Ph.D. thesis), Niedersächsische Staats- und Universitätsbibliothek Göttingen. Available online at: webdoc.sub.gwdg.de/diss/2008/levina/levina.pdf

Levina, A., Ernst, U., and Herrmann, J. M. (2007a). Criticality of avalanche dynamics in adaptive recurrent networks. *Neurocomputing* 70, 1877–1881. doi: 10.1016/j.neucom.2006.10.056

Levina, A., Herrmann, J. M., and Geisel, T. (2007b). Dynamical synapses causing self-organized criticality in neural networks. *Nat. Phys.* 3, 857–860. doi: 10.1038/nphys758

Levina, A., Herrmann, J. M., and Geisel, T. (2006). "Dynamical synapses give rise to a power-law distribution of neuronal avalanches," in *Advances in Neural Information Processing Systems 18*, eds Y. Weiss, B. Schölkopf, and J. Platt (Cambridge, MA: MIT Press), 771–778.

Levina, A., Herrmann, J. M., and Geisel, T. (2009). Phase transitions towards criticality in a neural system with adaptive interactions. *Phys. Rev. Lett.* 102:118110. doi: 10.1103/PhysRevLett.102.118110

Lin, M., and Chen, T.-L. (2005). Self-organized criticality in a simple model of neurons based on small-world networks. *Phys. Rev. E* 71:016133. doi: 10.1103/PhysRevE.71.016133

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between pyramidal neurons. *Nature* 382, 807–810. doi: 10.1038/382807a0

Natschlaeger, T., Maass, W., and Markram, H. (2002). The "liquid computer": a novel strategy for real-time computing on time series. *Telematik* 8, 39–43.

Otter, R. (1949). The multiplicative process. *Ann. Math. Statist.* 20, 248–263. doi: 10.1214/aoms/1177730031

Rabinovich, M., Huerta, R., and Laurent, G. (2008). Transient dynamics for neuronal processing. *Science* 321, 48–50. doi: 10.1126/science.1155564

Ramo, P., Kaufmann, S., Kesseli, J., and Yliharja, O. (2007). Measures for information propagation in boolean networks. *Physica D* 227, 100–104. doi: 10.1016/j.physd.2006.12.005

Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W. (1997). *Spikes: Exploring the Neural Code.* Cambridge, MA: MIT Press.

Romani, S., Amit, D. J., and Mongillo, G. (2006). Mean-field analysis of selective persistent activity in presence of short-term synaptic depression. *J. Comput. Neurosci.* 20, 201–217. doi: 10.1007/s10827-006-6308-x

Schrobsdorff, H., Ihrke, M., Behrendt, J., Herrmann, J., Geisel, T., and Levina, A. (2009). Are age-related cognitive effects caused by optimization? *BMC Neurosci.* 10(Suppl. 1), P9. doi: 10.1186/1471-2202-10-S1-P9

Shew, W. L., and Plenz, D. (2012). The functional benefits of criticality in the cortex. *Neuroscientist* 19, 88–100. doi: 10.1177/1073858412445487

Teramae, J., and Fukai, T. (2007). Local cortical circuit model inferred from power-law distributed neuronal avalanches. *J. Comput. Neurosci.* 22, 301–312. doi: 10.1007/s10827-006-0014-6

Treves, A. (2005). Frontal latching networks: a possible neural basis for infinite recursion. *Cogn. Neuropsychol.* 22, 276–291. doi: 10.1080/02643290442000329

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835. doi: 10.1162/089976698300017502

Tsodyks, M. V. (1989). Associative memory in neural networks with the Hebbian leanring rule. *Mod. Phys. Lett. B* 3, 555–560. doi: 10.1142/S021798498900087X

Tsodyks, M. V., and Feigel'man, M. V. (1988). The enhanced storage capacity in neural networks with low activity level. *Europhys. Lett.* 6, 101–105. doi: 10.1209/0295-5075/6/2/002

Tsodyks, M. V., and Markram, H. (1996). Plasticity of neocortical synapses enables transitions between rate and temporal coding. *Lecture Notes Comp. Sci.* 1112, 445–450. doi: 10.1007/3-540-61510-5_77

Tsodyks, M. V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *PNAS* 94, 719–723. doi: 10.1073/pnas.94.2.719

Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., and Nelson., S. B. (1998). Activity-dependent scaling of quantal amplitude in neocortical pyramidal neurons. *Nature* 391, 892–896. doi: 10.1038/36103

Short term synaptic depression improves information transfer in perceptual multistability

Zachary P. Kilpatrick*

Department of Mathematics, University of Houston, Houston, TX, USA

Competitive neural networks are often used to model the dynamics of perceptual bistability. Switching between percepts can occur through fluctuations and/or a slow adaptive process. Here, we analyze switching statistics in competitive networks with short term synaptic depression and noise. We start by analyzing a ring model that yields spatially structured solutions and complement this with a study of a space-free network whose populations are coupled with mutual inhibition. Dominance times arising from depression driven switching can be approximated using a separation of timescales in the ring and space-free model. For purely noise-driven switching, we derive approximate energy functions to justify how dominance times are exponentially related to input strength. We also show that a combination of depression and noise generates realistic distributions of dominance times. Unimodal functions of dominance times are more easily told apart by sampling, so switches induced by synaptic depression induced provide more information about stimuli than noise-driven switching. Finally, we analyze a competitive network model of perceptual tristability, showing depression generates a history-dependence in dominance switching.

Keywords: binocular rivalry, neural field, ring model, bump attractor, short term depression

INTRODUCTION

Ambiguous sensory stimuli with two interpretations can produce perceptual rivalry (Blake and Logothetis, 2002). For instance, presenting two orthogonal gratings to either eye results in perception switching between gratings repetitively—binocular rivalry (Leopold and Logothetis, 1996). Perceptual rivalry can also be triggered by a single stimulus with two interpretations, like the Necker cube (Orbach et al., 1963). The switching process in perceptual rivalry is considerably stochastic—a histogram of the dominance times of each percept spreads across a broad range (Fox and Herrmann, 1967). Senses other than vision also exhibit perceptual rivalry. When two different odorants are presented to the two nostrils, a similar phenomenon occurs with olfaction, termed "binaral" rivalry (Zhou and Chen, 2009). Similar experiences have been evoked in the auditory (Deutsch, 1974; Pressnitzer and Hupé, 2006) and tactile (Carter et al., 2008) system.

Several principles govern the relationship between the strength of ambiguous stimuli and the mean switching statistics in perceptual rivalry (Levelt, 1965). "Levelt's propositions" relate stimulus contrast to the *mean dominance times*: (1) increasing the contrast of one stimulus increases the proportion of time that stimulus is dominant; (2) increasing the contrast of one stimulus does not affect its average dominance time; (3) increasing the contrast of one stimulus increases the rivalry alternation rate; and (4) increasing the contrast of both stimuli increases the rivalry alternation rate. Properties of the input also affect the stochastic variation in the dominance times (Brascamp et al., 2006). For instance, a histogram of dominance times is well fit by a gamma distribution (Fox and Herrmann, 1967; Lehky, 1995; van Ee, 2009).

The fact that dominance times are not exponentially distributed suggests some background slow adaptive process plays a role in providing a non-zero peak in the dominance histograms (Shpiro et al., 2009). Two commonly proposed mechanisms for this adaptation are spike frequency adaptation and short term synaptic depression (Laing and Chow, 2002; Wilson, 2003; Shpiro et al., 2007). A stronger case can be made for the existence of adaptation in perceptual processing networks by examining results of experiments on *perceptual tristability* (Hupe, 2010). Here, perception alternates between three possible choices and subsequent switches are determined by the previous switch (Naber et al., 2010). This memory suggests switches in perceptual multistability are not purely noise-driven (Moreno-Bote et al., 2007).

Most theoretical models of perceptual rivalry employ two pools of neurons, each selective to one percept, coupled to one another by mutual inhibition (Matsuoka, 1984; Laing and Chow, 2002; Shpiro et al., 2007; Seely and Chow, 2011). With no other mechanisms at work, such architectures lead to *winner-take-all* states, where one pool of neurons inhibits the other indefinitely (Wang and Rinzel, 1992). However, switches between the dominance of one pool and the other can be initiated with the inclusion of fluctuations (Moreno-Bote et al., 2007) or an adaptive process (Laing and Chow, 2002; Shpiro et al., 2007). Combining the two mechanisms leads to dominance times that are distributed according to the gamma distribution (Laing and Chow, 2002; Shpiro et al., 2009; van Ee, 2009). Thus, slow adaptation and noise allow sampling of the stimulus through changes in network activity.

In light of these observations, we wish to consider the role adaptive mechanisms play in properly sampling ambiguous

stimuli in a mutual inhibitory network. Two stimuli of different orientations are presented to the network (Levelt, 1965). The network outputs a time-dependent, orientation-dependent firing rate, whose peak switches between two locations determined by the two stimuli. We think of the information output by the network as a series of dominance times. We will study how well the relative strength of the two stimuli (information) is encoded by the amount of time each subpopulation remains active during a dominance period (Levelt, 1965; Moreno-Bote et al., 2007). Purely fluctuation driven switching provides a noisy sample of the two percepts, but adaptation driven switching provide an extremely reliable sampling of percept contrast (Shpiro et al., 2009). As the level of adaptation is increased and noise is decreased, mutual inhibitory networks encode information about ambiguous stimuli better. We focus specifically on the adaptive mechanism of short term synaptic depression (Tsodyks and Markram, 1997).

Using parameterized models, we will explore how synaptic depression improves the ability of a network to extract stimulus contrasts. First, we study how much information can be determined about the contrast of each of the two percepts of an ambiguous stimulus. In the case of a *winner-take-all* solution, only information about a single percept can be known, since the pool of neurons encoding the other percept is quiescent. We will study this using an anatomically motivated neural field model of an orientation column with synaptic depression (York and van Rossum, 2009; Kilpatrick and Bressloff, 2010a). Increasing the strength of synaptic depression leads to a bifurcation which produces rivalrous oscillations. When rivalrous switching occurs through a combination of depression and noise, we show stronger depression improves the transfer of information. We also analyze a reduced network model with depression and noise to help study the combined effects of noise and depression on perceptual switching. Finally, we study perceptual tristability as oscillations generated in a three population network, where each population spends time in dominance. This shows depression generates a history dependence in switching that would not arise in the network with purely noise-driven switching.

MATERIALS AND METHODS

RING MODEL WITH SYNAPTIC DEPRESSION

As a starting point, we consider a model for processing the orientation of visual stimuli (Ben-Yishai et al., 1995; Bressloff and Cowan, 2002) which also includes short term synaptic depression (York and van Rossum, 2009; Kilpatrick and Bressloff, 2010a). Since GABAergic inhibition is much faster than AMPA-mediated excitation (Kawaguchi and Kubota, 1997), we assume that inhibition is slaved to excitation as in Amari (1977). Reduction this disynaptic pathway and assuming depression acts on excitation (Tsodyks and Markram, 1997), we then have the model

$$\tau_m \dot{u} = -u(x, t) + w * (qf(u)) + I(x) + \xi(x, t), \quad (1a)$$

$$\tau \dot{q} = 1 - q(x, t) - \beta q(x, t) f(u(x, t)). \quad (1b)$$

Here $u(x, t)$ measures the synaptic input to the neural population with stimulus preference $x \in [-\pi/2, \pi/2]$ at time t, evolving

on the timescale τ_m. Synaptic interactions are described by the integral term

$$w * (qf(u)) = \int_{-\pi/2}^{\pi/2} w(x - y) q(y, t) f(u(y, t)) dy,$$

so $w(x - y)$ describes the strength (amplitude of w) and net polarity (sign of w) of synaptic interactions from neurons with stimulus preference y to those with preference x. The modulation of the synaptic strength is given by the cosine

$$w(x - y) = \cos(2(x - y)), \quad (2)$$

so neurons with similar orientation preference excite one another and those with dissimilar orientation preference disynaptically inhibit one another (Ben-Yishai et al., 1995; Ferster and Miller, 2000). The factor $q(x, t)$ measures of the fraction of available presynaptic resources, which are depleted at a rate βf (Tsodyks and Markram, 1997), and are recovered on a timescale specified by the time constant τ (Chance et al., 1998). Firing rates are given by taking the gain function $f(u)$ of the synaptic input, which we usually proscribe to be (Wilson and Cowan, 1973)

$$f(u) = \frac{1}{1 + e^{-\gamma(u - \kappa)}}, \quad (3)$$

and often take the $\gamma \to \infty$, so (Amari, 1977)

$$f(u) = H(u - \kappa) = \begin{cases} 0 : u < \kappa, \\ 1 : u \geq \kappa. \end{cases} \quad (4)$$

External input, representing flow from upstream in the visual system is prescribed by the time-independent function $I(x)$ (Ben-Yishai et al., 1995; Bressloff and Cowan, 2002). For the majority of our study of Equation (1), we employ the bimodal stimulus

$$I(x) = -I_0 \cos(4x) + I_a \sin(2x), \quad (5)$$

representing stimuli at the two orthogonal angles $-\pi/4$ and $\pi/4$ and I_0 controls the mean of each peak and I_a controls the level of asymmetry between the peaks. Effects of noise are described by the stochastic process $\langle \xi(x, t) \rangle$ with $\langle \xi(x, t) \rangle = 0$ and $\langle \xi(x, t) \xi(y, s) \rangle = C(x - y) \delta(t - s)$, and spatial correlations are take to have a cosine profile $C(x) = \pi \cos(x)$.

We assume units of time t to be 10 ms each. Excitatory synaptic time constants are roughly 10 ms (Häusser and Roth, 1997), so we set $\tau_m = 1$ (10 ms). Experimental observations have shown synaptic resources specified q are recovered on a timescale of 200–800 ms (Tsodyks and Markram, 1997), so we require τ is between 20 and 80, usually setting it to be $\tau = 50$. Our parameter β can then be varied independently to adjust the effective depletion rate of synaptic depression. In our numerical simulations, we typically use the winner-take-all state as the initial condition.

IDEALIZED COMPETITIVE NEURAL NETWORK

We also study space-free competitive neural networks with synaptic depression (Shpiro et al., 2007). As a general model of

networks connected by mutual inhibition, we consider the system (Laing and Chow, 2002; Moreno-Bote et al., 2007; Shpiro et al., 2007)

$$\dot{u}_R = -u_R(t) + f(I_R - q_L(t)u_L(t)) + \xi_1(t), \tag{6a}$$

$$\dot{u}_L = -u_L(t) + f(I_L - q_R(t)u_R(t)) + \xi_2(t), \tag{6b}$$

$$\tau\dot{q}_R = 1 - q_R(t) - \beta u_R(t)q_R(t), \tag{6c}$$

$$\tau\dot{q}_L = 1 - q_L(t) - \beta u_L(t)q_L(t), \tag{6d}$$

where $u_j(t)$ represents the firing rate of the $j = L, R$ population. The resource usage rate by synapse projecting from population $j = L, R$ is specified by $\beta u_j q_j$ and the resource recovery timescale is τ. Fluctuations are introduced into population j with the independent white noise processes ξ_j with $\langle x_j(t) \rangle = 0$ and $\langle \xi_j(t)\xi_j(s) \rangle = \varepsilon\delta(t-s)$. Units of time are taken to be 10 ms each. In numerical simulations, $u_j(0)$ are initialized by randomly drawing from a uniform distribution on $[0, 1]$; $q_j(0)$ are initialized by randomly drawing from a uniform distribution on $[1/(1+\beta), 1]$.

NUMERICAL SIMULATION OF STOCHASTIC DIFFERENTIAL EQUATIONS

The spatially extended model (Equation 1) is simulated using an Euler–Maruyama method with a timestep `dt = 10⁻⁴`, using Riemann integration on the convolution term with 2000 spatial grid points. A population is considered dominant if the peak of its activity bump is higher than the other; switches occur when the other bump attains a higher peak. The reduced network (Equation 6) was also simulated using Euler–Maruyama with a timestep `dt = 10⁻⁶`. Population j is considered dominant when $u_j > u_k$ $(j \neq k)$; switches occur when the inequality switches direction. To generate histograms of dominance times, we simulated systems for 20,000s.

FITTING DOMINANCE TIME DISTRIBUTIONS

To generate the theoretical curves presented for exponentially distributed dominance times, we simply take the mean of dominance times and use it as the scaling in the exponential (Equation 28). For those densities that we presume are gamma distributed, we solve a linear system to fit the constants c_1, c_2, and c_3 of

$$f(T) = e^{c_1} T^{c_2} e^{-c_3 T} \tag{7}$$

an alternate form of Equation (30). Upon taking the logarithm of Equation (7), we have the linear sum

$$\ln f(T) = c_1 + c_2 \ln T - c_3 T. \tag{8}$$

Then, we select three values of the numerically generated distribution $p^n(T^n)$ along with its associated dominance times: (T_1^n, p_1^n); (T_2^n, p_2^n); (T_3^n, p_3^n) where $p_j^n = p^n(T_j^n)$. We always choose $T_2^n = \arg\max_T p^n(T)$ as well as $T_1^n = T_2^n/2$ and $T_3^n = 3T_2^n/2$. It is then straightforward to solve the linear system

$$\begin{pmatrix} 1 & \ln T_1^n & -T_1^n \\ 1 & \ln T_2^n & -T_2^n \\ 1 & \ln T_3^n & -T_3^n \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \ln p_1^n \\ \ln p_2^n \\ \ln p_3^n \end{pmatrix}$$

using the\command in `MATLAB`.

RESULTS

We now present results that reveal the importance of synaptic depression in preserving information about bimodal stimuli. No previous work, to our knowledge, has studied how activity in a ring model with depression (Equation 1) can be collapsed to a low dimensional oscillation. The oscillation results from a combination of depression and mutual inhibition, which produces population dominance times and can thus be sampled to give information about the strength of the stimulus that produced them. Once noise is added to these low dimensional oscillations, dominance time distributions still remain relatively tight, which can be sampled to infer relative contrasts of each input. We contrast this with a previous cue orientation selective model which used a heterogeneous population of spiking neurons with lateral inhibition and slow adaptation, so chaos rather than noise produced apparent stochasticity in dominance times (Laing and Chow, 2002). We can use an energy function for a reduced system to approximate the relative effect of depression and noise on dominance times. These energy methods are also useful in the study of perceptual tristability, where we also show depression introduces a history dependence in dominance transitions.

DETERMINISTIC SWITCHING IN THE RING MODEL

To start we consider the ring model with depression (Equation 1) in the absence of noise, so $\xi \equiv 0$. In previous work, noise-free versions of Equation (1) have been analyzed to explore how synaptic depression can generate traveling pulses (York and van Rossum, 2009; Kilpatrick and Bressloff, 2010b), self-sustained oscillations (Kilpatrick and Bressloff, 2010b), and spiral waves in two-dimensions (Kilpatrick and Bressloff, 2010c). Here, we will extend previous work that explored input-driven oscillations in two-layer networks possessing statistics matching binocular rivalry (Kilpatrick and Bressloff, 2010a). We think of Equation (1) as a model of *monocular rivalry*, since oscillations can be due to competition between representations in a single orientation column (Ben-Yishai et al., 1995). Competition between ocular dominance columns (Kilpatrick and Bressloff, 2010a) is not necessary for our theory. For exposition, we will employ specific functional forms: cosine weight (Equation 2); a Heaviside firing rate function (Equation 4); and a bimodal input (Equation 5).

Winner take all state

We now look for winner-take-all solutions, as shown in **Figure 1A**. These states consist of a single activity bump arising in the network, representing only one of the two percepts contained in the bimodal stimulus (Equation 5). These are stationary in time, so $u_t = q_t = 0$, implying $u = U(x)$ and $q = Q(x)$. Also, they are single bump solutions, so there is a single region $x \in (\pi/4 - a, \pi/4 + a)$ that is superthreshold ($U(x) > \kappa$). The parameter a is the half-width of the bump. We assume the right stimulus is represented by a bump, although we can derive analogous results when the left stimulus is represented. The steady state solution is then determined

$$U(x) = \int_{\pi/4-a}^{\pi/4+a} \cos(2(x-y))Q(y)\mathrm{d}y - I_0 \cos(4x)$$

$$+ I_a \sin(2x), \tag{9}$$

$$Q(x) = [1 + \beta H(U(x) - \kappa)]^{-1}, \quad (10)$$

so by plugging Equation (10) into (9) and using $\cos(2(x - y)) = \cos(2x)\cos(2y) + \sin(2x)\sin(2y)$ we have

$$U(x) = A\cos(2x) + (B + I_a)\sin(2x) - I_0\cos(4x),$$

where the constants A, B can be computed

$$A = \frac{1}{1 + \beta} \int_{\pi/4 - a}^{\pi/4 + a} \cos(2x)\mathrm{d}x = 0,$$

$$B = \frac{1}{1 + \beta} \int_{\pi/4 - a}^{\pi/4 + a} \sin(2x)\mathrm{d}x = \frac{\sin(2a)}{1 + \beta}.$$

Therefore, by simplifying the threshold condition, $U(\pi/4 \pm a) = \kappa$, we have

$$U(\pi/4 \pm a) = \frac{\sin(4a)}{2(1 + \beta)} + I_0\cos(4a) + I_a\cos(2a) = \kappa. \quad (11)$$

The implicit Equation (11) can be solved numerically using root finding algorithms. For symmetric inputs ($I_a = 0$), we can solve (Equation 11) explicitly

$$a = \frac{1}{2}\tan^{-1}\left[\frac{1 \pm \sqrt{1 + 4(1 + \beta)^2(I_0^2 - \kappa^2)}}{2(1 + \beta)(I_0 + \kappa)}\right], \quad (12)$$

and winner-take-all solutions take the form

$$U(x) = \frac{\sin(2a)}{1 + \beta}\sin(2x) - I_0\cos(4x) + I_a\sin(2x). \quad (13)$$

With this solution, we can relate the parameters of the model to the existence of the winner-take-all state. To do so, we need to look at a second condition that must be satisfied, $U(x) < \kappa$ for all $x \notin (\pi/4 - a, \pi/4 + a)$. Since the function (Equation 13) is bimodal across $(-\pi/2, \pi/2)$, we check the other possible local

maximum at $x = -\pi/4$ as

$$U(\pi/4) = I_0 - I_a - \frac{\sin(2a)}{1 + \beta} < \kappa. \quad (14)$$

At the point in parameter space where the Equation (14) is violated, a bifurcation occurs, so the winner-take-all state ceases to exist. This surface in parameter space is given by the equation

$$I_0 = \kappa + I_a + \frac{\sin(2a)}{1 + \beta}, \quad (15)$$

along with the explicit formula for the bump half-width (Equation 12). Beyond the bifurcation boundary (Equation 15), one of two behaviors can occur. Either there is a symmetric two-bump solution that exists, the fusion state (Wolfe, 1986; Blake, 1989; Shpiro et al., 2007), or rivalrous oscillations (Levelt, 1965; Blake and Logothetis, 2002).

Fusion state

Experiments on ambiguous stimuli have shown sufficiently strong contrast rivalrous stimuli can be perceived as a single fused image (Blake, 1989; Buckthought et al., 2008). This should not be surprising, considering stereoscopic vision and audition behave in exactly this way (Wolfe, 1986). However, the contrast necessary to evoke this state with dissimilar images is much higher than with similar images (Blake and Logothetis, 2002). The fusion state (**Figure 1C**) is represented as two disjoint bumps. Therefore

$$U(x) = \frac{1}{1 + \beta}\left[\int_{-\pi/4 - b}^{-\pi/4 + b} + \int_{\pi/4 - a}^{\pi/4 + a}\right]\cos(2(x - y))\mathrm{d}y$$
$$- I_0\cos(4x) + I_a\sin(2x).$$

Computing the integrals, we find

$$U(x) = \frac{\mathcal{S}(x, a) - \mathcal{S}(x, b)}{1 + \beta} - I_0\cos(4x) + I_a\sin(2x), \quad (16)$$

FIGURE 1 | Three possible active states of the noise-free stimulus driven ring model with depression (Equation 1). (A) Winner take all ($I_0 = 0.6$) defined by the bump half-width a; **(B)** Rivalrous oscillations ($I_0 = 0.84$); **(C)** Fusion ($I_0 = 1$), where initial condition is (Equation 16). Other parameters are $\kappa = 0.5$, $\beta = 1$, and $\tau = 50$.
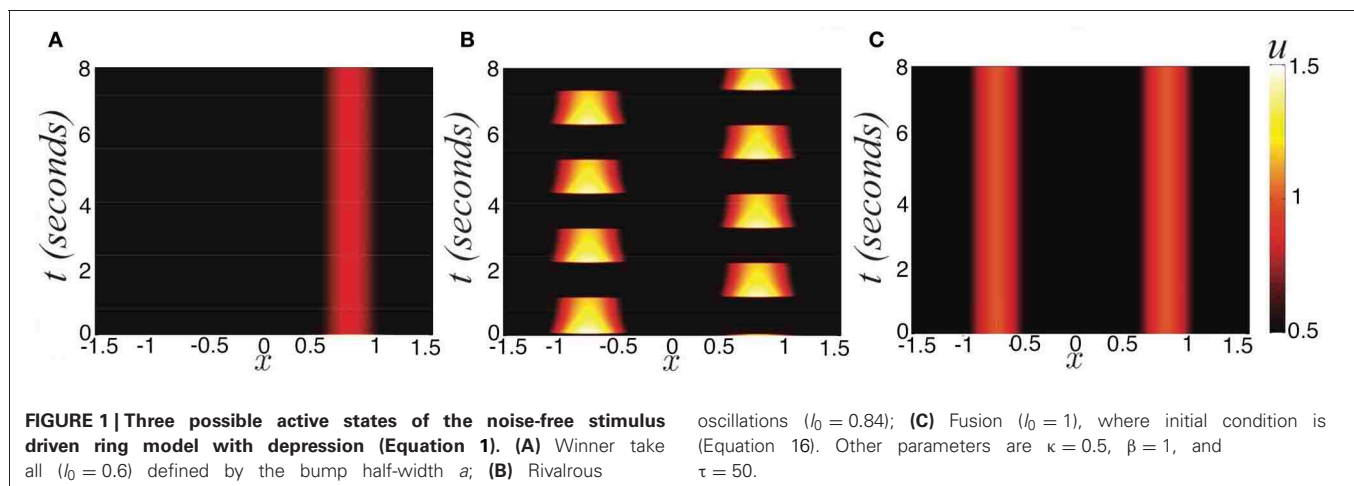
where $\mathcal{S}(x, y) = \sin^2(x + y) - \sin^2(x - y)$. Requiring the threshold conditions $U(-\pi/4 \pm b) = U(\pi/4 \pm a) = \kappa$ are satisfied,

$$\frac{\mathcal{C}(a, b)}{1 + \beta} + I_0 \cos(4a) + I_a \cos(2a) = \kappa,$$

$$\frac{\mathcal{C}(b, a)}{1 + \beta} + I_0 \cos(4b) - I_a \cos(2b) = \kappa,$$

where $\mathcal{C}(x, y) = \cos(2x)[\sin(2x) - \sin(2y)]$, which implicitly relates parameters to the half-widths a, b of each bump. We will now study rivalrous oscillations by simply constructing them using a fast-slow analysis.

Rivalrous oscillations

Oscillations can occur, where the two bump locations trade dominance successively (**Figure 1B**). We will show Levelt's proposition (i) holds; increasing the contrast of a stimulus (**Figures 2A–C**) increases the proportion of time that stimulus is dominant (**Figures 2D–F**). This information is not revealed when the system is stuck in a winner-take-all state. Thus, synaptic depression can unmask otherwise hidden stimuli. We will also examine how well the noise-free version of Equation (1) recapitulates Levelt's other propositions concerning the mean dominance of percepts.

To study oscillations, we assume that the timescale of synaptic depression $\tau \gg \tau_m$, is long enough that we can decompose (Equation 1), with $\xi \equiv 0$, into a fast and slow system

FIGURE 2 | Dependence of rivalry dominance times on the amplitudes of the bimodal input (Equation 5). (A–C) Various profiles of the external input $I(x)$, showing only positive part. Increasing I_0 increases both peaks; increasing I_a decreases the left and increases the right peak. **(D–F)** Rivalrous oscillations in the neural activity $u(x, t)$ corresponding to the input in **(A–C)**. Dominance times decrease from **(D)** to **(E)** since the input amplitude increases from **(A)** to **(B)**. **(F)** Dominance time of right input (red bar : $T_R \approx 0.9$ s) is longer than left (blue bar : $T_L \approx 0.6$ s) for asymmetric input in **(C)**. **(G)** Increasing the strength of the symmetric ($I_a = 0$) bimodal input (Equation 5) decreases the dominance time T of both populations. Our theory (black) computed from fast-slow analysis (Equation 19) fits results of numerical simulations (blue) well. **(H)** For asymmetric inputs ($I_a \neq 0$), we find that varying $I_R = I_0 + I_a$ while keeping $I_L = I_0 - I_a$ fixed changes the dominance times of the left percept T_L (blue) much more than that of the right percept T_R (red). Other parameters are $\kappa = 0.5$, $\beta = 1$, and $\tau = 50$.

(Laing and Chow, 2002; Kilpatrick and Bressloff, 2010a). Synaptic input u then tracks the slowly varying state of the synaptic scaling term q. We have also verified in simulations q is essentially piecewise constant in space, in the case of the Heaviside non-linearity (Equation 4), which yields

$$u(x, t) \approx \int_{-\pi/2}^{\pi/2} \cos(2(x - y))q(y, t)H(u(y, t) - \kappa)dy$$
$$- I_0 \cos(4x), \qquad (17)$$

and q is governed by Equation (1b). To start, we will also assume a symmetric bimodal input ($I_a = 0$). This way, we can simply track q in the interior of one of the bumps, given $q_i(t) = q(\pi/4, t)$. Solving the resulting piecewise system of differential equations, we can derive an implicit formula for

$$q_0 = \frac{1}{1 + \beta} + \frac{\beta}{1 + \beta}e^{-T/\tau} - (1 - q_0)e^{-2T/\tau}, \qquad (18)$$

the value of the synaptic depression variable inside a bump just prior to a switch. We can rearrange (Equation 18) to yield a formula for the dominance time

$$T = \tau \ln \left[\frac{\beta + \sqrt{\beta^2 - 4(1 + \beta)(1 - q_0)[(1 + \beta)q_0 - 1]}}{2(1 + \beta)q_0 - 2} \right], \qquad (19)$$

so that we now must specify the value q_0. We can examine the fast Equation (17), solving for the form of the slowly narrowing right bump during its dominance phase

$$u(x, t) = q_i(t) \left[\sin^2(x + a(t)) - \sin^2(x - a(t)) \right]$$
$$- I_0 \cos(4x). \qquad (20)$$

We solve for the slowly changing width $a(t)$ by enforcing the threshold condition $u(\pi/4 \pm a(t), t) = \kappa$ and using trigonometric identities to find

$$a(t) = \frac{1}{2} \tan^{-1} \left[\frac{q_i(t) + \sqrt{q_i(t)^2 + 4(I_0^2 - \kappa^2)}}{2(I_0 + \kappa)} \right]. \qquad (21)$$

We can also identify the maximal value of $q_i(t) = q_0$ which still leads to the right bump suppressing the left. Once $q_i(t)$ falls below q_0, the other bump escapes suppression, flipping the dominance of the current bump. This is the point at which the other bump of Equation (20) rises above threshold, as defined by the equation $I_0 - q_0 \sin(2a_0) = \kappa$. Combining this with Equation (21) and solving the resulting algebraic equation, we find

$$q_0 = \frac{2I_0 \sqrt{(I_0 - \kappa)(3I_0 + \kappa)}}{3I_0 + \kappa}. \qquad (22)$$

The amplitude of synaptic depression is excluded from Equation (22), but we know $q_0 \in ([1 + \beta]^{-1}, 1)$. This establishes a bounded region of parameter space in which we can expect to find rivalrous oscillations, which we use to construct a partitioning of parameter space in **Figure 3**. We can also now approximate the dominance time using Equation (19) with (22), as shown in **Figure 2G**.

In the case of an asymmetric bimodal input ($I_a > 0$), we can also solve for explicit approximations to the dominance times of the right T_R and left T_L populations. Following the same formalism as for the symmetric input case

$$T_R = \tau \ln \left[\frac{Q_+ + \sqrt{Q_+^2 - B_R}}{2(1 + \beta)q_R - 2} \right], \qquad (23)$$

$$T_L = \tau \ln \left[\frac{Q_- + \sqrt{Q_-^2 - B_L}}{2(1 + \beta)q_L - 2} \right], \qquad (24)$$

where $Q_\pm = \beta \pm (1 + \beta)(q_R - q_L)$ and $B_{R,L} = 4(1 + \beta)(1 - q_{L,R})[(1 + \beta)q_{R,L} - 1]$, in terms of the local values q_L and q_R of the synaptic scaling in the right and left bump immediately prior to their suppression. Notice when $q_L = q_R$, then $q_d = 0$ and Equations (23) and (24) reduce to Equation (19). We now need to examine the fast Equation (17) to identify these two values. This is done by generating two implicit equations for the half-width of the right bump a_R and q_R at the time of a switch

$$\frac{q_R}{2} \sin(4a_R) + I_0 \cos(4a_R) + I_a \cos(2a_R) = \kappa,$$

$$I_0 - I_a - q_R \sin(2a_R) = \kappa,$$

FIGURE 3 | Partition of parameter space (β, I_0) into various stimulus-induced states of (Equation 1) when $\xi \equiv 0$, $\kappa = 0.5$, and $\tau = 50$.
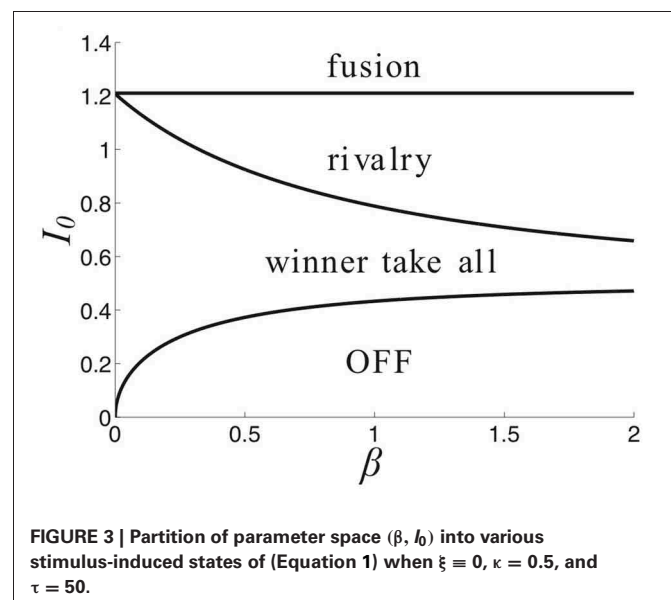
which we can solve explicitly for

$$a_R = \frac{1}{2}\cos^{-1}\left[\frac{\kappa}{2I_0} + \frac{1}{2}\right],$$

and

$$q_R = \frac{2I_0(I_L - \kappa)}{\sqrt{(3I_0 + \kappa)(I_0 - \kappa)}}, \qquad (25)$$

where $I_L = I_0 - I_a$ is the strength of input to the left side of the network. Likewise, we can find the value of the synaptic scaling in the left bump immediately prior to its suppression

$$q_L = \frac{2I_0(I_R - \kappa)}{\sqrt{(3I_0 + \kappa)(I_0 - \kappa)}}, \qquad (26)$$

where $I_R = I_0 + I_a$ is the strength of input to the right side of the network. Using the expressions (25) and (26) we can now compute the dominance time formulae (23) and (24), showing the relationship between inputs and dominance times in **Figure 2H**. Notice that all of Levelt's propositions are essentially satisfied. Changing the strength of the right stimulus I_R has a very weak effect on the dominance time of the right percept. Thus, dominance times obey the classic description of Levelt's second proposition (Levelt, 1965). Recent evidence does suggest this only holds at high contrast (Brascamp et al., 2006), and our study is consistent with this since inputs are high contrast here, since it lies just below a fusion state. This is characteristic of competitive networks whose switches occur via an escape mechanism (Wang and Rinzel, 1992; Shpiro et al., 2007), whereby the suppressed population comes on and overtakes the previously dominant population.

Finally, we demonstrate how the strength of a symmetric input I_0 and strength of depression β lead to different behaviors of the network (Equation 1) in **Figure 3**. For weaker synaptic depression strength β, there is a narrower range of stimulus strengths I_0 for which rivalrous oscillations exist. When synaptic depression is sufficiently strong, the range of I_0 that leads to a winner-take-all state narrows. For sufficiently strong I_0, increasing β leads to a network that reveals a piece of the stimulus that would otherwise be kept hidden. As we will show, synaptic depression helps the network reveal stimulus information in a way that is much more reliable than noise.

PURELY STOCHASTIC SWITCHING IN THE RING MODEL

We will now study rivalrous switching brought about by fluctuations. In particular, we ignore depression and examine the noisy system

$$\dot{u}(x, t) = -u(x, t) + w * f(u) + I(x) + \xi(x, t). \qquad (27)$$

where $\langle \xi(x, t) \rangle = 0$ and $\langle \xi(x, t)\xi(y, s) \rangle = \varepsilon C(x - y)\delta(t - s)$ defines the spatiotemporal correlations of the system. Since there is no synaptic depression in the model (Equation 27), no deterministic mechanisms will generate switches between one

winner-take-all state and another. Thus, consider the effects of introducing a small amount of noise ($0 < \varepsilon \ll 1$), reflective of synaptic fluctuations, with spatial correlation function $C(x) = \cos(x)$. Noise generates switches in between the two dominant states (**Figure 4A**). Activity of neurons not driven by the stimulus remains close to zero even during dominance switching. There will be no mixing of the two inputs in the networks representation of the stimulus. Dominance switching occurs via an escape mechanism (Wang and Rinzel, 1992), whereby noise drives the suppressed population on, which in turn suppresses the dominant population. As opposed to depression-induced switching, there is an exponential spread in the possible dominance times for a given set of parameters (**Figure 4B**). By sampling two dominance times back to back, it may be difficult to tell if the input strengths are roughly the same or not.

We now explore the task of discerning the relative contrasts of the two stimuli I_R and I_L based on samples of the dominance time distributions. Notice in **Figure 5** that the likelihood assigned to $I_R > I_L$ approaches 1/2 as the number of observations n increases.
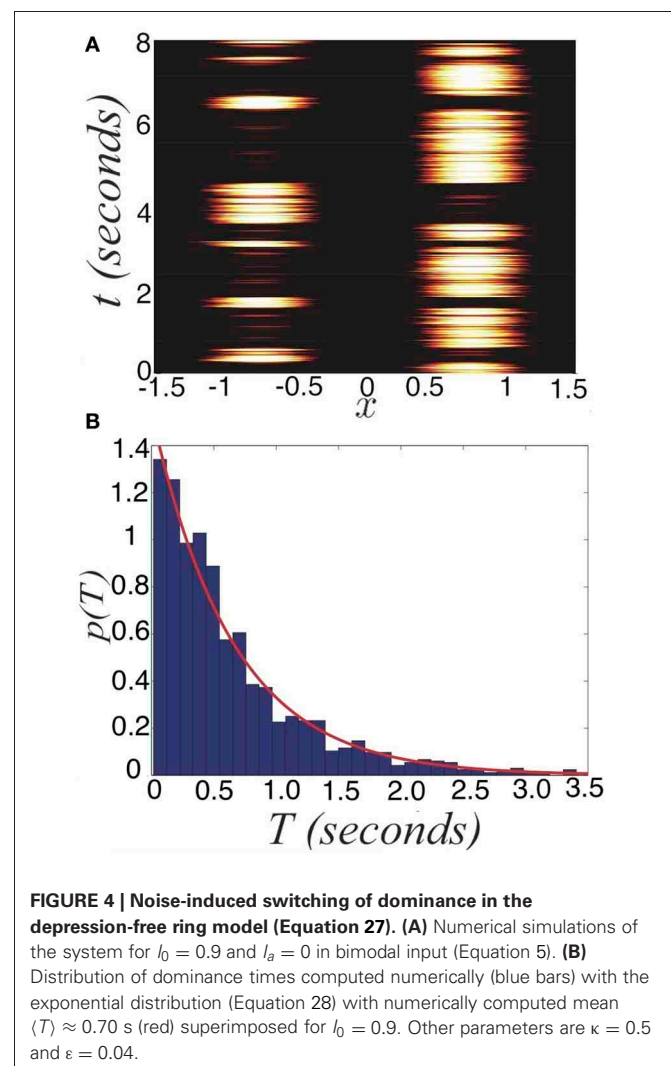
FIGURE 4 | Noise-induced switching of dominance in the depression-free ring model (Equation 27). (A) Numerical simulations of the system for $I_0 = 0.9$ and $I_a = 0$ in bimodal input (Equation 5). **(B)** Distribution of dominance times computed numerically (blue bars) with the exponential distribution (Equation 28) with numerically computed mean $\langle T \rangle \approx 0.70$ s (red) superimposed for $I_0 = 0.9$. Other parameters are $\kappa = 0.5$ and $\varepsilon = 0.04$.

We compute $p[I_R > I_L | T^*(n)]$, the predicted probability $I_R > I_L$ based on sampling dominance time pairs from n cycles $T^*(n) = \{T_R^{(1)}, T_L^{(1)}; T_R^{(2)}, T_L^{(2)}; \ldots; T_R^{(n)}, T_L^{(n)}\}$. As $n \to \infty$, the exponential distributions approximately defining the identical probability
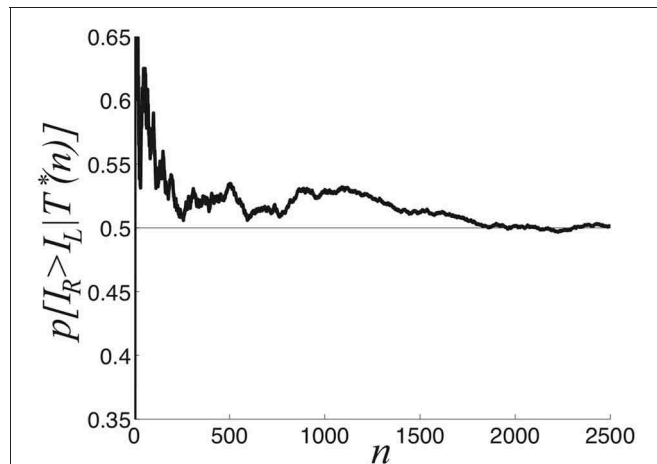
FIGURE 5 | Predicted probability right input I_R is higher than the left input I_L, based on the sampling n cycles ($2n$ switches between percepts), for symmetric inputs $I_L = I_R = 0.9$. After 2000 cycles, $p[I_R > I_L | T^*(n)] \approx 0.5$. Other parameters are $\kappa = 0.5$ and $\varepsilon = 0.04$.

densities $p_R(T_R) = p_L(T_L) = p(T)$ are fully sampled and $p(I_R > I_L | T^*(\infty)) = 1/2$, as in **Figure 5**.

We explore this further in the case of asymmetric inputs, showing dominance times are still specified by exponential distributions as shown in **Figure 6**. Despite the fact $I_R > I_L$, the exponential distributions $p(T_R)$ and $p(T_L)$ still have substantial overlap, so sampling from these distributions can yield $T_R < T_L$. Using such a sample to guess the ordering of amplitudes I_R and I_L would yield $I_R < I_L$, rather than the correct $I_R > I_L$. In terms of conditional probabilities, we expect situations where $p(I_R > I_L | T^*(n)) < 1/2$ for finite n, even though $I_R > I_L$. We can quantify this effect numerically, as shown in **Figure 6B**. Since the marginal distributions are approximately exponential

$$p_j(T_j) = e^{-T_j/\langle T_j \rangle}/\langle T_j \rangle \quad j = L, R, \tag{28}$$

we can approximate the conditional probability

$$\begin{aligned} p[I_R > I_L | T^*(\infty)] &= \int_0^\infty \int_0^x p_R(x) p_L(y) \mathrm{d}y \mathrm{d}x \\ &= \frac{\langle T_R \rangle}{\langle T_R \rangle + \langle T_L \rangle}. \end{aligned} \tag{29}$$

Using Equation (29), we can estimate the limit $p(I_R > I_L | T^*(\infty))$ (**Figure 6B**). Recent psychophysical experiments suggest humans
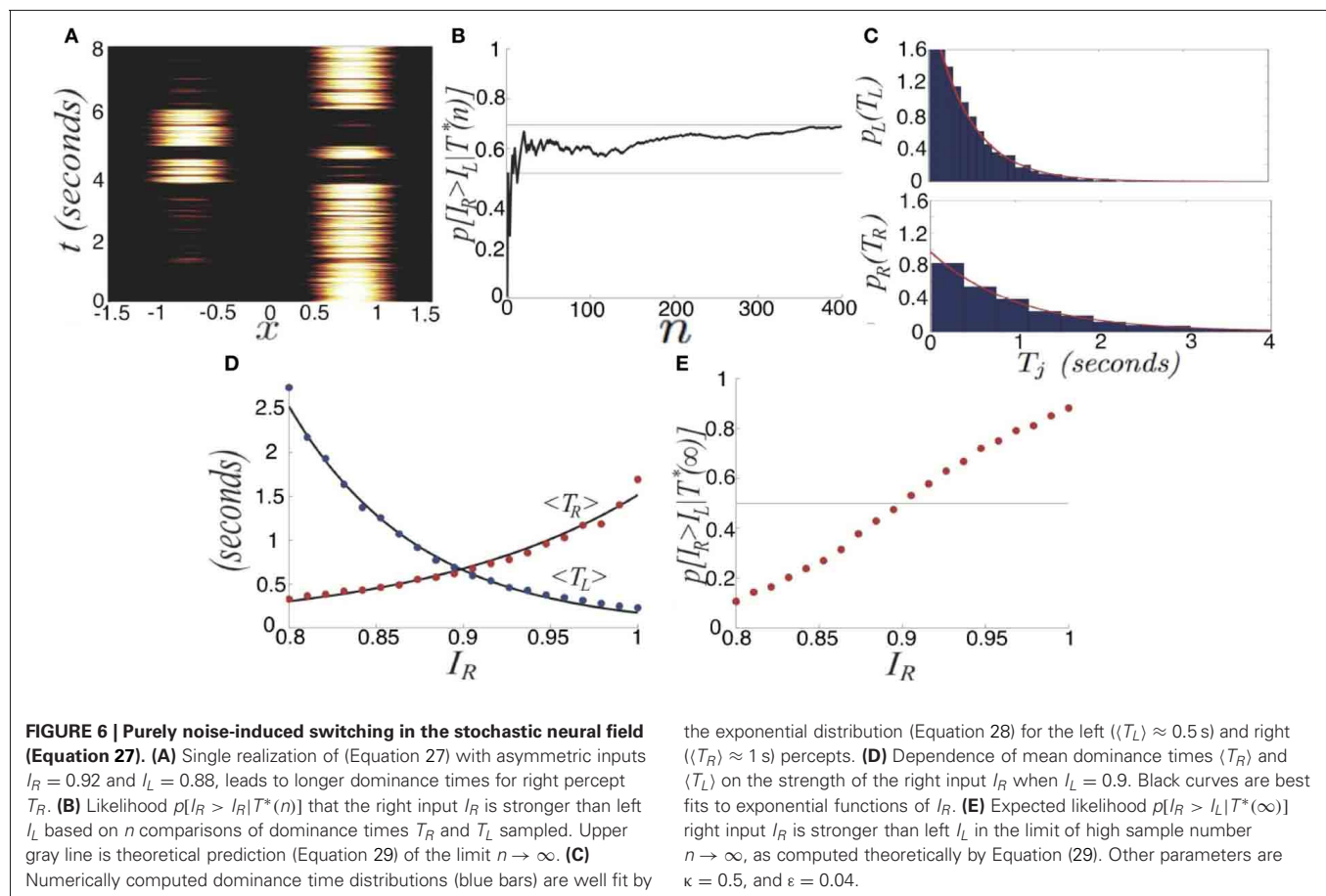
FIGURE 6 | Purely noise-induced switching in the stochastic neural field (Equation 27). (A) Single realization of (Equation 27) with asymmetric inputs $I_R = 0.92$ and $I_L = 0.88$, leads to longer dominance times for right percept T_R. **(B)** Likelihood $p[I_R > I_L | T^*(n)]$ that the right input I_R is stronger than left I_L based on n comparisons of dominance times T_R and T_L sampled. Upper gray line is theoretical prediction (Equation 29) of the limit $n \to \infty$. **(C)** Numerically computed dominance time distributions (blue bars) are well fit by the exponential distribution (Equation 28) for the left ($\langle T_L \rangle \approx 0.5$ s) and right ($\langle T_R \rangle \approx 1$ s) percepts. **(D)** Dependence of mean dominance times $\langle T_R \rangle$ and $\langle T_L \rangle$ on the strength of the right input I_R when $I_L = 0.9$. Black curves are best fits to exponential functions of I_R. **(E)** Expected likelihood $p[I_R > I_L | T^*(\infty)]$ right input I_R is stronger than left I_L in the limit of high sample number $n \to \infty$, as computed theoretically by Equation (29). Other parameters are $\kappa = 0.5$, and $\varepsilon = 0.04$.

would perform this task of contrast differentiation of bistable images in this way (Moreno-Bote et al., 2011).

We also see the mean dominance times still obey Levelt's propositions (**Figure 6D**). Thus, comparing the mean dominance times $\langle T_R \rangle$ and $\langle T_L \rangle$ provides very precise information about the ordering of contrasts I_R and I_L. However, when comparing successive dominance times, accurately discerning the relative input contrasts is more difficult. This becomes more noticeable when the input contrasts are quite close to one another, as in **Figure 6E**. We will explore now how introducing depression along with noise improves discernment of the input contrasts by an observer using simple comparison of dominance times.

SWITCHING THROUGH COMBINED DEPRESSION AND NOISE

We now study the effects of combining noise and depression in the full ring model of perceptual rivalry (Equation 1). Numerical simulations of Equation (1) reveal that noise-induced switches occur robustly, even in parameter regimes where the noise-free system supports no rivalrous oscillations, as shown in **Figure 7**. Rather than dominance times being distributed exponentially, they roughly follow a gamma distribution (Fox and Herrmann, 1967; Lehky, 1995)

$$p_j(T_j) = \frac{1}{\sigma^k \Gamma(k)} T_j^k \exp\left[-T_j/\sigma\right], \quad k > 1, \quad (30)$$

which is peaked away from zero at $T_j = k\sigma$, the mean of the distribution. We show two gamma distributions of dominance times with different means can be more easily discerned than two exponential distributions. Gamma distributions with different means are better separated than two exponential distributions. We summarize how this separation improves the inference of relative contrast in **Figure 8**. As the strength β of depression is increased discernment of relative contrast from sampling dominance time distributions is improved. The likelihood assigned to I_R being greater than I_L is a sigmoidal function of I_R whose steepness increases with β. For no noise, the likelihood function is simply a step function $H(I_R > I_L)$, implying perfect discernment.

ANALYZING SWITCHING IN A REDUCED MODEL

We now perform similar analysis on a reduced network model (Equation 6) and extend some of the results for the ring model. We can construct an energy function (Hopfield, 1984), which provides us with intuition as to the exponential dependence of mean dominance times on input strengths in the noise-driven case. In particular, we analyze Equation (6) where the firing rate function is Heaviside (Equation 4), starting with the case of no noise

$$\dot{u}_R = -u_R + H(I_R - q_L u_L), \quad (31a)$$

$$\dot{u}_L = -u_L + H(I_L - q_R u_R) \quad (31b)$$
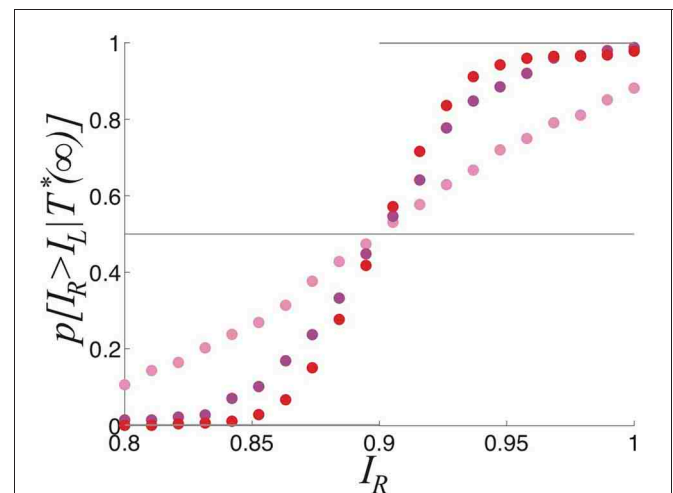
FIGURE 8 | Comparing the probability densities of dominance times in the stochastic ring model with depression (Equation 1). Expected likelihood $p[I_R > I_L | T^*(\infty)]$ the right input I_R is stronger than the left I_L based in the limit of an infinite number of samples of the dominance times T_R and T_R for the parameters: $\beta = 0$, $\varepsilon = 0.04$ (pink); $\beta = 0.2$, $\varepsilon = 0.01$ (magenta); and $\beta = 0.4$ and $\varepsilon = 0.0025$ (red). Other parameters are $\tau = 50$ and $\kappa = 0.5$.
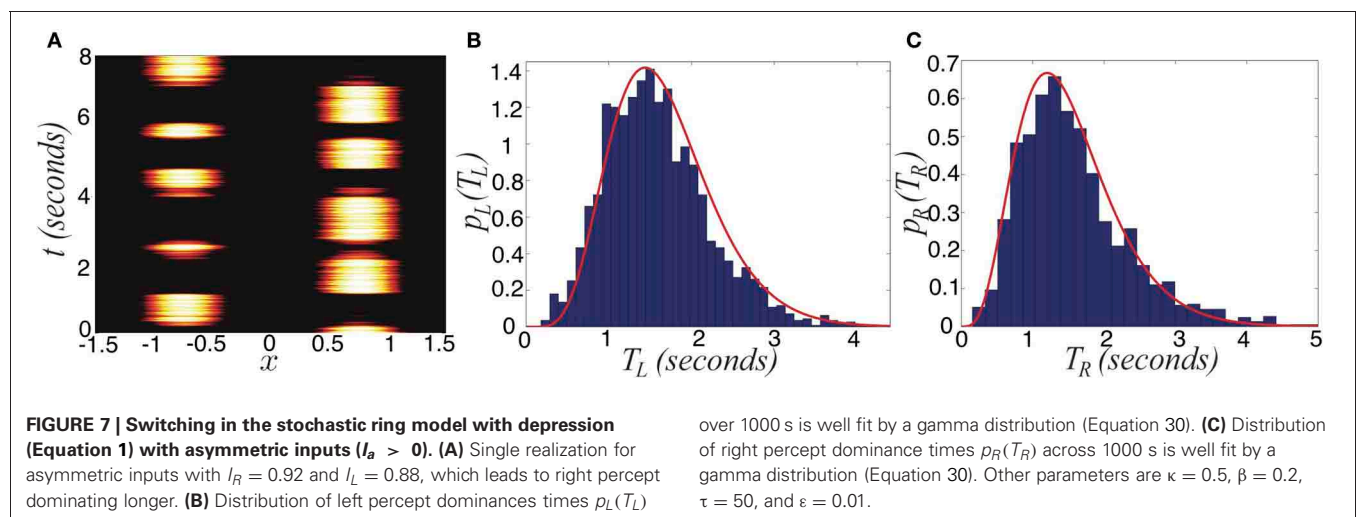
FIGURE 7 | Switching in the stochastic ring model with depression (Equation 1) with asymmetric inputs ($I_a > 0$). (A) Single realization for asymmetric inputs with $I_R = 0.92$ and $I_L = 0.88$, which leads to right percept dominating longer. **(B)** Distribution of left percept dominances times $p_L(T_L)$ over 1000 s is well fit by a gamma distribution (Equation 30). **(C)** Distribution of right percept dominance times $p_R(T_R)$ across 1000 s is well fit by a gamma distribution (Equation 30). Other parameters are $\kappa = 0.5$, $\beta = 0.2$, $\tau = 50$, and $\varepsilon = 0.01$.

$$\tau \dot{q}_R = 1 - q_R - \beta u_R q_R, \qquad (31c)$$

$$\tau \dot{q}_L = 1 - q_L - \beta u_L q_L. \qquad (31d)$$

First, we note Equation (31) has a stable winner-take-all solution in the jth population ($j = R, L$) for $I_j > 0$ and $I_k < 1/(1 + \beta)$ ($k \neq j$). Second, a stable fusion state exists when both $I_L, I_R > 1/(1 + \beta)$. Coexistent with the fusion state, there may be rivalrous oscillations, as we found in the spatially extended system (Equation 1). To study these, we make a similar fast-slow decomposition of the model (Equation 31), assuming $\tau \gg \tau_m$ to find u_j's possess the quasi-steady state

$$u_R = H(I_R - q_L u_L), \quad u_L = H(I_L - q_R u_R). \qquad (32)$$

so we expect $u_j = 0$ or 1 almost everywhere. Therefore, we can estimate the dominance time of each stimulus using a piecewise equation for the slow subsystem

$$\tau q_j = \begin{cases} 1 - q_j - \beta q_j : u_j = 1, \\ 1 - q_j \qquad : u_j = 0, \end{cases} \quad j = L, R. \qquad (33)$$

Combining the slow subsystem (Equation 33) with the quasi-steady state (Equation 32), we can use self-consistency to solve for the dominance times T_R and T_L of the right and left populations. We simply note that switches will occur through escape, when cross-inhibition is weakened enough by depression such that the suppressed population's (j) input becomes superthreshold, so $I_j = q_k$. Using Equation (33), we find

$$T_R = \tau \ln \left[\frac{Q_- + \sqrt{Q_-^2 - 4B_R}}{2(1 + \beta)I_L - 2} \right], \qquad (34)$$

$$T_L = \tau \ln \left[\frac{Q_+ + \sqrt{Q_+^2 - 4B_L}}{2(1 + \beta)I_R - 2} \right], \qquad (35)$$

where $Q_{\pm} = \beta \pm (1 + \beta)[I_R - I_L]$ and $B_{R,L} = (1 - I_{R,L})(1 + \beta)[(1 + \beta)I_{L,R} - 1]$. For symmetric stimuli, $I_L = I_R = I$, both Equations (34) and (35) reduce to

$$T = \tau \ln \left[\frac{\beta + \sqrt{\beta^2 - 4(1 - I)(1 + \beta)[(1 + \beta)I - 1]}}{2(1 + \beta)I - 2} \right], (36)$$

using which we can solve for the critical input strength I above which only the fusion state exists, $I = (2 + \beta)/[2(1 + \beta)]$, in the case of symmetric inputs. We show in **Figure 9** that this asymptotic approximations Equations (34) and (35) of the dominance times match well with the results of numerical simulations, recapitulating Levelt's propositions.

Next, we show that the network with depression and noise generates activity oscillations with dominance times that are gamma distributed (Fox and Herrmann, 1967; Lehky, 1995; Brascamp et al., 2006). We now provide some analytic intuition as to how gamma distributed dominance times may arise in the fast-slow system. First, we display as single realization of the network
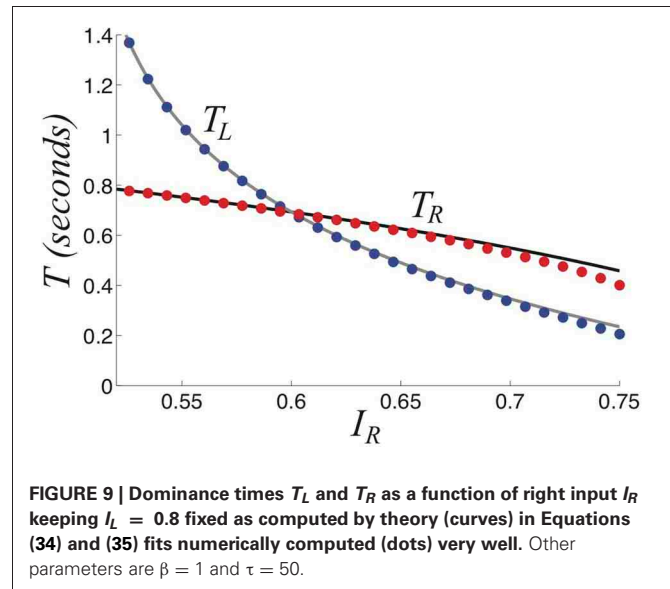
FIGURE 9 | Dominance times T_L and T_R as a function of right input I_R keeping $I_L = 0.8$ fixed as computed by theory (curves) in Equations (34) and (35) fits numerically computed (dots) very well. Other parameters are $\beta = 1$ and $\tau = 50$.

(Equation 6) in **Figure 10A**. An approximate energy function for Equation (6) can be computed in the limit of slow depression recovery time $\tau \gg \tau_m$ by assuming we can augment the energy of the depression-free ($\beta = 0$) network (Hopfield, 1984)

$$E[u_R, u_L] = H(I_L - u_R)H(I_R - u_L)$$
$$- I_L H(I_L - u_R) - I_R H(I_R - u_L),$$

by the synaptic scalings imposed by q_R and q_L (Mejias et al., 2010), so

$$E[u_R, u_L, q_R, q_L] = H(I_L - q_R u_R)H(I_R - q_L u_L)$$
$$- \frac{I_L}{q_R}H(I_L - q_R u_R) - \frac{I_R}{q_L}H(I_R - q_L u_L).$$

A similar energy function was previously used in a model with spike frequency adaptation (Moreno-Bote et al., 2007). Here, we are able to derive the energy function from the model (Equation 6). Therefore, the energy gap between a winner-take-all state and the fusion state will be time-dependent, varying as the synaptic scaling variables q_R and q_L change. The energy difference between the right dominant state and fusion is

$$\Delta E_R(t) = 1 - \frac{I_L}{q_R(t)}, \quad \Delta E_L(t) = 1 - \frac{I_R}{q_L(t)},$$

for the right and left population, respectively.

Notice that dominance times of stochastic switching (**Figures 10B,C**) in Equation (6) are distributed roughly according to a gamma distribution (Equation 30). Superimposing the probability density of right (left) dominance times on the left (right) probability density, we see they are reasonably separated. Using the analysis we performed for the spatially extended system, we could also show that depression improves discernment of the input contrast difference. Mainly here, we

FIGURE 10 | Switching induced by noise and depression. (A) Single realization of the network (Equation 6) with depression and noise. Activity variables u_R (black) and u_L (blue) stay close to attractors at 0 and 1, aside from depression or noise induced switching. Depression variables q_R (red) and q_L (green) slowly exponentially change in response to the states of u_R and u_L. **(B)** Right and **(C)** left dominance time distributions fit with gamma distributions (Equation 30), in the network (Equation 6) with depression and noise in the case of asymmetric inputs $I_R = 0.82$ and $I_L = 0.78$, sampled over 1000 s. The right population has a longer mean dominance time. Other parameters are $\beta = 0.2$, $\tau = 50$, and $\varepsilon = 0.036$.

wanted to provide a justification as to the relationship between input strength and mean dominance times. Using energy arguments, we have provided reasoning behind why Levelt's propositions are still preserved in this model, when noise is included, even when switches are noise-induced. Increasing one input leads to a reduction in the energy barrier between the *other* population's winner-take-all state and the fusion state. This leads to the *other* population's dwell time being shorter.

SWITCHING BETWEEN THREE PERCEPTS

Finally, we will compare the transfer of information in competitive networks that process more than two inputs. Recently, experiments have revealed that perceptual multistability can switch between three or four different percepts (Fisher, 1968; Burton, 2002; Naber et al., 2010; Hupé and Pressnitzer, 2012). In particular, the work of Naber et al. (2010) characterized some of the switching statistics during the oscillations of perceptual tristability. **Figure 11A** shows an example of a tristable percept. Since dominance times are gamma distributed and there is memory evident in the ordering of percepts (Naber et al., 2010), the process is also likely governed by some slow adaptive process in addition to fluctuations.

We study perceptual tristability in a competitive neural network model with only depression, to start, with a Heaviside firing rate (Equation 4), and symmetric inputs $I_1 = I_2 = I_3 = I$, we study the system

$$\dot{u}_1 = -u_1 + H(I - q_2 u_2 - q_3 u_3), \tag{37a}$$

$$\dot{u}_2 = -u_2 + H(I - q_1 u_1 - q_3 u_3), \tag{37b}$$

$$\dot{u}_3 = -u_3 + H(I - q_1 u_1 - q_2 u_2), \tag{37c}$$

$$\tau \dot{q}_j = 1 - q_j - \beta u_j q_j, \quad j = 1, 2, 3. \tag{37d}$$

We are interested in rivalrous oscillations, which do arise in this network (**Figure 11B**). Once again, we can perform a fast-slow decomposition of our system, assuming $\tau \gg \tau_m$ to compute

the dominance time T of a population as it depends on input strength I. We find

$$T = \tau \ln \left[\frac{B + \sqrt{B[3I(1 + \beta) + \beta - 3]}}{2[(1 + \beta)I - 1]} \right],$$

where $B = (1 - I)(1 + \beta)$, which compares very well with numerically computed dominance times in **Figure 12**. Recent experimental observations have suggested relationships between mean dominance time and input contrast in perceptual tristability may be similar to the two percept case (Hupé and Pressnitzer, 2012). In our model, we see that as the input strength is increased, dominance times decrease. One other important point is that percept dominance occurs in the same order every time (**Figure 11B**): one, two, three. There are no "switchbacks." We will show that switchbacks can occur in the noisy regime, which degrades history dependence.

Now, we study how noise alters the switching behavior when added to the deterministic network (Equation 37). Thus, we discuss the three population competitive network with noisy in activity

$$\dot{u}_1 = -u_1 + H(I - q_2 u_2 - q_3 u_3) + \xi_1, \tag{38a}$$

$$\dot{u}_2 = -u_2 + H(I - q_1 u_1 - q_3 u_3) + \xi_2, \tag{38b}$$

$$\dot{u}_3 = -u_3 + H(I - q_1 u_1 - q_2 u_2) + \xi_3, \tag{38c}$$

$$\tau \dot{q}_j = 1 - q_j - \beta u_j q_j, \quad j = 1, 2, 3, \tag{38d}$$

where ξ_j are identical independent white noise processes with variance ε. In **Figure 13**, we show the noise in Equation (38) degrades two pieces of information carried by dominance switches: the switching time and the direction of switching. Notice that adding noise spreads out the distribution of dominance times (**Figure 13B**). Thus, there is a less precise characterization of the input strength in the network. Concerning the direction of switching, the introduction of noise makes "switch backs" more likely. We define a "switch back" as a series of three
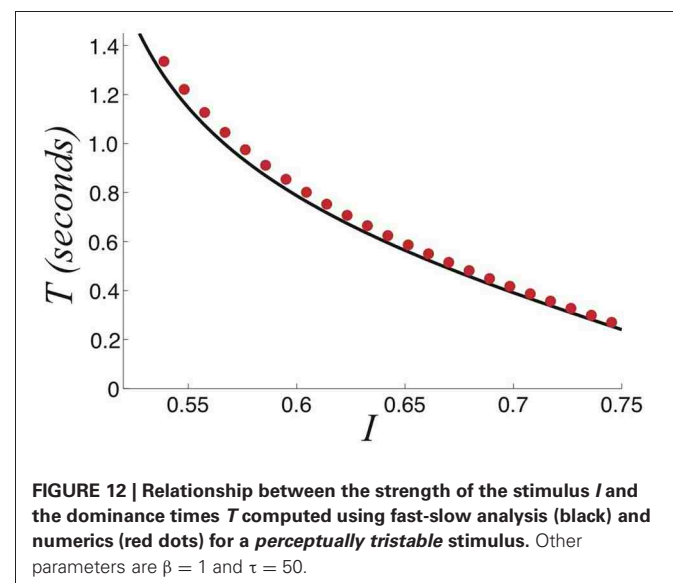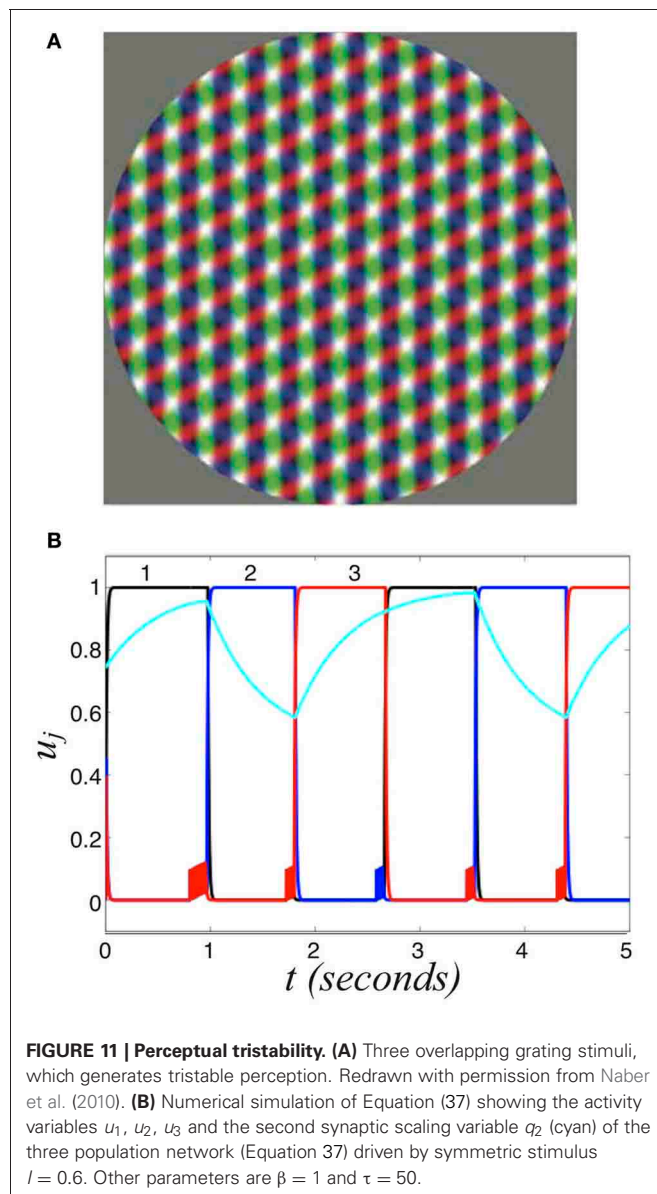
percepts that contains the same percept twice (e.g., 1 → 3 → 1). This is opposed to a "switch forward," which contains all three percepts (e.g., 1 → 3 → 2). Statistics like these were analyzed from psychophysical experiments of perceptual tristability, using an image like **Figure 11A** (Naber et al., 2010). The main finding of Naber et al. (2010) concerning this property is that switch forwards occur more often than chance would suggest. Therefore, they proposed that some slow process may be providing a memory of the previous image. Memory in perceptual rivalry has also been observed in experiments where ambiguous stimuli are presented intermittently (Leopold et al., 2002; Pastukhov and Braun, 2008; Gigante et al., 2009). We suggest short term depression as a candidate substrate for this memory. As seen in **Figure 13B**, the bias in favor of switching forward persists even for non-zero levels of noise. The idea of short term plasticity as a substrate of working memory was also recently proposed in Mongillo et al. (2008).
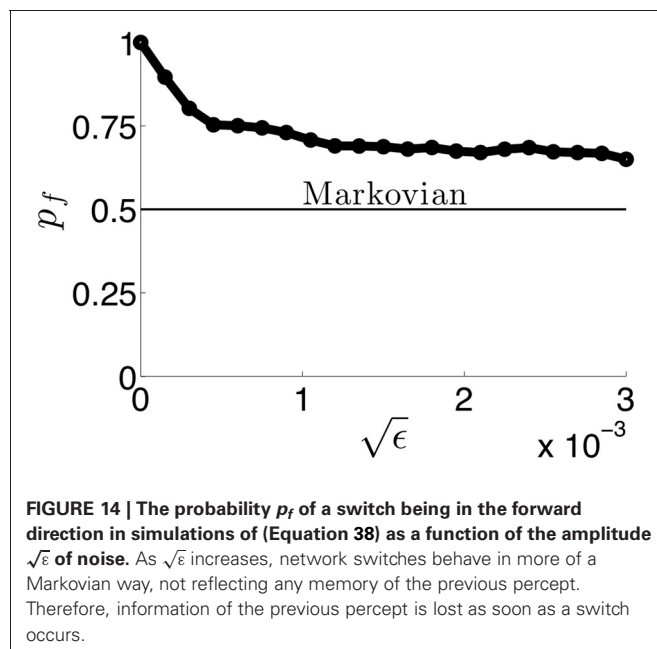
Our results extend this idea, suggesting synaptic mechanisms of working memory may be useful in visual perception tasks, such as understanding ambiguous images. In **Figure 14**, we show that the process of dominance switching becomes more Markovian, less history dependent, as the level of noise $\sqrt{\epsilon}$ is increased. In the limit of large noise, the likelihoods of "switch forwards" and "switch backs" are the same, making the ordering of switching purely Markovian.

DISCUSSION

Mechanisms underlying stochastic switching in perceptual rivalry have been explored in a variety of psychophysical (Fox and Herrmann, 1967; Lehky, 1995; Brascamp et al., 2006), physiological (Leopold and Logothetis, 1996; Blake and Logothetis, 2002), and theoretical studies (Matsuoka, 1984; Laing and Chow, 2002; Moreno-Bote et al., 2007). Since psychophysical data is widely accessible, it can be valuable to use the hallmarks of its statistics as benchmarks for theoretical models. For instance, the fact that dominance time distributions are unimodal functions peaked away from zero suggests that some adaptive process must underlie switching in addition to noise (Laing and Chow, 2002; Brascamp et al., 2006; Shpiro et al., 2009). In addition, Moreno-Bote et al. (2011) information about bistable images may be extracted by sampling a posterior distribution associated with the dominance fraction of each percept. This type of sampling can be well modeled by attractor networks analogous to those presented here (Moreno-Bote et al., 2007). Thus, many dominance time statistics from perceptual rivalry experiments can be employed as points of reference for physiologically based models of visual perception. New data now exists concerning tristable images showing this process also is likely guided by a slow adaptive process in addition to fluctuations (Naber et al., 2010).

We have studied various aspects of competitive neuronal network models of perceptual multistability that include short term synaptic depression. First, we were able to analyze the onset of

FIGURE 11 | Perceptual tristability. (A) Three overlapping grating stimuli, which generates tristable perception. Redrawn with permission from Naber et al. (2010). (B) Numerical simulation of Equation (37) showing the activity variables u_1, u_2, u_3 and the second synaptic scaling variable q_2 (cyan) of the three population network (Equation 37) driven by symmetric stimulus $I = 0.6$. Other parameters are $\beta = 1$ and $\tau = 50$.

FIGURE 12 | Relationship between the strength of the stimulus I and the dominance times T computed using fast-slow analysis (black) and numerics (red dots) for a *perceptually tristable* stimulus. Other parameters are $\beta = 1$ and $\tau = 50$.

FIGURE 13 | Noise degrades two sources of information provided by dominance switches. (A) In the absence of noise, switches always move "forward," so that the previous percept perfectly predicts the subsequent percept. Dominance times accumulate at a single value too. **(B)** For

non-zero noise ($\sqrt{\epsilon} = 0.003$), "switch backs" can occur where the subsequent percept is the same as the previous percept. Also, the distribution of dominance times spreads. Other parameters are $I = 0.6$, $\beta = 1$, and $\tau = 50$.

FIGURE 14 | The probability p_f of a switch being in the forward direction in simulations of (Equation 38) as a function of the amplitude $\sqrt{\epsilon}$ of noise. As $\sqrt{\epsilon}$ increases, network switches behave in more of a Markovian way, not reflecting any memory of the previous percept. Therefore, information of the previous percept is lost as soon as a switch occurs.

rivalrous oscillations in a ring model with synaptic depression (York and van Rossum, 2009; Kilpatrick and Bressloff, 2010a). Stimulating the network with a bimodal input leads to winner-take-all solutions, in the form of single bumps, in the absence

of synaptic depression. As the strength of synaptic depression is increased, the network undergoes a bifurcation which leads to slow oscillations whose timescale is set by that of synaptic depression. Each stimulus peak is represented in the network by a bump whose dominance time is set by the height of each peak. When noise is added, dominance time histograms obey a gamma distribution. We considered the simple task of an upstream network inferring the relative contrast of stimuli based on partial and whole observations of the dominance time distribution. Thus, we study how well the dominance times (information output) of the network reflect the relative stimulus contrasts (information input). Sampling dominance times better identifies contrast differences when switches are more depression-driven and less noise-driven. Thus, short term depression improves information transfer of networks that process ambiguous images in multiple ways. To our knowledge, no previous studies have explored how sampling dominance time distributions might be used by upstream neurons to infer relative stimulus contrast.

We also used energy methods in reduced models to understand how a combination of noise and depression interact to produce switching. Using the energy function derived by Hopfield (1984) for analog neural networks, we justify the exponential dependence of dominance times upon input strength in purely noise-driven switching. Studying an adiabatically derived energy function for the case of slow depression, we also show how depression works to reduce the energy barrier between winner-take-all

states, leading to the slow timescale that defines the peak in depression-noise generated switches. Finally, using a three population space-clamped neural network, we analyzed depression and noise generated switching that may underlie perceptual tristability. We found this network also sustained some of the same relationships between input contrast and dominance times as the two population network. When switches are generated by depression there is an ordering to the population dominance that is lost when switches are noise generated. This is due to the memory generated by short term depression (Mongillo et al., 2008), so the switching process is non-Markovian due to the inherent slow timescale in the background. Dynamical variability must be weak enough to not totally wash out the non-Markovian character of switches. To our knowledge, neither short term depression or adaptation has been proposed before as a mechanism for history dependence in the switching between tristable stimuli. Also, no previous authors have used the history dependence of switching observed in Naber et al. (2010) as a bench mark for a perceptually tristable network model. As opposed to tristability, perceptual bistability generally does not demonstrate strong history dependence in dominance time statistics, behaving more as a renewal process (Lehky, 1995; Laing and Chow, 2002). However, there is some recent evidence that suggests there may be very minor serial correlations in dominance times (van Ee, 2009), likely arising as a signature of a slow adaptive process partially responsible for switching.

Mutual inhibitory rate models with terms representing only spike frequency adaptation (Wilson, 2003; Moreno-Bote et al., 2007) or only short term depression (Kilpatrick and Bressloff, 2010b; Bressloff and Webber, 2012) or both adaptation and depression (Laing and Chow, 2002; Shpiro et al., 2007; Seely and Chow, 2011) have been analyzed in several previous studies. Both mechanisms, when they are included in rate models, can generate dominance time statistics that correspond well with the stimulus contrast dependencies of Levelt (1965), if placed in the right parameter regime. One subtle difference is that if the firing rate function is steep enough in models with depression only, there are no parameter regimes where dominance times increase with contrast (Seely and Chow, 2011). Even if the firing rate function is not very steep, rate models with only depression favor parameter regimes where dominance times decrease with contrast. The effect is not seen in mutually inhibitory rate models with only adaptation (Shpiro et al., 2007). Since Levelt (1965) observed that dominance times decrease with contrast, this suggests depression may be a more suitable choice of slow negative feedback in models of perceptual multistability. On the other hand, it has been demonstrated that gamma distributed dominance time distributions also emerge in perceptual rivalry models with spike frequency adaptation (Shpiro et al., 2009), so it seems the models may often yield similar results (see Shpiro et al., 2007). Note, we have demonstrated a combination of mutual inhibition and depression can generate ordered switching that may be a substrate of perceptual tristability. We presume these results would also extend to a model with mutual inhibition and spike frequency adaptation.

Spatially extended neural field models are a useful tool for understanding complex dynamics that emerge in networks connected by synapses that are stimulus preference dependent (Wilson and Cowan, 1973; Amari, 1977; Bressloff and Cowan, 2002). Processes underlying perceptual rivalry can evolve with a characteristic spatiotemporal structure, as has been found in experiments where observers report waves of visual dominance sweeping one percept over another (Wilson et al., 2001). Bressloff and Webber (2012) and Webber and Bressloff (2013) recently modeled this using a two spatially extended populations coupled to one another by mutual inhibition, where short term depression leads to switches in the direction of activity wave propagation. Our work is distinct from this in several ways. First, we are concerned with non-propagating activity whose switches are abrupt, not gradual as in Bressloff and Webber (2012). In addition, we compute dominance time distributions whereas Bressloff and Webber (2012) compute mean first passage time distributions for their traveling wave. Finally, we have demonstrated phenomena that only require a single cortical layer, and their results require one layer for each percept.

Note to analytically study the relationship between dominance times and input contrast in the noisy system, we resorted to a simple space-clamped neural network. In future work, we plan to develop energy methods for spatially extended systems like Equation (27). Such methods have seen success in analyzing stochastic partial differential equation models such as Ginzburg-Landau models (E et al., 2004). Energy functions have recently been developed for neural field models, but have mostly been studied as a means of determining global stability in deterministic systems (Wu et al., 2002). The fact that pure noise does lead to exponentially distributed dominance times suggests it may be possible to develop a large deviations theory for switching in the system (Equation 27), using techniques like those of E et al. (2004). We propose that by deriving the specific potential energy of spatially extended neural fields, it may be possible to approximate the transition rates of solutions from the vicinity of one attractor to another. In the system (Equation 27), there should be some separatrix between the two winner-take-all states that must be crossed in order for a transition to occur. The least action principle states that there is even a specific point on this separatrix through which the dynamics most likely flows (E et al., 2004). Finding this point using an energy function would allow us to relate the parameters of the model to the distribution of dominance times. This would provide a theoretical framework for interpreting data concerning rivalry of spatially extended images, such as those that produce waves (Wilson et al., 2001). We could also extend this work to analyze interocular grouping Lee and Blake (2004), the phenomenon by which partial images split between either eye are grouped together in perception and rival. Thus, we would need to consider several orientation columns associated with each eye. Columns driven by similarly oriented stimuli would excite one another, overriding weak inhibition between columns in different eyes. Our fast-slow analysis could be useful for analyzing how system dynamics might collapse to group images together in perception.

REFERENCES

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87. doi: 10.1007/BF00337259

Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3844–3848. doi: 10.1073/pnas.92.9.3844

Blake, R. (1989). A neural theory of binocular rivalry. *Psychol. Rev.* 96, 145–167. doi: 10.1037/0033-295X.96.1.145

Blake, R., and Logothetis, N. (2002). Visual competition. *Nat. Rev. Neurosci.* 3, 1–11. doi: 10.1038/nrn701

Brascamp, J. W., van Ee, R., Noest, A. J., Jacobs, R. H. A. H., and van den Berg, A. V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *J. Vis.* 6, 1244–1256. doi: 10.1167/6.11.8

Bressloff, P. C., and Cowan, J. D. (2002). An amplitude equation approach to contextual effects in visual cortex. *Neural Comput.* 14, 493–525. doi: 10.1162/089976602317250870

Bressloff, P. C., and Webber, M. A. (2012). Neural field model of binocular rivalry waves. *J. Comput. Neurosci.* 32, 233–252. doi: 10.1007/s10827-011-0351-y

Buckthought, A., Kim, J., and Wilson, H. R. (2008). Hysteresis effects in stereopsis and binocular rivalry. *Vis. Res.* 48, 819–830. doi: 10.1016/j.visres.2007.12.013

Burton, G. (2002). Successor states in a four-state ambiguous figure. *Psychon. Bull. Rev.* 9, 292–297. doi: 10.3758/BF03196284

Carter, O., Konkle, T., Wang, Q., Hayward, V., and Moore, C. (2008). Tactile rivalry demonstrated with an ambiguous apparent-motion quartet. *Curr. Biol.* 18, 1050–1054. doi: 10.1016/j.cub.2008.06.027

Chance, F. S., Nelson, S. B., and Abbott, L. F. (1998). Synaptic depression and the temporal response characteristics of v1 cells. *J. Neurosci.* 18, 4785–4799.

Deutsch, D. (1974). An auditory illusion. *Nature* 251, 307–309. doi: 10.1038/251307a0

E, W., Ren, W., and Vanden-Eijnden, E. (2004). Minimum action method for the study of rare events. *Commun. Pure Appl. Math.* 57, 637–656. doi: 10.1002/cpa.20005

Ferster, D., and Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.* 23, 441–471. doi: 10.1146/annurev.neuro.23.1.441

Fisher, G. H. (1968). Mother, father, and daughter: a three-aspect ambiguous figure. *Am. J. Psychol.* 81, 274–277. doi: 10.2307/1421277

Fox, R., and Herrmann, J. (1967). Stochastic properties of binocular rivalry alternations. *Atten. Percept. Psychophys.* 2, 432–436. doi: 10.3758/BF03208783

Gigante, G., Mattia, M., Braun, J., and Del Giudice, P. (2009). Bistable perception modeled as competing stochastic integrations at two levels. *PLoS Comput. Biol.* 5:e1000430. doi: 10.1371/journal.pcbi.1000430

Häusser, M., and Roth, A. (1997). Estimating the time course of the excitatory synaptic conductance in neocortical pyramidal cells using a novel voltage jump method. *J. Neurosci.* 17, 7606–7625.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. U.S.A.* 81, 3088–3092. doi: 10.1073/pnas.81.10.3088

Hupe, J.-M. (2010). Dynamics of ménage à trois in moving plaid ambiguous perception. *J. Vis.* 10:1217. doi: 10.1167/10.7.1217

Hupé, J.-M., and Pressnitzer, D. (2012). The initial phase of auditory and visual scene analysis. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 942–953. doi: 10.1098/rstb.2011.0368

Kawaguchi, Y., and Kubota, Y. (1997). Gabaergic cell subtypes and their synaptic connections in rat frontal cortex. *Cereb. Cortex* 7, 476–486. doi: 10.1093/cercor/7.6.476

Kilpatrick, Z. P., and Bressloff, P. C. (2010a). Binocular rivalry in a competitive neural network with synaptic depression. *SIAM J. Appl. Dyn. Syst.* 9, 1303–1347. doi: 10.1137/100788872

Kilpatrick, Z. P., and Bressloff, P. C. (2010b). Effects of synaptic depression and adaptation on spatiotemporal dynamics of an excitatory neuronal network. *Physica D* 239, 547–560. doi: 10.1016/j.physd.2009.06.003

Kilpatrick, Z. P., and Bressloff, P. C. (2010c). Spatially structured oscillations in a two-dimensional excitatory neuronal network with synaptic depression. *J. Comput. Neurosci.* 28, 193–209. doi: 10.1007/s10827-009-0199-6

Laing, C. R., and Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *J. Comput. Neurosci.* 12, 39–53. doi: 10.1023/A:1014942129705

Lee, S.-H., and Blake, R. (2004). A fresh look at interocular grouping during binocular rivalry. *Vis. Res.* 44, 983–991. doi: 10.1016/j.visres.2003.12.007

Lehky, S. R. (1995). Binocular rivalry is not chaotic. *Proc. R. Soc. Lond. B Biol. Sci.* 259, 71–76. doi: 10.1098/rspb.1995.0011

Leopold, D. A., and Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379, 549–553. doi: 10.1038/379549a0

Leopold, D. A., Wilke, M., Maier, A., and Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nat. Neurosci.* 5, 605–609. doi: 10.1038/nn0602-851

Levelt, W. J. M. (1965). *On Binocular Rivalry*. Soesterberg: Institute for Perception RVO–TNO.

Matsuoka, K. (1984). The dynamic model of binocular rivalry. *Biol. Cybern.* 49, 201–208. doi: 10.1007/BF00334466

Mejias, J. F., Kappen, H. J., and Torres, J. J. (2010). Irregular dynamics in up and down cortical states. *PLoS ONE* 5:e13651. doi: 10.1371/journal.pone.0013651

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science* 319, 1543–1546. doi: 10.1126/science.1150769

Moreno-Bote, R., Knill, D. C., and Pouget, A. (2011). Bayesian sampling in visual perception. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12491–12496. doi: 10.1073/pnas.1101430108

Moreno-Bote, R., Rinzel, J., and Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *J. Neurophysiol.* 98, 1125–1139. doi: 10.1152/jn.00116.2007

Naber, M., Gruenhage, G., and Einhäuser, W. (2010). Tristable stimuli reveal interactions among subsequent percepts: rivalry is biased by perceptual history. *Vis. Res.* 50, 818–828. doi: 10.1016/j.visres.2010.02.004

Orbach, J., Ehrlich, D., and Heath, H. A. (1963). Reversibility of the necker cube. I. an examination of the concept of "satiation of orientation." *Percept. Mot. Skills* 17, 439–458. doi: 10.2466/pms.1963.17.2.439

Pastukhov, A., and Braun, J. (2008). A short-term memory of multi-stable perception. *J. Vis.* 8, 7.1–7.14. doi: 10.1167/8.13.7

Pressnitzer, D., and Hupé, J.-M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Curr. Biol.* 16, 1351–1357. doi: 10.1016/j.cub.2006.05.054

Seely, J., and Chow, C. C. (2011). Role of mutual inhibition in binocular rivalry. *J. Neurophysiol.* 106, 2136–2150. doi: 10.1152/jn.00228.2011

Shpiro, A., Curtu, R., Rinzel, J., and Rubin, N. (2007). Dynamical characteristics common to neuronal competition models. *J. Neurophysiol.* 97, 462–473. doi: 10.1152/jn.00604.2006

Shpiro, A., Moreno-Bote, R., Rubin, N., and Rinzel, J. (2009). Balance between noise and adaptation in competition models of perceptual bistability. *J. Comput. Neurosci.* 27, 37–54. doi: 10.1007/s10827-008-0125-3

Tsodyks, M. V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U.S.A.* 94, 719–723. doi: 10.1073/pnas.94.2.719

van Ee, R. (2009). Stochastic variations in sensory awareness are driven by noisy neuronal adaptation: evidence from serial correlations in perceptual bistability. *JOSA A* 26, 2612–2622. doi: 10.1364/JOSAA.26.002612

Wang, X., and Rinzel, J. (1992). Alternating and synchronous rhythms in reciprocally inhibitory model neurons. *Neural Comput.* 4, 84–97. doi: 10.1162/neco.1992.4.1.84

Webber, M. A., and Bressloff, P. C. (2013). The effects of noise on binocular rivalry waves: a stochastic neural field model. *J. Stat. Mech. Theor. Exp.* 2013:P03001. doi: 10.1088/1742-5468/2013/03/P03001

Wilson, H. R. (2003). Computational evidence for a rivalry hierarchy in vision. *Proc. Natl. Acad. Sci. U.S.A.* 100, 14499–14503. doi: 10.1073/pnas.2333622100

Wilson, H. R., Blake, R., and Lee, S. H. (2001). Dynamics of travelling waves in visual perception. *Nature* 412, 907–910. doi: 10.1038/35091066

Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biol. Cybern.* 13, 55–80. doi: 10.1007/bf00288786

Wolfe, J. M. (1986). Stereopsis and binocular rivalry. *Psychol. Rev.* 93,

269–282. doi: 10.1037/0033-295X. 93.3.269

Wu, S., Amari, S.-I., and Nakahara, H. (2002). Population coding and decoding in a neural field: a computational study. *Neural. Comput.* 14, 999–1026. doi: 10.1162/089976602753633367

York, L. C., and van Rossum, M. C. W. (2009). Recurrent networks with short term synaptic depression. *J.*

Comput. Neurosci. 27, 607–620. doi: 10.1007/s10827-009-0172-4

Zhou, W., and Chen, D. (2009). Binaral rivalry between the nostrils and in the cortex. *Curr. Biol.* 19, 1561–1565. doi: 10.1016/j.cub.2009.07.052

Signal enhancement in the output stage of the basal ganglia by synaptic short-term plasticity in the direct, indirect, and hyperdirect pathways

Mikael Lindahl[1]*, Iman Kamali Sarvestani[1], Örjan Ekeberg[1] and Jeanette Hellgren Kotaleski[1,2]

[1] *Department of Computational Biology, School of Computer Science and Communication, KTH Royal Institute of Technology, Stockholm, Sweden*
[2] *Department of Neuroscience, Karolinska Institute, Stockholm, Sweden*

Many of the synapses in the basal ganglia display short-term plasticity. Still, computational models have not yet been used to investigate how this affects signaling. Here we use a model of the basal ganglia network, constrained by available data, to quantitatively investigate how synaptic short-term plasticity affects the substantia nigra reticulata (SNr), the basal ganglia output nucleus. We find that SNr becomes particularly responsive to the characteristic burst-like activity seen in both direct and indirect pathway striatal medium spiny neurons (MSN). As expected by the standard model, direct pathway MSNs are responsible for decreasing the activity in SNr. In particular, our simulations indicate that bursting in only a few percent of the direct pathway MSNs is sufficient for completely inhibiting SNr neuron activity. The standard model also suggests that SNr activity in the indirect pathway is controlled by MSNs disinhibiting the subthalamic nucleus (STN) via the globus pallidus externa (GPe). Our model rather indicates that SNr activity is controlled by the direct GPe-SNr projections. This is partly because GPe strongly inhibits SNr but also due to depressing STN-SNr synapses. Furthermore, depressing GPe-SNr synapses allow the system to become sensitive to irregularly firing GPe subpopulations, as seen in dopamine depleted conditions, even when the GPe mean firing rate does not change. Similar to the direct pathway, simulations indicate that only a few percent of bursting indirect pathway MSNs can significantly increase the activity in SNr. Finally, the model predicts depressing STN-SNr synapses, since such an assumption explains experiments showing that a brief transient activation of the hyperdirect pathway generates a tri-phasic response in SNr, while a sustained STN activation has minor effects. This can be explained if STN-SNr synapses are depressing such that their effects are counteracted by the (known) depressing GPe-SNr inputs.

Keywords: substantia nigra pars reticulata, short-term plasticity, basal ganglia, network model, subthalamic nucleus, globus pallidus, facilitation, depression

INTRODUCTION

An important question in neuroscience is to understand how synaptic signaling contributes to network function in the brain. The synapse, as a basic communication channel between neurons, has classically been viewed as providing information of whether a pre-synaptic neuron has spiked or not. However, the effect of the synaptic signal varies with previous activity pattern either at one or at both sides of the synapse, and these modifications include short-term- to long-term plasticities, which together span from milliseconds up to months (Abbott and Regehr, 2004). The activity history of the synapse thus becomes important in determining its current function in neural circuits. The ability of synapses to perform non-linear transformations of signals over time makes them crucial components enabling a diverse set of circuit functions in the nervous system such as gain control, information filtering, coincident detection, short term- and long term memory (Abbott and Regehr, 2004; Deng and Klyachko, 2011).

Synapses with short-term plasticity are frequent in the basal ganglia, a group of subcortical nuclei involved in action selection and procedural learning (Mink, 1996; Redgrave et al., 1999; Grillner et al., 2005), but still the functional role of these synapses remains poorly understood. Synapses that undergo frequency dependent facilitation and depression on the time scale of hundred milliseconds can be found in several parts of the basal ganglia (Hanson and Jaeger, 2002; Sims et al., 2008; Connelly et al., 2010; Gittis et al., 2010; Planert et al., 2010). Many computational models of the basal ganglia exist. However, with regard to how synaptic connectivity is represented they can roughly be divided into two categories, those without synaptic plasticity and those with long term synaptic plasticity (see, e.g., Bar-Gad et al., 2000; Terman et al., 2002; Humphries et al., 2006; Leblois et al., 2006; O'Reilly, 2006; Houk et al., 2007; Kumar et al., 2011). Although synaptic short-term plasticity is prominent in the basal ganglia, it has not been included in computational models of the basal ganglia.

The basal ganglia nuclei have been suggested to be involved in action selection, working memory representation, sequence learning, and reinforcement learning of appropriate actions (Chakravarthy et al., 2010; Kamali Sarvestani et al., 2011). The excitatory input to striatum, the basal ganglia main input stage, arrives from nearly all parts of cerebral cortex (Gerfen and Bolam, 2010) as well as midline, intralaminar, mediodorsal and ventral lateral, and anterior thalamus (Groenewegen, 1988; Smith et al., 2004). The basal ganglia output targets are also midline, intralaminar and mediodorsal thalamus as well as ventral lateral thalamus, involved in cortical planning and execution of motor behavior (Smith et al., 2004). Another major output are areas in the brainstem such as the superior colliculus, which generates eye and head movements, and pedunculopontine nucleus, involved in orienting of body movements (Gerfen and Bolam, 2010) and muscle tone control (Takakusaki et al., 2004). A third important output from substantia nigra reticulata (SNr) is to neighboring neurons in the substantia nigra compacta (SNc) were SNr efficiently controls the activity of SNc dopaminergic neurons (Tepper and Lee, 2007). Three major pathways, converging on the basal ganglia output stages have been described, the direct, indirect and hyperdirect pathways (Nambu, 2008). Specifically the output nuclei receive inputs from striatal medium spiny neurons expressing dopamine receptor D1 (MSN D1) in striatum (the direct pathway) and from MSNs expressing dopamine receptor D2 (MSN D2) in striatum via globus pallidus externa (GPe) and the subthalamic nucleus (STN) (the indirect pathway), and directly from cortex via the STN (the hyperdirect pathway). The temporal and spatial integration of these three pathways onto the output nuclei determine the ultimate effect basal ganglia signaling has on the behavioral response.

The relative contribution of signals from striatum, GPe and STN to activity changes in basal ganglia output nuclei, such as SNr, is not understood in detail, nor how changes in SNr activity facilitates or inhibits spiking behavior in target areas. SNr has an inhibitory control of thalamic and brainstem areas (Deniau et al., 2007) and a standard view is that decreased SNr activity promote actions whereas increase activity suppress actions (Mink, 1996; Redgrave et al., 1999). Recent experimental data support this view and show how SNr neurons increase and decrease their activity in relation to actions (Fan et al., 2012). SNr activity can potentially be decreased by either increased inputs from MSN D1 or GPe, whereas the SNr activity can be increased either through disinhibition via GPe or by increased excitatory input from STN. It still remains an open question which inputs are responsible for the observed increases and decreases in activity in SNr seen in experiments (Fan et al., 2012). Most of these inputs to SNr are in addition displaying short term plasticity and are thus modulated with activity over time.

Here we build a quantitative computational model of the striatal, pallidal, and subthalamic inputs to the basal ganglia output stage, SNr, assuming biologically plausible neuron dynamics, synaptic conductances and projection patterns, as well as appropriate firing patterns in the pre-synaptic neurons. We quantify the relative contribution of the direct, indirect and hyperdirect pathways for increasing and decreasing the activity in SNr as

well as for the temporal integration of the inputs. We hypothesize that facilitating striato-nigral and depressing pallido-nigral and subthalamo-nigral synapses in a significant way determine the relationship between timing and strength of input signals in SNr. We find that the direct pathway is responsible for decreased activity in SNr whereas pauses in GPe are preferentially responsible for the increased activity in SNr neurons. By assuming that STN synapses are depressing we can explain experiments showing that STN input, on a slower time scale, act as less potent source for changing activity in SNr compared to brief transient (ms) STN activity. Simulations are used to investigate how the rate coding may change with duration of the input signal and the proportion of active neurons. We also show how facilitating and depressing synapses buffer against fluctuations in input background activity.

MATERIALS AND METHODS

NEURONAL FIRING RATES
The characteristic of MSN activity *in vivo* (in both anesthetized and un-anesthetized preparations) is a low frequency firing interrupted by bursts (Wilson, 1993). The basal firing rate for MSNs ranged in simulations between 0.01 and 2.0 Hz while spike frequency during the bursts ranged between 17 and 48 Hz (Miller et al., 2008). The length of the burst was set to 500 ms which is in line with experiments showing that MSNs usually burst for 100–1000 ms (Miller et al., 2008; Gage et al., 2010).

GPe neurons fire tonically at high frequency, interrupted by bursts and pauses (Jaeger and Kita, 2011; Kita and Kita, 2011) and have been reported to fire, *in vivo* in rodents, at 17 Hz (Gage et al., 2010), 26 Hz (Walters et al., 2007), 29 Hz (Kita and Kita, 2011), 32 Hz (Urbain et al., 2000), 36 Hz (Ruskin et al., 1999), and 52 Hz (Celada et al., 1999). Here the GPe basal firing rate is required to be around 30 Hz.

STN neurons were required to have a basal firing rate around 10 Hz which is in accordance with *in vivo* recordings in rat: 6 Hz (Walters et al., 2007), 10 Hz (Farries et al., 2010), 11 Hz (Fujimoto and Kita, 1993), and 13 Hz (Paz et al., 2005).

The basal firing rate of SNr neurons, with MSN input arriving at 0.1 Hz, GPe input arriving at around 30 Hz and an STN background of 10 Hz, was required to be around 30 Hz which is in the range of reported values from *in vivo* recordings in rat: 22 Hz (Zahr et al., 2004), 24 Hz (Walters et al., 2007), 24–27 Hz (Maurice et al., 2003), and 29 Hz (Gernert et al., 2004).

NEURON MODELING APPROACH
To model the SNr, GPe, and STN neurons we have chosen the adaptive exponential integrate and fire model (Brette and Gerstner, 2005). It has few parameters, simplifying the estimation of them from limited amount of experimental data, as compared to more complicated biophysical models with up to hundred or more parameters. The model can capture the spike initiation and upstroke, as well as subthreshold resonance and adaptation of neural activity. It can be tuned to reproduce simulated subthreshold and spiking behaviors that are very similar to *in vitro* and *in vivo* neuronal voltage responses. The model equations are explained below, where V is the membrane

potential and w is the contribution of the neurons slow currents:

$$C\frac{dV}{dt} = -g_L(V - E_L) + g_L\Delta_T \exp\left(\frac{V - V_T}{\Delta_T}\right) - w + I$$

$$\tau_w \frac{dw}{dt} = a(V - E_L) - w \qquad (1)$$

$$\text{if } V > V_{peak} \text{ then } V = V_r \text{ and } w = w + b$$

Here C is the capacitance, g_L is the leak conductance, E_L and V_T are the resting and threshold potentials, Δ_T is the slope factor of the spike upstroke, I is a current source and represents injected current I_{inj} and/or synaptic contributions I_{syn}, τ_w and a are respectively the time constant and the subthreshold adaptation of the recovery current w. When the membrane potential V reaches the cut off V_{peak} it is reset to V_r and then the recovery current w is increased with b.

SNr NEURON MODEL

Without any synaptic input SNr neurons fire tonically at membrane potentials above −54 mV (Richards et al., 1997; Atherton and Bevan, 2005; Chuhma et al., 2011). The autonomous firing is caused by a sodium dependent TTX insensitive inward current activated above −60 mV and a TTX sensitive current activated close to spike threshold. It also has an outward SK channel mediated current responsible for the spike afterhyperpolarization and the precise regular autonomous spiking (Atherton and Bevan, 2005; Zhou et al., 2008). Below we list the quantitative properties of the SNr neuron that are captured with the model:

1. Current voltage relation in the range −80 to −65 mV to be compatible with an input resistance in the range of 80–400 MΩ (see **Figure 1A**; Nakanishi et al., 1997; Richards et al., 1997; Lee and Tepper, 2007a,b; Chuhma et al., 2011).
2. Current frequency relation to be 0.08–0.2 Hz/pA in the range of 0–300 pA (see **Figure 1B**; Nakanishi et al., 1987b; Richards et al., 1997).
3. From holding potential at just below spike threshold, small changes around 5 pA in injected current are sufficient to bring the neuron from silent to repetitive firing (see **Figure 1C**; Atherton and Bevan, 2005).
4. Silent below −54 mV (Richards et al., 1997; Atherton and Bevan, 2005; Chuhma et al., 2011).
5. Rebound spike upon release from hyperpolarization (see **Figure 1C**; Nakanishi et al., 1987b, 1997).
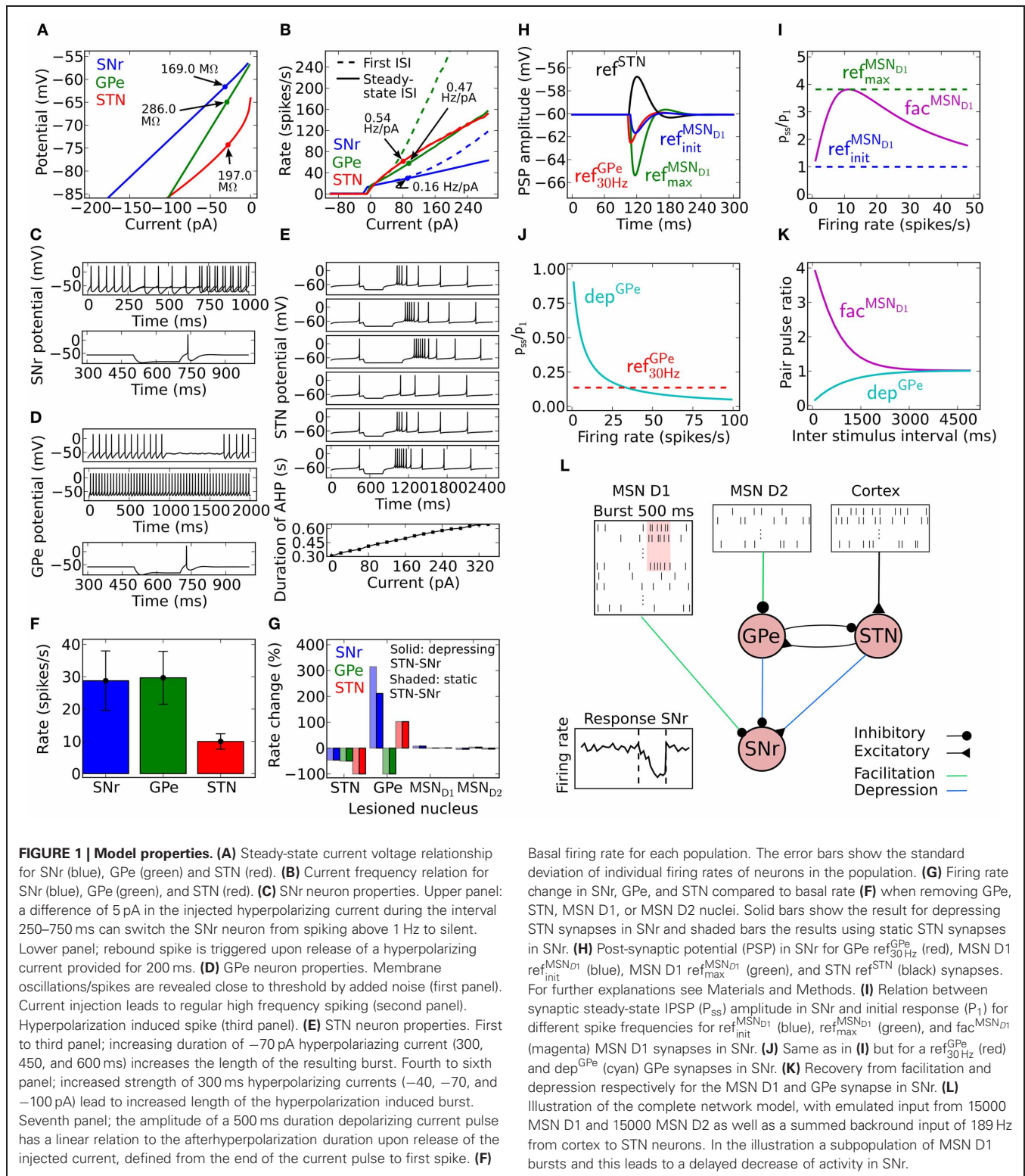
The resulting SNr neuron model parameters are listed in **Table 1**. To capture the rebound spike induced after injection of a hyperpolarizing current (Nakanishi et al., 1987b, 1997) the level of subthreshold adaptation a was set to 3 nS and the time constant τ_w to 20 ms. This also contributed to achieving a model with characteristic afterhypolarization (Atherton and Bevan, 2005) and a positive a ensured that the modeled SNr neuron went from silent to spiking at above 1 Hz by a small change in injected current (Atherton and Bevan, 2005). The SNr neuron's steady-state I–V relation was then produced by setting g_L to 3 nS (Nakanishi et al., 1987b; Richards et al., 1997; Atherton and Bevan, 2005;

Zhou et al., 2008). Near spike initiation the adaptive exponential integrate and fire model can approximate the upstroke and thus the voltage speed/acceleration of the action potential (Platkiewicz and Brette, 2010). For the modeled SNr neuron to go from silent to spiking at approximately −54 mV (Richards et al., 1997; Atherton and Bevan, 2005; Chuhma et al., 2011) and having spike threshold at −52 mV (Richards et al., 1997), defined as when the rate of rise is 10.2 mV/ms, the resting and threshold potentials and slope factor, E_L, V_T, and Δ_T were respectively estimated to −55.8, −55.2, and 1.8 mV. Note, the action potential threshold was measured when the rate of rise was 5% of max in Richards et al. (1997) which we estimated to 10.2 mV/ms from a sigmoid fit of the upstroke of an action potential. The capacitance C was set to 80 pF (Nakanishi et al., 1997) and the summed recovery current contribution, b, at spike reset was set to 200 pA to get the frequency acceleration and the spike frequency adaptation (Nakanishi et al., 1987b; Richards et al., 1997) of the SNr neuron. With the spike voltage reset, V_r, at −65 mV and spike cut off, V_{peak}, at 20 mV we got an after hyperpolarization and spike amplitude in accordance with literature (Lee and Tepper, 2007b). $I_{inj} = I_{in\ vitro}$ was set to 15 pA to shift the current- voltage and frequency curves along the current axis, such that the neuron fired without any synaptic input around 14 Hz (see **Figures 1A,B**) which is in range of measured mean values in experiments with rat/mice slice preparations 7 Hz (Richards et al., 1997), 9–13 Hz (Atherton and Bevan, 2005), 16 Hz (Nakanishi et al., 1997), 16 Hz (Chuhma et al., 2011), and 16–20 Hz (Lee and Tepper, 2007b). To obtain the current-frequency and voltage curves in **Figures 1A,B** $I_{in\ vitro}$ was successively changed. In the network simulations $I_{inj} = I_{in\ vivo}$ was set to 254 pA to obtain around 30 Hz base line firing rate with full synaptic connectivity in the network model (see **Figure 1F**).

GPe NEURON MODEL

Several different types of neurons in GPe have been reported. They have been classified into subgroups based on electrophysiological properties such as rebound firing, membrane resistance, current-frequency relation, hyperpolarizing induced sag, and firing patterns (Kita and Kitai, 1991; Nambu and Llinaś, 1994; Cooper and Stanford, 2000; Bugaysen et al., 2010). However, in an exhaustive modeling and experimental study, it was showed that the properties of the GPe neurons vary in a continuous space without any clear division into subtypes (Günay et al., 2008). Thus, it is not clear how to come up with one model of the GPe neuron. Our approach was to create a GPe neuron model which showed general dominant characteristics of GPe neurons stated below:

1. Current voltage relation in the range −80 mV to −65 mV to be compatible with an input resistance in the range of 90–560 MΩ (see **Figure 1A**; Cooper and Stanford, 2000; Bugaysen et al., 2010; Chuhma et al., 2011).
2. Current frequency relation to be 0.2–0.6 Hz/pA in the range of 0–300 pA (see **Figure 1B**; Cooper and Stanford, 2000; Bugaysen et al., 2010).
3. Membrane oscillations close to spike threshold causing irregular firing and regular firing at higher depolarizing currents (see **Figure 1D**; Nambu and Llinaś, 1994; Cooper and Stanford, 2000).

FIGURE 1 | Model properties. (A) Steady-state current voltage relationship for SNr (blue), GPe (green) and STN (red). **(B)** Current frequency relation for SNr (blue), GPe (green), and STN (red). **(C)** SNr neuron properties. Upper panel: a difference of 5 pA in the injected hyperpolarizing current during the interval 250–750 ms can switch the SNr neuron from spiking above 1 Hz to silent. Lower panel; rebound spike is triggered upon release of a hyperpolarizing current provided for 200 ms. **(D)** GPe neuron properties. Membrane oscillations/spikes are revealed close to threshold by added noise (first panel). Current injection leads to regular high frequency spiking (second panel). Hyperpolarization induced spike (third panel). **(E)** STN neuron properties. First to third panel; increasing duration of −70 pA hyperpolariazing current (300, 450, and 600 ms) increases the length of the resulting burst. Fourth to sixth panel; increased strength of 300 ms hyperpolarizing currents (−40, −70, and −100 pA) lead to increased length of the hyperpolarization induced burst. Seventh panel; the amplitude of a 500 ms duration depolarizing current pulse has a linear relation to the afterhyperpolarization duration upon release of the injected current, defined from the end of the current pulse to first spike. **(F)**

Basal firing rate for each population. The error bars show the standard deviation of individual firing rates of neurons in the population. **(G)** Firing rate change in SNr, GPe, and STN compared to basal rate **(F)** when removing GPe, STN, MSN D1, or MSN D2 nuclei. Solid bars show the result for depressing STN synapses in SNr and shaded bars the results using static STN synapses in SNr. **(H)** Post-synaptic potential (PSP) in SNr for GPe $ref^{GPe}_{30\,Hz}$ (red), MSN D1 $ref^{MSN_{D1}}_{init}$ (blue), MSN D1 $ref^{MSN_{D1}}_{max}$ (green), and STN ref^{STN} (black) synapses. For further explanations see Materials and Methods. **(I)** Relation between synaptic steady-state IPSP (P_{ss}) amplitude in SNr and initial response (P_1) for different spike frequencies for $ref^{MSN_{D1}}_{init}$ (blue), $ref^{MSN_{D1}}_{max}$ (green), and $fac^{MSN_{D1}}$ (magenta) MSN D1 synapses in SNr. **(J)** Same as in **(I)** but for a $ref^{GPe}_{30\,Hz}$ (red) and dep^{GPe} (cyan) GPe synapses in SNr. **(K)** Recovery from facilitation and depression respectively for the MSN D1 and GPe synapse in SNr. **(L)** Illustration of the complete network model, with emulated input from 15000 MSN D1 and 15000 MSN D2 as well as a summed backround input of 189 Hz from cortex to STN neurons. In the illustration a subpopulation of MSN D1 bursts and this leads to a delayed decrease of activity in SNr.

4. Silent below −53 mV (Bugaysen et al., 2010; Chuhma et al., 2011).
5. Rebound spike upon release from hyperpolarization (see **Figure 1D**; Nambu and Llinaś, 1994; Cooper and Stanford, 2000).

The resulting GPe neuron model parameters are listed in **Table 2**. The hyperpolarization triggered spike (Nambu and Llinaś, 1994; Cooper and Stanford, 2000) was captured by setting the sub-threshold adaptation a to 2.5 nS and the time constant τ_w to 20 ms. Steady-state current voltage relation of the GPe neuron

Table 1 | SNr neuron model parameters.

Name	Value	Description
a	3 nS	Subthreshold adaptation
b	200 pA	Spike-triggered adaptation
C	80 pF	Membrane capacitance
Δ_T	1.8 ms	Slope factor of spike upstroke
E_L	−55.8 mV	Leak reversal potential
g_L	3 nS	Leak conductance
$I_{in\ vitro}$	15 pA	I_{inj} to obtain in vitro firing rate without synaptic input
$I_{in\ vivo}$	254 pA	I_{inj} to obtain in vivo firing rate with synaptic input
τ_w	20 ms	Adaptation time constant
V_{peak}	20 mV	Spike cut off
V_r	−65mV	Spike reset
V_T	−55.2 mV	Threshold potential

Table 2 | GPe neuron model parameters.

Name	Value	Description
a	2.5 nS	Subthreshold adaptation
b	70 pA	Spike-triggered adaptation
C	40 pF	Membrane capacitance
Δ_T	1.7 ms	Slope factor of spike upstroke
E_L	−55.1 mV	Leak reversal potential
g_L	1 nS	Leak conductance
$I_{in\ vitro}$	5 pA	I_{inj} to obtain in vitro firing rate without synaptic input
$I_{in\ vivo}$	47 pA	I_{inj} to obtain in vivo firing rate with synaptic input
τ_w	20 ms	Adaptation time constant
V_{peak}	15 mV	Spike cut off
V_r	−60 mV	Spike reset
V_T	−54.7 mV	Threshold potential

was then produced by letting g_L be 1.0 nS (Cooper and Stanford, 2000; Bugaysen et al., 2010). The capacitance C was set to 40 pF (Cooper and Stanford, 2000). Note that with these parameters a GPe neuron exhibits subthreshold oscillations close to rheobase current (the minimal current necessary to elicit a spike). Touboul and Brette (2008) showed that whether an adaptive exponential integrate and fire neuron model exhibit oscillations close to spike threshold depends on the parameters a, C, g_L, and τ_w and occurs when equations 2 and 3, with $\tau_m = C/g_L$, are fulfilled. For the modeled GPe neuron to go from silent to spiking at approximately −53 mV (Bugaysen et al., 2010; Chuhma et al., 2011) and having a spike threshold at −43 mV (Bugaysen et al., 2010), defined as when the acceleration of the membrane potential reaches 50% of its max, estimated to 1270 mV/ms² from Bugaysen et al. (2010), the resting and threshold potentials and the slope factor, E_L, V_T and Δ_T were set to respectively −55.1, −54.7, and 1.7 mV. The summed recovery current contribution, b, at spike reset was set to 70 pA, to mimic the frequency acceleration and the spike frequency adaptation of the GPe neuron (Nambu and Llinás, 1994; Cooper and Stanford, 2000; Bugaysen et al., 2010). With the spike voltage reset, V_r, at −60 mV and spike cut off, V_{peak}, at 15 mV we got an after hyperpolarization and spike amplitude in accordance with literature

(Cooper and Stanford, 2000). $I_{inj} = I_{in\ vitro}$ was set to 5 pA to move the current- voltage- and frequency-curves along the current axis, such that the neuron fired around 15 Hz without any synaptic input (see **Figures 1A,B**) which is in range of measured mean values in experiments with rate slice preparations 8–14 Hz (Cooper and Stanford, 2000) and 4–18 Hz (Bugaysen et al., 2010). To get the current- frequency and voltage curves in **Figures 1A,B** $I_{in\ vitro}$ was successively changed. In the network simulations $I_{inj} = I_{in\ vivo}$ was set to 47 pA to obtain around 30 Hz base line firing rate with full synaptic connectivity in the network model (see **Figure 1F**).

$$0 > \frac{\tau_m}{\tau_w} - \frac{a}{g_L} \tag{2}$$

$$0 > \frac{\tau_m}{4\tau_w}\left(1 - \frac{\tau_w}{\tau_m}\right)^2 - \frac{a}{g_L} \tag{3}$$

STN NEURON MODEL

The parameters for the model of the STN neuron were chosen such that it got some of the characteristic properties of STN neurons (Bevan and Wilson, 1999; Bevan et al., 2000). *In vitro* and in the absence of synaptic input, STN neurons exhibit autonomous rhythmic single-spike activity that is generated by voltage-dependent Na (Nav) channels and can fire at 250 Hz following current injection (Bevan and Wilson, 1999). We requested the following quantitative properties of the STN neurons:

1. Current voltage relation in the range −80 to −70 mV to be compatible with an input resistance in the range of 150–250 MΩ (see **Figure 1A**; Nakanishi et al., 1987a; Beurrier et al., 1999; Loucif et al., 2008).
2. Current frequency relation to be 0.4–0.8 Hz/pA in the range of 0–300 pA (see **Figure 1B**; Bevan and Wilson, 1999; Hallworth et al., 2003).
3. Duration of afterhypolarization after a brief depolarization around 500 ms should depend upon injected current strength (see **Figure 1E**; Bevan and Wilson, 1999).
4. Silent below −64 mV (Kass and Mintz, 2006).
5. Depolarizing the neuron below −70 mV for a certain period should lead to a rebound burst (**Figure 1E**; Bevan et al., 2000; Hallworth et al., 2003).

The resulting STN neuron model parameters are listed in **Table 3**. To account for the hyperpolarization activated inwards current responsible for rebound bursts, the subthreshold adaptation a was set to 0.3 nS below −70 mV with τ_w to 333 ms, such that $333\dot{w} = 0.3\ (V + 70) - w$, and to get minimal spike frequency adaptation (Bevan and Wilson, 1999) a was was set to 0 nS above −70 mV, such that $333\dot{w} = -w$. The STN neuron's steady-state current-voltage relation was captured by setting g_L to 10.0 nS (Nakanishi et al., 1987a; Beurrier et al., 1999). To get resting membrane potential at −64 mV (Kass and Mintz, 2006) and a spike threshold at −35 mV, when the acceleration of membrane potential is 50 mV/ms² (Farries et al., 2010), the resting and threshold potentials, and the slope factor, E_L, V_T, and Δ_T were respectively set to −80.2, −64.0, and 16.2 mV. To capture the characteristic delayed afterhypolarization caused by

Table 3 | STN neuron model parameters.

Name	Value	Description
a	0.3 nS	Subthreshold adaptation (below −70) otherwise equal to 0
b	0.05 pA	Spike-triggered adaptation
C	60 pF	Membrane capacitance
Δ_T	16.2 ms	Slope factor of spike upstroke
E_L	−80.2 mV	Leak reversal potential
g_L	10 nS	Leak conductance
$I_{in\ vitro}$	6 pA	I_{inj} to obtain in vitro firing rate without synaptic input
$I_{in\ vivo}$	6 pA	I_{inj} to obtain in vivo firing rate with synaptic input
τ_w	333 ms	Adaptation time constant
V_{peak}	15 mV	Spike cut off
V_r	−70 mV	Spike reset
V_T	−64.0 mV	Threshold potential

Table 4 | Summary of network properties.

Name	Value	Description
$N_{MSN\,D1}$	15,000	Number of MSN D1 inputs
$N_{MSN\,D2}$	15,000	Number of MSN D2 inputs
N_{SNr}	300	Size of SNr population
N_{GPe}	300	Size of GPe population
N_{STN}	100	Size of STN population
$v_{MSN\,D1}$	0–47 Hz	Firing rate interval of individual MSN D1 neurons
$v_{MSN\,D2}$	0–47 Hz	Firing rate interval of individual MSN D2 neurons
v_{STN}	189 Hz	Baser rate of external poisson type excitatory input to STN
$N_{MSN\,D1-SNr}$	500	Number of MSN D1 connecting to each SNr neuron
$N_{GPe-SNr}$	32	Number of GPe connecting to each SNr neuron
$N_{STN-SNr}$	30	Number of STN connecting to each SNr neuron
$N_{MSN\,D2-GPe}$	500	Number of MSN D2 connecting to each GPe neuron
$N_{STN-GPe}$	30	Number of STN connecting to each GPe neuron
$N_{GPe-GPe}$	30	Number of GPe reciprocal connections
$N_{GPe-STN}$	30	Number of GPe connecting to each STN neuron

increased current injection (Bevan and Wilson, 1999) as well as the spike frequency acceleration (Bevan and Wilson, 1999; Hallworth et al., 2003) the capacitance, C, the summed recovery current contribution, b, at spike reset and the spike voltage reset, V_r, was respectively set to 60 pF, 0.05 pA, and −70 mV. The hyperpolarization induced bursts (**Figure 1E**; Bevan et al., 2000; Hallworth et al., 2003) were captured by resetting V following a spike to $V_r + \max(w \times -10, 10)$ if $w < 0$ and else to V_r. A similar modification to the spike reset point has been done by Izhikevich (2003). With the spike cut off, V_{peak}, at 15 mV we got a spike amplitude in accordance with literature (Beurrier et al., 1999). $I_{inj} = I_{in\ vitro}$ was set to 6 pA to shift the current-voltage and frequency curves along the current axis, such that the neuron fired without any synaptic input around 10 Hz (see **Figures 1A,B**) which is in range of measured mean values in experiments with rate slice preparations, 6 Hz (Baufreton et al., 2005), 8 Hz (Wilson et al., 2004), 8 Hz (Loucif et al., 2008), 10 Hz (Farries et al., 2010) 12 Hz (Hallworth et al., 2003). To obtain the current- frequency and voltage curves in **Figures 1A,B** $I_{in\ vitro}$ was successively changed. In the network simulations $I_{inj} = I_{in\ vivo}$ was also set to 6 pA to obtain around 10 Hz base line firing rate with full synaptic connectivity in the network model (see **Figure 1F**).

NETWORK MODEL
The model network consists of a population of SNr, GPe, and STN neurons receiving emulated inhibitory synaptic inputs from MSN D1, MSN D2 and cortex with a spike frequency as seen in experiments. The temporal distribution of the spikes was assumed to derive from an uncorrelated Poisson process. The synaptic inputs and neuron population sizes used are listed in **Table 4**, and are in accordance with experiments (Oorschot, 1996). To account for variability in mean firing rate of neurons, seen in experiments, the firing rate of neurons in SNr, GPe, and STN were Gaussian distributed with a standard deviation of 0.2 of respectively each nucleus mean in vitro firing rate. The distributions were created by varying the injected current for each of the neurons in a population.

CONNECTIVITY IN THE NETWORK
Synaptic parameters such as conductances and projection patterns are constrained by experimental data (**Tables 4, 5**). Below we first estimate the connectivity in the network starting with MSN D1 to SNr.

1. Striatal fibers entering SNr follow the dendritic course of single SNr neurons (Rinvik and Grofová, 1970; Schwyn and Fox, 1974; Tokuno et al., 1990) and the axons arborize in clusters along the way (Wu et al., 2000). Based on Miller (2007) (pp 21–28) we assume that a single axon from an MSN makes 20 synaptic contacts upon a single SNr neuron, similarly as for globus pallidus interna (GPi). We modeled this by assuming that the synaptic efficacy of an MSN-SNr connection in the model is the sum of the efficiency of all synaptic contacts that a pre-synaptic neuron makes onto a post-synaptic neuron.
2. The upper bound of the number of synapses an MSN gives off in SNr is 192 (Wu et al., 2000). By dividing 192 by 20, which was the number of synaptic contacts upon one SNr, we estimate that an MSN on average contacts around 10 SNr neurons.
3. Striatum in rat contains 2.8 million MSNs (Oorschot, 1996) and half of these, 1.4 million, belong to the direct pathway and project to SNr (Gerfen et al., 1990) with a subpopulation also sending collaterals to the much smaller endopeduncular nucleus (EP) (homologous to GPi in rat) (Wu et al., 2000).
4. SNr contains 26,000 neurons and EP contains 3200 (Oorschot, 1996) and the ratio between number of MSNs and SNr neurons becomes around 50 (divide 1.4 million by 26,000; assuming EP only receives SNr collaterals)

5. Combining the information in 2 and 4 suggests that each SNr can receive input from up to 500 MSNs.

To estimate the connectivity between GPe and SNr we use the following:

1. GPe axons form baskets around target SNr neurons giving rise to multiple large synaptic boutons (Smith et al., 1998) and activation of a single GPe neuron evokes large IPSPs with a conductance estimated as 76 nS (Connelly et al., 2010). This indicates that the GPe neurons exert a strong inhibitory control over SNr neurons through multiple synaptic contacts on the GPe neuron.

Table 5 | Basic synaptic model parameters.

Name	Value	Source
$\tau_{gaba}^{MSN_{D1}-SNr}$	5.2 ms	Connelly et al., 2010
$g_0^{MSN_{D1}-SNr}$	2 nS	constrained by Connelly et al. (2010)
$t_{delay}^{MSN_{D1}-SNr}$	7 ms	Connelly et al., 2010
$E_{rev}^{MSN_{D1}-SNr}$	−80 mV	Connelly et al., 2010
$\tau_{gaba}^{GPe-SNr}$	2.1 ms	Connelly et al., 2010
$g_0^{GPe-SNr}$	76 nS	Connelly et al., 2010
$t_{delay}^{GPe-SNr}$	3 ms	Nakanishi et al., 1991
$E_{rev}^{GPe-SNr}$	−72 mV	Connelly et al., 2010
$\tau_{ampa}^{STN-SNr}$	12 ms	n.d. assume as for STN to GPe Hanson and Jaeger (2002)
$g_0^{STN-SNr}$	0.91 nS	fitted to model constrains and in range of Shen and Johnson (2006)
$t_{delay}^{STN-SNr}$	4.5 ms	Shen and Johnson (2006) and Ammari et al. (2010)
$E_{rev}^{STN-SNr}$	0 mV	n.d.
$\tau_{gaba}^{MSN_{D2}-GPe}$	6 ms	Shen et al., 2008
$g_0^{MSN_{D2}-GPe}$	2 nS	constrained by Shen et al. (2008)
$t_{delay}^{MSN_{D2}-GPe}$	7 ms	Park et al., 1982
$E_{rev}^{MSN_{D2}-GPe}$	−65 mV	Rav-Acha et al., 2005
$\tau_{ampa}^{STN-GPe}$	12 ms	Hanson and Jaeger, 2002
$g_0^{STN-GPe}$	0.35 nS	fitted to model constrains and in range of Hanson and Jaeger (2002)
$t_{delay}^{STN-GPe}$	5 ms	Ammari et al., 2010
$E_{rev}^{STN-GPe}$	0 mV	n.d.
$\tau_{gaba}^{GPe-GPe}$	5 ms	Shen et al., 2008
$g_0^{GPe-GPe}$	1.3 nS	fitted to model constrains and in range of Hanson and Jaeger (2002)
$t_{delay}^{GPe-GPe}$	1 ms	n.d.
$E_{rev}^{GPe-GPe}$	−65 mV	n.d. assumed as for MSN D1
$\tau_{ampa}^{CTX-STN}$	4 ms	Baufreton et al., 2005
$g_0^{CTX-STN}$	0.25 nS	n.d.
$t_{delay}^{CTX-STN}$	2.5 ms	Fujimoto and Kita, 1993
$E_{rev}^{CTX-STN}$	0 mV	n.d.
$\tau_{gaba}^{GPe-STN}$	8 ms	Baufreton et al., 2005
$g_0^{GPe-STN}$	0.08 nS	n.d. fitted to model constrains
$t_{delay}^{GPe-STN}$	5 ms	Baufreton et al., 2005
$E_{rev}^{GPe-STN}$	−84 mV	Baufreton et al., 2009

2. Pharmacologically induced inhibition of GPe leads to a large increase of firing rate at more than 300% of basal SNr activity (Celada et al., 1999). We tuned the SNr neuron in the network, by injecting current (254 pA) and adding STN input (at 10 Hz), to fire at above 300% of GPe base firing rate without input from GPe. Note that STN activity have been reported to increase to 20 Hz without GPe input (Farries et al., 2010), thus maintaining STN at 10 Hz might seem to be the wrong thing to do. However, experiments (Moran et al., 2011; Rosenbaum et al., 2012a) and model predictions (see Results below) suggest that the synapses between STN and SNr are depressing. Thus, when tuning the model with static synapses between STN and SNr we did not change the activity of STN in order to avoid overestimating the effect of STN to SNr. We found that emulated input from 32 GPe neurons, each with firing frequency around 30 Hz and depressive synapses with 76 nS (Connelly et al., 2010) as the max conductance strength, were needed to decrease the firing rate of the SNr neuron close to 30 Hz.

To estimate the connectivity between GPe and STN we use the following:

1. GPe has sparse but selective connectivity with STN, with an estimate of 300 synapses per GPe neuron (Baufreton et al., 2009). It has also been estimated that single GPe axons make multiple synaptic contacts with one STN cell (Smith et al., 1990).
2. We assume that a single GPe makes on average 10 synaptic contacts with each STN, then, given 1 above, we estimate that each GPe makes 30 connections in STN.
3. STN neurons increase their firing rate with 100% when removing GPe input (Farries et al., 2010) whereas GPe firing rate decrease with 50% when removing STN input (Féger and Robledo, 1991). The synaptic weight between GPe and STN was tuned such that this was fulfilled.

To estimate the connectivity between STN to GPe and SNr we use the following:

1. STN terminals spread evenly over perikarya and dendrites of GPe and SNr neurons (Smith et al., 1998), and have a synaptic conductance around 1 nS for GPe and SNr (Hanson and Jaeger, 2002; Shen and Johnson, 2006).
2. STN fires at 10 Hz *in vivo* (Fujimoto and Kita, 1993; Paz et al., 2005; Walters et al., 2007; Farries et al., 2010), and silencing the nucleus leads to a 50% decrease of activity in GPe and SNr (Féger and Robledo, 1991).
3. Assuming that STN neurons make 30 connections in GPe or SNr we found that with the synaptic weights at 0.35 nS and 0.91 nS for respectively STN to GPe and STN to SNr connections we were in range of criteria 1 and fulfilled criteria 2.

The MSN D2 type makes synaptic contact preferentially on distal dendrites in GPe similarly to MSN D1 in SNr (Smith et al., 1998). Given that MSNs innervate their target in a similar way we

assumed that the number of connections between MSN D2 and GPe equal the number of connection between MSN D1 and SNr. Estimation of GPe collaterals:

1. GPe collaterals innervate the soma and proximal dendrites, transmitting information reliably (Sadek et al., 2007; Sims et al., 2008) with an estimate around 500 synapses per neuron (Sadek et al., 2007).
2. We assume that a single GPe neuron makes on average 17 synaptic contacts with each GPe, then, given 1 above, we estimate that each GPe neuron makes 30 connections in GPe.
3. GPe firing rate increases with 55% when the MSN and collateral GPe inputs are removed (Celada et al., 1999).
4. We found that 3 was fulfilled with the conductance of pallidal synapse set to 1.3 nS, which is in line with Sims et al. (2008).

Estimation of synaptic input rate between cortex and STN:

1. STN neurons fire at around 10 Hz *in vitro* (Fujimoto and Kita, 1993; Paz et al., 2005; Walters et al., 2007; Farries et al., 2010).
2. Without inhibitory input STN neurons fire at 20 Hz (Farries et al., 2010).
3. Assuming a conductance of 0.25 nS we set the cortical input rate to 189 Hz to fulfill 1 and 2.

The resulting connectivity parameters are listed in **Table 4** and the mentioned synaptic conductances in **Table 5**. See **Figure 1G** for the effect on network base firing rate following different lesions.

SYNAPSE MODELS

In order to reveal how activity dependent synapses differentially shape post-synaptic neuron firing frequencies, all simulation results are also compared with the case when static reference (i.e., frequency independent) synapses are used instead. To model the simpler static synapse, a standard conductance based exponential decay model (Equation 4) is used.

$$\frac{dg}{dt} = -\frac{g}{\tau_{syn}} + g_o \times \delta(t - t_{spike}) \qquad (4)$$

Here g is the conductance, τ_{syn} (syn = ampa/gaba) the synaptic time constant, g_o the maximal conductance for a synaptic event, t_{spike} the time of the synaptic event and δ is the Dirac delta function. When a pre-synaptic spike arrives, the conductance g is updated with g_0 and then, in between the spikes, the conductance decays toward zero with time constant τ_{syn}. The post-synaptic current is given by $I_{syn} = g \times (E_{rev} - V)$.

To model a frequency dependent synapse, the Tsodyks model (Tsodyks et al., 1998) was used (Equations 5 and 6) with the common FD formalism (Abbott et al., 1997; Dittman et al., 2000; Abbott and Regehr, 2004; Puccini et al., 2007). The FD formalism dictates that the synaptic strength is updated by the product of facilitating (F) and depressing (D) variables/factors. This description shows quantitatively good approximations of experimentally measured synapse dynamics (Tsodyks and Markram, 1997; Markram et al., 1998; Planert et al., 2010; Klaus et al.,

2011). The model formalism assumes a finite pool of synaptic resources in active (y), inactive (z) and recovered (x) states. At rest y and z are 0 and x is 1. Depression occurs because some of the resources remain for a while in the inactive state before entering the recovered state with a rate determined by the recovery time constant τ_{rec}. The facilitation is modeled by u which is a variable that is step-wise increased at each spike with the product of the utilization factor U and $1 - u$ (U is between 0 and 1) and decays exponentially toward 0 with time constant τ_{fac} in between spikes (Equation 5). The resources in the active state y are increased with the product of the variables x and u (capturing depression and facilitation respectively) and are then quickly inactivated by decaying toward zero with time constant τ_{syn} (Equation 6). The post-synaptic conductance is proportional to the fraction of resources in the active state and is given by $g = g_0 \times y$ with the resulting post-synaptic current $I_{syn} = g \times (E_{rev} - V)$.

$$\frac{du}{dt} = -\frac{u}{\tau_{fac}} + U \times (1 - u) \times \delta(t - t_{spike}) \qquad (5)$$

$$\frac{dx}{dt} = \frac{z}{\tau_{rec}} - u \times x \times \delta(t - t_{spike})$$

$$\frac{dy}{dt} = -\frac{y}{\tau_{syn}} + u \times x \times \delta(t - t_{spike})$$

$$\frac{dz}{dt} = \frac{y}{\tau_{syn}} - \frac{z}{\tau_{rec}} \qquad (6)$$

The value and source of the basic synaptic parameters, τ_{syn} (syn = ampa/gaba), g_o, t_{delay} and E_{rev}, for both plastic and static synapse models are listed in **Table 5**. In simulations the synaptic weights and delays were randomly drawn from a uniform interval $\pm 50\%$ of peak conductances g_0 and delays t_{delay}. We created two static reference synapses from MSN D1 data; a weak static synapse $ref_{init}^{MSN_{D1}}$ representing the initial non-facilitated peak conductance, $g_0^{MSN_{D1}-SNr}$, and a strong static synapse $ref_{max}^{MSN_{D1}}$ representing the maximally facilitated peak conductance, $4 \times g_0^{MSN_{D1}-SNr}$, during steady-state (see also **Figure 1I**). The unitary conductive strength $g_0^{MSN_{D1}-SNr}$ of a striato-nigral synapse could not be established by Connelly et al. (2010). From their data we however, estimate the conductance to 2 nS, assuming it to be 50% of the measured mean conductance strength evoked by minimal stimulation of MSNs inputs. The mean conductance was calculated by dividing the measured peak of the first inhibitory post-synaptic current, 300 pA, with the driving force, 75 mV (GABA high chloride reversal potential at 5 mV and holding potential is at −70 mV). For GPe we have one reference synapse $ref_{30\,Hz}^{GPe}$ with conductance $0.15 \times g_0^{GPe-SNr}$ which is the steady-state strength of the depressing synapse at 30 Hz activation (a typical *in vivo* frequency). The unitary conductive strength of $g_0^{GPe-SNr}$ was set to 76 nS as measured by Connelly et al. (2010). The static synapse STN synapse in SNr was named ref^{STN} and had the synaptic strength $g_0^{STN-SNr}$. In **Figure 1H** are the dynamics of the static synapses onto SNr displayed.

For facilitating and depressing synapses in SNr we use two data sets collected from the published material by Connelly et al. (2010) for tuning of the synapse models. The first data

set describes the relative synaptic current increase over 10 successive spikes at 10, 50, and 100 Hz and the second data set shows the relative size of a recovery spike after 5 pulses at 100 Hz and measured after 60, 160, 560, 3000, and 9000 ms. For facilitating synapse in GPe we used one data set from Sims et al. (2008) with the relative synaptic current increase over 10 successive spikes at 20 and 50 Hz. We fitted parameters for the Tsodyks synapse in Matlab using a least square method minimizing the squared error between experimental and model current pair pulse data. To find the solution we used the *fminserach* method in Matlab which implements the Nelder-Mead Simplex method (Lagarias et al., 1998). The resulting parameters for the facilitating MSN D1 synapse, $fac^{MSN_{D1}}$, and depressing GPe synapse, dep^{GPe}, in SNr, and facilitating MSN D2 synapse, $fac^{MSN_{D2}}$, in GPe, are listed in **Table 6** and the resulting behavior of the dynamic synapses onto SNr, $fac^{MSN_{D1}}$ and dep^{GPe}, are displayed in **Figures 1I–K**. The weights of the dynamical synapses were tuned such that the conductance of the first spike equaled $g_0^{MSN_{D1}-SNr}$ and $g_0^{MSN_{D2}-GPe}$ for the MSN synapses onto SNr or GPe, and $g_0^{GPe-SNr}$ for the GPe depressing synapse onto SNr. Finally Moran et al. (2011) and Rosenbaum et al. (2012b) suggest that STN connects with depressing synapses to the basal ganglia output nucleus SNr. For the STN synapse in SNr we assumed standard depressing synaptic parameters (Tsodyks and Markram, 1996) with $U = 0.35$ and $\tau_{ref} = 800$, with a peak conductance of $3.64 \times g_0^{STN-SNr}$. This ensured that the synaptic efficacy of the depressing STN synapse, at 10 Hz activation, was equal to $g_0^{STN-SNr}$.

DEFINITION OF "THRESHOLD CODING" AND "RATE CODING" IN SNr USED IN THIS STUDY

Striatal MSNs show firing rate changes with respect to the behavioral choice or according to the reward or the reward expectancy for certain actions (Ito and Doya, 2009). SNr neurons likewise change their activity and are modulated by duration and contingency of actions (Fan et al., 2012). Neurons in SNr can potentially code for action on/off or for a graded action-value/salience. In tasks where the basal ganglia are assumed to be involved in action selection (Albin et al., 1989; DeLong, 1990; Mink, 1996; Redgrave et al., 1999) an action is selected when a threshold is passed and consequently an action is either on or off. We call this "*threshold coding*" and in accordance with earlier work, we define that an action is signaled/selected as the firing rate of an SNr neuron drops below 5 Hz (Chevalier and Deniau, 1990; Humphries et al., 2006). Furthermore the basal ganglia might play a role in coding for different action-values (Samejima et al., 2005) or action saliences (Redgrave et al.,

1999). Studies in monkeys suggest that action-value, independent of resulting actions, is coded in the firing rate of striatal neurons (Samejima et al., 2005; Lau and Glimcher, 2007, 2008; Pasquereau et al., 2007). Also SNr neurons show graded increases and decreases in firing rate in relation to action duration and likelihood (Fan et al., 2012). We call this "*rate coding*" and we thus also investigate how well changes in input rates, filtered by activity dependent synapses, can be picked up in the output nuclei.

IMPLEMENTATION

The simulations were run using the NEST simulator (Gewaltig and Diesmann, 2007) and the network was built using PyNest which is a Python-interface to the NEST simulator. Model fitting of dynamical synapses were done in Matlab. The scripts necessary to run the model are available for download at ModelDB (http://senselab.med.yale.edu/ModelDB/).

RESULTS

CHARACTERISTICS OF THE DERIVED MODEL NEURONS AND THEIR SYNAPTIC INPUTS

The SNr, GPe and STN neuron models were tuned to exhibit properties that are characteristic of the firing of these neurons *in vitro,* exhibiting realistic membrane resistances (**Figure 1A**) and current frequency relationships (**Figure 1B**). The SNr neuron model was tuned to exhibit a switch from silence to spiking above 1 Hz at −54 mV (**Figure 1C** upper panel) and in addition it showed hyperpolarization induced rebound spikes (**Figure 1C** lower panel). The GPe neuron exhibited noise induced oscillations close to spike threshold (**Figure 1D** first trace), and then fired regular at higher current input intensities (**Figure 1D** second trace). It also showed rebound spikes upon release from hyperpolarization (**Figure 1D** third trace). The STN neuron model mimics the characteristic hyperpolarization induced burst, where the length of the burst depends both on the duration (**Figure 1E** first-third trace) and the magnitude (**Figure 1E** fourth-sixth trace). It also showed a dependency on time to first spike after a depolarizing 500 ms current induced high frequency discharge (**Figure 1E** seventh trace). To get the spontaneous activity seen in *in vitro* experiments for the SNr (7–20 Hz), the GPe (7–17 Hz), and the STN (8–12 Hz) neuron model, the parameter $I_{in\ vitro}$ (see **Tables 1–3**) was respectively set to 15, 5, and 6 pA.

Synaptic conductances in the model (**Table 5**) where picked such that they would be in agreement with *in vitro* experiments. A few of the parameters in the model were tuned (see Materials and Methods) within biological realistic ranges, such that the steady-state firing rate of SNR, GPe and STN populations in control and lesion experiments were in agreement with literature (**Figures 1F,G**). The model of the facilitating striato-nigral and striato-pallidal, and depressing pallido-nigral synapses are fitted to data from *in vitro* experiments (**Table 6**). The dynamics of the plastic synapse types onto SNr is shown in **Figures 1I–K**. The facilitating MSN D1 to SNr synapse with peaking synaptic steady state strength at 10 Hz is around four times the resting state (base) conductance (**Figure 1I**), and a fast depressing GPe-SNr synapse which at 30 Hz has a steady state conductance around 15% of the resting state base line (**Figure 1J**). Depressing STN synapses in

Table 6 | Parameters for facilitating and depressing Tsodyks synapse models.

Synapse	U	τ_{rec} (ms)	τ_{fac} (ms)
$fac^{MSN_{D1}}$	0.0192	623	559
dep^{GPe}	0.196	969	0
$fac^{MSN_{D2}}$	0.24	11	73
dep^{STN}	0.35	800	0

SNr were assumed to have standard depressing synaptic parameters (Tsodyks and Markram, 1996). Our full model constituted a network of SNr, GPe, and STN neurons, with connection parameters listed in **Table 4**, and the network was activated with emulated patterns of activity from respectively MSN D1, MSN D2, and Cortex (**Figure 1L**).

DELAYED SNr INHIBITION DUE TO SYNAPTIC FACILITATION IN THE DIRECT PATHWAY

The presence of facilitating synapses in the striato-nigral pathway can significantly delay the suppression of SNr firing following activation of only a few pre-synaptic MSNs spiking at moderate burst frequency. The decrease in the SNr firing rate and the temporal changes during the burst period differ when the input arrives through the static $ref_{init}^{MSN_{D1}}$, $ref_{max}^{MSN_{D1}}$ vs. $fac^{MSN_{D1}}$ synapses (**Figures 2A–C**). In the example, 4% of the MSNs are bursting at 20 Hz. If assuming threshold coding, the threshold passing occurs in the simulations with the $ref_{max}^{MSN_{D1}}$ and facilitating synapse model, whereas with the $ref_{init}^{MSN_{D1}}$ synapse model the SNr neuron is not effectively suppressed. The facilitating synapse in the striato-nigral pathway needs, however, about 200 ms before it reaches the same conductive strength as when the $ref_{max}^{MSN_{D1}}$ static synapse is used. Threshold passing is thus delayed for 200 ms when only a few pre-synaptic MSNs are active, showing that the communicated inhibitory signal is successively increasing over time before it suppresses the SNr neuron.

SYNAPTIC DEPRESSION IN THE INDIRECT PATHWAY ALLOWS DETECTION OF IRREGULAR GPe ACTIVITY
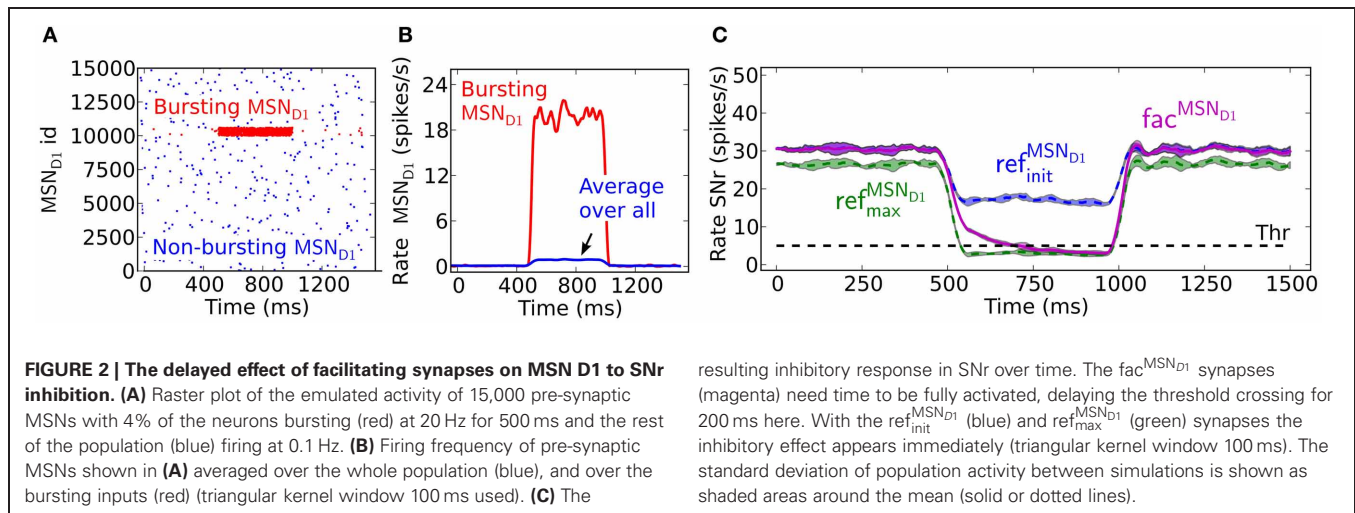
A burst in MSN D2 subpopulations is most effective in disinhibiting SNr when this leads to pauses in GPe subpopulations (**Figures 3A–C**). GPe neurons have a peculiar firing pattern *in vivo*. They fire tonically at high frequency around 30 Hz *in vivo*, interrupted by bursts and pauses (Jaeger and Kita, 2011; Kita and Kita, 2011). During dopamine depleted condition the number of bursts and pauses increase, but still the same mean firing rate is maintained. The increased irregular activity of GPe neurons under dopamine depleted conditions have been hypothesized to disturb the information processing in basal ganglia output nuclei (Kita and Kita, 2011). Here we investigate how depressing GPe synapses convey the irregular GPe activity to SNr. We test this by setting up two scenarios. The first scenario is when both the pre-synaptic bursting and non-bursting MSN D2 subpopulations project in a diffuse way to all post-synaptic GPe neurons, such that the population of GPe neurons only sense the average change of MSN input (**Figure 3C**). A burst in an MSN D2 subpopulation then leads to a minor homogenous decrease in the GPe population. Simulations show that the resulting disinhibition in SNr will be stronger with static synapses, $ref_{30\,Hz}^{GPe}$, than with depressing, dep^{GPe}, synapses (**Figure 3D**) because the depressing GPe synapses in SNr recover their inhibitory strength over time as a result of the decreased GPe spike frequency, and thus the firing rate in SNr is higher in the beginning of the burst. Thus, in this scenario depressing synapses are responsible for producing a transient disinhibition of SNr following a burst in MSN D2. The second scenario is when striatal bursting

and non-bursting MSN D2 project in a non-diffuse way (i.e., topographic) to post-synaptic GPe neurons. Here the GPe neurons receiving input from the bursting pre-synaptic MSN D2 become almost silent and the GPe population receiving input from the non-bursting pre-synaptic MSNs increase their firing further (due to reduced inhibition from the directly inhibited GPe neurons) (see **Figure 3C**). This situation is more effective in disinhibiting SNr over the whole burst (**Figure 3D**), even though the number of synaptic events/s from the total pool of pre-synaptic GPe neurons are the same as above (**Figure 3C** solid magenta vs. dotted blue line). The explanation is that the synapses of the subpopulation of the already tonically firing GPe neurons, which further increase their firing, become even more depressed and therefore do not compensate for the removed inhibition from the subpopulation which becomes quiet. Note that when the MSN D2 to GPe inhibition suddenly is released the synapses of the previously silenced GPe subpopulation have recovered in strength and are responsible for a transient inhibitory response in SNr (see discussion for a hypothetical effect of this). The present simulations thus indicate that irregular activity in GPe subpopulations leads to increased spiking in SNr despite no change in GPe to SNr mean synaptic activation frequency. This might contribute to the disturbed signaling through the basal ganglia output nuclei during Parkinson's disease.

DETECTION OF MSN D1 BURSTING SUBPOPULATIONS IN THE DIRECT PATHWAY

Facilitating synapses selectively enhance input arriving at high frequency rates as in *in vivo* experiments. This is likely important because the number of simultaneously bursting MSNs in striatum is estimated to be low at any given time point (Wilson, 1993). The activation of only a few percent of pre-synaptic direct pathway MSNs, which burst with physiologically realistic burst frequencies, 17–48 Hz (Miller et al., 2008), results in robust inhibition of SNr during steady-state (**Figure 4A**). At lower MSN D1 spike frequencies, action signaling, if assumed to require threshold coding, becomes more resource demanding requiring activation of significantly higher numbers of pre-synaptic MSNs. As indicated in **Figure 2** above, facilitation increases the response to pre-synaptic signals over time, with the result that fewer neurons are required to sustain the same amount of inhibition if the burst is sustained a few 100 ms (**Figure 4A**). Synaptic facilitation thus enables signal amplification of sustained bursts in the striato-nigral pathway. Such amplification due to synaptic facilitation has also been observed in hippocampus (Klyachko and Stevens, 2006), where facilitating synapses enhance the input during epochs of high frequency discharge associated with hippocampal place fields, suggesting that this might be a general function of facilitating synapses.

Facilitating synapses filter out low frequency input possibly preventing unspecific modulation of SNr firing rate due to a fluctuation in background MSN D1 activity. Facilitating synapses stay weak (as for simulation with $ref_{init}^{MSN_{D1}}$) when activated at low input rates, limiting the inhibitory effect of such a signal (**Figure 4B**). Simulations suggest that threshold passing in SNr is not occurring with an increase in background activity of the whole pre-synaptic MSN D1 pool up to 1.2 Hz. Facilitating

FIGURE 2 | The delayed effect of facilitating synapses on MSN D1 to SNr inhibition. (A) Raster plot of the emulated activity of 15,000 pre-synaptic MSNs with 4% of the neurons bursting (red) at 20 Hz for 500 ms and the rest of the population (blue) firing at 0.1 Hz. **(B)** Firing frequency of pre-synaptic MSNs shown in **(A)** averaged over the whole population (blue), and over the bursting inputs (red) (triangular kernel window 100 ms used). **(C)** The resulting inhibitory response in SNr over time. The $fac^{MSN_{D1}}$ synapses (magenta) need time to be fully activated, delaying the threshold crossing for 200 ms here. With the $ref_{init}^{MSN_{D1}}$ (blue) and $ref_{max}^{MSN_{D1}}$ (green) synapses the inhibitory effect appears immediately (triangular kernel window 100 ms). The standard deviation of population activity between simulations is shown as shaded areas around the mean (solid or dotted lines).

synapses thus disregard low frequency input and buffer effectively against fluctuations in the basal activity.

Another way to quantify how the facilitating synapses can detect high frequency input, but buffer against changes in background firing is illustrated in **Figure 4C**, where significantly fewer synaptic events/s (400 compared to 600 synaptic events/s) are required to suppress the SNr when the input is arriving though pre-synaptic subpopulations with high frequency discharge rather than an unspecific increase in MSN D1 firing rate in the whole striatal pool (arrow indicates the intensity used in **Figure 2C**).

The above results show that facilitating synapses enable the post-synaptic neuron to differentiate between bursting- and non-bursting MSN D1 activity patterns, even though there are a constant number of pre-synaptic events. Increasing the number of high frequency firing direct pathway MSNs, and at the same time decreasing the background firing rate of the rest of the MSN D1 pool, such that the number of synaptic events is kept constant in post-synaptic SNr neurons will give a constant total inhibitory effect if $ref_{init}^{MSN_{D1}}$ or $ref_{max}^{MSN_{D1}}$ static synapses are assumed (blue and green **Figure 4D**). However, with facilitating synapses detection of the changed pre-synaptic firing pattern is seen as a decrease in SNr firing rate with increasing contrast in spike frequency between the pre-synaptic neurons (magenta **Figure 4D**).
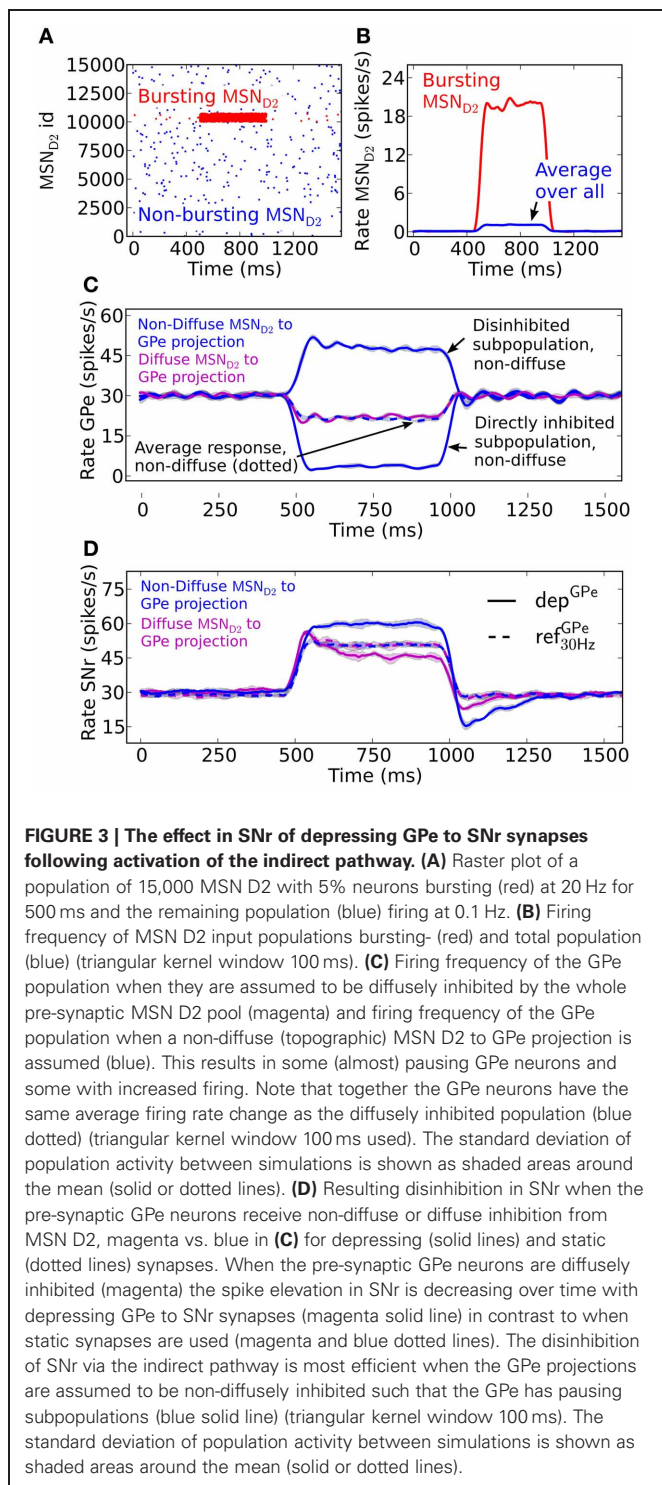
EFFECTS OF DEPRESSING STN-SNr AND GPe-SNr SYNAPSES FOR SIGNALING THROUGH THE INDIRECT AND HYPERDIRECT PATHWAYS

An increased activity of STN may excite SNr directly and/or inhibit SNr through GPe. If both the GPe and STN synapses in SNr were static one would expect that they counteract each other, e.g., they might even cancel each other out such that increased activity in STN only leads to very small activity changes in SNr (**Figure 5A**, blue dotted line). But, since GPe synapses in SNr are depressing (Connelly et al., 2010), the activity from STN would come to dominate the response in SNr such that increased activity in STN leads to increased activity in SNr (**Figure 5A**, blue solid). This happens since depressing synapses tend to converge toward a constant post-synaptic current with increased firing rate (Tsodyks and Markram, 1996), thus the effect of the inhibitory

signal through the depressing GPe-SNr synapses would saturate while the excitatory input from STN would continue to increase with frequency. Experimental studies in rat and monkey, however, contradict such scenarios, and rather suggest that increased activity in STN will not lead to increases in the basal ganglia output nuclei GPi, the analog to SNr (Maurice et al., 2003; Kita et al., 2005; Moran et al., 2011). Such results are well explained by published (Moran et al., 2011) and unpublished (Rosenbaum et al., 2012b) work suggesting that STN is assumed to connect to SNr with depressing synapses. With standard depressing STN-SNr synaptic parameters (Tsodyks and Markram, 1996) (**Table 6**) with $U = 0.35$, $\tau_{fac} = 0$ and $\tau_{rec} = 800$, our simulation results are in accordance with experimental results, i.e., that the excitatory control of SNr by STN is weak (**Figure 5A**, solid green). This suggests that STN is not a major contributor to increased activity in SNr if the input is channeled in parallel via GPe.

In contrast with the above prediction that steady state activation of the hyperdirect pathway leads to only small effects in SNr, the indirect pathway enhances SNr firing when activated from MSN D2 populations (**Figure 5B**). SNr is disinhibited in a (sub) linear fashion following sustained elevated MSN D2 background activity. Increased MSN D2 inhibition of GPe will indirectly increase STN firing through disinhibition, in turn increasing SNr firing significantly if STN-SNr synapses are static (**Figure 5B**, blue lines). When assuming depressing STN-SNr synapses a more moderate disinhibition through the indirect pathway is seen during steady state (**Figure 5B**, green curve).

From these results, achieved for steady-state activation of the hyperdirect and indirect pathways, one would predict that mainly the indirect pathway plays a significant role for controlling the SNr activity level. However, if the temporal effects are considered during e.g., different parts of a 500 ms burst, another scenario emerges. If assuming non-depressing STN-GPe synapses the STN input would indirectly excite SNr more and more during a 500 ms burst because of the GPe-SNr synaptic depression (solid lines in **Figure 5C**). We note, however, that with depressing synapses between both STN and SNr (Rosenbaum et al., 2012b) as well as between GPe and SNr the excitatory effect is not seen (dotted lines **Figure 5C**). The explanation is that the excitatory effect of
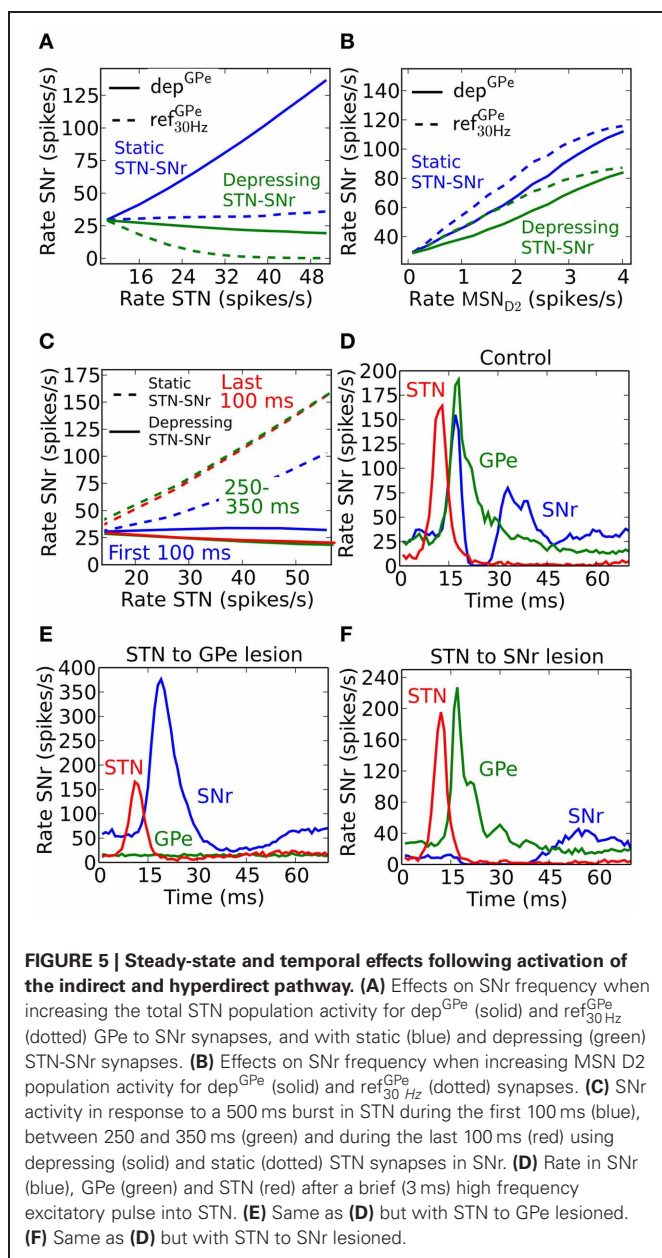
FIGURE 3 | The effect in SNr of depressing GPe to SNr synapses following activation of the indirect pathway. (A) Raster plot of a population of 15,000 MSN D2 with 5% neurons bursting (red) at 20 Hz for 500 ms and the remaining population (blue) firing at 0.1 Hz. **(B)** Firing frequency of MSN D2 input populations bursting- (red) and total population (blue) (triangular kernel window 100 ms). **(C)** Firing frequency of the GPe population when they are assumed to be diffusely inhibited by the whole pre-synaptic MSN D2 pool (magenta) and firing frequency of the GPe population when a non-diffuse (topographic) MSN D2 to GPe projection is assumed (blue). This results in some (almost) pausing GPe neurons and some with increased firing. Note that together the GPe neurons have the same average firing rate change as the diffusely inhibited population (blue dotted) (triangular kernel window 100 ms used). The standard deviation of population activity between simulations is shown as shaded areas around the mean (solid or dotted lines). **(D)** Resulting disinhibition in SNr when the pre-synaptic GPe neurons receive non-diffuse or diffuse inhibition from MSN D2, magenta vs. blue in **(C)** for depressing (solid lines) and static (dotted lines) synapses. When the pre-synaptic GPe neurons are diffusely inhibited (magenta) the spike elevation in SNr is decreasing over time with depressing GPe to SNr synapses (magenta solid line) in contrast to when static synapses are used (magenta and blue dotted lines). The disinhibition of SNr via the indirect pathway is most efficient when the GPe projections are assumed to be non-diffusely inhibited such that the GPe has pausing subpopulations (blue solid line) (triangular kernel window 100 ms). The standard deviation of population activity between simulations is shown as shaded areas around the mean (solid or dotted lines).

FIGURE 4 | Effects of synaptic facilitation in the direct pathway during steady-state. (A) The number of MSN D1 bursting with a certain frequency (7–48 Hz) which are needed for action selection, defined as decreasing SNr firing under a certain threshold. If facilitated synapses are used (magenta), only a few MSNs are needed when bursting in the interval 17–48 Hz, and with performance closer to $ref_{max}^{MSN_{D1}}$ (green) synapses than to $ref_{init}^{MSN_{D1}}$ (blue) synapses during the last 100 ms of the 500 ms burst. **(B)** Steady-state firing rate in post-synaptic SNr cells when all pre-synaptic MSN D1 successively increase their firing. Facilitating synapses (magenta) allow background activity to increase up to 1.2 Hz before suppressing SNr to action signal threshold. **(C)** SNr neuron activity when increasing the total number of MSN D1-SNr synaptic events (#/s). Significantly fewer synaptic events are necessary to bring SNr below threshold if the pre-synaptic inputs come from a subpopulation of bursting MSN D1. Arrow corresponds to the synaptic event intensity used in **Figure 2C**. **(D)** Example of SNr activity as a function of number of bursting pre-synaptic MSN D1 when keeping the total number of synaptic events constant (450 events/s). The facilitating synapse (magenta) enables the SNr neuron to detect a change in input patterns resulting from a few bursting MSN D1.

the STN-SNr pathway is balanced by the inhibitory effect of the STN-GPe-SNr pathway.

To see an excitatory STN effect in the simulations when assuming both depressing STN-SNr and GPe-SNr synapses one needs to focus on an even finer time scale of a few tens of ms. The response following a very brief activation of STN generates a fast increase in activity followed by an inhibition and then a second increase
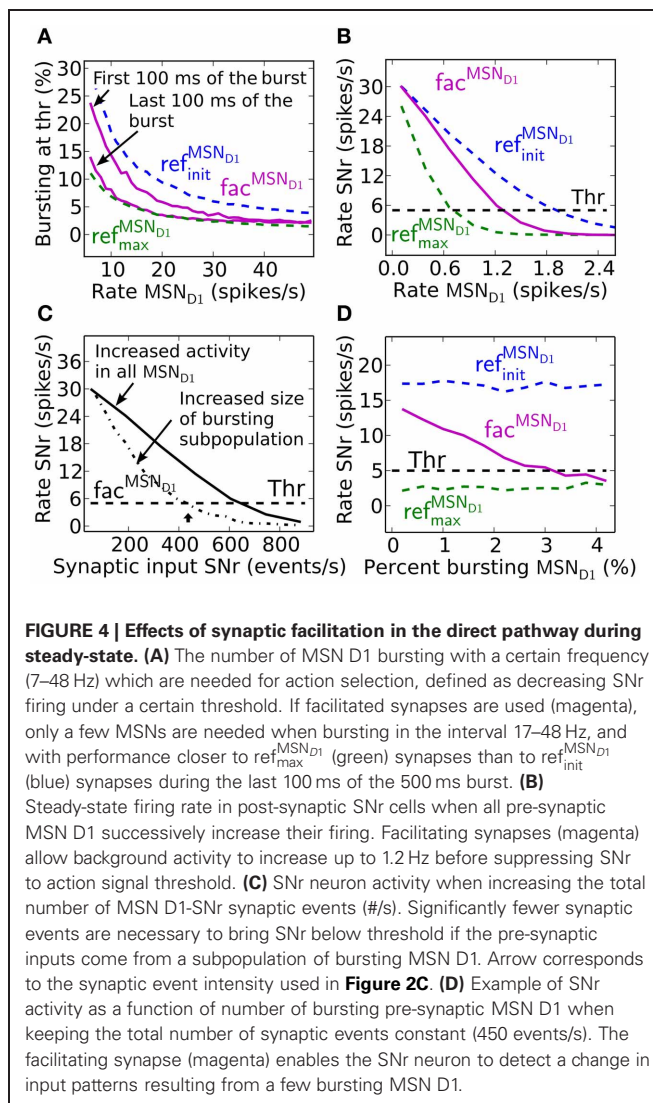
in SNr (**Figure 5D**). This is in accordance with experiments in rat and monkey where such a triphasic response is evoked by a short pulse directly in STN or in cortex (Maurice et al., 2003; Kita et al., 2005; Jaeger and Kita, 2011). Note that in the simulations an activation of STN alone is sufficient to explain the triphasic response, even though the recruitment of the direct and indirect pathways are likely contributing during *in vivo* like conditions when stimulating in cortex. The inhibitory response in SNr following the brief STN activation can be extinguished by removing STN to GPe connections (**Figure 5E**), which also could be interpreted as if GPe and STN do not converge on the same post-synaptic SNr neurons and STN activation would excite those SNr populations over a longer time. The result is supported by experiments which show how application of Gabazine in GPi (homolog to SNr) in monkeys extinguishes the inhibitory and late excitatory response in GPi following cortical activation *in vivo* Tachibana et al. (2008). As expected, STN will indirectly inhibit SNr via GPe for a longer period when the connections between STN and SNr instead are removed (**Figure 5F**). This is also supported by

FIGURE 5 | Steady-state and temporal effects following activation of the indirect and hyperdirect pathway. (A) Effects on SNr frequency when increasing the total STN population activity for dep^{GPe} (solid) and ref$^{GPe}_{30\,Hz}$ (dotted) GPe to SNr synapses, and with static (blue) and depressing (green) STN-SNr synapses. **(B)** Effects on SNr frequency when increasing MSN D2 population activity for dep^{GPe} (solid) and ref$^{GPe}_{30\,Hz}$ (dotted) synapses. **(C)** SNr activity in response to a 500 ms burst in STN during the first 100 ms (blue), between 250 and 350 ms (green) and during the last 100 ms (red) using depressing (solid) and static (dotted) STN synapses in SNr. **(D)** Rate in SNr (blue), GPe (green) and STN (red) after a brief (3 ms) high frequency excitatory pulse into STN. **(E)** Same as **(D)** but with STN to GPe lesioned. **(F)** Same as **(D)** but with STN to SNr lesioned.

experiments where blocking AMPA receptors in GPi in monkeys gives rise to a prolonged inhibition in GPi followed by a short period of elevated activity Tachibana et al. (2008). Simulations thus predict that for a brief activation of the hyperdirect pathway, a tri-phasic excitation-inhibition-excitation response pattern in SNr is seen if GPe and STN converge onto the SNr neurons. For longer STN bursts synaptic depression in both STN-SNr and GPe-SNr synapses prevents sustained effects in SNr. Thus, one could say that the presence of depressing synapses explain the somewhat puzzling experimental finding that STN for brief inputs excites SNr, but for longer activation has no effect or even decreases the firing rate in SNr (Maurice et al., 2003; Tachibana et al., 2008; Moran et al., 2011). Note that a burst in STN can still have a transient excitatory effect in SNr, controlled by the dynamics

of the depressing STN synapses, if STN-SNr and STN-GPe-SNr pathways do not converge in SNr.

SYNAPTIC INTEGRATION AND NEURAL CODING IN SNr

Striatal MSNs show firing rate changes with respect to the behavioral choice. Neurons which change firing rate according to reward probability for action candidates, are present in basal ganglia (Ito and Doya, 2009). SNr neurons likewise change their activity and are modulated by duration and contingency of actions (Fan et al., 2012). Neurons in SNr can thus potentially code for graded action-values/saliences (rate coding). To determine how synaptic facilitation and depression influence rate coding we quantify this by measuring the slope ($\Delta_{SNr}/\Delta_{MSN}$) of a linear fit to the frequency curves of MSN D1 or MSN D2 and SNr, and for different numbers of bursting MSNs. The slope factor indicates how well MSN input rates are sensed in SNr. A small slope factor shows that the activity level in SNr is only moderately controlled by the burst frequency of MSNs, whereas a large slope factor shows that MSN input frequencies are well represented in SNr.

Facilitating synapses allow better detection of MSN D1 firing rate changes in SNr during the first part of a burst (**Figure 6A**). This is further illustrated in the bottom panel in **Figure 6A** where the magnitude of the slope differ with a factor of 3 during the first 100 ms compared to the last 100 ms of a 500 ms burst. This result suggests that an MSN D1 subpopulation better signal rate coded action-values during an initial brief time window immediately following striatal activation. This is explained by the shape of the steady-steady activation curve of the facilitating MSN D1 synapse (**Figure 1I**). At longer time intervals the effective inhibition on SNr (spike frequency times the facilitation) levels off.

Depressing GPe-SNr synapses can enable rate coding of pre-synaptic MSN D2 populations during the whole burst interval (**Figure 6B**). The size of the pre-synaptic bursting MSN D2 population decides when such rate coding is most optimal. The optimal size of the MSN D2 subpopulation for rate coding is slightly increased over a 500 ms burst (**Figure 6B**, bottom panel).

CO-ACTIVATION OF THE DIRECT- AND INDIRECT- OR HYPERDIRECT-PATHWAY

SNr neurons increase and decrease their activity in relation to actions (Sato and Hikosaka, 2002; Basso and Sommer, 2011; Fan et al., 2012). SNr receives input from MSN D1, GPe, and STN and can potentially be decreased by either increased activity in striatal MSN D1 input or by increased activity in GPe input, whereas the activity in SNr can be increased either by disinhibition via GPe or by increased excitation from STN. It is not obvious which input is responsible for increases and decreases in activity in SNr seen in behavioral experiments (Fan et al., 2012). Our results suggest that it is the inhibitory input arriving from MSN D1 that is responsible for inhibition in SNr whereas di-synaptic input from STN through GPe only have a significant effect for very brief inputs (compare **Figure 5**). Conversely we found that MSN D2 can produce an increase in SNr activity through disinhibition via GPe, and that STN has only little effects on SNr activity. Combining inputs onto SNr we see how recruitment of MSN D2 can increase the activity in SNr, potentially suppressing an action
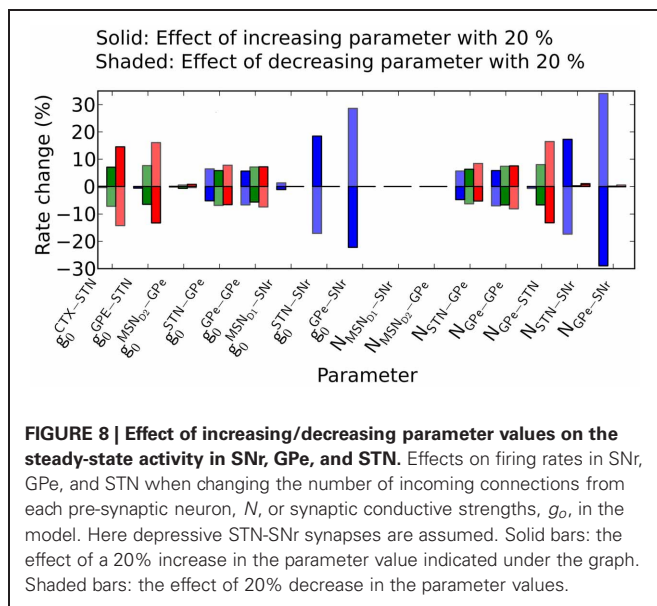
FIGURE 6 | Rate coding in SNr during a sustained burst in striatal populations. (A) Upper panel; effect on SNr firing rate if 2, 4, or 6% of the pre-synaptic MSN D1 pool burst. The result is shown during the first 100 ms of a 500 ms long burst. Middle panel; same as upper panel but during the last 100 ms of the burst. Lower panel shows the slope of linear fits to traces such as in upper and middle panel for three intervals during a 500 ms burst: for the first 100 ms, between 250 and 350 ms and for the last 100 ms. The slope is plotted against the percent of bursting MSN D1. The standard deviation is shown as shaded areas around the mean. **(B)** Upper panel; effect on SNr firing rate if 3, 7, or 11% of the pre-synaptic MSN D2 pool is bursting. The result for the first 100 ms of a 500 ms long burst is shown. Middle panel; Same as upper panel but during the last 100 ms of the burst. Lower panel shows the slope of linear fits to traces such as upper and middle panel for three intervals: the first 100 ms (blue trace), between 250 and 350 ms (green trace) and the last 100 ms (red trace). The result is plotted against percent of bursting MSN D2 populations. Diffuse MSN D2-GPe projections are assumed here (compare **Figure 3**). The standard deviation is shown as shaded areas around the mean.

FIGURE 7 | The result of convergent and non-convergent striato-nigral, pallido-nigral and subthalmo-nigral inputs. (A) Scenario when STN and GPe converge onto SNr neurons. 500 ms 30 Hz bursts in 3% of the MSN D1 pool: alone (blue), combined with 4% bursting MSN D2 (green), combined with elevated STN input induced by doubling the backround cortical drive to $2 \times v_{STN}$ Hz and using static (red) or depressive STN to SNr synapses (magenta). The standard deviation is shown as shaded areas around the mean. **(B)** Scenario were STN and GPe do not converge in SNr. GPe recieves independent poisson input, instead of input from STN neurons, at 10 Hz. 500 ms 30 Hz bursts in 3% of the MSN D1 pool: alone (blue), combined with elevated STN input induced by doubling the backround cortical drive to $2 \times v_{STN}$ Hz and using static (red) or depressive STN to SNr synapses (magenta). In **(B)** it is assumed that the SNr neurons measured from receive increased STN inputs in combination with a constant basal level of GPe inhibition. The standard deviation is shown as shaded areas around the mean.

signal initiated via MSN D1, especially during its initial phase of a 500 ms burst (**Figure 7A**, green line). Note that for a smaller proportion of bursting MSN D2 we would maybe get a delayed action signal when MSN D1-SNr synapses successively facilitate. A similar observation holds when the hyperdirect pathway is recruited. If the synapses between STN and SNr are assumed to be static (**Figure 7A**; red line) they counteract (or delay) an action selection signaling induced through the direct pathway. However, following our prediction that STN-SNr synapses are depressing then the excitatory control of SNr by STN is negligible (**Figure 7A**; magenta line, compare also **Figure 5** above). Finally we tested how increased activity in STN influences SNr if GPe and STN do not converge in SNr neurons. Now, when simulating with depressing
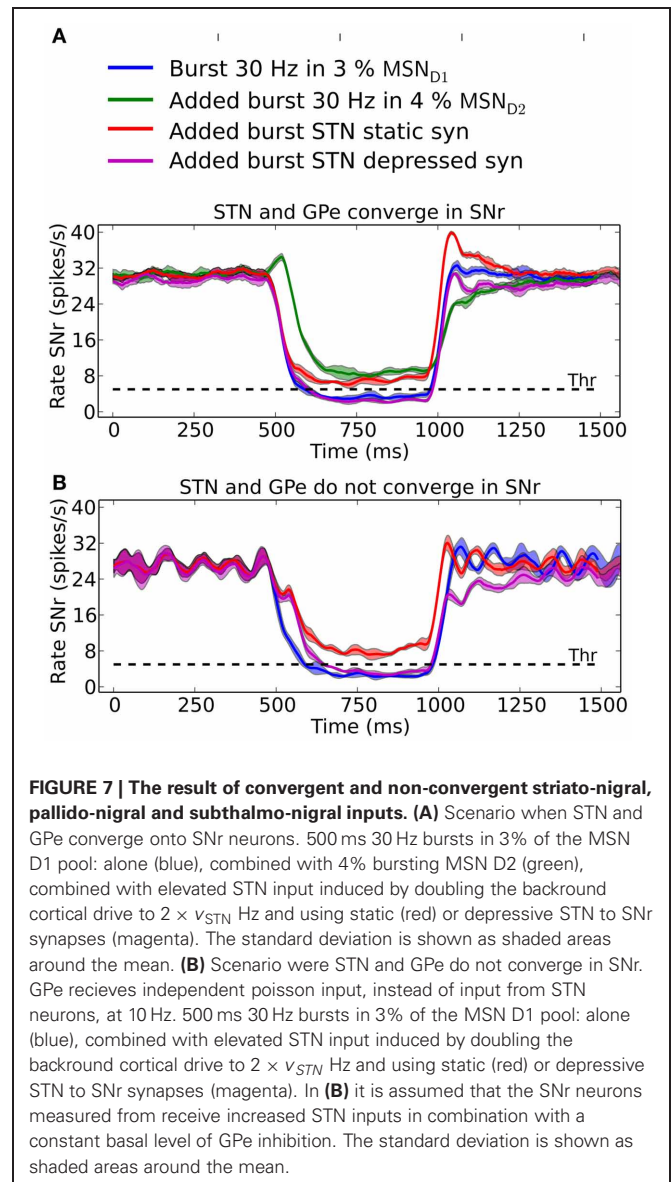
STN-SNr synapses onto the SNr neurons, we see (**Figure 7B**; magenta line) how STN can delay an action signal induced by MSN D1 activity for a period of 100–200 ms, a delay directly determined by the dynamics of the depressing synapses. Thus, the patterns of convergence of the direct, indirect and hyperdirect pathway determine the effect a signal though either of the pathways can have.

HOW PARAMETER PERTUBATIONS INFLUENCE THE BASAL FIRING RATES

Simulations predict that parameter changes in GPe-SNr and STN-SNr connections affect the firing rate in SNr the most. The model have many parameters and one natural question is how robust the model behavior is to parameter changes. We tried to address this by varying the conductances and number of incoming connections from each pre-synaptic neuron with 20% up/down while measuring the change in the basal

FIGURE 8 | Effect of increasing/decreasing parameter values on the steady-state activity in SNr, GPe, and STN. Effects on firing rates in SNr, GPe, and STN when changing the number of incoming connections from each pre-synaptic neuron, N, or synaptic conductive strengths, g_o, in the model. Here depressive STN-SNr synapses are assumed. Solid bars: the effect of a 20% increase in the parameter value indicated under the graph. Shaded bars: the effect of 20% decrease in the parameter values.

rates in SNr, GPe, and STN. We find that the rate in SNr is most sensitive to parameter changes in the pallido-nigral and subthalamo-nigral pathways (**Figure 8**) ($g_o^{GPe-SNr}$, $g_0^{STN-SNr}$, $N_{GPe-SNr}$ and $N_{STN-SNr}$). Specifically we see a superlinear change in firing rate in SNr when changing the paramters in the pallido-nigral pathway. The reason for the superlinear increase is the high inhibitory influence GPe has on SNr at basal firing rate. SNr neurons increases their firing rate with more than 300% (see Materials and Methods) when removing GPe (i.e., decreasing GPe activity with 100%), thus increasing the conductance or number of connections between GPe and SNr will have a strong effect. The firing rate in GPe and STN nuclei are significantly less effected and are more robust against changes in parameter values.

DISCUSSION

The present study has important implications for how to think about the role of basal ganglia pathways, and further contributes to the understanding of which combinations of pathways in basal ganglia are responsible for the signaling in basal ganglia output stages.

We have investigated how dynamical synapses in the direct, indirect and hyperdirect pathways quantitatively shape the activity in SNr neurons over time. The frequency dependencies of the synapses play a significant role in producing the response of SNr neurons to characteristic *in vivo* spike patterns from MSN D1, MSN D2, and cortex. Simulations predict that only bursting activity in a few percent of the direct or indirect pathways MSNs are sufficient to respectively substantially decrease or increase the activity in SNr. For the indirect pathways the model predict that, due to depressing synapses, irregular activity in GPe is more effective in increasing the SNr activity. We hypothesize that synapses between STN and SNr are depressing and thus could explain experiments showing that prolonged activation of STN has a weak effect on SNr firing rate whereas a brief STN input leads to a tri-phasic response in SNr. The

prediction that STN-SNr synapses are depressing together with the result that GPe has a strong inhibitory control of SNr suggest that the signaling in the indirect pathway through either striatum-GPe-SNr or striatum-GPe-STN-SNr is functionally dominated by the former. Our findings further indicate that a rate code, signaling action-values or saliences, in striato-nigral pathways is optimal during the initial part of at 500 ms burst in a striatal subpopulations. For the indirect pathway the simulation showed that the input-output frequency separation could be obtained during most parts of the burst. Simulations suggest that for optimal rate coding only a low number of active pre-synaptic MSNs (a few percent) need to be activated in the direct and indirect pathways. We also show that facilitating MSN D1-SNr synapses enhance action signaling caused by increased activity in a small subpopulation of pre-synaptic MSN D1 and at the same time the presence of facilitating synapses buffer against non-specific action signaling due to fluctuation in striatal background activity. Likewise non-specific steady-state changes in background activity in MSN D2 are ignored as a result of depressing GPe-SNr synapses. In summary, the quantitative effects of the frequency dependent synapses on basal ganglia output stages seen in this study highlight the role of short term plasticity in the basal ganglia for signaling, and ultimately, for control of behavior.

In addition to controlling action selection, SNr also influences SNc. SNc provides the main dopaminergic input to the striatum and cerebral cortex. Loss of neurons in SNc is the major pathology behind the Parkinson's disease. Since a major source of GABAergic control of SNc is the neighboring SNr (Tepper and Lee, 2007), the temporal profile of activity in SNr, can effectively shape the activity of SNc over time. For example, our results suggest that when striatal inhibition is lifted from GPe, reactivated GPe synapses can inhibit SNr for a short interval since the GPe-SNr synapses are depressing. This transient inhibition of SNr may result in a short excitation in SNc. The duration of this activity (compare **Figure 3D**) in SNc (100–200 ms) is equal to the reported phasic dopaminergic signals (Redgrave and Gurney, 2006). Whether this chain of influence is at all involved in the generation of phasic dopamine signals is, however, to be elucidated in the future.

MODEL ASSUMPTIONS

The qualitative results of the model are more robust to parameter changes compared to the quantitative results. For example, the finding that the detection of subpopulations of bursting or pausing neurons in the basal ganglia nuclei occurs while changes in background fluctuations are buffered against, are qualitative phenomena enabled by short-time plasticity. They are not dependent on the exact model connectivity or synaptic strength used. This also applies to the result of how short term plasticity in the pathways through the basal ganglia qualitatively shape the output signal over time. However, changes in parameters will e.g., affect the predicted proportion of striatal populations that need to be activated to significantly affect the basal ganglia output stages. Thus, to improve the quantitative properties of the model, it is necessary to successively update model parameters based on new data produced.

We have included important aspects of the basal ganglia circuitry with regard to the output stage, but in the present model the input from the striatum and cortex are emulated. By including GPe and STN we have tried to account for their important interactions. In future versions of the network it would be interesting to incorporate a striatal module and its interactions with GPe (Mallet et al., 2012).

Some recent papers have questioned the value of using a deterministic synapse model, and instead argued for moving to models which take into account the stochasticity of synaptic signaling (De la Rocha and Parga, 2005; Merkel and Lindner, 2010; Rosenbaum et al., 2012a). These studies showed that when one takes into account the trial-to-trial variability in synaptic release events, the resulting post-synaptic response can differ considerable on individual trials. However, considering that it is probably a population of neurons in basal ganglia output nuclei that are coding for a specific message, then averaging over the population likely represent the outcome. One future direction could, however, be to use a stochastic synaptic model and investigate how this affects the variability of signaling.

THE ROLE OF STN IN BASAL GANGLIA

Several computational studies have tried to find the role for STN in basal ganglia signaling. Frank (2006) suggests that STN reduces premature behavioral responses by excitation of the basal ganglia output nuclei and thus dynamically adjusts the response threshold there. In Leblois et al. (2006) loops though STN-SNr/GPi-thalamus-cortex are assumed to compete with loops though striatum-SNr/GPi-thalamus-cortex in SNr/GPi, allowing the system to control action selection. In Humphries et al. (2006) inputs to STN have an excitatory effect in basal ganglia output nuclei setting an appropriate contrast level for action selection. All these models assume that activating STN results in increased activity in SNr. Experiments suggest that STN can control the firing rate in SNr following brief synchronized inputs, but not following prolonged activations. In reproducing these observations our simulations predict that STN makes depressing synapses in SNr. Our results further suggest that the effect STN can have on signaling in SNr depends on the convergence pattern of GPe and the exact dynamics of the synaptic depressions in GPe-SNr and STN-SNr synapses. We speculate that the hyperdirect pathway filter incoming signals such that transient brief signals are let through while longer sustained signals are disregarded. Brief excitations of SNr by STN could then possibly signal start or stop of actions. However, the role of such an STN filtering mechanism has to be settled by future work.

Recent work by Mallet et al. (2012) provides an alternative hypothesis for the role of STN in the basal ganglia network. Their study suggests that a subset of neurons in GPe are driven by STN, and each one of these GPe neurons in turn gives off over 10,000 GABAergic synapses in striatum and thus potentially have a significant inhibitory control of striatum. Thus, STN could serve an important role in regulating the activity of striatal neurons and gate the cortical and thalamic input activity at the striatal level. In line with the present study, such mechanisms of increasing or decreasing the number of activated striatal MSNs might significantly control signaling in basal ganglia output stages.

AUTHOR CONTRIBUTIONS

ACKNOWLEDGMENTS

REFERENCES

Abbott, L. F., and Regehr, W. G. (2004). Synaptic computation. *Nature* 431, 796–803. doi: 10.1038/nature03010

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224. doi: 10.1126/science.275.5297.221

Albin, R. L., Young, A. B., and Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends Neurosci.* 12, 366–375.

Ammari, R., Lopez, C., Bioulac, B., Garcia, L., and Hammond, C. (2010). Subthalamic nucleus evokes similar long lasting glutamatergic excitations in pallidal, entopeduncular and nigral neurons in the basal ganglia slice. *Neuroscience* 166, 808–818. doi: 10.1016/j.neuroscience.2010.01.011

Atherton, J. F., and Bevan, M. D. (2005). Ionic mechanisms underlying autonomous action potential generation in the somata and dendrites of GABAergic substantia nigra pars reticulata neurons *in vitro*. *J. Neurosci.* 25, 8272–8281. doi: 10.1523/JNEUROSCI.1475-05.2005

Bar-Gad, I., Havazelet-Heimer, G., Goldberg, J. A., Ruppin, E., and Bergman, H. (2000). Reinforcement-driven dimensionality reduction–a model for information processing in the basal ganglia. *J. Basic Clin. Physiol. Pharmacol.* 11, 305–320. doi: 10.1515/JBCPP.2000.11.4.305

Basso, M. A., and Sommer, M. A. (2011). Exploring the role of the substantia nigra pars reticulata in eye movements. *Neuroscience* 198, 205–212. doi: 10.1016/j.neuroscience.2011.08.026

Baufreton, J., Atherton, J. F., Surmeier, D. J., and Bevan, M. D. (2005). Enhancement of excitatory synaptic integration by GABAergic inhibition in the subthalamic nucleus. *J. Neurosci.* 25, 8505–8517. doi: 10.1523/JNEUROSCI.1163-05.2005

Baufreton, J., Kirkham, E., Atherton, J. F., Menard, A., Magill, P. J., Bolam, J. P., et al. (2009). Sparse but selective and potent synaptic transmission from the globus pallidus to the subthalamic nucleus. *J. Neurophysiol.* 102, 532–545. doi: 10.1152/jn.00305.2009

Beurrier, C., Congar, P., Bioulac, B., and Hammond, C. (1999). Subthalamic nucleus neurons switch from single-spike activity to burst-firing mode. *J. Neurosci.* 19, 599–609.

Bevan, M. D., and Wilson, C. J. (1999). Mechanisms underlying spontaneous oscillation and rhythmic firing in rat subthalamic neurons. *J. Neurosci.* 19, 7617–7628.

Bevan, M. D., Wilson, C. J., Bolam, J. P., and Magill, P. J. (2000). Equilibrium potential of GABA(A) current and implications for rebound burst firing in rat subthalamic neurons *in vitro*. *J. Neurophysiol.* 83, 3169–3172.

Brette, R., and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.* 94, 3637–3642. doi: 10.4249/scholarpedia.8427

Bugaysen, J., Bronfeld, M., Tischler, H., Bar-Gad, I., and Korngreen, A. (2010). Electrophysiological characteristics of globus pallidus neurons. *PLoS ONE* 5:e12001. doi: 10.1371/journal.pone.0012001

Celada, P., Paladini, C. A., and Tepper, J. M. (1999). GABAergic control of rat substantia nigra dopaminergic neurons: role of globus pallidus and substantia nigra pars reticulata. *Neuroscience* 89, 813–825. doi: 10.1016/S0306-4522(98)00356-X

Chakravarthy, V. S., Joseph, D., and Bapi, R. S. (2010). What do the basal ganglia do? A modeling perspective. *Biol. Cybern.* 103, 237–253. doi: 10.1007/s00422-010-0401-y

Chevalier, G., and Deniau, J.-M. M. (1990). Disinhibition as a basic process in the expression of striatal functions. *Trends Neurosci.* 13, 277–280. doi: 10.1016/0166-2236(90)90109-N

Chuhma, N., Tanaka, K. F., Hen, R., and Rayport, S. (2011). Functional connectome of the striatal medium spiny neuron. *J. Neurosci.* 31, 1183–1192. doi: 10.1523/JNEUROSCI.3833-10.2011

Connelly, W., Schulz, J., and Lees, G. (2010). Differential short-term plasticity at convergent inhibitory synapses to the substantia nigra pars reticulata. *J. Neurosci.* 30, 14854–14861. doi: 10.1523/JNEUROSCI.3895-10.2010

Cooper, A. J., and Stanford, I. M. (2000). Electrophysiological and morphological characteristics of three subtypes of rat globus pallidus neurone in vitro. *J. Physiol.* 527(Pt 2), 291–304. doi: 10.1111/j.1469-7793.2000.t01-1-00291.x

De la Rocha, J., and Parga, N. (2005). Short-term synaptic depression causes a non-monotonic response to correlated stimuli. *J. Neurosci.* 25, 8416–8431. doi: 10.1523/JNEUROSCI.0631-05.2005

DeLong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.* 13, 281–285. doi: 10.1016/0166-2236(90)90110-V

Deng, P.-Y., and Klyachko, V. A. (2011). The diverse functions of short-term plasticity components in synaptic computations. *Commun. Integr. Biol.* 4, 543–548. doi: 10.4161/cib.4.5.15870

Deniau, J.-M., Mailly, P., Maurice, N., and Charpier, S. (2007). The pars reticulata of the substantia nigra: a window to basal ganglia output. *Prog. Brain Res.* 160, 151–172. doi: 10.1016/S0079-6123(06)60009-5

Dittman, J. S., Kreitzer, A. C., and Regehr, W. G. (2000). Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *J. Neurosci.* 20, 1374–1385.

Fan, D., Rossi, M. A., and Yin, H. H. (2012). Mechanisms of action selection and timing in substantia nigra neurons. *J. Neurosci.* 32, 5534–5548. doi: 10.1523/JNEUROSCI.5924-11.2012

Farries, M. A., Kita, H., and Wilson, C. J. (2010). Dynamic spike threshold and zero membrane slope conductance shape the response of subthalamic neurons to cortical input. *J. Neurosci.* 30, 13180–13191. doi: 10.1523/JNEUROSCI.1909-10.2010

Féger, J., and Robledo, P. (1991). The effects of activation or inhibition of the subthalamic nucleus on the metabolic and electrophysiological activities within the pallidal complex and substantia nigra in the rat. *Eur. J. Neurosci.* 3, 947–952.

Frank, M. J. (2006). Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Netw.* 19, 1120–1136. doi: 10.1016/j.neunet.2006.03.006

Fujimoto, K., and Kita, H. (1993). Response characteristics of subthalamic neurons to the stimulation of the sensorimotor cortex in the rat. *Brain Res.* 609, 185–192. doi: 10.1016/0006-8993(93)90872-K

Gage, G. J., Stoetzner, C. R., Wiltschko, A. B., and Berke, J. D. (2010). Selective activation of striatal fast-spiking interneurons during choice execution. *Neuron* 67, 466–479. doi: 10.1016/j.neuron.2010.06.034

Gerfen, C. R., and Bolam, J. P. (2010). "The neuroanatomical organization of the basal ganglia," in *Handbook of Basal Ganglia Structure and Function*, eds H. Steiner and K. Y. Tseng (London: Elsevier), 3–28.

Gerfen, C. R., Engber, T., Mahan, L., Susel, Z., Chase, T., Monsma, F. Jr., et al. (1990). D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* 250, 1429. doi: 10.1126/science.2147780

Gernert, M., Fedrowitz, M., Wlaz, P., and Löscher, W. (2004). Subregional changes in discharge rate, pattern, and drug sensitivity of putative GABAergic nigral neurons in the kindling model of epilepsy. *Eur. J. Neurosci.* 20, 2377–2386. doi: 10.1111/j.1460-9568.2004.03699.x

Gewaltig, M.-O., and Diesmann, M. (2007). NEST (NEural Simulation Tool). *Scholarpedia* 2, 1430.

Gittis, A. H., Nelson, A. B., Thwin, M. T., Palop, J. J., and Kreitzer, A. C. (2010). Distinct roles of GABAergic interneurons in the regulation of striatal output pathways. *J. Neurosci.* 30, 2223–2234. doi: 10.1523/JNEUROSCI.4870-09.2010

Grillner, S., Hellgren, J., Ménard, A., Saitoh, K., and Wikström, M. A. (2005). Mechanisms for selection of basic motor programs–roles for the striatum and pallidum. *Trends Neurosci.* 28, 364–370. doi: 10.1016/j.tins.2005.05.004

Groenewegen, H. J. (1988). Organization of the afferent connections of the mediodorsal thalamic nucleus in the rat, related to the mediodorsal-prefrontal topography. *Neuroscience* 24, 379–431. doi: 10.1016/0306-4522(88)90339-9

Günay, C., Edgerton, J. R., and Jaeger, D. (2008). Channel density distributions explain spiking variability in the globus pallidus: a combined physiology and computer simulation database approach. *J. Neurosci.* 28, 7476–7491. doi: 10.1523/JNEUROSCI.4198-07.2008

Hallworth, N. E., Wilson, C. J., and Bevan, M. D. (2003). Apamin-sensitive small conductance calcium-activated potassium channels, through their selective coupling to voltage-gated calcium channels, are critical determinants of the precision, pace, and pattern of action potential generation in rat subthalamic nu. *J. Neurosci.* 23, 7525–7542.

Hanson, J. E., and Jaeger, D. (2002). Short-term plasticity shapes the response to simulated normal and parkinsonian input patterns in the globus pallidus. *J. Neurosci.* 22, 5164–5172.

Houk, J. C., Bastianen, C., Fansler, D., Fishbach, a, Fraser, D., Reber, P. J., et al. (2007). Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1573–1583. doi: 10.1098/rstb.2007.2063

Humphries, M. D., Stewart, R. D., and Gurney, K. N. (2006). A physiologically plausible model of action selection and oscillatory activity in the basal ganglia. *J. Neurosci.* 26, 12921–12942. doi: 10.1523/JNEUROSCI.3486-06.2006

Ito, M., and Doya, K. (2009). Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J. Neurosci.* 29, 9861–9874. doi: 10.1523/JNEUROSCI.6157-08.2009

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Trans. Neural Netw.* 14, 1569–1572. doi: 10.1109/TNN.2003.820440

Jaeger, D., and Kita, H. (2011). Functional connectivity and integrative properties of globus pallidus neurons. *Neuroscience* 198, 44–53. doi: 10.1016/j.neuroscience.2011.07.050

Kamali Sarvestani, I., Lindahl, M., Hellgren-Kotaleski, J., and Ekeberg, Ö. (2011). The arbitration-extension hypothesis: a hierarchical interpretation of the functional organization of the Basal Ganglia. *Front. Syst. Neurosci.* 5:13. doi: 10.3389/fnsys.2011.00013

Kass, J. I., and Mintz, I. M. (2006). Silent plateau potentials, rhythmic bursts, and pacemaker firing: three patterns of activity that coexist in quadristable subthalamic neurons. *Proc. Natl. Acad. Sci. U.S.A.* 103, 183–188. doi: 10.1073/pnas.0506781102

Kita, H., and Kita, T. (2011). Role of striatum in the pause and burst generation in the globus pallidus of 6-OHDA-treated rats. *Front. Syst. Neurosci.* 5:42. doi: 10.3389/fnsys.2011.00042

Kita, H., and Kitai, S. T. (1991). Intracellular study of rat globus pallidus neurons: membrane properties and responses to neostriatal, subthalamic and nigral stimulation. *Brain Res.* 564, 296–305. doi: 10.1016/0006-8993(91)91466-E

Kita, H., Tachibana, Y., and Nambu, A. (2005). Balance of monosynaptic excitatory and disynaptic inhibitory responses of the globus pallidus induced after stimulation of the subthalamic nucleus in the monkey. *J. Neurosci.* 25, 8611–8619. doi: 10.1523/JNEUROSCI.1719-05.2005

Klaus, A., Planert, H., Hjorth, J. J. J., Berke, J. D., Silberberg, G., and Kotaleski, J. H. (2011). Striatal fast-spiking interneurons: from firing patterns to postsynaptic impact. *Front. Syst. Neurosci.* 5:57. doi: 10.3389/fnsys.2011.00057

Klyachko, V. A., and Stevens, C. F. (2006). Excitatory and feed-forward inhibitory hippocampal synapses work synergistically as an adaptive filter of natural spike trains. *PLoS Biol.* 4:e207. doi: 10.1371/journal.pbio.0040207

Kumar, A., Cardanobile, S., Rotter, S., and Aertsen, A. (2011). The role of inhibition in generating and controlling Parkinson's disease oscillations in the Basal Ganglia. *Front. Syst. Neurosci.* 5:86. doi: 10.3389/fnsys.2011.00086

Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E. (1998).

Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM J. Optim.* 9, 112–147. doi: 10.1137/S1052623496303470

Lau, B., and Glimcher, P. W. (2007). Action and outcome encoding in the primate caudate nucleus. *J. Neurosci.* 27, 14502–14514. doi: 10.1523/JNEUROSCI.3060-07.2007

Lau, B., and Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–463. doi: 10.1016/j.neuron.2008.02.021

Leblois, A., Boraud, T., Meissner, W., Bergman, H., and Hansel, D. (2006). Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. *J. Neurosci.* 26, 3567–3583. doi: 10.1523/JNEUROSCI.5050-05.2006

Lee, C. R., and Tepper, J. M. (2007a). A calcium-activated nonselective cation conductance underlies the plateau potential in rat substantia nigra GABAergic neurons. *J. Neurosci.* 27, 6531–6541. doi: 10.1523/JNEUROSCI.1678-07.2007

Lee, C. R., and Tepper, J. M. (2007b). Morphological and physiological properties of parvalbumin- and calretinin-containing gamma-aminobutyric acidergic neurons in the substantia nigra. *J. Comp. Neurol.* 500, 958–972. doi: 10.1002/cne.21220

Loucif, A. J. C., Woodhall, G. L., Sehirli, U. S., and Stanford, I. M. (2008). Depolarisation and suppression of burst firing activity in the mouse subthalamic nucleus by dopamine D1/D5 receptor activation of a cyclic-nucleotide gated non-specific cation conductance. *Neuropharmacology* 55, 94–105. doi: 10.1016/j.neuropharm.2008.04.025

Mallet, N., Micklem, B. R., Henny, P., Brown, M. T., Williams, C., Bolam, J. P., et al. (2012). Dichotomous organization of the external globus pallidus. *Neuron* 74, 1075–1086. doi: 10.1016/j.neuron.2012.04.027

Markram, H., Gupta, A., Uziel, A., Wang, Y., and Tsodyks, M. (1998). Information processing with frequency-dependent synaptic connections. *Neurobiol. Learn. Mem.* 70, 101–112. doi: 10.1006/nlme.1998.3841

Maurice, N., Thierry, A.-M., Glowinski, J., and Deniau, J.-M. (2003). Spontaneous and evoked activity of substantia nigra pars reticulata neurons during high-frequency stimulation of the subthalamic nucleus. *J. Neurosci.* 23, 9929–9936.

Merkel, M., and Lindner, B. (2010). Synaptic filtering of rate-coded information. *Phys. Rev. E* 81:041921. doi: 10.1103/PhysRevE.81.041921

Miller, B. R., Walker, A. G., Shah, A. S., Barton, S. J., and Rebec, G. V. (2008). Dysregulated information processing by medium spiny neurons in striatum of freely behaving mouse models of Huntington's disease. *J. Neurophysiol.* 100, 2205–2216. doi: 10.1152/jn.90606.2008

Miller, R. (2007). *A Theory of the Basal Ganglia and Their Disorders (Conceptual Advances in Brain Research)*. London: CRC Press.

Mink, J. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* 50, 381–425. doi: 10.1016/S0301-0082(96)00042-1

Moran, A., Stein, E., Tischler, H., Belelovsky, K., and Bar-Gad, I. (2011). Dynamic stereotypic responses of Basal Ganglia neurons to subthalamic nucleus high-frequency stimulation in the parkinsonian primate. *Front. Syst. Neurosci.* 5:21. doi: 10.3389/fnsys.2011.00021

Nakanishi, H., Kita, H., and Kitai, S. T. (1987a). Electrical membrane properties of rat subthalamic neurons in an *in vitro* slice preparation. *Brain Res.* 437, 35–44. doi: 10.1016/0006-8993(87)91524-1

Nakanishi, H., Kita, H., and Kitai, S. T. (1987b). Intracellular study of rat substantia nigra pars reticulata neurons in an *in vitro* slice preparation: electrical membrane properties and response characteristics to subthalamic stimulation. *Brain Res.* 437, 45–55. doi: 10.1016/0006-8993(87)91525-3

Nakanishi, H., Kita, H., and Kitai, S. T. (1991). Intracellular study of rat entopeduncular nucleus neurons in an *in vitro* slice preparation: response to subthalamic stimulation. *Brain Res.* 549, 285–291. doi: 10.1016/0006-8993(91)90469-C

Nakanishi, H., Tamura, A., Kawai, K., and Yamamoto, K. (1997). Electrophysiological studies of rat substantia nigra neurons in an *in vitro* slice preparation after middle cerebral artery occlusion. *Neuroscience* 77, 1021–1028. doi: 10.1016/S0306-4522(96)00555-6

Nambu, A. (2008). Seven problems on the basal ganglia. *Curr. Opin. Neurobiol.* 18, 595–604. doi: 10.1016/j.conb.2008.11.001

Nambu, A., and Llinaœ, R. (1994). Electrophysiology of globus pallidus

neurons *in vitro. J. Neurophysiol.* 72, 1127–1139.

O'Reilly, R. C. (2006). Biologically based computational models of high-level cognition. *Science* 314, 91–94. doi: 10.1126/science.1127242

Oorschot, D. E. (1996). Total number of neurons in the neostriatal, pallidal, subthalamic, and substantia nigral nuclei of the rat basal ganglia: a stereological study using the cavalieri and optical disector methods. *J. Comp. Neurol.* 366, 580–599. doi: 10.1002/(SICI)1096-9861(19960318)366:4<580::AID-CNE3>3.0.CO;2-0

Park, M. R., Falls, W. M., and Kitai, S. T. (1982). An intracellular HRP study of the rat globus pallidus. I. Responses and light microscopic analysis. *J. Comp. Neurol.* 211, 284–294. doi: 10.1002/cne.902110307

Pasquereau, B., Nadjar, A., Arkadir, D., Bezard, E., Goillandeau, M., Bioulac, B., et al. (2007). Shaping of motor responses by incentive values through the basal ganglia. *J. Neurosci.* 27, 1176–1183. doi: 10.1523/JNEUROSCI.3745-06.2007

Paz, J. T., Deniau, J.-M., and Charpier, S. (2005). Rhythmic bursting in the cortico-subthalamo-pallidal network during spontaneous genetically determined spike and wave discharges. *J. Neurosci.* 25, 2092–2101. doi: 10.1523/JNEUROSCI.4689-04.2005

Planert, H., Szydlowski, S. N., Hjorth, J., Grillner, S., and Silberberg, G. (2010). Dynamics of synaptic transmission between fast-spiking interneurons and striatal projection neurons of the direct and indirect pathways. *J. Neurosci.* 30, 3499–3507. doi: 10.1523/JNEUROSCI.5139-09.2010

Platkiewicz, J., and Brette, R. (2010). A threshold equation for action potential initiation. *PLoS Comput. Biol.* 6:e1000850. doi: 10.1371/journal.pcbi.1000850

Puccini, G. D., Sanchez-Vives, M. V., and Compte, A. (2007). Integrated mechanisms of anticipation and rate-of-change computations in cortical circuits. *PLoS Comput. Biol.* 3:e82. doi: 10.1371/journal.pcbi.0030082

Rav-Acha, M., Sagiv, N., Segev, I., Bergman, H., and Yarom, Y. (2005). Dynamic and spatial features of the inhibitory pallidal GABAergic synapses. *Neuroscience* 135, 791–802. doi: 10.1016/j.neuroscience.2005.05.069

Redgrave, P., and Gurney, K. N. (2006). The short-latency dopamine signal: a role in discovering novel actions? Nature reviews. *Neuroscience* 7, 967–975. doi: 10.1038/nrn2022

Redgrave, P., Prescott, T. J., and Gurney, K. N. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* 89, 1009–1023. doi: 10.1016/S0306-4522(98)00319-4

Richards, C. D., Shiroyama, T., and Kitai, S. T. (1997). Electrophysiological and immunocytochemical characterization of GABA and dopamine neurons in the substantia nigra of the rat. *Neuroscience* 80, 545–557. doi: 10.1016/S0306-4522(97)00093-6

Rinvik, E., and Grofová, I. (1970). Observations on the fine structure of the substantia nigra in the cat. *Exp. Brain Res.* 11, 229–248.

Rosenbaum, R., Rubin, J., and Doiron, B. (2012a). Short term synaptic depression imposes a frequency dependent filter on synaptic information transfer. *PLoS Comput. Biol.* 8:e1002557. doi: 10.1371/journal.pcbi.1002557

Rosenbaum, R., Zimnik, A., Turner, R., Doiron, B., and Rubin., J. (2012b). "Synaptic neurotransmitter depletion as a mechanism for the reduction of Parkinsonian beta oscillations during deep brain stimulation," in *2012 Neuroscience Meeting Planner*. New Orleans, LA: Society for Neuroscience.

Ruskin, D. N., Bergstrom, D. A., Kaneoke, Y., Patel, B. N., Twery, M. J., and Walters, J. R. (1999). Multisecond oscillations in firing rate in the basal ganglia: robust modulation by dopamine receptor activation and anesthesia. *J. Neurophysiol.* 81, 2046–2055.

Sadek, A. R., Magill, P. J., and Bolam, J. P. (2007). A single-cell analysis of intrinsic connectivity in the rat globus pallidus. *J. Neurosci.* 27, 6352–6362. doi: 10.1523/JNEUROSCI.0953-07.2007

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340. doi: 10.1126/science.1115270

Sato, M., and Hikosaka, O. (2002). Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement. *J. Neurosci.* 22, 2363–2373.

Schwyn, R. C., and Fox, C. A. (1974). The primate substantia nigra: a Golgi and electron microscopic study. *J. Hirnfors.* 15, 95–126.

Shen, K.-Z., and Johnson, S. W. (2006). Subthalamic stimulation evokes complex EPSCs in the rat substantia nigra pars reticulata *in vitro*. *J. Physiol.* 573, 697–709. doi: 10.1113/jphysiol.2006.110031

Shen, W., Flajolet, M., Greengard, P., and Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321, 848–851. doi: 10.1126/science.1160575

Sims, R. E., Woodhall, G. L., Wilson, C. L., and Stanford, I. M. (2008). Functional characterization of GABAergic pallidopallidal and striatopallidal synapses in the rat globus pallidus *in vitro*. *Eur. J. Neurosci.* 28, 2401–2408. doi: 10.1111/j.1460-9568.2008.06546.x

Smith, Y., Bevan, M. D., Shink, E., and Bolam, J. P. (1998). Microcircuitry of the direct and indirect pathways of the basal ganglia. *Neuroscience* 86, 353–387. doi: 10.1016/S0306-4522 (98)00004-9

Smith, Y., Bolam, J. P., and Von Krosigk, M. (1990). Topographical and synaptic organization of the GABA-containing pallidosubthalamic projection in the rat. *Eur. J. Neurosci.* 2, 500–511.

Smith, Y., Raju, D. V., Sidibé, M., and Pare, J.-F. (2004). The thalamostriatal system: a highly specific network of the basal ganglia circuitry. *Trends Neurosci.* 27, 520–527. doi: 10.1016/j.tins.2004.07.004

Tachibana, Y., Kita, H., Chiken, S., Takada, M., and Nambu, A. (2008). Motor cortical control of internal pallidal activity through glutamatergic and GABAergic inputs in awake monkeys. *Eur. J. Neurosci.* 27, 238–253. doi: 10.1111/j.1460-9568.2007.05990.x

Takakusaki, K., Saitoh, K., Harada, H., and Kashiwayanagi, M. (2004). Role of basal ganglia-brainstem pathways in the control of motor behaviors. *Neurosci. Res.* 50, 137–151. doi: 10.1016/j.neures.2004.06.015

Tepper, J. M., and Lee, C. R. (2007). GABAergic control of substantia nigra dopaminergic neurons. *Prog. Brain Res.* 160, 189–208. doi: 10.1016/S0079-6123(06)60011-3

Terman, D., Rubin, J. E., Yew, A. C., and Wilson, C. J. (2002). Activity patterns in a model for the subthalamopallidal network of the basal ganglia. *J. Neurosci.* 22, 2963–2976.

Tokuno, H., Nakamura, Y., Kudo, M., and Kitao, Y. (1990). Laminar organization of the substantia nigra pars reticulata in the cat. *Neuroscience* 38, 255–270. doi: 10.1016/0306-4522(90)90390-P

Touboul, J., and Brette, R. (2008). Dynamics and bifurcations of the adaptive exponential integrate-and-fire model. *Biol. Cybern.* 99, 319–334. doi: 10.1007/s00422-008-0267-4

Tsodyks, M., and Markram, H. (1996). Plasticity of neocortical synapses enables transitions between rate and temporal coding. *Artif. Neural Netw.* 1112, 445–450. doi: 10.1007/3-540-61510-5_77

Tsodyks, M., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U.S.A.* 94, 719–723.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835.

Urbain, N., Gervasoni, D., Soulière, F., Lobo, L., Rentéro, N., Windels, F., et al. (2000). Unrelated course of subthalamic nucleus and globus pallidus neuronal activities across vigilance states in the rat. *Eur. J. Neurosci.* 12, 3361–3374. doi: 10.1046/j.1460-9568.2000.00199.x

Walters, J. R., Hu, D., Itoga, C. A., Parr-Brownlie, L. C., and Bergstrom, D. A. (2007). Phase relationships support a role for coordinated activity in the indirect pathway in organizing slow oscillations in basal ganglia output after loss of dopamine. *Neuroscience* 144, 762–776. doi: 10.1016/j.neuroscience.2006.10.006

Wilson, C. J. (1993). The generation of natural firing patterns in neostriatal neurons. *Prog. Brain Res.* 99, 277–297.

Wilson, C. J., Puntis, M., and Lacey, M. G. (2004). Presynaptic GABAB and adenosine A1 receptors regulate synaptic transmission to rat substantia nigra reticulata neurones. *Neuroscience* 123, 187–200.

Wu, Y., Richard, S., and Parent, A. (2000). The organization of the striatal output system: a single-cell juxtacellular labeling study in the rat. *Neurosci. Res.* 38, 49–62. doi: 10.1016/S0168-0102(00)00140-1

Zahr, N. M., Martin, L. P., and Waszczak, B. L. (2004). Subthalamic nucleus lesions alter basal and dopamine agonist stimulated electrophysiological output from the rat basal ganglia. *Synapse* 54, 119–128. doi: 10.1002/syn.20064

Zhou, F.-W., Matta, S. G., and Zhou, F.-M. (2008). Constitutively active TRPC3 channels regulate basal ganglia output neurons. *J. Neurosci.* 28, 473–482. doi: 10.1523/JNEUROSCI.3978-07.2008

Resolution enhancement in neural networks with dynamical synapses

C. C. Alan Fung[1], He Wang[1], Kin Lam[1†], K. Y. Michael Wong[1]* and Si Wu[2]*

[1] Department of Physics, The Hong Kong University of Science and Technology, Hong Kong, China
[2] State Key Laboratory of Cognitive Neuroscience and Learning, School of Cognitive Neuroscience, Lab of Neural Information Processing, Beijing Normal University, Beijing, China

Conventionally, information is represented by spike rates in the neural system. Here, we consider the ability of temporally modulated activities in neuronal networks to carry information extra to spike rates. These temporal modulations, commonly known as population spikes, are due to the presence of synaptic depression in a neuronal network model. We discuss its relevance to an experiment on transparent motions in macaque monkeys by Treue et al. in 2000. They found that if the moving directions of objects are too close, the firing rate profile will be very similar to that with one direction. As the difference in the moving directions of objects is large enough, the neuronal system would respond in such a way that the network enhances the resolution in the moving directions of the objects. In this paper, we propose that this behavior can be reproduced by neural networks with dynamical synapses when there are multiple external inputs. We will demonstrate how resolution enhancement can be achieved, and discuss the conditions under which temporally modulated activities are able to enhance information processing performances in general.

Keywords: continuous attractor neural network, neural field model, short-term synaptic depression, short-term synaptic plasticity, transparent motion

1. INTRODUCTION

An important issue in computational neuroscience is how information is represented in the neural system. It was widely accepted that spike rates of neurons carry information. This notion was further illustrated in *population codes*, in which the a group of neurons encode information and even represent uncertainties therein through their collective activities (Zemel and Dayan, 1999; Pouget et al., 2000). Consequently, population coding has been successfully applied to describe the encoding of spatial and directional information, such as orientation (Ben-Yishai et al., 1995), head direction (Zhang, 1996), and spatial location (Samsonovich and McNaughton, 1997). They are also used to explain information processing in the recently discovered grid cells (Fuhs and Touretzky, 2006).

An interesting question arises, namely, whether information can be encoded in other aspects of population coding besides spike rates. For example, can extra information be carried by the coding if the spikes are modulated in time, so that different spike trains modulated differently may convey different messages even though their spike rates appear to be the same. Given this possibility, the information content of population coding can be much richer than its superficial appearance as spike rates.

In this paper, we will explore the ability of population spikes to carry information extra to spike rates. Population spikes are temporal modulations of the population neuronal activity, and are also known as ensemble synchronizations, representing extensively coordinated rises and falls in the discharge of many neurons (Loebel and Tsodyks, 2002; Holcman and Tsodyks, 2006). The population spikes are due to the presence of short-term depression (STD) of the synapses, referring to the reduction of synaptic efficacy of a neuron after firing due to the depletion of neurotransmitters (Stevens and Wang, 1993; Markram and Tsodyks, 1996; Dayan and Abbott, 2001). This adds to a recently expanding list of the roles played by STD in neural information processing. For example, STD was recently suggested to be useful in expanding the dynamic range of the system (Abbott et al., 1997; Tsodyks and Markram, 1997), estimating the information of the presynaptic membrane potential (Pfister et al., 2010), and stabilizing the self-organized critical behavior for optimal computational capabilities (Levine et al., 2007). STD was also found to be useful in enhancing the mobility of the network state in tracking moving stimuli (Fung et al., 2012a), and hence was recently proposed to be a foundation of a potential anticipation mechanism (Fung et al., 2012b).

Previously, population spikes were found to be global synchronizations of neuronal activities. However, in order for them to encode spatial information, the population spikes that will be considered in this paper are localized ones. We will use the case of transparent motion as an example. This example illustrates the possibility that the modulation by population spikes enables the neural system to refine the resolution of direction for multiple stimuli. The prediction by the proposed mechanism has an excellent agreement with experimental results (Treue et al., 2000).

Transparent motion is one of the most well-known experiments in the psychophysical community. In the experiment, the stimulus usually contains moving dots with different directions. So, there are multiple moving directions transparently superimposed on one another. In the nervous system, the middle temporal (MT) area was found to be responsible for detecting moving directions of objects (Maunsell and Van Essen, 1983). Here, it was recently found that the neurons are heterogeneous, with some neurons responding to the pattern of moving stimuli, while others responding to the components of composite moving patterns (Rust et al., 2006). In 2000, Treue et al. found that if the directions of two groups of moving dots differ by an angle larger than the tuning width of the neurons, the observed neuronal response profile begins to split (Treue et al., 2000). However, subjects can still distinguish the two directions if their difference is as small as about 10° (Mather and Moulden, 1980), while the average direction tuning width of neurons is about 96°.

To resolve this paradox, Treue et al. proposed that when the resultant neuronal response is too board for a single direction, the perception can identify the two directions by considering the resultant neuronal response as a superposition of two individual neuronal responses of each direction. However, when the two directions differ by an angle less than the tuning width, it becomes difficult to resolve the peaks of the two superposed responses, if the curvature of the average neural activity profile is not taken into account This difficulty was also observed in simulations with distributional population codes (Zemel and Dayan, 1999). The mechanism of enhanced resolution remained unknown, and coding by firing rates may not reveal the complete picture.

In a recently proposed model on motion transparency, the enhanced resolution was achieved (Raudies et al., 2011). Two mechanisms held the key to this advance. First, as in standard neural field models, there is a local center-surround competition in the space of motion directions. Although this is not sufficient to explain the enhanced resolution, there is the second mechanism, namely, the modulatory feedback signals from higher stages of processing in the area medial superior temporal (MST) area. Motion attraction (that is, under-estimation of the directional difference) at small angular difference, and motion repulsion (that is, over-estimation) at larger angles were successfully explained. Perception repulsion can also be found in a Bayesian inference explanation on identification of audiovisual stimulus (Sato and Toyoizumi, 2007).

Here, we propose a novel mechanism for resolution enhancement based on the temporal modulation inherent in population coding. To focus on the generic issue of whether information carried in the temporal modulation of population coding can be usefully applied in a processing task, we consider a simplified model of transparent motion. We assume that inputs from different locations of the receptive field have been integrated, the directional information has been filtered, and the processing of input information can proceed without the assistance of feedback modulations. Thus our working model reduces to a single network. The working principle is a continuous attractor neural network (CANN) with dynamical synapses. Continuous attractor neural networks, also known as neural field models, are models used for describing phenomena and features observed in some brain regions where localized attractor neuronal responses are used to represent continuous information. Due to short-range excitatory interactions and long-range/global inhibitory interactions, bump-shaped neuronal response profiles are attractors of CANNs. Since the response profiles are easy to shift their positions in the space of continuous information, they are useful in tracking moving stimuli (Amari, 1977; Ben-Yishai et al., 1995; Wu et al., 2008; Fung et al., 2010) and their drifting behaviors have been studied (Itskov et al., 2011). In contrast to these studies of tracking, we will focus on stationary stimuli and their time-dependent neuronal responses.

Dynamical synapses are found to enrich the dynamical behaviors of CANNs (York and van Rossum, 2009; Fung et al., 2012a). Short-term synaptic depression (STD) can degrade the synaptic efficacies between neurons, depending temporally on the activity history of the presynaptic neuron (Tsodyks et al., 1998). In the presence of an external stimulus, the bumps can remain temporally stable if STD were absent. However, with STD, the population activity may drop after it reaches a maximum, since neurotransmitters have been consumed. After the drop, neurotransmitters are recovered and the neuronal population is ready to respond to the external stimulus again. This results in periodic bursts of local neuronal responses, referred to as population spikes. As we shall see, the temporal modulation induced by STD, together with input fluctuations, enable the system to reduce the angle of resolution in transparent motion down to one-fourth to one-third of the tuning width of the neuron.

In the rest of this paper, we will begin with an introduction of the CANN model and its basic properties. After that, we will discuss simulation results showing that our model is able to represent acute difference in transparent stimuli. At the end, there is a discussion section concluding our proposed mechanism.

2. MODEL AND METHOD

In the continuous attractor neural network model, we specify the dynamics and the state of the system by the neuronal current. For neurons with preferred stimulus x in the range $-L/2 \leq x \leq L/2$, its neuronal current is denoted by $u(x, t)$. The dynamics of $u(x, t)$ is given by Fung et al. (2012a)

$$\tau_s \frac{du}{dt}(x, t) = -u(x, t) + I^{\text{ext}}(x, t) + \rho \int dx' J(x - x') p(x', t) r(x', t).$$

(1)

τ_s is the timescale of $u(x, t)$. It is usually of the order of the magnitude of 1 ms. ρ is the density of neurons over the space spanned

by $\{x\}$. $J(x - x')$ is a translational invariant excitatory coupling given by

$$J(x - x') = \frac{J_0}{\sqrt{2\pi}a} \exp\left(-\frac{|x - x'|^2}{2a^2}\right), \qquad (2)$$

where a is the range of excitatory connection and J_0 is the average strength of the coupling. $r(x, t)$ is the neural activity related to $u(x, t)$ by

$$r(x, t) = \Theta\left[u(x, t)\right] \frac{u(x, t)^2}{B(t)}. \qquad (3)$$

Here, Θ is a step function centered at 0. The denominator, $B(t) \equiv 1 + k\rho \int dx' u(x', t)^2$, in this formula is the global inhibition, controlled by the inhibition parameter k. This type of global inhibition can be achieved by shunting inhibition (Heeger, 1992; Hao et al., 2009). $I^{\text{ext}}(x, t)$ is the external input to the system, which will be defined in the latter part of this section.

In the integral of Equation (1), $p(x, t)$ is the available fraction of neurotransmitters of the presynaptic neurons. Neurotransmitters are consumed when a neuron sends chemical signals to its postsynaptic neurons. However, the recovery time of the neurotransmitters is considerably longer than τ_s. This process can be modeled by Tsodyks et al. (1998) and Fung et al. (2012a)

$$\tau_d \frac{dp}{dt}(x, t) = -p(x, t) + 1 - \tau_d \beta p(x, t) r(x, t). \qquad (4)$$

τ_d is the timescale of recovery process of neurotransmitters. The recovery process usually takes 25–100 ms. Here, we choose $\tau_d = 50\tau_s$. These two differential equations, Equations (1) and (4), are found to be consistent with the model proposed by Tsodyks et al. (1998).

The stimulus fed to the system consists of n components, each with a Gaussian profile and a time-dependent fluctuation in strength. It is given by

$$I_0^{\text{ext}}(x, t) = \sum_{i=1}^{n} [A_0 + \delta A_i(t)] \exp\left(-\frac{|x - z_i|^2}{2a_I^2}\right). \qquad (5)$$

Here, z_i's are the peak positions of the components, and a_I is the width of the Gaussian profiles. If not specified, it was assumed to be the same as the synaptic interaction range a used in Equation (2). A_0 is the average relative magnitude of one input component, while $\delta A_i(t)$ is a random fluctuation with standard deviation σ_A in amplitude of input components.

Note that when the Gaussian profiles have strong overlaps, the components cannot be resolved, as illustrated in **Figures 1A–C**. We consider the amplitude fluctuations of each component to be independent of each other, i.e., $\langle \delta A_i \delta A_j \rangle = 0$, where the average is over time. These fluctuations provide a cue for the system to distinguish different components (**Figure 1D**). This is consistent with the psychophysical experiment which showed that spatial and temporal randomness is important for perception of motion transparency (Qian et al., 1994). Since the fluctuations
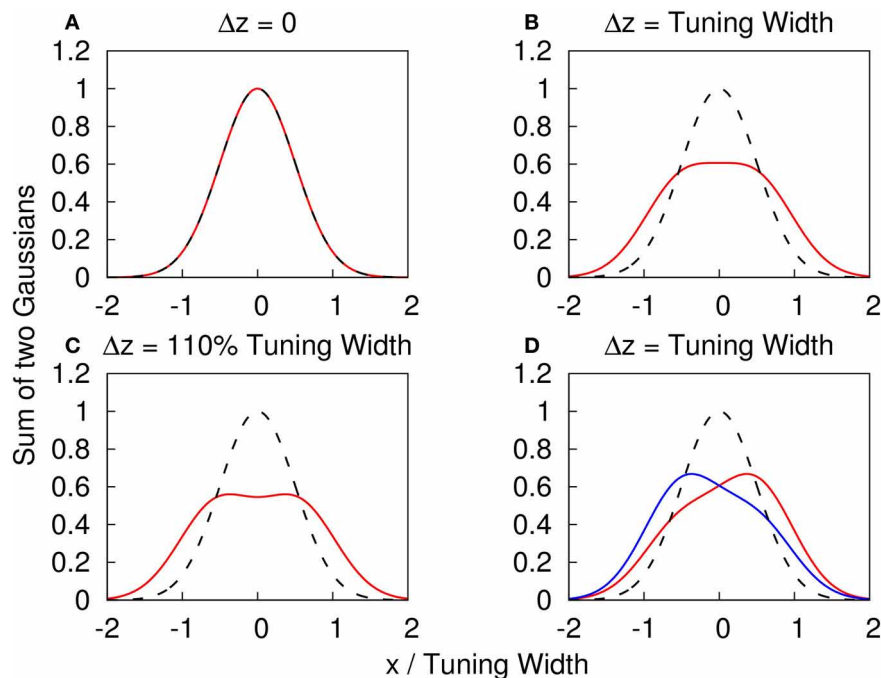
FIGURE 1 | (A–C) The profile of two superposed Gaussian functions with the same height. $f(x) \equiv \{\exp[(x - \Delta z/2)^2/(2a^2)] + \exp[(x + \Delta z/2)^2/(2a^2)]\}/2$. Red solid line: $y = f(x)$ with different Δz. Dashed line: $y = f(x)$ with $\Delta z = 0$ as a reference. **(A)** $\Delta z = 0$. **(B)** $\Delta z = $ tuning width $= 2a$. **(C)** $\Delta z = 110\%$ tuning width $= 2.2a$. **(D)** The profile of two superposed Gaussian functions with

different heights to illustrate how the amplitude fluctuations provide a cue to distinguish the components. $g(x) \equiv \{A_0 \exp[(x - \Delta z/2)^2/(2a^2)] + A_1 \exp[(x + \Delta z/2)^2/(2a^2)]\}$. Dashed line: $y = f(x)$ with $\Delta z = 0$ as a reference. Red solid line: $y = g(x)$ with $\Delta z = $ tuning width, $A_0 = 0.4$ and $A_1 = 0.6$. Blue solid line: $y = g(x)$ with $\Delta z = $ tuning width, $A_0 = 0.6$ and $A_1 = 0.4$.

vanish when averaged over time, a system responding only to time-averaged inputs is unable be able to detect the components. Here, the role of STD is to modulate the network state, so that it responds to one input component once a time.

To model the situation that the maximum strength of the input profile is invariant, we consider the input in Equation (1) to be

$$I^{\text{ext}}(x, t) = \frac{A}{\max_x \left[I_0^{\text{ext}}(x, t) \right]} I_0^{\text{ext}}(x, t), \tag{6}$$

where A is the fixed maximum magnitude of the external input. As the external input profile is set to have a constant maximum, only the ratio σ_A/A_0, rather than the magnitudes of A_0 and σ_A, is relevant in our studies.

It is convenient to rescale the dynamical variables as follows. We first consider the case without STD when $\beta = 0$, and the synaptic interaction range $a \ll L$. In this case, $p(x, t) = 1$ in Equation (1). For $k \leq k_c \equiv \rho J_0^2/(8\sqrt{2\pi}a)$, the network holds a continuous family of Gaussian-shaped stationary states when $I^{\text{ext}}(x, t) = 0$. These stationary states are

$$\tilde{u}(x) = \tilde{u}_0 \exp\left(-\frac{|x - z|^2}{4a^2}\right), \tag{7}$$

and

$$\tilde{r}(x) = \tilde{r}_0 \exp\left(-\frac{|x - z|^2}{2a^2}\right). \tag{8}$$

where $\tilde{u}(x)$ is the rescaled variable $\rho J_0 u(x)$, and $\tilde{u}_0$ is the rescaled bump height. The parameter z, i.e., the center of the bump, is a free parameter, implying that the stationary state of the network can be located anywhere in the space x. In this paper, we assume that the variable is represented solely by the peak position of the neural activity profile. This assumption is one of the most direct ways to interpret the population code. However, there are other ways to interpret population codes. For example, Treue et al. (2000) proposed that the curvature of the average of the neural activity carries information represented by the neural population code, although the mechanism achieving this objective is not clear (Treue et al., 2000). On the phenomenological level, distributional population coding and double distributional population coding were proposed to represent information in population coding with more sophistication (Zemel and Dayan, 2000; Sahani and Dayan, 2003).

The tuning width of a neuron, defined as the standard deviation of the firing rate profile multiplied by 2, is therefore $2a$. In the present work, we rescale the neuronal current as $\tilde{u}(x, t) \equiv \rho J_0 u(x, t)$, together with the corresponding rescaling of other variables given by $\tilde{A} \equiv \rho J_0 A$, $\tilde{k} \equiv k/k_c$, $\tilde{\beta} \equiv \tau_d \beta/\left(\rho^2 J_0^2\right)$. By using these rescaling rules, the dynamics of the system should only depend on $\tilde{k}$, $\tilde{\beta}$, τ_d/τ_s, σ_A/A_0, z_i's and $\tilde{A}$. Below, only these parameters will be specified.

In each simulation, the variables $u(x, t)$ are modeled to be located at N discrete positions uniformly distributed in the space of preferred stimuli $\{x\}$. To do massive simulations, all simulation results are generated by using $N = 80$. We have verified that the dynamics of the system is independent to N, and the number of neurons should not affect the conclusion. The boundary condition of the space is periodic. The range of the network is $360°$ and the tuning width of the neurons is $96°$, following the experimental estimates in Treue et al. (2000). To solve differential equations in Equations (1) and (4), we used the Runge-Kutta Prince-Dormand (8,9) method provided by the GNU Scientific Library. Initial conditions of $u(x, t)$'s is zero, while $p(x, t)$'s are initially 1. The local error of each evolution step is less than 10^{-6}. The random number generator used to generate the Gaussian random number is the generator proposed by Lüscher (1994). The Gaussian fluctuation is updated every $50\tau_s$.

3. RESULTS

3.1. POPULATION SPIKES

We first consider the response of the network when the input consists of one component. We explore the network behavior by varying the parameters $\tilde{k}$, $\tilde{\beta}$, and $\tilde{A}$. We found a rich spectrum of behaviors including population spikes, static bumps, and moving bumps. The full picture will be reported elsewhere. For the purpose of the present paper, we fix $\tilde{k}$ and $\tilde{\beta}$ at a typical value and consider the behavior when $\tilde{A}$ increases. As shown in the top panel of **Figure 2**, the network cannot be triggered to have significant activities when the input is weak. In the bottom panel, the input is so strong that the network response is stabilized to a static bump with time-independent amplitude. An interesting case arises in the middle panel for moderately strong input, where population spikes can be observed. Population spikes are the consequence of the presence of STD. They are caused by a rapid rise of neuronal activity due to the external stimulus. Then in a time of the order of τ_d, the neurotransmitters are consumed, leading to a rapid drop in neuronal activity. When the neurotransmitters recover, the neurons become ready for the next population spike, resulting in the interesting periodic behavior. Population spikes have been found before as synchronization of neuronal activities, and their potential role in processing information was appreciated, but no specific context of such applications was identified (Loebel and Tsodyks, 2002), Here, we will present an example that spatially localized population spikes endow the neural system a capacity of reading-out input components.

3.2. NETWORK ACTIVITIES FOR TWO STIMULI

Next, we consider inputs with two components separated by $\Delta z > 0$ and study the network behavior when Δz gradually increases. Without loss of generality, we choose $z_1 = \Delta z/2$ and $z_2 = -\Delta z/2$. The relative fluctuation is $\sigma_A/A_0 = 0.3$.

When the separation is small, the positions of the population spikes fluctuate around the mid-position of the two stimuli, as illustrated in **Figure 3A**. The two components cannot be resolved.

When the separation increases to the extent that the two components remain barely resolved, an interesting change in the spiking pattern occurs as shown in **Figure 3B**. The positions of the population spike peaks begin to center around the two input components, although the shoulders of the population spikes remain overlapping considerably. Note that in this regime, the profile of the neuronal activities remain unresolved when they are averaged over time. However, due to the presence of STD, it is likely that a population spike is produced at the position of the component which happens to be higher due to height fluctuations. Hence in

FIGURE 2 | Firing rates $\tilde{r}(x, t)$ (A, D, and G), available fraction of neurotransmitters $p(x, t)$ (B, E, and H) and corresponding input (C, F, and I) for various magnitudes of single-peaked external inputs. (A–C) $\tilde{A} = 0.4$, (D–F) $\tilde{A} = 0.8$, and (G–I) $\tilde{A} = 2.0$. Other parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0.24$, $a = 48\pi/180$, and $\tau_d = 50\tau_s$.

FIGURE 3 | Raster plot of firing rates $\tilde{r}$ for (A) $\Delta z = 0.5$, (B) $\Delta z = 1.0$, and (C) $\Delta z = 2.0$. White dashed lines: positions of stimuli. Parameters: other parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0.24$, $a = 48\pi/180$, $\tilde{A} = 0.8$, $\sigma_{\delta A_i}/A_0 = 0.3$, and $\tau_d = 50\tau_s$.

this regime, the population spike peaks are no longer aligned at the center. Rather, they are arranged in two rows, each around the two components. Furthermore, the two rows of population spikes tend to fire alternately. This implies that although it is hard to resolve the two components by considering the time-averaged signals, the temporal modulation by the alternating population spikes may be utilized for resolution enhancement.

When the separation increases further, the population spikes form two groups clearly, as shown in **Figure 3C**. The two components are clearly resolved.

To compare our model with experimental results, we measure the time average of neuronal activities as a function of preferred stimuli of neurons and the separation of the two stimuli, shown in **Figure 4A**. We found that this result is very similar to the experimental results reported by Treue et al. [Figure 2C in Treue et al. (2000)]. The peak of the average profile of neuronal activities splits near $\Delta z \sim 1.0\times$ tuning width. However, the time-averaged data cannot explain why subjects can resolve separations much less than the tuning width.

3.3. EXTRACTION OF MODULATED INFORMATION

To demonstrate that the neuronal activities carry the information about two stimuli, we collect statistics on the peak positions of the population spikes. Here the peak position is calculated by $\max_x \tilde{r}(x)$. In **Figure 4C**, we present the contour plot of the distribution of peak positions in the space of the preferred stimuli of neurons and separation between the two stimuli in units of the tuning width. To focus on peaks with significant information only, we counted only population spikes with maximum amplitudes above an appropriately chosen threshold. Each column in **Figure 4C** is a normalized histogram with 80 bins. In order to obtain a relatively smooth distribution, the sampling process lasted for 100,000 τ_s. The mean of the separation between peak positions is plotted in **Figure 5** as a function of Δz. We found that in this setting, the system can detect the input separation down to one-fourth of the tuning width. We note that in **Figure 4C**, when the difference between the components is too small, $\Delta z \lesssim 1/4$ tuning width, population spikes occur at the middle of the net external input profile with a relatively small variance. However, when the network starts to resolve the two components, there are notable variances on positions of the population spikes in each component. The standard deviation of the positions of the population spikes in each component is roughly of the order of 0.1 times the tuning width, which is roughly 20°, as shown in **Figure 5**.

To investigate whether the statistics with long sampling period is applicable to sampling periods in actual experiments, we have also collected statistics for 500τ_s. (In the experiment done by

FIGURE 4 | (A) Time average of firing rates $\tilde{r}$ as a function of the preferred stimuli of neurons, x, and the separation between the two stimuli, Δz. Contour lines: $\langle \tilde{r} \rangle_t = 1$ (dotted-dashed line), $\langle \tilde{r} \rangle_t = 2$ (dashed line), $\langle \tilde{r} \rangle_t = 3$ (solid line), $\langle \tilde{r} \rangle_t = 4$ (dotted line). Parameters: same as **Figure 3**. **(B)** The average neural activity recorded by Treue et al. (2000) (with license number 3125800919243 for the reuse purpose). **(C)** Contours of the distribution of peak positions higher than 6.2 as a function of preferred stimuli, x, and the separation between the two stimuli, Δz. White dashed line: positions of the two stimuli. L_1, one-third of the tuning width. L_2, tuning width. Parameters, same as **Figure 3**.

Treue et al., subjects took 500 ms to perform the discriminational task.) The result is shown in **Figure A2A** in Appendix. Although the distribution is rougher because of the relatively small sampling size, enhanced resolution down to 0.3 tuning width is still visible.

Furthermore, when the separation between the two stimuli lies between one-third and three-halves of the tuning width, the system slightly overestimated the separation of the two profiles. If we take the tuning width to be 96° (Treue et al., 2000), this range will be approximately from 30° to 140°. This is consistent with the experimental results of Braddick et al. (2002), in which subjects overestimated some moving direction difference in transparent motion experiments. However, it was reported in Figure 4 in Treue et al. (2000) that the perceived separation of movement direction starts to underestimate the truth when the stimulus separation increases above 40°. Since the range corresponding to "motion repulsion" reported by Braddick et al. (2002). is different from that reported by Treue et al., it seems that the range of differences between stimuli corresponding to "motion repulsion" is different for different experimental settings.

We have also tested the effects of choosing the widths of external input components to be different from the tuning width of the neuronal response. We found that the results for different stimulus strengths in **Figure A1** in Appendix are qualitatively the same as that in **Figure 4C**.

The result shown in **Figure 4C** is not particular for the chosen set of parameters. In **Figure 6**, there is a phase diagram along with some selected parameters. In **Figure 6A**, the colored region is the region for population spikes with one stimulus. If $\tilde{A}$ and $\tilde{\beta}$ are chosen from this region, as far as we have observed, similar

FIGURE 5 | The mean separation of peak positions of $r(x, t)$ shown in Figure 4C. Symbols, simulation. Dashed line, diagonal line representing perfect distinguishability.
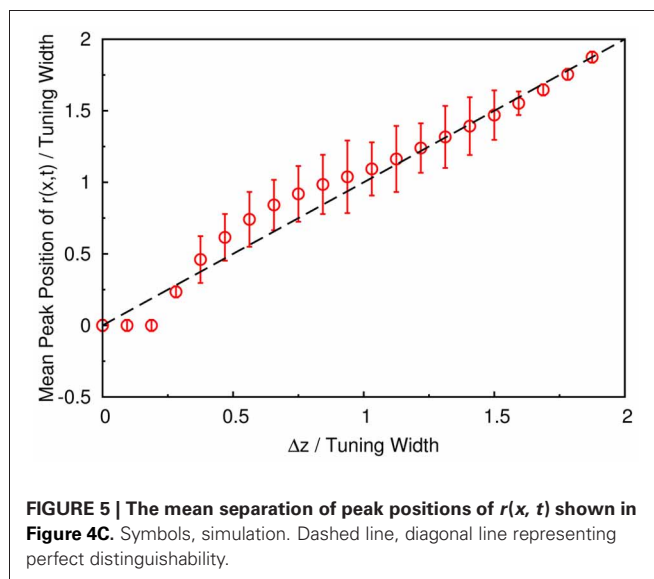
results can be obtained by choosing appropriate thresholds. If $\tilde{A}$ and $\tilde{\beta}$ are outside the colored region, no matter what the threshold was, the result shown in **Figure 4C** cannot be reproduced. This result suggests that population spikes are important to resolution enhancement.

3.4. NETWORK RESPONSE WITH MULTIPLE STIMULI

We further test the response of our model to more than two stimuli. **Figure 7** shows the case for three stimuli of equal amplitude,
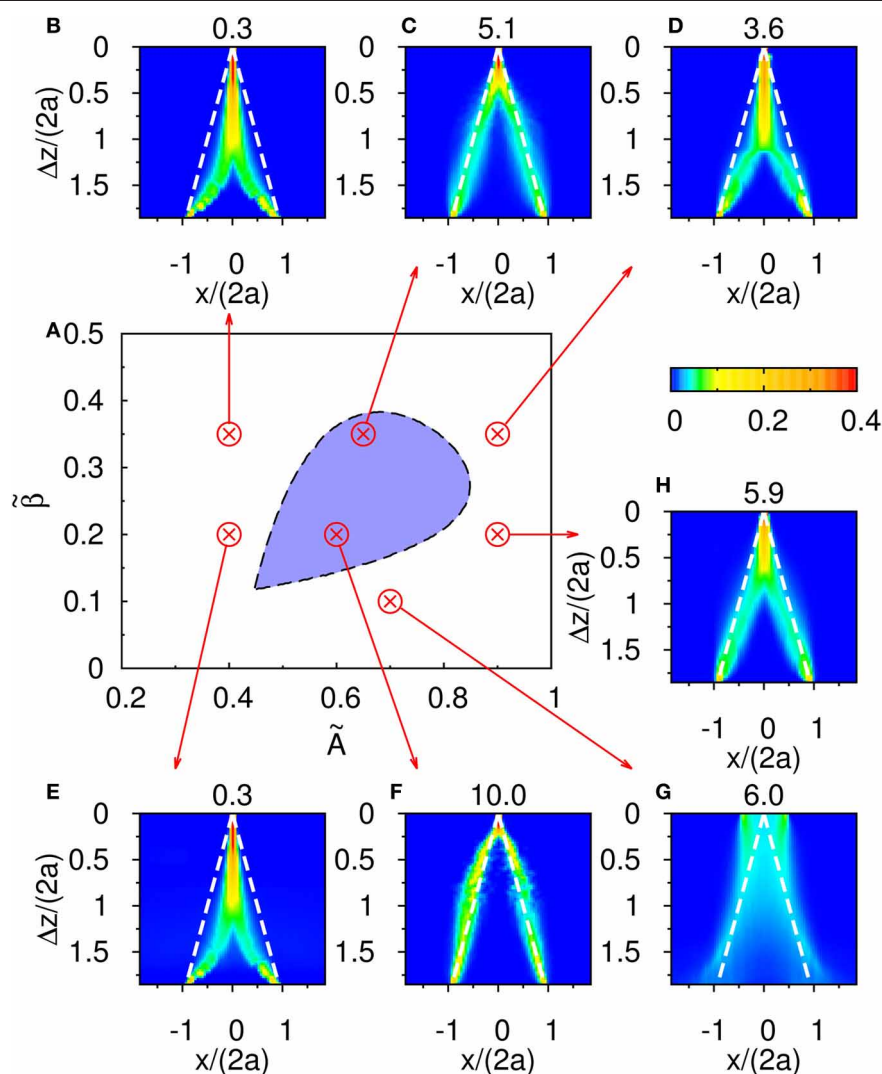
FIGURE 6 | (A) The phase diagram of population spikes over the parameter space spanned by $(\tilde{A}, \tilde{\beta})$ with the parameter $\tilde{k} = 0.5$ and $\tau_d/\tau_s = 50$. **(B–H)** are distributions of the occurence of peak positions as function of Δz. The numbers at the top of **(B–H)** are thresholds used to sample peak positions. Parameters: **(B)** $\tilde{A} = 0.4$ and $\tilde{\beta} = 0.35$. **(C)** $\tilde{A} = 0.65$ and $\tilde{\beta} = 0.35$. **(D)** $\tilde{A} = 0.9$ and $\tilde{\beta} = 0.35$. **(E)** $\tilde{A} = 0.4$ and $\tilde{\beta} = 0.2$. **(F)** $\tilde{A} = 0.6$ and $\tilde{\beta} = 0.2$. **(G)** $\tilde{A} = 0.7$ and $\tilde{\beta} = 0.1$. **(H)** $\tilde{A} = 0.9$ and $\tilde{\beta} = 0.2$.

whose peak positions are labeled by the white dashed lines. However, the contours of the distribution of population spikes are double-peaked, similar to those in **Figure 5**. This result suggests that, if there are three stimuli overlapped together, the network response should give only two groups of neuronal responses. Also, it predicts that peaks of population spikes should occur at positions that underestimate the separation between the outermost stimuli. A similar result for shorter sampling periods comparable to actual experiments can be found in **Figure A2B** in Appendix.

We found that the experimental result of multiple stimuli reported by Treue et al. is consistent with this prediction. In their paper, it was reported that, when there were three groups of moving dots moving at directions ±50° and 0°, the subjects would

report that there were only two moving directions at ±40°. This consistency is shown in **Figure 7**, where the vertical dotted line L labels the position that the outermost stimuli are directed at ±50° when the tuning width is 96°, and the pair of horizontal dashed lines labels ±40° correspondingly.

4. CONDITIONS FOR RESOLUTION ENHANCEMENT

We have demonstrated the phenomenon of resolution enhancement due to modulations of population spikes. To see whether this picture can be generalized to other cases and what alternative models are to be excluded, we summarize the general conditions of its occurrence. To appreciate the significance of each condition, we will consider the alternative scenarios in the presence and absence of the various conditions.

4.1. SHORT-TERM SYNAPTIC DEPRESSION

Without the STD, the steady state of the neuronal activity profile becomes centered at either one of the two input stimuli. In **Figure 8A**, when the difference between the input profiles is large, $\Delta z/a = 3.7$ for instance, the neuronal activity is trapped by the input profile near $x = 1.55$. This case is not consistent with experiments, because when the separation between the input profiles is large enough, the neuronal activity should be able to identify both stimuli. This shows that STD plays the following roles in this phenomenon.

First, STD gives rise to the temporal modulation characterized by the population spikes, in which rapid rises in population activities alternate periodically with drops due to the consumption of neurotransmitters. Spiking activities enable the activity profile to jump from one stimulus position to another easily.

Second, the presence of STD enhances the mobility of the activity profiles. Due to the consumption of neurotransmitters in the active region, the profile tends to relocate itself to less active regions. This is the cause of the increased mobility when the activity profile tracks the movement of external stimuli, as well as their anticipatory tracking as a possible mechanism for delay compensation (Fung et al., 2012a,b). In the parameter regime where the stationary profile becomes unstable in its position, and population spikes become the attractor state, the network tends to establish a population spike in new locations, preventing itself from being trapped by one stimulus. This results in population spikes centered at alternating stimuli and hence the temporal modulation.

For example, if the two stimuli are strongly overlapped, the average neuronal response concentrates at the in-between region of the two stimuli, as shown in **Figure 3B**. In this case, the time-average profile of the dynamical variable $p(x, t)$ has a dip centered at the midpoint between two stimuli, as shown in **Figure 9**. Since, in our model, there are fluctuations of the magnitude of each component of the external input, population spikes occur near the positions of the stimuli, labeled by the blue lines in **Figure 9**. Since the synaptic efficacies of the presynaptic neurons are stronger in the side region further away from the other stimulus, population spikes are more likely to happen in the outer region rather than the inner region. So, the separation between the two groups of population spikes can be larger than the separation between the two stimuli. This is also the reason why only two groups of population spikes can be observed in the case with three stimuli (**Figure 7**). STD also explains the slight over-estimation of the perceived positions when the separation of the stimuli is around the tuning width.

Third, when STD is not sufficiently strong, we observe that sloshers rather than population spikes are formed (Folias, 2011). These sloshers are bumps that oscillate back and forth around the external stimuli, as shown in **Figure 10**. The height of the bumps is highest when they slosh to the extreme positions, but due to the weaker STD, the height variation in a cycle is not as extreme as those in the population spikes. The positional extent of their oscillations is mainly determined by the restoring attraction from the external input, and is effectively insensitive to the stimulus profile. Hence in the task of resolving the stimulus directions, the
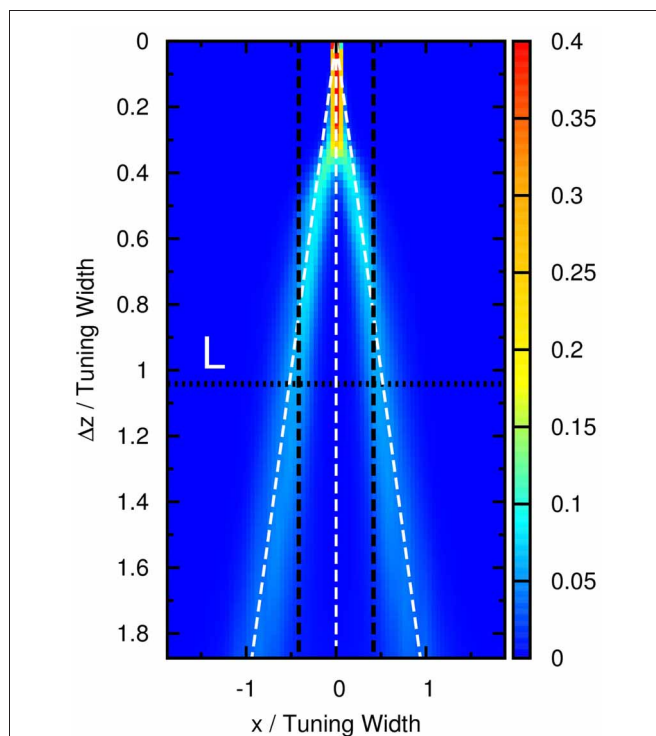
FIGURE 7 | Contours of the distribution of peak positions higher than 6.2 as a function of preferred stimuli, *x*, and the separation between the two outermost stimuli, Δ*z*, in the case of three equally strong stimuli. White dashed line: positions of three stimuli. Horizontal dotted line: the case comparable to the three-stimulus experiment reported by Treue et al., 2000. Vertical dashed lines: perception (±40°) reported by subjects in the experiment in units of the tuning width (96°). Parameters: same as **Figure 5**.

performance is degraded by the very flat part of the curve of the perceived separation when the stimuli have strong overlaps, as shown in **Figure 6G**.

There are also other variants of the model that demonstrate the significance of STD in similar ways. For example, in recurrent networks with local inhibition, we may replace $B(t)$ in Equation (3) by $B'(x, t)$ given by

$$B'(x, t) = 1 + \rho k \int dx' \exp\left(-\frac{|x - x'|^2}{2b^2}\right) u(x', t)^2 . \quad (9)$$

To stabilize the neural activity, the range of the local inhibition, b, has to be larger than the range of excitatory connection, a. However, if a is as large as 48°, this local inhibition can be fairly replaced by $B(t)$ with appropriate $\tilde{k}$. In the presence of STD, the discrimination performance is comparable to that in **Figure 5**, but the resolution is poor otherwise.

4.2. SUITABLY STRONG INPUT PROFILES

Suitably strong input magnitude is needed to produce the temporally modulated patterns, as illustrated in **Figure 6**. First, when the magnitude of the external input is too small, no significant system-driven neuronal activity can be observed. Fluctuations
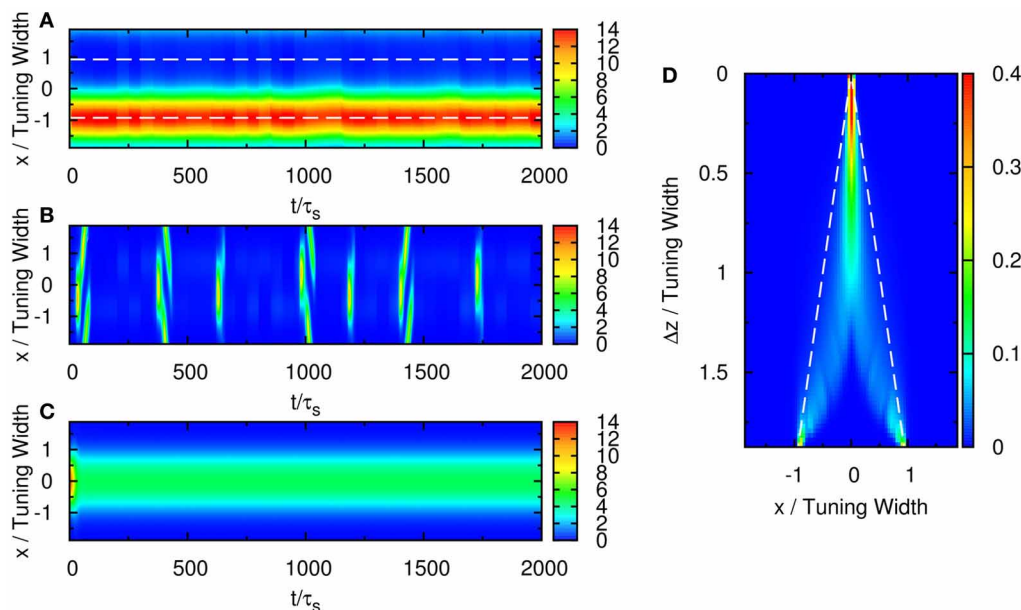
FIGURE 8 | **(A)** Raster plot of firing rate $\tilde{r}$ of the network with two stimuli and without STD. Parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0$, $\tilde{A} = 0.8$, $a = 48\pi/180$, $\sigma_A/A_0 = 0.3$, and $\Delta z = 3.1$. **(B)** Rastor plot of firing rate $\tilde{r}$ of the network with two stimuli with weak net input profile. Parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0.24$, $\tilde{A} = 0.4$, $a = 48\pi/180$, $\sigma_A/A_0 = 0.3$, and $\Delta z = 2.5$. **(C)** Rastor plot of firing rate $\tilde{r}$ of the network with two

stimuli without height fluctuations in the external input profile. Parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0.24$, $\tilde{A} = 0.8$, $a = 48\pi/180$, $\sigma_A/A_0 = 0$, and $\Delta z = 1.67$. **(D)** Contours of the distribution of peak positions for all peak heights. White dashed line: positions of the two stimulus components. Parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0.24$, $\tilde{A} = 0.8$, $a = 48\pi/180$, $\sigma_A/A_0 = 0.3$, and $\Delta z = 1.0$.
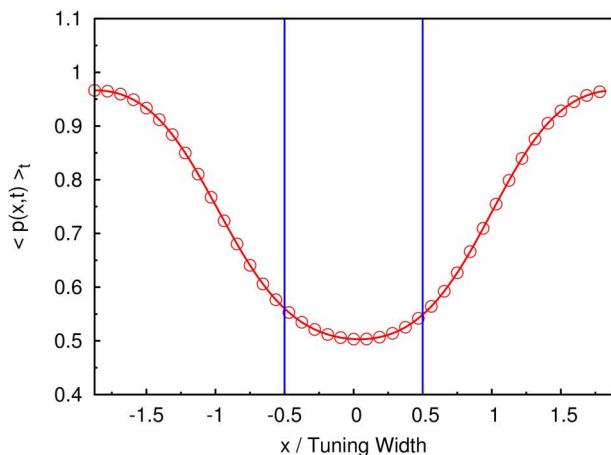
FIGURE 9 | The time-averaged dynamical variable $p(x, t)$. Symbols and red line: measurement from the simulation. Blue lines: positions of two stimuli. Parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0.24$, $a = 48\pi/180$, $\tau_d/\tau_s = 50$, $\tilde{A} = 0.8$, and $\Delta z = $ tuning width of attractor states.

of external input components cannot stimulate the population spike, as the activation by input profiles was not strong enough. Second, even when the magnitude of the external input is larger, population spikes can be produced but the stimulus is too weak to pin them at the position of the stimuli. Since the mobility of the population spikes is enhanced by STD, moving population spikes are formed, as illustrated in **Figure 8B**. Since the

population spikes move away from the stimulus positions after their formation, they cannot be used to encode the stimulus positions and also become part of the noisy background affecting the recognition of the stimulus positions. When the stimulus is too strong, population spikes cannot be generated and the resolution degrades.

4.3. FLUCTUATIONS IN INPUT PROFILES

Fluctuations on external input components is important to the behavior in **Figure 3**. If there were no fluctuations in the input profiles, the net input profile will have only one peak for $\Delta z < 2a$. As a result, there is effectively one bell-shaped input profile if the difference between two stimuli is too small, and the network response will also be single-peaked, as shown in **Figure 8C**. Hence fluctuations in the external input play the role of rendering the components distinguishable. As shown in **Figure 11**, recognition of input location always follows a strong input on the same side at the current step, and a strong input on the other side in the previous step, suggesting that a sudden shift in input bias provides condition for reliable recognition. In fact, the noise fluctuations act as the signals themselves, without which the single-peaked input provides little information about the components. Results in **Figure 11** also illustrate that, statistically, the system is able to give valid responses to stimulus changes in a single step. This explains why the network yields discrimination performance equally well for short and long sampling periods, as demonstrated in a comparison between **Figures 4C**, **A2A**.

The fluctuations may come from randomness in the inputs. Psychophysical experiments show that spatial and temporal
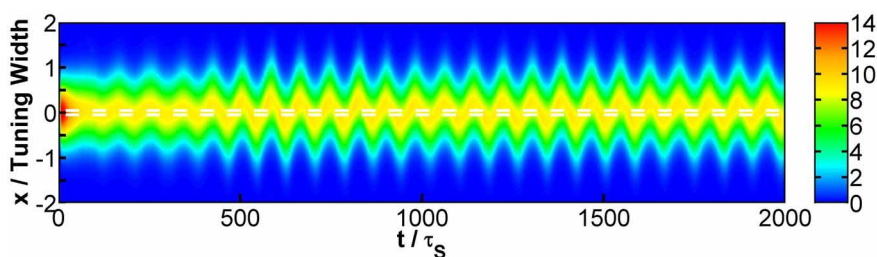
FIGURE 10 | Raster plot of firing rates $\tilde{r}$ at $\Delta z = 0.1$, showing a slosher. White dashed lines: positions of the stimuli. Other parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0.1$, $a = 48\pi/180$, $\tilde{A} = 0.8$, $\sigma_A/A_0 = 0.2$, and $\tau_d = 50\tau_s$.
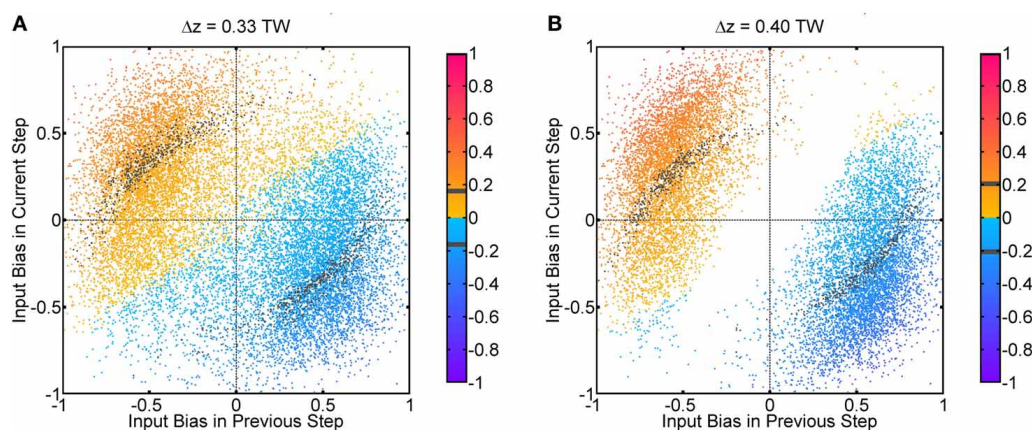
FIGURE 11 | Population spikes' positions conditional on input fluctuations. One step refers to $50\tau_s$, which is the temporal interval between every update in the Gaussian fluctuation $\delta A_i(t)$. Input bias is defined as $(\delta A_1(t) - \delta A_2(t))/\max(A_0 + \delta A_1(t), A_0 + \delta A_2(t))$. The color code indicates the average position of population spikes above threshold within one step in unit of the tuning width (TW). Gray color means the average position is within the true position of either input ± 0.01 TW. True positions of inputs: $z_1 = \Delta z/2$, $z_2 = -\Delta z/2$. **(A)** $\Delta z = 0.33$ TW. **(B)** $\Delta z = 0.40$ TW. Parameters: same as **Figure 3**.

randomness is important for perceptions of motion transparency. For example, regularly spaced lines moving in opposite directions do not give the perception of transparent motion, whereas randomly spaced lines are able to do so (Qian et al., 1994). The input signals come from different locations of the visual field, and fluctuations arise when the perceived objects move from one location to another. Fluctuations may also arise when feedback signals from advanced stages of processing guide the system to shift its attention from one specific component to another.

Functionally, fluctuations facilitate the resolution of the directional inputs in the following two aspects. Spatially, it breaks the symmetry of the input profile. Temporally, it provides the time-dependent signals that induce the population spikes centered at the component that happens to be strengthened by fluctuations. This enables the system to recognize the temporally modulated inputs. On the other hand, for systems processing only time-averaged inputs, the height fluctuations vanish when averaged over time, so that the components cannot be detected.

4.4. THRESHOLDING

Even after temporal modulation, resolution based on the network response can still carry large errors. As shown in **Figure 3C**,

there are obviously two groups centering around the positions of the two components, but in between the two components, there is a region with moderate neuronal activities. If the network includes neuronal activities of all magnitudes, the errors in estimating the component positions will be large, especially when Δz is small. Indeed, **Figure 8D** shows that without imposing any thresholds on the neuronal activities, the network cannot resolve the two components until the separation exceeds the tuning width.

In order to solve this problem, we introduce a threshold on the maximum firing rates. We collect statistics of the peak positions of the firing rate profile when their height exceeds the threshold. The result is shown in **Figure 5**, indicating a significant improvement of resolution compared with **Figure 8D**. The effects of the threshold value on the resolution performance are shown in **Figure 12**. When the threshold is low, the components are not resolved even at a separation of 0.5 times the tuning width. On the other hand, when the threshold is too high, the statistics of peak positions becomes too sparse to be reliable. In an intermediate range of thresholds that is not too narrow, the resolution of the components can be achieved down to separations of 0.3–0.4 times the tuning width.
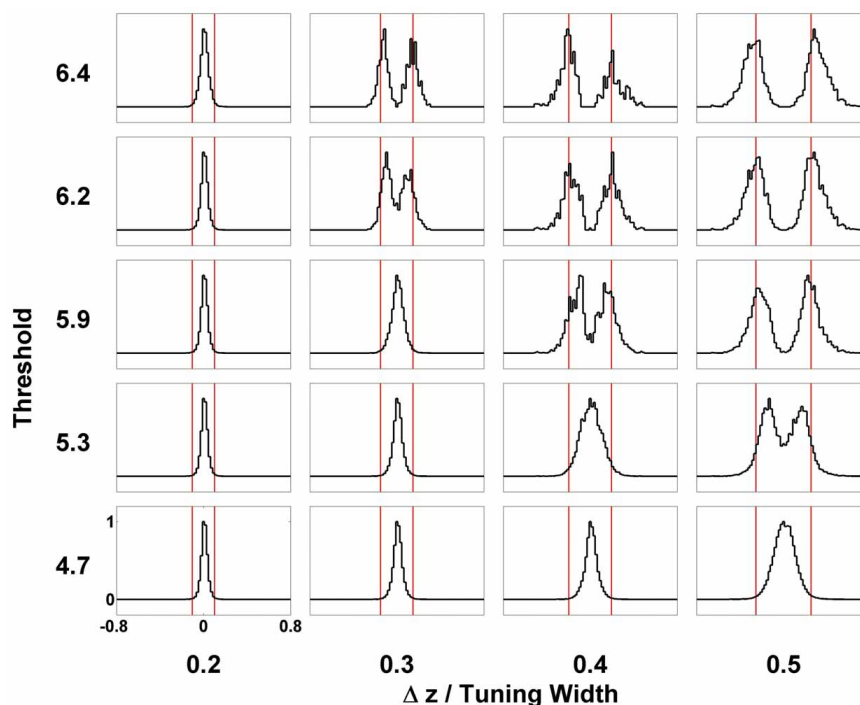
FIGURE 12 | Effect of thresholding on the statistics of peak positions. Distributions of peak positions higher than different thresholds (shown on the left) when two stimuli are separated by Δz (shown at the bottom) are plotted in patches. The scales of all the patches are the same, shown on the lower-left patch. Distributions are normalized to the maximum value. Red bars mark the positions of two stimuli. Other parameters: $\tilde{k} = 0.5$, $\tilde{\beta} = 0.1$, $a = 48\pi/180$, $\tilde{A} = 0.8$, $\sigma_A/A_0 = 0.2$, and $\tau_d = 50\tau_s$.

4.5. RECURRENT CONNECTIONS

Finally, we would like to stress the importance of recurrent connections in achieving resolution enhancement. With no recurrence, population spikes cannot be generated and the amplification of the difference between nearly overlapping inputs cannot be achieved. Let us consider a purely feedforward network, with weaker but spatially broader inhibition than excitation,

$$\tau_s \frac{du}{dt}(x, t) = -u(x, t) + \rho \int dx' \left[J_E \exp\left(-\frac{|x - x'|^2}{2a^2} \right) \right.$$
$$\left. - J_I \exp\left(-\frac{|x - x'|^2}{2b^2} \right) \right] p(x', t) I^{\text{ext}}(x', t) \tag{10}$$

$$\tau_d \frac{dp}{dt}(x, t) = -p(x, t) + 1 - \tau_d \beta p(x, t) I^{\text{ext}}(x, t) \tag{11}$$

$$r(x, t) = \Theta\left[u(x, t)\right] u(x, t), \tag{12}$$

where $J_E > J_I$ and $I^{\text{ext}}(x, t)$ is the same as that in recurrent network in Equation (6). Although in this feedforward network STD can still modulate the synaptic efficacy so that neuronal activities prefer the side region to the midpoint between two stimuli, temporal modulation, which is essential to population spikes, cannot be realized without feedback. As mentioned above, population spikes make it easier for the activity profile to switch off on one

side and grow up on the other. As shown in **Figure 13**, the resolution enhancement in the purely feedforward network is poor. In fact, the behavior is very similar to those in the non-spiking region even when the architecture is recurrent, as shown in **Figures 6A,E**.

5. DISCUSSION

In this paper, we have demonstrated how STD plays the role of generating population spikes that can carry information extra to spike rates. We have used the example of resolving transparent motion with two components in a continuous attractor neural network, and have shown that the temporal modulation of the firing rates enables the network to enhance the resolution of motion transparency, thereby providing a possible explanation to the longstanding mystery of resolving separations narrower than the tuning width of the neurons, and resulting in input-output relations that can have excellent agreement with experimental results (Treue et al., 2000). The role played by STD was further clarified by comparison with alternate scenarios under 4 general conditions.

First, the strength of STD should be sufficiently strong. Weaker STD may result in the network response being pinned by one of the two components, or slosher modes that span a range of positions effectively independent of the component separations. On the other hand, sufficiently strong STD can give rise to population spikes, endowing them the freedom to alternate between the two components. Equally important is the provision of temporal modulation by the population spikes, so that the firing patterns

FIGURE 13 | (A) Raster plot of firing rate r for purely feedforward network. $\Delta z = 0.5$ TW. White dashed line: positions of the two stimuli. **(B)** Contours of the distribution of peak positions higher than 0.45 as a function of preferred stimuli, x, and the separation between the two stimuli, Δz. White dashed line: positions of the two stimuli. Parameters: $J_I = 0.3J_E$, $b = 3a, \tau_d\beta/\rho J_E = 0.2$, $\rho J_E A = 0.8$, $\sigma_{\delta A_i}/A_0 = 0.3$, and $\tau_d = 50\tau_s$.

indeed contain information of the stimuli, even though the time-averaged firing rate can only resolve separations larger than the tuning width of neurons, as shown in **Figure 4** and found experimentally by Treue et al. (2000). The role played by temporally modulated signals in transparent motions can be tested in future experiments.

Second, the strength of the input should be sufficiently strong. Otherwise, no population spikes can be produced. Even for moderately strong input, the population spikes become moving ones, and fail to represent the stimulus positions.

Third, fluctuations in the input profiles are also important. They provide the temporally sensitive signals when the two components cannot be resolved in the time-averaged input. They correspond to the "unbalanced motion signals" in the detection of transparent motion with opposite moving directions (Qian et al., 1994).

Fourth, thresholds are needed to extract the information of the stimuli contained in the firing patterns, since they are able to truncate background activities that interfere the signals from the two components.

Our proposed model is not the first model or mechanism to explain the behavior of the discriminational task in transparent motion experiments. It was suggested that the curvature of the average neural activity may provide information of multiple stimuli, but the neural activity is wider than expected (Treue et al., 2000). Other proposals require more complex structures to achieve the task. For example, a population to encode uncertainty is needed to differentiate between multiplicity and uncertainty (Sahani and Dayan, 2003), and additional internal structures are needed to provide feedback information (Raudies et al., 2011). While admittedly involving additional structures and layers can augment the functionality of the brain, our work shows that it is possible to achieve with little additional structure the performance consistent with experiments in Treue et al. (2000) and Braddick et al. (2002). An interesting future direction is to consider whether firing rates multiplexed with temporal modulations can be an instrument to achieve the differentiation between multiplicity and uncertainty posed in Sahani and Dayan (2003).

The ability of STD to generate temporally modulated response is also applicable to other brain tasks, such as switching between percepts in competitive neural networks (Kilpatrick, 2012). Compared with other conventional neural network models processing time-averaged or static neuronal response profiles, the temporal component provides an extra dimension to encode acute stimuli, so that information processing performance can be significantly enhanced.

REFERENCES

Abbott, L. F., Varela, J. A., Sen., K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224. doi: 10.1126/science.275.5297.221

Amari, S. I. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87.

Ben-Yishai, B., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3844–3848. doi: 10.1073/pnas.92.9.3844

Braddick, O. J., Wishart, K. A., and Curran, W. (2002). Directional performance in motion transparency. *Vis. Res.* 42, 1237–1248. doi: 10.1016/S0042-6989(02)00018-4

Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience.* Cambridge; London: MIT Press.

Folias, S. E. (2011). Nonlinear analysis of breathing pulses in a synoptically coupled neural network. *SIAM J. Appl. Dyn. Syst.* 10, 744–787. doi: 10.1137/100815852

Fuhs, M. C., and Touretzky, D. S. (2006). A spin glass model of path integration in rat medial entorhinal cortex. *J. Neurosci.* 26, 4266–4276. doi: 10.1523/JNEUROSCI.4353-05.2006

Fung, C. C. A., Wong, K. Y. M., Wang, H., and Wu, S. (2012a). Dynamical synapses enhance neural information processing: gracefulness, accuracy and mobility. *Neural Compute.* 24, 1147–1185. doi: 10.1162/NECO_a_00269

Fung, C. C. A., Wong, K. Y. M., and Wu, S. (2012b). "Delay compensation with dynamical synapses," in *Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012* (New York, NY: Curran Associates, Inc.).

Fung, C. C. A., Wong, K. Y. M., and Wu, S. (2010). A moving

bump in a continuous manifold: a comprehensive study of the tracking dynamics of continuous attractor neural networks. *Neural Compute.* 22, 752–792. doi: 10.1162/neco.2009.07-08-824

Hao, J., Wang, X., Dan, Y., Poor, M., and Zhang, X. (2009). An arithmetic rule for spatial summation of excitatory and inhibitory inputs in pyramidal neurons. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21906–21911. doi: 10.1073/pnas.0912022106

Heeger, D. (1992). Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* 9, 181–197.

Holcman, D., and Tsodyks, M. (2006). The emergence of up and down states in cortical networks. *PLoS Comput. Biol.* 2, 174–181. doi: 10.1371/journal.pcbi.0020023

Itskov, V., Hansel, D., and Tsodyks, M. (2011). Short-term facilitation may stabilize parametric working memory trace. *Front. Comput. Neurosci.* 5, 1–19. doi: 10.3389/fncom.2011.00040

Kilpatrick, Z. P. (2012). Short term synaptic depression improves information transfer in perceptual multistability. Available online at: http://arxiv.org/abs/1212.0076

Levine, A., Herrmann, J. M., and Giesel, T. (2007). Dynamical synapses causing self-organized criticality in neural networks. *Nat. Phys.* 3, 857–860. doi: 10.1038/nphys758

Loebel, A., and Tsodyks, M. (2002). Computation by ensemble synchronization in recurrent networks with synaptic depression. *J. Compute. Neurosci.* 13, 111–124.

Lüscher, M. (1994). A portable high-quality random number generator for lattice field theory simulations. *Compute. Phys. Commun.*

79, 100–110. doi: 10.1016/0010-4655(94)90232-1

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382, 807–810. doi: 10.1038/382807a0

Mather, G., and Moulden, B. (1980). A simultaneous shift in apparent direction: further evidence for a "distribution-shift" model of direction coding. *Q. J. Exp. Psychol.* 32, 325–333. doi: 10.1080/14640748008401168

Maunsell, J. H., and Essen Van, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J. Neurophysiol.* 49, 1127–1147.

Pfister, J.-P., Dayan, P., and Lengyel, M. (2010). Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials. *Nat. Neurosci.* 13, 1271–1275. doi: 10.1038/nn.2640

Pouget, A., Dayan, P., and Zemel, R. (2000). Information processing with population codes. *Nat. Rev. Neurosci.* 1, 125–132. doi: 10.1038/35039062

Qian, N., Andersen, R. A., and Adelson, E. H. (1994). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. *J. Neurosci.* 14, 7357–7366.

Raudies, F., Mingolla, E., and Neumann, H. (2011). A model of motion transparency processing with local center-surround interactions and feedback. *Neural Comput.* 23, 2868–2914. doi: 10.1162/NECO_a_00193

Rust, N. C., Mante, V., Simoncelli, E. P., and Movshon, J. A. (2006). How MT cells analyze the motion

of visual patterns. *Nat. Neurosci.* 9, 1421–1431. doi: 10.1038/nn1786

Sahani, M., and Dayan, P. (2003). Doubly distributional population codes: simultaneous representation of uncertainty and multiplicity. *Neural Comput.* 15, 2255–2279. doi: 10.1162/089976603322362356

Samsonovich, A., and McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *J. Neurosci.* 17, 5900–5920.

Sato, Y., and Toyoizumi, T. (2007). Bayesian inference explains perception of unity and ventriloquism aftereffect: identification of common sources of audiovisual stimuli. *Neural Comput.* 19, 3335–3355. doi: 10.1162/neco.2007.19.12.3335

Stevens, C. F., and Wang, Y. (1995). Facilitation and depression at single central synapses. *Neuron* 14, 795–802. doi: 10.1016/0896-6273(95)90223-6

Treue, S., Hol, K., and Dauber, H.-J. (2000). Seeing multiple directions of motion—physiology and psychophysics. *Nat. Neurosci.* 3, 270–276. doi: 10.1038/72985

Tsodyks, M., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U.S.A.* 94, 719–723.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835. doi: 10.1162/089976698300017502

Wu, S., Hamaguchi, K., and Amari, S. (2008). Dynamics and computation of continuous attractors. *Neural Comput.* 20, 994–1025. doi: 10.1162/neco.2008.10-06-378

York, L. C., and van Rossum, M. C. W. (2009). Recurrent networks with short term synaptic depression. *J. Comput. Neurosci.* 27, 607–620. doi: 10.1007/s10827-009-0172-4

Zemel, R. S., and Dayan, P. (1999). *Distributional Population Codes and Multiple Motion Models.* Advances in NIPS, Vol. 11. Cambridge, MA: MIT Press, 174–180.

Zemel, R. S., and Dayan, P. (2000). *Distributional Population Codes and Multiple Motion Models.* Advances in Neural Information Processing Systems, Vol. 11. Cambridge, MA: MIT Press.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J. Neurosci.* 16, 2112–2126.

APPENDIX

FIGURE A1 | Contours of the distribution of peak positions higher than (A) 6.2 and (B) 5.5 as a function of preferred stimuli, *x*, and the separation between the two stimuli, Δ*z*. White dashed line: positions of two stimuli. Parameters: same as **Figure 3**, except $a_l = 0.95a$ for **(A)** and $a_l = 1.05a$ for **(B)**.
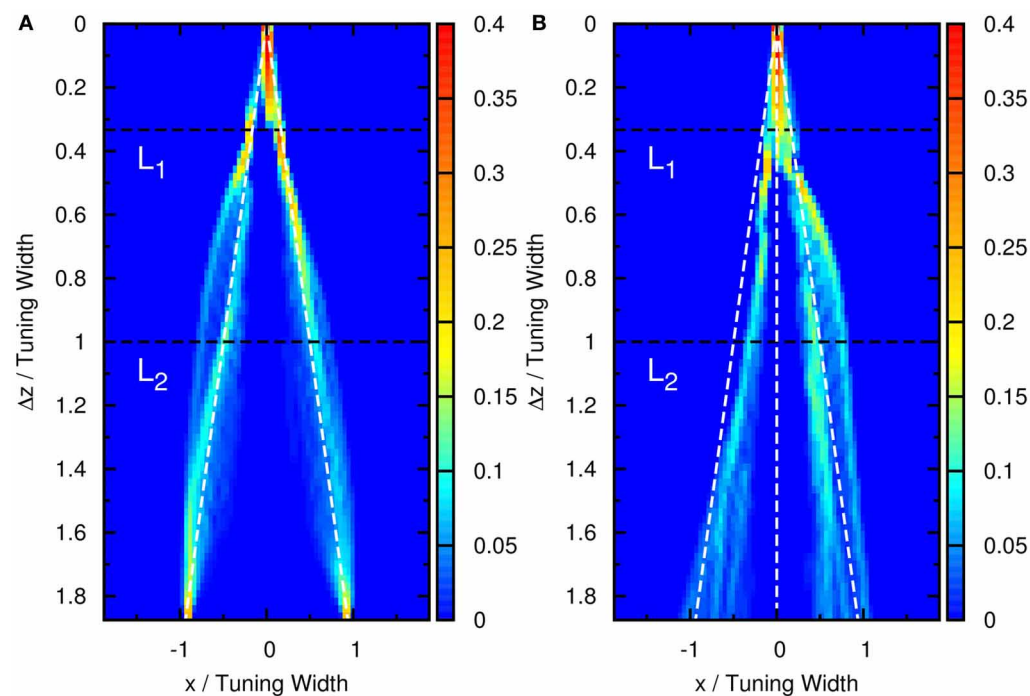
FIGURE A2 | The population spike occurrence counted within 500τ_s, which is comparable to the timescale in typical experiments. (A) Situations that there are two stimuli. **(B)** Situations with three stimuli. Other parameters: **(A)** same as **Figure 4**, **(B)** same as **Figure 7**.

Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits

Rui P. Costa[1]*, P. Jesper Sjöström[2] and Mark C. W. van Rossum[3]

[1] Neuroinformatics Doctoral Training Centre, Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, Edinburgh, UK
[2] Department of Neurology and Neurosurgery, The Research Institute of the McGill University Health Centre, McGill University, Montreal, QC, Canada
[3] Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, Edinburgh, UK

Short-term synaptic plasticity is highly diverse across brain area, cortical layer, cell type, and developmental stage. Since short-term plasticity (STP) strongly shapes neural dynamics, this diversity suggests a specific and essential role in neural information processing. Therefore, a correct characterization of short-term synaptic plasticity is an important step towards understanding and modeling neural systems. Phenomenological models have been developed, but they are usually fitted to experimental data using least-mean-square methods. We demonstrate that for typical synaptic dynamics such fitting may give unreliable results. As a solution, we introduce a Bayesian formulation, which yields the posterior distribution over the model parameters given the data. First, we show that common STP protocols yield broad distributions over some model parameters. Using our result we propose a experimental protocol to more accurately determine synaptic dynamics parameters. Next, we infer the model parameters using experimental data from three different neocortical excitatory connection types. This reveals connection-specific distributions, which we use to classify synaptic dynamics. Our approach to demarcate connection-specific synaptic dynamics is an important improvement on the state of the art and reveals novel features from existing data.

Keywords: short-term synaptic plasticity, probabilistic inference, neocortical circuits, experimental design, parameter estimation

1. INTRODUCTION

Synaptic plasticity is thought to underlie learning and information processing in the brain. Short-term plasticity (STP) refers to transient changes in synaptic efficacy, in the range of tens of milliseconds to several seconds or even minutes (Zucker and Regehr, 2002). It is highly heterogeneous and is correlated with developmental stage (Reyes and Sakmann, 1999), cortical layer (Reyes and Sakmann, 1999), brain area (Wang et al., 2006; Cheetham and Fox, 2010), and postsynaptic cell-type (Markram et al., 1998; Reyes et al., 1998; Scanziani et al., 1998; Tóth and McBain, 2000; Rozov et al., 2001; Sun and Dobrunz, 2006). For instance, short-term depression predominates in the juvenile brain, whereas more mature circuits have a preponderance for short-term facilitation (Pouzat and Hestrin, 1997; Reyes and Sakmann, 1999). Similarly, synapses from neocortical pyramidal cells (PCs) impinging on other PCs are depressing, whereas those onto specific interneurons can be strongly facilitating (Markram et al., 1998; Reyes et al., 1998).

STP has been proposed to shape information processing in neural networks in multiple ways (Abbott and Regehr, 2004; Fung et al., 2012), to enable cortical gain control (Abbott et al., 1997), pattern discrimination (Carvalho and Buonomano, 2011), input filtering (Markram et al., 1998), adaptation (van Rossum et al., 2008), spike burst detection (Maass and Zador, 1999), synchronization (Tsodyks et al., 2000), and to maintain the balance of excitation and inhibition in local circuits (Galarreta and Hestrin, 1998).

To model short-term depression, Tsodyks and Markram (1997) introduced a phenomenological model based on vesicle depletion, here referred to as the Tsodyks–Markram (TM) model. This model was later extended to include short-term facilitation (Markram et al., 1998; Tsodyks et al., 1998). Although several other STP models have been developed (Abbott et al., 1997; Varela et al., 1997; Dittman et al., 2000; Loebel et al., 2009; Pan and Zucker, 2009), the TM model has become particularly popular, probably because of its combination of appealing simplicity and biophysically relevant parameters (Markram et al., 1998; Richardson et al., 2005; Le Bé and Markram, 2006; Wang et al., 2006; Rinaldi et al., 2008; Ramaswamy et al., 2012; Testa-Silva et al., 2012; Romani et al., 2013).

Typically, STP models are numerically fitted to electrophysiological data by least-mean-square algorithms, which yield the parameter values that minimize the error between data and model. However, such fitting algorithms can get stuck in local optima and may provide little information about the certainty of the parameter values. As shown below, such fits may produce inaccurate results and may lead to unreliable clustering. Bayesian inference is a natural alternative, because it yields a *distribution* of parameter values rather than a single outcome. Bayesian inference has recently been applied to neurophysiological data analysis. McGuinness et al. (2010) used this to estimate large and small action potential-evoked Ca^{2+} events, while Bhumbra and Beato (2013) used a Bayesian framework of quantal analysis to estimate synaptic parameters, which required far fewer trials compared to

traditional methods. Here, we introduce a Bayesian approach to obtain the posterior distribution of TM model parameters. This enabled us to take into account the uncertainty inherent to experimental data, which provided a more complete description of STP data.

Our approach has several advantages. First, it allowed us to infer the distribution of synaptic parameters for individual connections and propose a better protocol to extract these parameters. Second, we found that parameter distributions extracted from cortical data are specific to different connection types. Third, we showed that we can automatically cluster the parameters of synaptic dynamics to at least partially classify postsynaptic cell types. We also performed model selection to determine which variant of the TM model best captures the synaptic dynamics of the connection type at hand.

2. MATERIALS AND METHODS

2.1. SHORT-TERM PLASTICITY PHENOMENOLOGICAL MODEL

The extended TM model (eTM) is a phenomenological model of short-term plasticity defined by the following ODEs (Markram et al., 1998; Tsodyks et al., 1998)

$$\frac{dR(t)}{dt} = \frac{1 - R(t)}{D} - u(t^-)R(t^-)\delta(t - t_{AP}) \tag{1}$$

$$\frac{du(t)}{dt} = \frac{U - u(t)}{F} + f[1 - u(t^-)]\delta(t - t_{AP}) \tag{2}$$

The first equation models the vesicle depletion process, where the number of vesicles $R(t)$ is decreased with $u(t)R(t)$ after release due to a presynaptic spike at time t_{AP}, modeled by a Dirac delta distribution $\delta(t)$. Between spikes $R(t)$ recovers to 1 with a depression timeconstant D. The second equation models the dynamics of the release probability $u(t)$ which increases with $f[1 - u(t)]$ after every presynaptic spike, decaying back to baseline release probability U with a facilitation timeconstant F. The notation t^- indicates that these functions should be evaluated in the limit approaching the time of the action potential from below (as would be natural in forward Euler integration).

By varying the four parameters $\vec{\theta} = \{D, F, U, f\}$ one can obtain depressing, combined facilitating-depressing and facilitating synaptic dynamics. We note that for some data a three parameter model [setting $f = U$, denoted the TM with facilitation model] or even a two parameter depression model with only Equation (1) [setting $u(t) = U$, denoted the TM model] is sufficient. This, however, is not generally the case, as shown below.

To speed up the numerical implementation we integrated the above equations between spikes n and $n + 1$, a time Δt_n apart, yielding

$$R_{n+1} = 1 - [1 - R_n(1 - u_n)] \exp\left(-\frac{\Delta t_n}{D}\right) \tag{3}$$

$$u_{n+1} = U + [u_n + f(1 - u_n) - U] \exp\left(-\frac{\Delta t_n}{F}\right) \tag{4}$$

As we assumed that at time zero the synapse has not been recently activated, we set $R_0 = 1$ and $u_0 = U$.

The postsynaptic potential PSP_n is given by

$$PSP_n = AR_n u_n \tag{5}$$

where A is an amplitude factor that includes the number of release sites, the properties and number of postsynaptic receptors, and cable filtering.

The steady-state values R_∞ and u_∞ in response to prolonged periodic stimulation with rate ρ are

$$R_\infty(\rho) = \frac{1 - \exp\left(-\frac{1}{\rho D}\right)}{1 - [1 - u_\infty(\rho)] \exp\left(-\frac{1}{\rho D}\right)} \tag{6}$$

$$u_\infty(\rho) = \frac{U + (f - U) \exp\left(-\frac{1}{\rho F}\right)}{1 - (1 - f) \exp\left(-\frac{1}{\rho F}\right)} \tag{7}$$

2.2. SIMULATED DATA

For the simulated data we used five sets of STP parameters, ranging from depression to facilitation, see **Table 1**.

As the commonly used paired-pulse ratio, $PPR = PSP_2/PSP_1$, only takes the first two pulses into account, we introduce the Every Pulse Ratio (EPR) as a more comprehensive measure of STP dynamics. It is defined as

$$EPR = \frac{1}{(n-1)} \sum_{i=1}^{n-1} \frac{PSP_{i+1}}{PSP_i} \tag{8}$$

This index measures the average amplitude change from the i to the $i + 1$ response normalized to the i response in the train. EPR is used in **Table 1** and elsewhere to quantify the average degree of depression (EPR < 1) or facilitation (EPR > 1). Using these parameters we calculated the synaptic responses with Equations (3, 4) to a spike train of five pulses at 30 Hz (**Figures 2**, **4**).

2.3. BAYESIAN FORMULATION

The posterior distribution of the synaptic parameters follows from Bayes' theorem as $P(\vec{\theta}|\vec{d}) \propto P(\vec{\theta})P(\vec{d}|\vec{\theta})$, where $\vec{d}$ is a vector of mean postsynaptic potential peaks extracted from simulated or experimental data and $\vec{\theta}$ is a vector encompassing the model parameters. Many factors contribute to variability in the

Table 1 | The five parameter sets used for simulated data.

Synaptic dynamics regime	D(s)	F(s)	U	f	EPR
Strong depression	1.70	0.02	0.7	0.05	0.45
Depression	0.50	0.05	0.5	0.05	0.64
Facilitation-depression	0.20	0.20	0.25	0.3	0.94
Facilitation	0.05	0.50	0.15	0.15	1.26
Strong facilitation	0.02	1.70	0.1	0.11	1.43

*EPR was calculated by simulating the eTM model with 5 pulses at 30 Hz as shown in **Figure 2**.*

measured EPSPs, including stochastic vesicle release and experimental noise. A typical noise model of synaptic transmission is a binomial distribution (Zucker and Regehr, 2002). However, we found that our data is well described by a Gaussian noise model (see below). Therefore, we write the likelihood of the data as

$$P(\vec{d}|\vec{\theta}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\left(d_i - \mathrm{STP}\left(\mathrm{PSP}_i|\vec{\theta}\right)\right)^2 / 2\sigma_i^2\right] \quad (9)$$

where $\mathrm{STP}(\mathrm{PSP}_i|\vec{\theta})$ is the voltage response from the eTM model for $i = 1 \ldots N$ runs over the data points in the pulse train. We set the noise σ_i independently for each pulse. For the data we extracted the CV for each pulse, while for the simulated data a fixed coefficient of variation (CV = 0.5) was assumed, based on **Figure 1**. Note that we did not include a model of stochastic vesicle release. This would be a possible extension of our model. A stochastic release model also leads to correlations between subsequent events, and Equations (4, 3) would thus have to be extended to their history-dependent variances, which would complicate our model. We did confirm that parameters from a simulated stochastic release model, were inferred correctly using the above noise model, although the posterior distributions were somewhat widened.

The priors were modeled as independent non-informative flat distributions over the model parameters

$$P(\vec{\theta}) = \begin{cases} P(D) = P(F) = \mathrm{Uniform}[0, 2] \\ P(U) = P(f) = \mathrm{Uniform}[0, 1] \end{cases} \quad (10)$$

which limits the posterior distribution within reasonable values.

Bhumbra and Beato (2013) sampled their bidimensional posterior probability using a brute-force grid search. For higher dimensions this is computationally expensive. We therefore inferred the posterior distribution by sampling using the Slice Sampling Markov Chain Monte Carlo (MCMC) method (Neal, 2003). The width parameter w was set equal to the upper limit of the flat prior distributions (i.e., $\vec{w} = \{2, 2, 1, 1\}$) and each parameter is sampled sequentially in the four orthogonal directions. We discarded the first 2500 samples as burn-in and use the last 7500. For the numerical implementation we use the log-likelihood $\log P(\vec{d}|\vec{\theta})$. The convergence of the Markov chain to the equilibrium distribution was assessed through the Gelman–Rubin statistical method (Brooks and Gelman, 1998). However, this diagnostic of convergence can indicate lack of convergence, but does not confirm it. Therefore, in order to ensure convergence, we used multiple chains ($n = 3$) starting at different initial conditions to ensure that the outcome was independent on the initial condition (Gelman and Shirley, 2011). The maximum *a posteriori* (MAP) estimator of the synaptic parameters is given by

$$\vec{\theta}_{\mathrm{MAP}} = \mathrm{argmax}_{\vec{\theta}} P(\vec{\theta}) P(\vec{d}|\vec{\theta}) \quad (11)$$

The MAP estimate was obtained by keeping the most likely sample from multiple MCMC chains. In addition we also ran an optimizer to find the most precise MAP using the distribution

peak as a starting point. As both approaches gave equally good fits for the sake of simplicity we decided to use the former.

We compared our estimation method with a standard stochastic optimization method, simulation annealing (SA). The SA method minimizes the RMSE

$$\vec{\theta}_{SA} = \mathrm{argmin}_{\vec{\theta}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left[d_i - \mathrm{STP}\left(\mathrm{PSP}_i|\vec{\theta}\right)\right]^2} \quad (12)$$

while trying to avoid getting stuck in local minima. We ran the SA algorithm 200 times and selected the estimate with lowest RMSE. Using an objective function scaled by the variance gave similar results when compared to the non-scaled version; thus for the sake of comparison with previous literature, we used the non-scaled version. To compare the goodness of fit of both MAP and SA solutions with the data, we used the coefficient of determination R^2.

As the amplitude A is not relevant for the synaptic dynamics, we set $A = A^{\mathrm{MAP}}$,

$$A^{\mathrm{MAP}} = \frac{\sum_{i=1}^{N} d_i m_i / \sigma_i^2}{\sum_{i=1}^{N} m_i^2 / \sigma_i^2} \quad (13)$$

where $m_i = \mathrm{STP}(\mathrm{PSP}_i|\vec{\theta})$. We used this value to normalize the data. Its value does not affect the dynamics estimation, because A only scales the responses.

To estimate the posterior probability distributions, we used a kernel density estimation method (Ihler and Mandel, 2007). Unless otherwise stated, the code was implemented in Matlab (inference code is available online[1]).

2.3.1. Quantifying inference performance

To quantify which protocol allows for the most precise recovery of the true parameters of simulated STP data (**Figure 3A**), we computed the sample estimation error over $N = 22{,}500$ MCMC samples $\vec{\theta}$ to the true parameters $\vec{\theta}^*$, as $E = \langle \sum_{i=1}^{4} [(\theta_i - \theta_i^*)/\theta_i^*]^2 \rangle$, where the average is over all the runs and all five parameter sets (**Table 1**). To achieve similar weighting, the parameters were normalized to the true parameters. Alternatively, we normalized the estimated parameters on the upper limit of their priors, or we omitted normalization altogether. This yielded similar results. Note that in probabilistic spirit, this error also quantifies the spread in the distribution. A smaller E gives more peaked distributions, which correspond to tighter parameter estimates. Note that, although similar, this error measure does not follow the standard bootstrap approach.

2.3.2. Model selection

For model selection, we used the Akaike Information Criterion (AIC), which is a information-theoretic measure of the goodness of fit of a given statistical model. It is defined as $AIC = 2k - \log P(\vec{\theta}_{\mathrm{MAP}}|\vec{d})$, where k is the number of estimable parameters in the model and $\log P(\vec{\theta}_{\mathrm{MAP}}|\vec{d})$ the log-posterior of the MAP estimate on the normalized data. The AIC evaluates models

---

[1]https://senselab.med.yale.edu/modeldb/ShowModel.asp?model=149914

according to their parsimonious description of the data, and is particularly suitable for probabilistic inference. We used the evidence ratio, which is a relative ranking of the Akaike weights, to find the least complex model that best describes the data (Turkheimer et al., 2003; Nakagawa and Hauber, 2011).

2.4. ELECTROPHYSIOLOGY

Quadruple whole-cell recordings and extracellular stimulation were performed in acute visual cortex slices of young mice (P12–P20) as previously described (Buchanan et al., 2012). The stimulating electrode was positioned in layer 5 (L5). L5 Pyramidal cells (PCs) were targeted based on their characteristic pyramidal soma and thick apical dendrite. Basket cells (BCs) were targeted in transgenic mice genetically tagged for parvalbumin, while Martinotti cells (MCs) were targeted in mice genetically labeled for somatostatin (Markram et al., 2004; Buchanan et al., 2012). Cell identities were verified by cell morphology and rheobase firing pattern. Five spikes were elicited at 30 Hz using 5 ms long current injections (0.7–1.4 nA) every 18 s in all neurons throughout the experiment. Excitatory postsynaptic potentials (EPSPs) were averaged from 20–40 sweeps.

For each connection, a histogram was built from the EPSP amplitudes extracted with 1–2-ms window fixed approximately on the peak depolarization. EPSP distributions were fit with a Gaussian (Equation 9). Recordings with mean EPSPs smaller than

0.015 mV were discarded. Electrophysiological data analysis was carried out in Igor Pro (WaveMetrics Inc., Lake Oswego, OR).

Figure 1 shows typical EPSP distributions for each of the three neocortical excitatory connection types that we studied, PC–PC, PC–BC, and PC–MC. We tested whether the Gaussian noise model was a valid description of the data using the Kolmogorov–Smirnov (KS) normality test, and we found that the null hypothesis that samples were drawn from a normal distribution could not be rejected for 160 out of 170 EPSP distributions, with no connection-specific bias. This suggests that EPSPs were typically normally distributed, consistent with previously published results [e.g., Figure 5B in Markram et al. (1997)]. Due to noise, apparently negative EPSPs (**Figure 1**) were occasionally recorded. These are consistent with our Gaussian noise model and require no special treatment.

2.5. CLUSTERING AND CLASSIFICATION

Distributional clustering was introduced by Pereira et al. (1993). Here we applied a similar information-theoretic approach to cluster $P(\vec{\theta}|\vec{d})$. Instead of a "soft" clustering approach we used "hard" clustering, due to its simplicity, computation speed and comparison with standard clustering techniques. We used an agglomerative method [unweighted average distance method, Sokal (1958)] and an f-divergence metric. F-divergence metrics constitute a family of functions that measure the difference between two
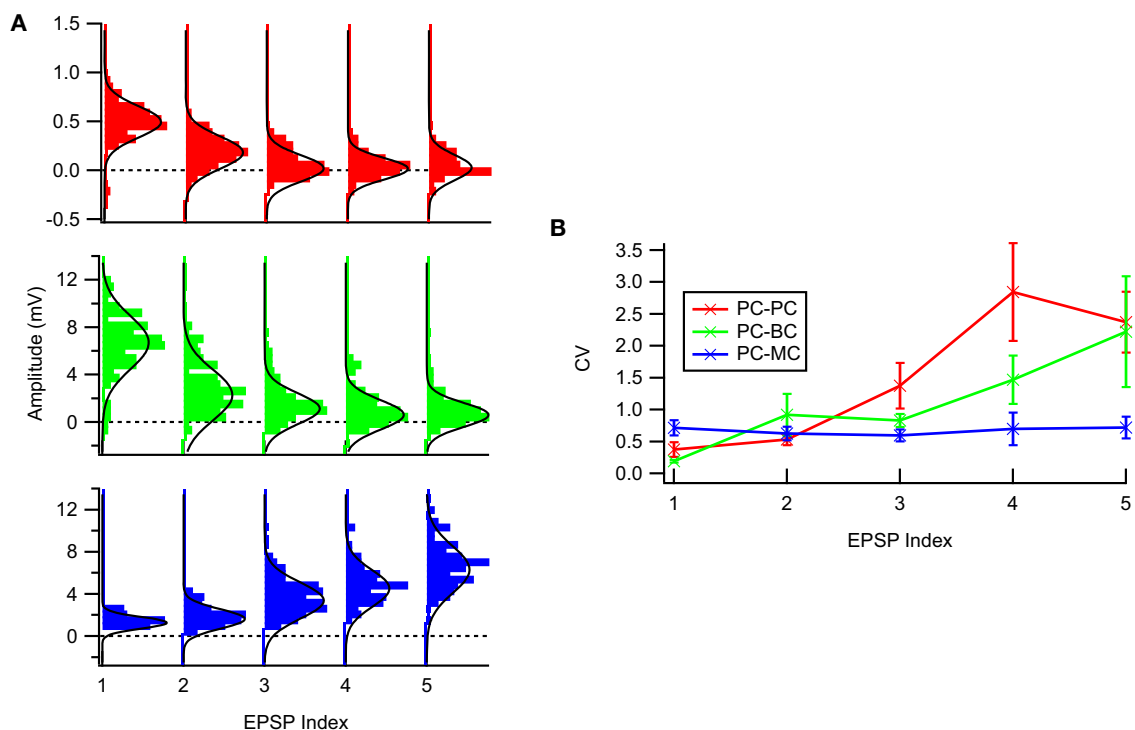
FIGURE 1 | The experimental STP data was well described by a Gaussian noise model. (A) Sample EPSP distributions for the three connection types: PC–PC (top, red), PC–BC (middle, green), and PC–MC (bottom, blue) with respective Gaussian fits (solid black line)—94% of the EPSP distributions were not statistically significant different from a

Gaussian distribution (see main text for more details). **(B)** Coefficient of variation analysis. While for facilitating synapses (PC–MC) it was more or less constant, for depressing synapses (PC–PC and PC–BC) we observed an approximately linear increase with EPSP amplitude. Error bars represent standard error of the mean.

probability distributions. Consider two discrete probability distributions P and Q both discretized into N bins. To compare any given pair of distributions we used two f-divergence metrics: (i) the symmetrized Kullback–Leibler divergence

$$\text{KL}_s(P, Q) = \frac{\text{KL}(P, Q) + \text{KL}(Q, P)}{2} \qquad (14)$$

with

$$\text{KL}(P, Q) = \sum_{i=1}^{N} P_i(\vec{\theta}|\vec{d}) \log \frac{P_i(\vec{\theta}|\vec{d})}{Q_i(\vec{\theta}|\vec{d})} \qquad (15)$$

and the (ii) Hellinger distance

$$\text{HL}(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{N} \left(\sqrt{P_i(\vec{\theta}|\vec{d})} - \sqrt{Q_i(\vec{\theta}|\vec{d})} \right)^2} \qquad (16)$$

Due to the high dimensionality of our problem, we approximated these two measures first marginalizing $P(\vec{\theta}|\vec{d})$ over the $d = 4$ dimensions and then computing the sKL and HL over the d marginal probabilities. We compared our posterior-based clustering with clustering based on the SA estimates. Here, we used the Euclidian distance on the z-scored parameters found with SA.

To estimate the number of clusters we used the Pseudo-F statistic (Caliński and Harabasz, 1974). The Pseudo-F statistic captures the tightness of clusters as the ratio of the mean sum of squares between clusters to the mean sum of squares within cluster

$$\text{Pseudo-F} = \frac{(T - P_G)/(G - 1)}{P_G/(n - G)} \qquad (17)$$

where $T = \sum_{i=1}^{n} (P_i - \overline{P})^2$ is the total sum of squares, $P_G = \sum_{i=1}^{G} \sum_{j=1}^{n_i} (P_i^j - \overline{P_i})^2$ is the within-cluster sum of squares, G is the number of clusters, and n the total number of items. A larger Pseudo-F usually indicates a better clustering solution. The Pseudo-F statistic has been found to give best performance in simulation studies when compared with 30 other methods (Milligan and Cooper, 1985).

To evaluate the clustering quality, we computed the dendrogram purity as described by Heller and Ghahramani (2005), where we considered two classes according to EPR: class 1 for EPR ≤ 1 and class 2 for EPR > 1. This threshold allows us to separate mostly depressing from mostly facilitating synaptic dynamics.

Finally, we also performed classification using the Naive Bayes Classifier: $P(C|\vec{\theta}) \propto P(C)P(\vec{\theta}|C)$, where $P(C)$ is the prior over the different synapse types C and $P(\vec{\theta}|C)$ the likelihood for a given class. Although information about connectivity rates could in principle be incorporated in the prior, we used a uniform prior over the classes. Our likelihood is given by the MCMC inference over the model parameters for a given training dataset d_C and synapse type C, i.e., $P(\vec{\theta}|C) = P(\vec{\theta}|d_C)$. As the Naive Bayes Classifier assumes independence between the different classes, we have one independent model per class with the maximum a posteriori decision rule $\text{argmax}_{(c \in C)} P(C = c) P(\vec{\theta}_{\text{MAP}}|C = c)$. We

estimated the performance of our classifier with K-cross validation ($K = 7$, i.e., ~80% for PC–PC ($n = 9$) and PC–MC ($n = 9$), and ~60% for PC–BC ($n = 12$)), where we sampled over K data points (i.e., connections) for each synapse-type to obtain our likelihood model and then test the classifier with the remaining data points. This process was repeated until all possible different K partitions of the data have been used. Accuracy is defined as the percentage of correct classifications for a given connection type.

3. RESULTS

3.1. PARAMETER INFERENCE CERTAINTY IS SYNAPTIC DYNAMICS DEPENDENT

We first checked our method in extracting STP parameters from simulated data with a standard stimulus train of five spikes at 30 Hz (see Materials and Methods). We simulated data with pre-defined parameter sets ranging from strong depression to strong facilitation. This was achieved by decreasing the baseline release probability U and the depression timeconstant D, while increasing the facilitation rate f and the facilitation timeconstant F (see Materials and Methods, **Table 1**). The resulting dynamics are shown in **Figure 2A**.

Figure 2B shows the inferred parameter distributions for the various parameter settings. As the full posterior distribution is four dimensional, we plotted the marginals only. The inferred parameter distributions showed varying behavior: The distributions for U were well-tuned to values close to the true parameter values. For the D parameter the shifts in the distributions followed the changes in the true parameter, becoming broader for depressing dynamics. Both F and f were not narrowly tuned to the true parameter. Although f was tuned to small values for facilitating synapses, its distribution became broader for depressing synapses. The F parameter was not particularly tuned to any value, being close to an uniform distribution for both depressing and facilitating synapses. We explored the possibility that the broadness of F depended on the prior boundary by extending it to 5 s and 10 s. However, the distribution remained uniform and merely grew wider, suggesting that the broad distribution was not caused by an improper choice of prior. In summary, the inference procedure shows that—depending on the dynamics—the inferred parameter distributions can be narrow or broad and that some parameters are much more tightly constrained than others.

3.2. IMPROVING EXPERIMENTAL PROTOCOL FOR PARAMETER INFERENCE

The fact that some of the inferred parameter distributions were broad suggested that the five pulse protocol did not yield enough information to reliably infer the true parameters. Therefore, we used our probabilistic formulation to find an experimental protocol that improves the inference quality (**Figure 3**). To this end, we compared the sample estimation error on the estimates (see Materials and Methods) for different spike trains: (1) a periodic train at 30 Hz, (2) a periodic train with recovery pulses, and (3) a Poisson train of 30 Hz (**Figure 3A**). We also varied the number of spikes in the train.

The widely used paired-pulse protocol to probe synaptic dynamics gave poor estimates even when coupled with nine

FIGURE 2 | Bayesian inference of short-term plasticity parameters using simulated data. (A) Simulated PSPs (filled circles in response to five pulses at 30 Hz) for five different synaptic parameter settings ranging from strong depression (yellow) to strong facilitation (dark red). The MAP solution of the inferred distribution is shown with diamonds. **(B)** Posterior marginalized distributions of the model parameters for the data in **(A)**. The true parameters are shown as filled circles and the MAP solutions as diamonds.

recovery pulses spaced exponentially across 4 s. Using five pulses in the spike train improved the performance only moderately. Some studies have inferred the TM model parameters with eight spikes and a single recovery pulse after 500 ms (e.g., Wang et al., 2006). This did not improve the recovery error when compared to a periodic spike train alone. A Poisson spike train, however, surpassed other protocols using only 20 spikes. Therefore, we propose a Poisson spike train with 20…100 spikes as a better protocol to obtain accurate estimates of the model parameters. However, also a spike train with eight periodic pulses and nine recovery pulses offers a good compromise, yielding a low recovery error in a reasonably short duration ($\approx$4.23 s). The distributions for these two protocols were more narrowly tuned to the true parameters (**Figures 3B,C**) compared to a periodic spike train without a full recovery phase (**Figure 3B**). Contrary to our intuition, the distributions for D were more narrowly tuned for facilitation (darker colors) than for depression (lighter colors). Although for the sake of simplicity, we do not show the results for a short periodic train followed by a Poisson train, such an approach would combine the ability to compute standard STP measures and recover information across frequencies. The reason for the poor performance of periodic trains even with many pulses is that the synapse quickly reaches steady-state, given by Equations (6, 7). Hence additional pulses do not increase information and the estimation error quickly reaches a plateau. In contrast, a random Poisson train allows the inference process to converge to the true parameter distributions in the limit of large spike trains.

Note, that both in **Figure 2B** and **Figures 3B,C**, the MAP solution is not always at the peak of the marginal distributions. The reason is that when there are dependencies in the parameters, the peak in the full distribution $P(\vec{\theta})$ does not need to coincide with the peaks of the marginals. Indeed, when we compared the log-posterior of the MAP estimate to the log-posterior of

the estimate given by the maximum of each marginal probability alone, the MAP approach yielded a much better estimate: $\log P(\vec{\theta}_{\text{MAP}}|\vec{d}) = -0.0038$, compared to the maximum of the marginal probabilities, $\log P(\vec{\theta}_{\text{marginals}}|\vec{d}) = -0.6588$.

3.3. PROBABILISTIC INFERENCE OF NEOCORTICAL DATA REVEALS CONNECTION-SPECIFIC DISTRIBUTIONS

Next, we performed Bayesian inference of the eTM parameters on experimental data from visual cortex L5. These data was recorded earlier using a standard five-pulse protocol, instead of the improved protocols suggested above. This means that the parameters may not be optimally constrained, but the overall findings should still hold. We inferred the posterior distributions of the parameters U, D, F, and f from PC–PC, PC–MC, and PC–BC connections (**Figure 4A**).

When comparing the Bayesian model inference of these three different synapse types (**Figure 4B**), the most salient difference was observed in the U parameter, i.e., the baseline probability of release. PC–MC connections had a small U, D and f. PC–PC connections had a medium U, medium to high D, a close to uniform F and a broad f with a preference for smaller values. PC–BC connections were similar to PC–PC connections, apart from a larger U (PC–BC: 0.72 ± 0.04, $n = 12$; PC–PC: 0.53 ± 0.05, $n = 9$; $p < 0.01$, Mann–Whitney test based on the MAP estimates). This higher value of U indicates that PC–BC synapses are generally more strongly depressing than PC–PC synapses. However, the EPRs for these two connection types were indistinguishable (PC–BC: 0.63 ± 0.04, $n = 12$; PC–PC: 0.69 ± 0.03, $n = 9$; $p = 0.21$, Mann–Whitney test), suggesting that the model inference is more sensitive than the EPR measure, and is therefore better suited for picking up connection-specific differences in STP.

We next used our Bayesian approach for synapse classification. We first clustered the data of the various connections
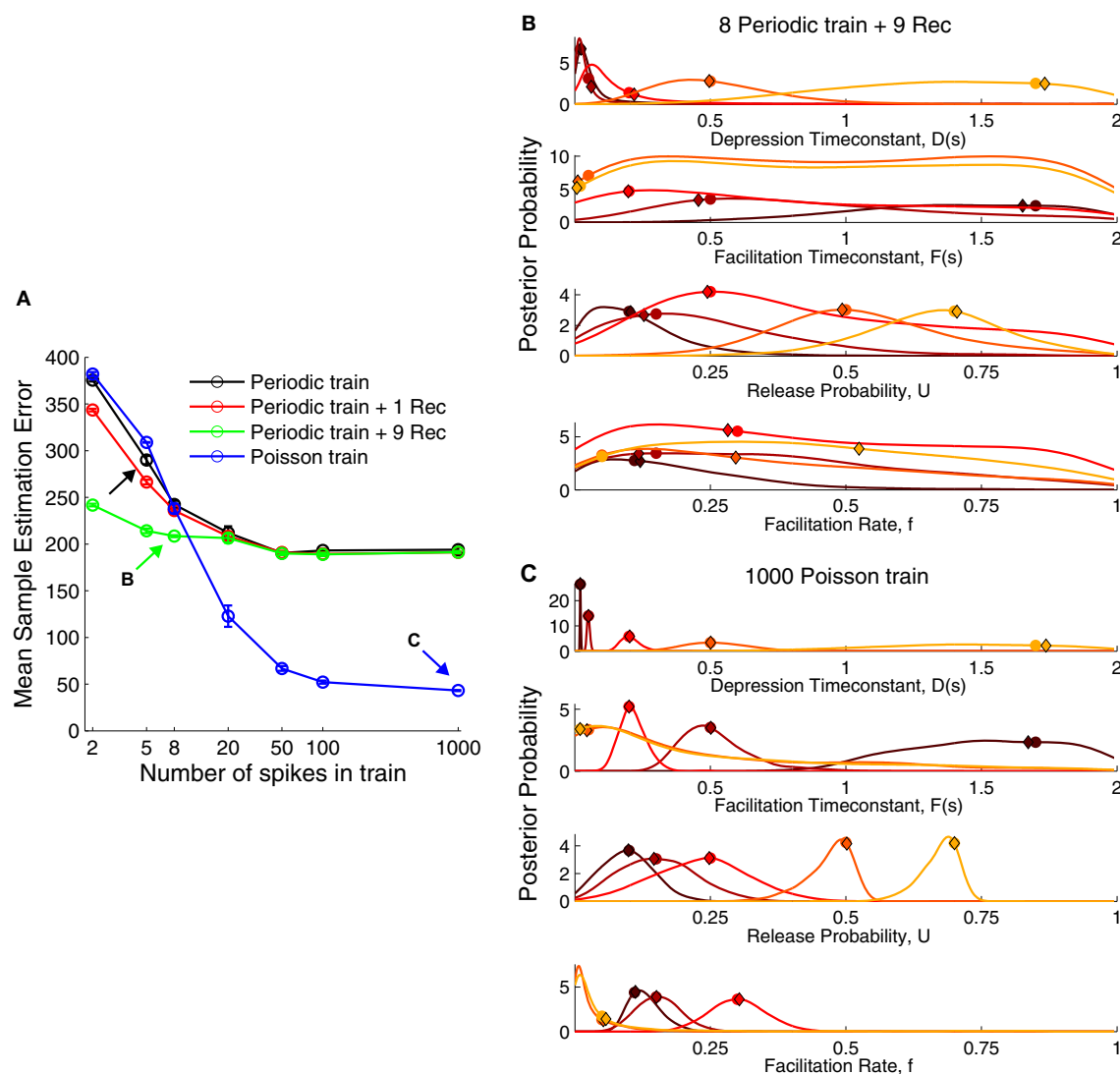
FIGURE 3 | The performance of various stimulation protocols to infer short-term plasticity parameters. (A) Comparison of mean sample estimation error using different stimulation trains. Black arrow corresponds to the protocol used in **Figure 2**. A periodic spike train at 30 Hz with eight pulses and nine recovery pulses [green arrow, (B)] already provided a better estimate of the STP parameters. However, a Poisson train provided a much smaller error when using more than 20 spikes with a close to zero error for 1000 spikes [blue arrow, (C)]. **(B)** Posterior distribution for a periodic train with nine recovery pulses (cf. **Figure 2B**). **(C)** Posterior distribution for a Poisson train with 1000 pulses. The true parameters are shown as filled circles and the MAP solutions as diamonds. For visualization the marginal probabilities were scaled by their standard deviation.

based on the model parameters found by SA, **Figure 5A**. We next clustered based on the marginalized posterior distributions, **Figure 5B** using the Hellinger distance (see Materials and Methods). Clustering analysis showed that the Bayesian approach improved the dendrogram purity (**Figure 5C**), as it split the data into two distinct clusters as assessed by the Pseudo-F statistic (**Figures 5B,D**).

With SA-based clustering, the Pseudo-F statistic suggested six clusters (**Figure 5D**) with a lower dendrogram purity (**Figure 5C**, 0.89 purity level), which indicates that these six clusters are spurious. Furthermore, with the Bayesian approach, the clusters map better to the EPR measure (**Figure 5B**, inset bottom), indicating that our approach captures the synaptic properties better than the

SA approach. The two clusters found by our approach correspond to synapses that are either chiefly depressing or facilitating. Still, the clusters did not correspond well to synapse type. In particular, PC–PC and PC–BC synapses were classified as the same type.

In an alternative approach, we also clustered the Bayesian posteriors using the symmetric KL-divergence (sKL). The sKL achieved 0.78 dendrogram purity and three clusters according to the Pseudo-F statistic; thus performing worse.

To determine how well the posterior distributions could be classified in keeping with the three connection types, we performed Naive Bayes classification with a 7-fold cross-validation (**Figure 5E**). We obtained 100% accuracy in PC–MC connection classification. Surprisingly, however, we also obtained a 72% and

FIGURE 4 | Inference of short-term plasticity parameters from experimental data from visual cortex. (A) Sample experimental STP traces are shown for PC–PC (red), PC–BC (green), and for PC–MC (blue) connections. **(B)** Marginalized posterior distributions obtained using slice sampling from these three different excitatory connections suggest that PC–MC ($n = 9$) connections are quite different from PC–PC ($n = 9$) and PC–BC ($n = 12$) connections. Light colored lines show individual connections, while dark colored lines correspond to their average.

75% classification accuracy for PC–PC and PC–BC connections, respectively. These results suggest that each synapse type can be to some extent separated from the other two types. The ability to separate the different connection types is likely to be mostly due to differences in the baseline release probability (cf. **Figure 4B**, parameter U).

3.4. COMPARISON TO TRADITIONAL FITTING METHODS
Above we found that for both the simulated and the experimental data, the marginalized posterior of the F parameter resembles an uniform distribution (**Figures 2B**, **4B**). This suggests that standard fitting techniques might not perform well and may become trapped in local minima, thus explaining why the SA-based clustering is not able to separate the different synaptic dynamics as well. To test this idea, we used SA on a depressing PC–PC connection and we found that this was indeed the case (**Figure 6**). Although the method found everytime good fits to the data (**Figure 6A**), the fit parameters were highly variable from one run to the next (**Figure 6B**). Although this variability could be used as a proxy for the parameter variance, there is no principled way in SA to estimate parameter variance. In contrast, with our Bayesian approach, the variability and exact distribution is captured in the posterior distribution. Similar observations were made by Varela

et al. (1997), who occasionally found an elongated error valley when fitting their particular STP model.

3.5. FINDING THE BEST MODEL USING PROBABILISTIC INFERENCE
The Bayesian approach offers a natural way to examine which model describes the data most parsimoniously. We performed model selection to identify which formulation of the TM model better described the data (see Materials and Methods). We compared three formulations of the TM model: (1) with depression only—only Equation (1) with D and U (two parameters)—, (2) depression and facilitation—Equations (1, 2) with D, F and U (three parameters)—and, (3) the full extended model used above. **Figure 7** shows that only the extended model is able to account for all the data from the three connection types. In contrast to Markram et al. (1998) and Richardson et al. (2005), we found that the TM-with-facilitation model does not fit the PC–MC connections well. Although for some recordings the three-parameter model was sufficient, it failed to fit other recordings (**Figure 7B**). This discrepancy might be due to experimental differences; our dataset was recorded in mice visual cortex L5 and included extracellular stimulation experiments, while Markram et al. (1998) and Richardson et al. (2005) recorded in the somatosensory cortex of the rat using paired recordings only.
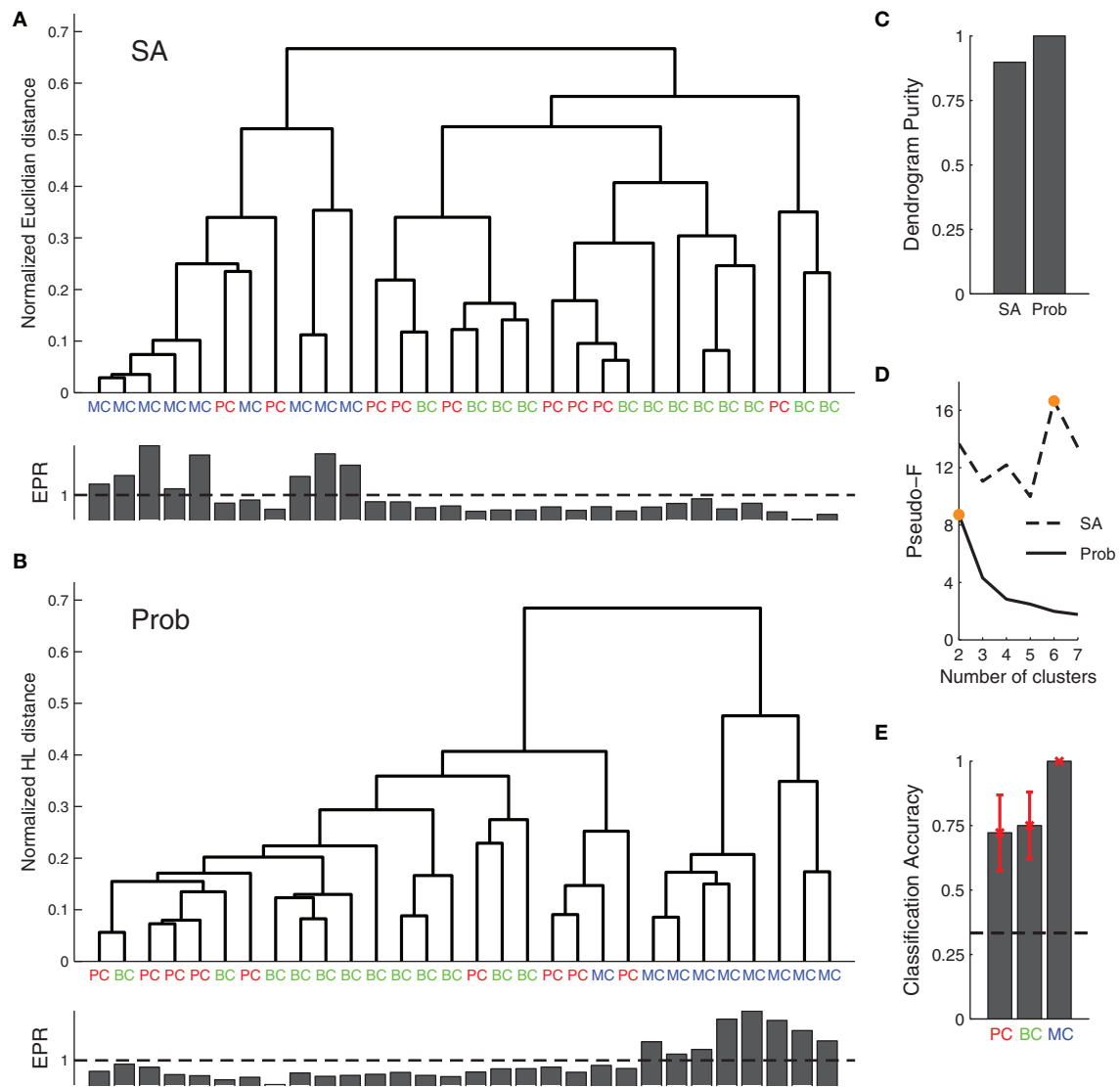
FIGURE 5 | Agglomerative clustering using posterior distributions improves synaptic dynamics clustering. (A) Clustering based on the synaptic parameters found by SA did not produce good clusters. **(B)** Clustering of posterior distributions using the probabilistic approach with the Hellinger distance gave rise to two clusters: one for short-term depression and the other for short-term facilitation (cf. EPR, inset bottom), with the first corresponding to both PC–PC and PC–BC connections, while the other roughly mapped onto PC–MC synapses. **(C)** EPR-based dendrogram purity with probability distribution-based clustering is higher than the purity from SA-based clustering. **(D)** Maximal Pseudo-F statistic suggests that the data contains two or six clusters when clustering the posterior distributions or SA-based clustering, respectively (orange filled circles). **(E)** A simple probabilistic classifier (Naive Bayes) achieved good performance for all the connection types, in particular for PC–MC connections (black dashed line represents chance level). Error bars represent standard error of the mean.

4. DISCUSSION

Past studies characterizing short-term synaptic dynamics have typically used traditional fitting methods. A Bayesian approach, however, turns out to be particularly advantageous for this problem, because accurate estimation of synaptic parameters is complicated. Here, we have shown that—depending on the synaptic dynamics and experimental protocol—some parameters are not narrowly tuned but broadly distributed. This insensitivity may cause traditional least-mean-square methods to get stuck in local minima.

When applied to experimental data, our method showed that different connections have different distributions. Such synapse-type specific plasticity supports the idea that different synapses perform different computations and subserve different functional roles in the local circuit. Our approach more robustly classifies synapses according to their synaptic dynamics than does clustering using simple point estimates of parameters obtained from standard optimization techniques. Our method might thus enable automatic and independent classification of synapses and cells taking into account the natural variability in the data. Future

FIGURE 6 | Comparison of Bayesian approach and traditional fitting methods. (A) STP models using either MAP or SA solutions (green and red crosses, respectively) provide good fits to the experimental data (black filled circles). **(B)** Marginalized posterior distributions for the depression and facilitation timeconstants (gray and black line, respectively). When fitting the data in **(A)** 10 times, SA yields widely different parameter values (red diamonds, all solutions provided good fits to the data $R^2 > 0.99$). The MAP solution is shown with green diamonds. The red arrows indicate the SA fit used in **(A)**.

FIGURE 7 | Extended TM model improves the fitting of PC–MC synapses. (A) The evidence ratio based on the Akaike Information Criterion is plotted on a logarithmic scale for the three excitatory synaptic types. Three formulations of the Tsodyks–Markram model were compared—with only depression (TM, two parameters), with a degree of facilitation (TM with facilitation, three parameters) and the extended version with full facilitation (eTM, four parameters). Error bars represent standard error of the mean. **(B)** Examples of normalized postsynaptic peak responses that can only be accurately fitted by the eTM model. **Top:** PC–PC recording with combined depression and facilitation. **Bottom:** PC–MC recording with attenuating facilitation. The postsynaptic peak responses (black filled circles) are given together with the MAP solutions from the two, three, and four parameters model, from dark to light green, respectively.

studies using larger datasets may better identify the synaptic properties that are specific to individual clusters. Furthermore, a model with a more detailed noise description could allow us to also infer the quantal parameters, which could in principle be combined with the Bayesian quantal analysis framework (Bhumbra and Beato, 2013).

We found that inference of the model parameters can be improved by having more pulses as well as by including a recovery

phase. The data used here, however, was collected using a standard STP electrophysiology protocol with five pulses at 30 Hz, which still enabled connection-specific clustering. To improve parameter inference further, we propose a combination of a periodic spike train and a Poisson spike train. More pulses add more information, which has an unsurprising positive impact on inference. Poisson trains cover the frequency space better without requiring excessively long experimental recordings. Indeed, Poisson trains add a considerable improvement as compared to the more standard protocol of using fixed-frequency trains (Markram and Tsodyks, 1996; Sjöström et al., 2003).

Experimentally STP has been observed to change with development (Reyes and Sakmann, 1999), drug wash-in (Buchanan et al., 2012), temperature changes (Klyachko and Stevens, 2006), and plasticity (Markram and Tsodyks, 1996; Sjöström et al., 2003). In such situations, it often becomes important to ascertain the particular parameter changes that occur. The Bayesian framework introduced here can be extended to elucidate which components of STP are affected by integrating prior knowledge, through an informative prior. For instance, inferred distributions can be tracked across development.

Our work can also be applied in constructing computer network models with STP using posterior distributions inferred from actual biological data as a generative model. This would yield models with richer dynamics without resorting to simplistic and unrealistic *ad-hoc* approaches to generate synaptic variability that are poorly grounded in biological data.

Our Bayesian approach promises improved computer models as well as a better and more nuanced understanding of biological data. Yet, this approach is not computationally intense, nor is it difficult to implement. We therefore fully expect probabilistic inference of STP parameters to become a widespread practice in the immediate future.

REFERENCES

Abbott, L. F., and Regehr, W. (2004). Synaptic computation. *Nature* 431, 796–803. doi: 10.1038/nature03010

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 221–224. doi: 10.1126/science.275.5297.221

Bhumbra, G. S., and Beato, M. (2013). Reliable evaluation of the quantal determinants of synaptic efficacy using Bayesian analysis. *J. Neurophysiol.* 109, 603–620. doi: 10.1152/jn.00528.2012

Brooks, S., and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* 7, 434–455. doi: 10.1080/10618600.1998.10474787

Buchanan, K. A., Blackman, A. V., Moreau, A. W., Elgar, D., Costa, R. P., Lalanne, T., et al. (2012). Target-specific expression of presynaptic NMDA receptors in neocortical microcircuits. *Neuron* 75, 451–466. doi: 10.1016/j.neuron.2012.06.017

Caliński, T., and Harabasz, J. (1974). A dendrite method for cluster analysis. *Commun. Stat. Theory Meth.* 3, 1–27. doi: 10.1080/03610927408827101

Carvalho, T., and Buonomano, D. (2011). A novel learning rule for long-term plasticity of short-term synaptic plasticity enhances temporal processing. *Front. Integr. Neurosci.* 5:20. doi: 10.3389/fnint.2011.00020

Cheetham, C. E. J., and Fox, K. (2010). Presynaptic development at L4 to l2/3 excitatory synapses follows different time courses in visual and somatosensory cortex. *J. Neurosci.* 30, 12566–12571. doi: 10.1523/JNEUROSCI.2544-10.2010

Dittman, J. S., Kreitzer, A. C., and Regehr, W. G. (2000). Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *J. Neurosci.* 20, 1374–1385.

Fung, C. C. A., Wong, K. Y. M., Wang, H., and Wu, S. (2012). Dynamical synapses enhance neural information processing: gracefulness, accuracy, and mobility. *Neural Comput.* 24, 1147–1185. doi: 10.1162/NECO_a_00269

Galarreta, M., and Hestrin, S. (1998). Frequency-dependent synaptic depression and the balance of excitation and inhibition in the neocortex. *Nat. Neurosci.* 1, 587–594. doi: 10.1038/2882

Gelman, A., and Shirley, K. (2011). "Inference from simulations and monitoring convergence," in *Handbook of Markov Chain Monte Carlo*, eds S. Brooks, A. Gelman, G. Jones, and X. L. Meng (Boca Raton, FL: CRC Press), 163–174.

Heller, K., and Ghahramani, Z. (2005). "Bayesian hierarchical clustering," in *Proceedings of the 22nd International Conference on Machine learning* (Bonn), 297–304.

Ihler, A., and Mandel, M. (2007). Kernel density estimation toolbox for matlab. Available online at: http://ics.udi.edu/ihler/code/kde.html

Klyachko, V. A., and Stevens, C. F. (2006). Temperature-dependent shift of balance among the components of short-term plasticity in hippocampal synapses. *J. Neurosci.* 26, 6945–6957. doi: 10.1523/JNEUROSCI.1382-06.2006

Le Bé, J., and Markram, H. (2006). Spontaneous and evoked synaptic rewiring in the neonatal neocortex. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13214. doi: 10.1073/pnas.0604691103

Loebel, A., Silberberg, G., Helbig, D., Markram, H., Tsodyks, M., and Richardson, M. J. E. (2009). Multiquantal release underlies the distribution of synaptic efficacies in the neocortex. *Front. Comput. Neurosci.* 3:27. doi: 10.3389/neuro.10.027.2009

Maass, W., and Zador, A. M. (1999). Dynamic stochastic synapses as computational units. *Neural Comput.* 11, 903–917. doi: 10.1162/089976699300016494

Markram, H., Lübke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol.* 500(Pt 2), 409–440.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nat. Rev. Neurosci.* 5, 793–807. doi: 10.1038/nrn1519

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature.* 382, 807–810. doi: 10.1038/382807a0

Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5323–5328.

McGuinness, L., Taylor, C., Taylor, R. D. T., Yau, C., Langenhan, T., Hart, M. L., et al. (2010). Presynaptic NMDARs in the hippocampus facilitate transmitter release at theta frequency. *Neuron* 68, 1109–1127. doi: 10.1016/j.neuron.2010.11.023

Milligan, G., and Cooper, M. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika* 50, 159–179. doi: 10.1007/BF02294245

Nakagawa, S., and Hauber, M. (2011). Great challenges with few subjects: statistical strategies for neuroscientists. *Neurosci. Biobeh. Rev.* 35, 462–473. doi: 10.1016/j.neubiorev.2010.06.003

Neal, R. (2003). Slice sampling. *Ann. Stat.* 31, 705–741. doi: 10.1214/aos/1056562461

Pan, B., and Zucker, R. (2009). A general model of synaptic transmission and short-term plasticity. *Neuron* 62, 539–554. doi: 10.1016/j.neuron.2009.03.025

Pereira, F., Tishby, N., and Lee, L. (1993). "Distributional clustering of

English words," in *Proceedings of the 31st Annual Meeting on Association for Computational Linguistics* (Columbus, OH: USA Association for Computational Linguistics).

Pouzat, C., and Hestrin, S. (1997). Developmental regulation of basket/stellate cell – Purkinje cell synapses in the cerebellum. *J. Neurosci.* 17, 9104–9112.

Ramaswamy, S., Hill, S. L., King, J. G., Schürmann, F., Wang, Y., and Markram, H. (2012). Intrinsic morphological diversity of thick-tufted layer 5 pyramidal neurons ensures robust and invariant properties of in silico synaptic connections. *J. Physiol.* 590, 737–752. doi: 10.1113/jphysiol.2011.219576

Reyes, A., Lujan, R., Rozov, A., Burnashev, N., Somogyi, P., and Sakmann, B. (1998). Target-cell-specific facilitation and depression in neocortical circuits. *Nat. Neurosci.* 1, 279–285. doi: 10.1038/1092

Reyes, A., and Sakmann, B. (1999). Developmental switch in the short-term modification of unitary EPSPs evoked in layer 2/3 and layer 5 pyramidal neurons of rat neocortex. *J. Neurosci.* 19, 3827–3835.

Richardson, M. J. E., Melamed, O., Silberberg, G., Gerstner, W., and Markram, H. (2005). Short-term synaptic plasticity orchestrates the response of pyramidal cells and interneurons to population bursts. *J. Comput. Neurosci.* 18, 323–331. doi: 10.1007/s10827-005-0434-8

Rinaldi, T., Silberberg, G., and Markram, H. (2008). Hyperconnectivity of local neocortical microcircuitry induced by prenatal exposure to valproic acid.

Cereb. Cortex 18, 763–770. doi: 10.1093/cercor/bhm117

Romani, A., Marchetti, C., Bianchi, D., Leinekugel, X., Poirazi, P., Migliore, M., et al. (2013). Computational modeling of the effects of amyloid-beta on release probability at hippocampal synapses. *Front. Comput. Neurosci.* 7:1. doi: 10.3389/fncom.2013.00001

Rozov, A., Burnashev, N., Sakmann, B., and NEHER, E. (2001). Transmitter release modulation by intracellular Ca2+ buffers in facilitating and depressing nerve terminals of pyramidal cells in layer 2/3 of the rat neocortex indicates a target cell-specific difference in presynaptic calcium dynamics. *J. Physiol.* 531, 807–826. doi: 10.1111/j.1469-7793.2001.0807h.x

Scanziani, M., Gähwiler, B., and Charpak, S. (1998). Target cell-specific modulation of transmitter release at terminals from a single axon. *Proc. Natl. Acad. Sci. U.S.A.* 95, 12004–12009. doi: 10.1073/pnas.95.20.12004

Sjöström, P. J., Turrigiano, G. G., and Nelson, S. B. (2003). Neocortical LTD via coincident activation of presynaptic NMDA and cannabinoid receptors. *Neuron* 39, 641–654. doi: 10.1016/S0896-6273(03)00476-8

Sokal, R. R. (1958). A statistical method for evaluating systematic relationships. *Univ. Kans. Sci. Bull.* 38, 1409–1438.

Sun, H., and Dobrunz, L. (2006). Presynaptic kainate receptor activation is a novel mechanism for target cell-specific short-term facilitation at Schaffer collateral synapses. *J. Neurosci.* 26, 10796. doi: 10.1523/JNEUROSCI.2746-06.2006

Testa-Silva, G., Loebel, A., Giugliano, M., de Kock, C. P. J., Mansvelder, H. D., and Meredith, R. M. (2012). Hyperconnectivity and slow synapses during early development of medial prefrontal cortex in a mouse model for mental retardation and autism. *Cereb. Cortex* 22, 1333–1342. doi: 10.1093/cercor/bhr224

Tóth, K., and McBain, C. J. (2000). Target-specific expression of pre- and postsynaptic mechanisms. *J. Physiol.* 525(Pt 1), 41–51. doi: 10.1111/j.1469-7793.2000.00041.x

Tsodyks, M., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U.S.A.* 94, 719–723.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835. doi: 10.1162/089976698300017502

Tsodyks, M., Uziel, A., and Markram, H. (2000). Synchrony generation in recurrent networks with frequency-dependent synapses. *J. Neurosci.* 20, RC50.

Turkheimer, F. E., Hinz, R., and Cunningham, V. J. (2003). On the undecidability among kinetic models: from model selection to model averaging. *J. Cereb. Blood Flow Metab.* 23, 490–498.

van Rossum, M. C. W., van der Meer, M. A. A., Xiao, D., and Oram, M. W. (2008). Adaptive integration in the visual cortex by depressing recurrent cortical circuits. *Neural Comput.* 20, 1847–1872. doi: 10.1162/neco.2008.06-07-546

Varela, J., Sen, K., Gibson, J., Fost, J., Abbott, L. F., and Nelson, S. (1997). A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J. Neurosci.* 17, 7926.

Wang, Y., Markram, H., Goodman, P. H., Berger, T. K., Ma, J., and Goldman-Rakic, P. S. (2006). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat. Neurosci.* 9, 534–542. doi: 10.1038/nn1670

Zucker, R. S., and Regehr, W. G. (2002). Short-term synaptic plasticity. *Annu. Rev. Physiol.* 64, 355–405. doi: 10.1146/annurev.physiol.64.092501.114547

Stimulus number, duration and intensity encoding in randomly connected attractor networks with synaptic depression

Paul Miller*

Volen National Center for Complex Systems and Department of Biology, Brandeis University, Waltham, MA, USA

Randomly connected recurrent networks of excitatory groups of neurons can possess a multitude of attractor states. When the internal excitatory synapses of these networks are depressing, the attractor states can be destabilized with increasing input. This leads to an itinerancy, where with either repeated transient stimuli, or increasing duration of a single stimulus, the network activity advances through sequences of attractor states. We find that the resulting network state, which persists beyond stimulus offset, can encode the number of stimuli presented via a distributed representation of neural activity with non-monotonic tuning curves for most neurons. Increased duration of a single stimulus is encoded via different distributed representations, so unlike an integrator, the network distinguishes separate successive presentations of a short stimulus from a single presentation of a longer stimulus with equal total duration. Moreover, different amplitudes of stimulus cause new, distinct activity patterns, such that changes in stimulus number, duration and amplitude can be distinguished from each other. These properties of the network depend on dynamic depressing synapses, as they disappear if synapses are static. Thus, short-term synaptic depression allows a network to store separately the different dynamic properties of a spatially constant stimulus.

Keywords: short-term plasticity, dynamic synapses, attractor networks, short-term memory, distributed coding, high-dimensional representation

INTRODUCTION

Circuits of reciprocally connected neurons have been long considered as a basis for the maintenance of persistent activity (Lorente de Nó, 1933). Such persistent neuronal firing that continues for many seconds after a transient input can represent a short-term memory of prior stimuli (Funahashi et al., 1991). Indeed, Hebb's famous postulate (Hebb, 1949) that causally correlated firing of connected neurons could lead to a strengthening of the connection, was based on the suggestion that the correlated firing would be maintained in a recurrently connected cell assembly beyond the time of a transient stimulus (Hebb, 1949). Since then, analytic and computational models have demonstrated the ability of such recurrent networks to produce multiple discrete attractor states (Brunel and Nadal, 1998), as in Hopfield networks (Hopfield, 1982, 1984), or to be capable of integration over time via a marginally stable network, often termed a line attractor (Zhang, 1996; Compte et al., 2000). Much of the work on these systems has assumed either static synapses, or considered changes in synaptic strength via long-term plasticity occurring on a much slower timescale than the dynamics of neuronal responses. Here we add some new results pertaining to the less well-studied effects of short-term plasticity—changes in synaptic strength that arise on a timescale of seconds, the same timescale as that of persistent activity—within recurrent discrete attractor networks.

The two long-established forms of short-term synaptic plasticity affect all synapses of the presynaptic cell according to its train of action potentials. Synaptic depression refers to a reduced synaptic efficacy in the few hundreds of milliseconds following a presynaptic spike, effectively weakening connections strengths as presynaptic firing rate increases (Markram and Tsodyks, 1996; Abbott et al., 1997). Such weakening of efficacy of the most active connections has an unavoidable destabilizing effect on any network state that depends on those active connections for its persistence. Synaptic facilitation is the opposite effect—a temporary enhancement of synaptic efficacy in the few hundreds of milliseconds following each spike (Markram et al., 1998), effectively strengthening connections to post-synaptic cells as presynaptic firing rate increases.

More recently described and information-theoretically more powerful than depression or facilitation, is an associative form of short-term plasticity (A-STP), which depends on both pre- and post-synaptic activity (Erickson et al., 2010). A-STP produces a temporary enhancement of synaptic efficacy between neurons after a short period of strong coactivity. Being a form of positive feedback, A-STP, like facilitation, is likely to stabilize states of persistent activity, but may have the added benefit of maintaining sequences of persistent firing states (Miller and Wingfield, 2010).

In this paper, we focus on short-term synaptic depression in randomly connected networks of discrete attractors (Rigotti et al., 2010). The attractors are formed by coupling multiple groups of neurons, each group rendered bistable through recurrent excitation. The destabilization of discrete attractor states by short-term synaptic depression produces a rich repertoire

of network responses, allowing it to encode and store multiple stimulus features.

Short-term depression arises from vesicle depletion (von Gersdorff and Matthews, 1997), which leads to a maximum, saturating rate of synaptic transmission—dependent on the rate of vesicle recycling. The temporary weakening of connection strengths from active cells tends to reduce the stability of active recurrent cell-groups. This can lead to more dynamic or itinerant activity states in recurrent networks. Here we show that in a network of randomly coupled cell-groups, the itinerancy produced by synaptic depression can cause the network to reach a state that depends on any of stimulus intensity, or stimulus duration or the number of successive identical stimuli presented. In the latter case, neurons can be tuned to a specific number of inputs, similarly to those recorded *in vivo*.

Counting of stimuli can be achieved without dynamic synapses in a network behaving as an integrator. Indeed, appropriate feedforward connections from an integrator can produce numerosity-tuned neurons (Verguts and Fias, 2004), with similar tuning curves to those found *in vivo* (Nieder and Miller, 2003; Tudusciuc and Nieder, 2007; Merten and Nieder, 2009; Nieder, 2013). However, an integrator, whether it arises from a finely tuned network with a continuous, line attractor (Seung et al., 2000; Miller et al., 2003; Machens et al., 2005), or more robustly from a series of discrete attractor states (Koulakov et al., 2002; Goldman et al., 2003), is not ideal as the input to a counter. While a perfect integrator does indeed produce distinct responses to successive identical stimuli, it conflates both amplitude and duration of the stimulus, with the number of stimuli, into a single response that only depends on the produce of these three quantities. Thus, an integrator's response to two stimuli of a given magnitude and duration is identical to that of a single stimulus with either twice the magnitude or twice the duration. Any non-linearities would remove such perfect scaling [which is essential in situations requiring perfect integration, such as from velocity to position (Zhang, 1996; Samsonovich and McNaughton, 1997; Song and Wang, 2005)] but would not remove the conflation of stimulus features, since an integrator's activity is confined to a one-dimensional surface—input amplitude, duration and number produce shifts along the same one-dimensional line. Thus, for an integrator to act as a counter, its inputs must be first scaled to a fixed duration and amplitude by upstream sensory processing.

Here we test whether any advantage over the integrator is offered by the high-dimensional space of attractor states produced by randomly connected bistable groups of neurons (Rigotti et al., 2010). In a group of cells with recurrent excitatory connections, the excitability of the cell-group—its ability to become rapidly active in response to input—increases with the effect strength of the internal connections. In a network with many such cell-groups, if they are predominantly coupled by cross-inhibition, those cell-groups most excited by the stimulus and activated most quickly, can suppress activity of other cell-groups. Short-term synaptic depression reduces the effective connection strengths between coactive neurons compared to those between quiescent neurons. Since the amplitude of synaptic depression is firing-rate dependent, and since internal randomness in the network causes cell-groups to respond with different amplitude-dependences of their firing rates, stimuli of different amplitudes are likely to affect the network differently. Moreover, dynamical synapses cause the network's response to depend on the temporal profile of stimuli, not just its temporal integration, so that two spaced stimuli could produce a different response from a single stimulus of twice the duration.

Therefore, we will vary three stimulus properties—number, duration and amplitude—both individually and together, to assess whether a randomly connected network with dynamic synapses, unlike an integrator, can dissociate these features. We first assess whether, when a stimulus is repeated, cell-groups active to its first presentation can be replaced by other active cell-groups during its second and later presentations. We then uncover how this process, in a randomly connected sparse recurrent network, depends on different qualities of the stimulus, such as its duration and intensity. Finally, we show these qualities interact with the number of stimuli in a non-trivial manner, often producing unique patterns of persistent activity as a function of number, duration and intensity of preceding stimuli.

METHODS

FIRING RATE MODEL WITH DEPRESSING SYNAPSES

To model the effects of synaptic depression in a network of coupled cells, we employ a firing rate model, which treats the mean input current, $I_i(t)$, the mean firing rate $r_i(t)$, the mean depression variable, $D_i(t)$ and the mean synaptic output, $S_i(t)$, of individual groups of neurons, labeled i, as continuous, time-dependent quantities. The formulation is appropriate for cells with Poisson spike statistics, as at fixed firing rates the depression variable and synaptic outputs approach the steady state values produced by Poisson spike trains, though with appropriate rate-dependent modifications to the effective time constants. Thus, the dynamics of the system is described by a set of coupled first order differential equations. The firing rate depends upon its input current according to a sigmoidal f–I curve, as:

$$\tau_r \frac{dr_i}{dt} = -r_i(t) + \frac{r_i^{\max}}{\exp\{[\Theta_i - I_i(t)]/\Delta_i\}} \tag{1}$$

where $\tau_r = 10$ ms is the time constant for, $r_i^{\max}$ is the maximum firing rate of that cell-group, Θ_i is the threshold, namely the level of input current required for half-maximal firing and Δ_i determines (with $r_i^{\max}$) the slope of the f–I curve.

The depression variable follows:

$$\tau_{Di} \frac{dD_i}{dt} = 1 - D_i(t) - p_0 r_i(t) \tau_{Di} D_i(t) \tag{2}$$

where p_0 is the fraction of docked vesicles released per spike and τ_{Di} is the recovery time to regain maximum transmission. Equation 2 is chosen so as to reach the steady state value produced by a Poisson spike train (Dayan and Abbott, 2001) of rate r_i:

$$D^{ss}(r_i) = \frac{1}{1 + p_0 r_i \tau_{Di}}, \tag{3}$$

if the rate were fixed, assuming each presynaptic spike at time t_s causes a reduction in the depression variable,

$D_i\left(t_s^+\right) = D_i\left(t_s^-\right)(1-p_0)$, due to loss of a proportion, p_0, of docked vesicles.

The synaptic gating variable follows:

$$\tau_s \frac{ds_i}{dt} = -s_i(t) + \tilde{\alpha} p_0 r_i(t) \tau_s D_i(t)\left[1 - s_i(t)\right] \qquad (4)$$

where τ_s is the synaptic time constant for decay of s_i to zero in the absence of synaptic transmission and $\tilde{\alpha}$ is the fraction of open receptors bound by maximal vesicle release—that is, the fractional increase in s for a given presynaptic spike at time t_s is $\tilde{\alpha} p_0 D_i(t_s^-)\left[1 - s_i(t_s^-)\right]$. Equation (3) reaches the steady state value for s_i produced by a Poisson train of releases with fixed D_i, at a rate r_i:

$$\tau_s S^{ss}\left(r_i, D_i\right) = \frac{\tilde{\alpha} p_0 D_i r_i \tau_s}{1 + \tilde{\alpha} p_0 D_i r_i \tau_s}. \qquad (5)$$

The connectivity matrix, $W_{i \to j}$ describes the connection strengths from each cell-group i to cell-group j, so determines the input current to a cell-group j via:

$$I_j(t) = \sum_i s_i(t) W_{i \to j} + I_j^{app}(t) + \sigma\eta(t) \qquad (6)$$

where $I_j^{app}(t)$ is the stimulus-dependent external, applied current to cell-group j and $\eta(t)$ is a white noise term which contributes fluctuations to each cell-groups current, with a standard deviation σ.

Full details of the simulation parameters are given in **Tables 1** and **2**.

NETWORK PROPERTIES, STIMULATION PROTOCOLS AND MEASUREMENTS

Our main results were achieved with a network of $N_E = 100$ excitatory cell-groups and a single inhibitory cell-group, though we tested the effects of using from $N_E = 20$ to $N_E = 400$ excitatory cell-groups. The dominant connections within the network

Table 1 | Components of the network simulations (Nordlie et al., 2009).

A. Model summary

Populations	100 excitatory (E), 1 inhibitory (I)
Connectivity	E-to-E: all-to-all with random strength; high self-excitation
Neuron model	Firing rate model with sigmoidal f–I curve
Synapse model	Single exponential with depression or facilitation
Plasticity	No long-term plasticity
Input	Fixed current pulses to all populations
Measurements	Persistent firing rates after current offset

B. Populations

Name	Elements	Size
Excitatory (E)	Firing-rate model	1 for each of 100 model cell-populations
Inhibitory (I)	Firing-rate model	1 for the single cell-population

(Continued)

Table 1 | Continued

C. Connectivity

Name	Source	Target	Pattern
EE (S)	E	Same E	Fixed at WEEs
EE (X)	E	All other E	All-to-all with weight a random iid in [0, 2WEEx]
EI	E	I	Fixed, constant at WEI
IE	I	E	Fixed, constant to all at WIE
II	I	I	Not Present

D. Neuron and synapse model

Name	Firing rate model with dynamical synapses
Type	Dynamic leaky integrate-and-fire, exponential conductance input
Input current	$I_j(t) = \sum_i s_i(t) W_{i \to j} + I_j^{app}(t) + \sigma\eta(t)$
Firing rate	$\tau_r \dfrac{dr_i}{dt} = -r_i(t) + \dfrac{r_i^{max}}{\exp\left\{\left[\Theta_i - I_i(t)\right]/\Delta_i\right\}}$
Depression variable	$\tau_{Di} \dfrac{dD_i}{dt} = 1 - D_i(t) + p_0 r_i(t) \tau_{Di} D_i(t)$
Synaptic transmission	$\tau_s \dfrac{ds_i}{dt} = -s_i(t) + \tilde{\alpha} p_0 r_i(t) \tau_s D_i(t)\left[1 - s_i(t)\right]$

E. Plasticity

No long-term plasticity present

F. Input

Type	Description
Applied current	Transient pulses of fixed current with number of pulses, amplitude of pulse and length of pulse varied across simulations. Current is identical to all excitatory populations and scaled by a constant factor to the inhibitory population

G. Measurements

Firing rates vectors	Mean rate per cell 750–1500 ms after stimulus onset
Correlations	Correlation between firing vectors for different stimuli
Confusability	Proportion of trials that response to a test stimulus is closer to the mean response produced by a target stimulus than to the mean response of any other target stimulus

were produced by strong self-excitation within each excitatory cell-group and strong cross-inhibition between all excitatory cell-groups via the inhibitory cell-group. The cell-groups were further coupled by all-to-all excitatory connections, with connection strength chosen randomly from a uniform distribution between zero and the maximum value. Such random cross-connections, even in sum, produced a weaker excitatory input than the within-group connection.

More specifically, the connection matrix, $W_{i \to j}$ (Equation 6) comprised four types of connection: fixed strength excitatory connections within an excitatory cell-group ($W_{i \to i} = W_{EE}^0$ for $1 \le i \le N_E$); random strength excitatory connections between excitatory cell-groups ($W_{i \to j} = \xi_{ij} W_{EE}^X/(N_E - 1)$, if $i \ne j$ and

Table 2 | Network simulation model parameters.

A. Firing rate model and input current

Population	τ_r	r_i^{max}	Θ_i	Δ_i	σ
Excitatory (E)	0.01 s	100 Hz	[6.3–6.5]	1	0.002, 0.005
Inhibitory (I)	0.01 s	200 Hz	12	3	0.002, 0.005

B. Depression and synaptic transmission

Connection	τ_D	p_0	τ_s	$\widetilde{\alpha}$
EE (S), EE (X), EI	0.5 s	1	0.05 s	1
IE	0.5 s	0.1	0.005 s	1

C. Connection strengths

Connection	W_{EE}^0	W_{EE}^X	W_{EI}	W_{IE}
Value	85	[0–0.4]	2.5	−300

D. Stimulus values

Property	N_{max}	T	I_0	Interval
Value	10 (8, 6, 1)	0.1 s (0.01 s–1 s)	2 (0.5–3)	1.5 s (2 s)

$1 \leq i, j \leq N_E$) and η_{ij} is a random number selected from a uniform distribution ($0 < \xi_{ij} < 1$); fixed strength excitatory connections to the inhibitory cell-group ($W_{i \to j} = W_{EI}/(N_E - 1)$ if $1 \leq i \leq N_E$ and $j = N_E + 1$); and fixed strength inhibitory connections to each excitatory cell-group ($W_{i \to j} = W_{IE}$ if $i = N_E + 1$ and $1 \leq j \leq N_E$). Values of these parameters are given in **Table 2**. Different versions of a network with the same parameters were generated by selecting a new set of random excitatory cross-connections through a new generation of the random matrix, ξ_{ij}. In contrast, repeated trials with the same network were produced with a fixed connection matrix, $W_{i \to j}$, but with a new instantiation of trial-specific random noise in the simulation, via $\eta(t)$ (Equation 6).

Stimuli were trains of transient current pulses, with each pulse producing the same current input to all excitatory cell-groups, as well as an input to the inhibitory cell-group. Depending on the protocol, current pulses ranged in number from 1 to 10, in duration from 10 ms to 1 s and in amplitude from 0.5 to 3 (in units where the firing threshold was in the range 6.3–6.5 for excitatory cells). Current pulses were delivered every 1.5 s in all protocols, except for those with varying stimulus duration, in which case delivery was every 2 s. While these current pulses could evoke immense changes in network activity, even the strongest inputs contributed only a small fraction of the total input to any cell-group, as the network is dominated by feedback within the circuit.

Mean network activity was calculated in all cases from at least 100 ms after stimulus offset until the onset of the subsequent stimulus. In the standard protocol, with a stimulus of 250 ms, rates of each cell were averaged from 375 to 1500 ms from stimulus onset (i.e., 125–1250 ms from stimulus offset) to determine the stimulus responses used in later analyses.

CONFUSABILITY MATRIX

To calculate a confusability matrix, we first simulated a set of 10 different random trials of the same network with different instances of noise via $\eta(t)$ (Equation 6). We used these initial trials to obtain the mean response in the delay period following each stimulus number or stimulus type, and defined these mean responses as the "target response." We then simulated a new set of 10 different random trials ("test trials") of the same network, for each test trial assessing which target response the delay activity most closely resembled. The confusability matrix gives the fraction of test trials, for which the response to one stimulus type and number most closely resembles the "target response" of a given stimulus type and number.

WEBER SCALING

To test for Weber's law, we produced 10 distinct networks, with 25 target trials and 25 test trials in each network. Importantly, across trials we allowed the level of noise to vary randomly, in this case according to a uniform distribution over the range $0.0015 < \sigma < 0.0075$. For each network, for a given test stimulus number, we calculated the mean and standard deviation of the target stimulus number the delayed activity most closely resembled. We then plot the mean standard deviation across networks versus the mean target reached in **Figure 2C**.

RESULTS

NUMEROSITY

Numerosity is the ability of a circuit to represent the number of transient stimuli. In the first task, we simply applied, repeatedly, a constant transient stimulus current to all cell-groups and assessed how reliably the resultant activity depended on the number of stimuli to date. Given appropriate parameters—in particular such that recurrent self-excitation within cell-groups was sufficient to maintain activity beyond the time of the transient stimulus (**Figure 1A**), but not so strong that it could not be suppressed by cross-inhibition arising from later activity in other cell-groups—the network could switch through stable, distributed activity states as shown in **Figure 1**. Moreover, when averaging single-cell responses during the delays between stimuli across 10 trials, many cells were tuned to individual numbers of stimuli (**Figure 1B1**). With increased noise, the observed tuning was broader for neurons selective to higher numbers (**Figure 1B2**). Similar tuning is seen in the neural activity of numerosity-selective neurons in primates (Nieder and Miller, 2003; Tudusciuc and Nieder, 2007, 2009), neurons which also respond to a temporal sequence of discrete stimuli (Nieder, 2012).

When analyzing the complete network response (**Figures 1C1,C2**) one notices that the overall pattern of activation is distributed: many cell-groups are active following any particular number of stimuli and any one cell-group can be active following multiply different stimuli. However, the activity patterns following particular numbers of stimuli are distinct from each other (**Figures 2A1,A2**). Indeed, the strongest effect of depression is to decorrelate subsequent stimuli from each other, so the lowest correlation is seen in a band surrounding the diagonal in **Figure 2A1**. Such an effect can be understood as
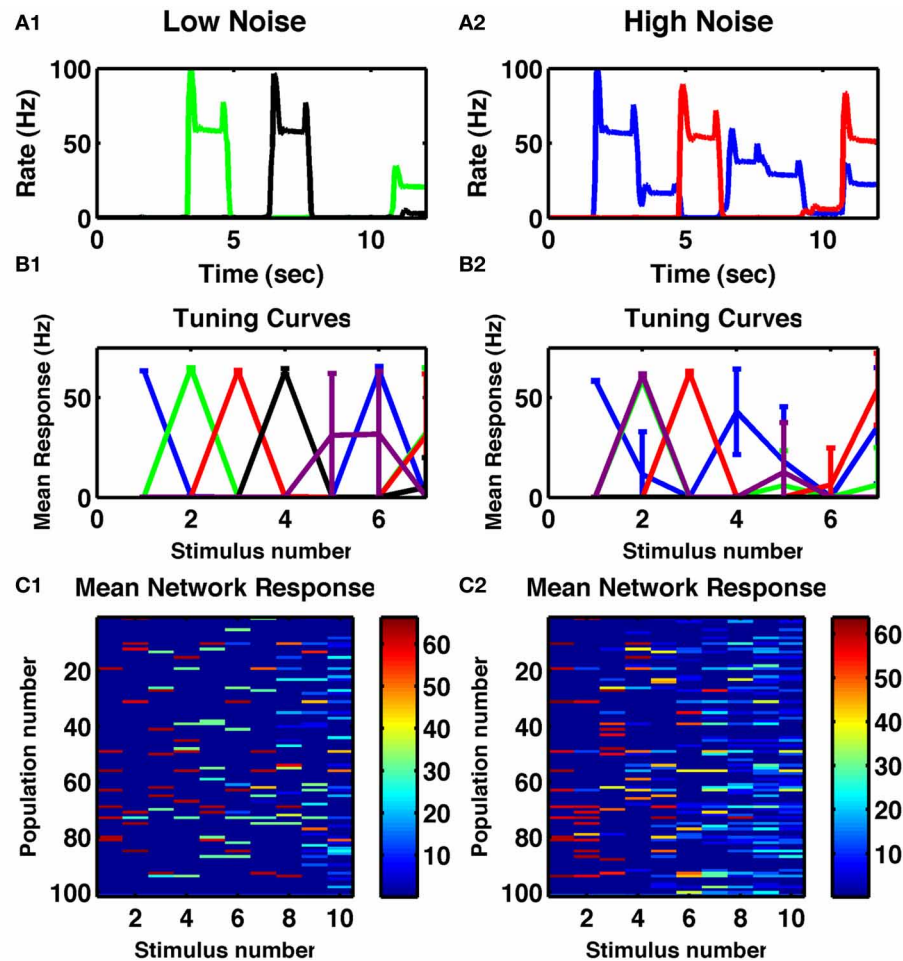
FIGURE 1 | Cell-group activity is tuned to the number of repeated stimuli. (A1) Responses of two different cells to a series of 250 ms constant inputs repeated every 1.5 s. **(A2)** Responses of two other cells to the same sequence with increased noise in the network. **(B1,B2)** Tuning curves of specific cell-groups in the network, each color represents a different cell, whose firing rate is plotted as a function of the number of successive identical stimuli. Colors matching those in **(A1,A2)** indicate the same cell. **(C1,C2)** Mean delay activity of all cell-groups to a train of ten identical stimuli, with color indicating firing rate. **(A1,B1,C1)** Internal noise, $\sigma = 0.002$. **(A2,B2,C2)** Increased internal noise, $\sigma = 0.005$. All panels: Mean responses averaged across ten trials of a single network, with error bars indicating standard deviations.

depression ensuring a group of cells is least likely to be active if it has just been active.

To assess how distinguishable were these different activity patterns from each other, we produced a set of 20 trials by using different instances of temporal noise. We took the mean responses of the first 10 trials to produce "target" responses. We then assessed for each of the next 10 "test" trials, which "target" representation the persistent activity was most similar to. If any two stimuli resulted in the same network response, then the test stimuli would be as often as similar to one as the other, producing a "confusability" of 0.5 to each pair. However, as we see (**Figure 2B1**), in the low noise case, we found 100% reproducibility of distinct activity patterns for the first 9 of 10 stimulus types. With increased noise, while the first three stimuli remained distinct with 100% reliability, the confusability increased with increasing stimulus count (**Figure 2B2**).

To quantify the variability in the response, in a separate experiment we selected a different level of noise in each trial used to simulate target responses then test responses. As in the calculation of the confusability matrix, for each stimulus number in a test trial, we treated the network's output as the stimulus number of the target response most correlated with the test response. Across the 10 test trials we calculated the standard deviation of these network outputs. We repeated across 10 different networks to produce the curve in **Figure 2C1**. With noise in the low range of $0.1 < \sigma < 0.3$, the responses to the first three stimuli are always precisely reproduced, so the variability is zero, but thereafter the standard deviation in the networks' responses increases linearly with stimulus number.

While our standard network comprised 100 excitatory cell-groups ($N_E = 100$), the qualitative behavior did not depend on this number. With increasing number of cell-groups, the effect of noise was decreased, with an approximate noise-scaling factor
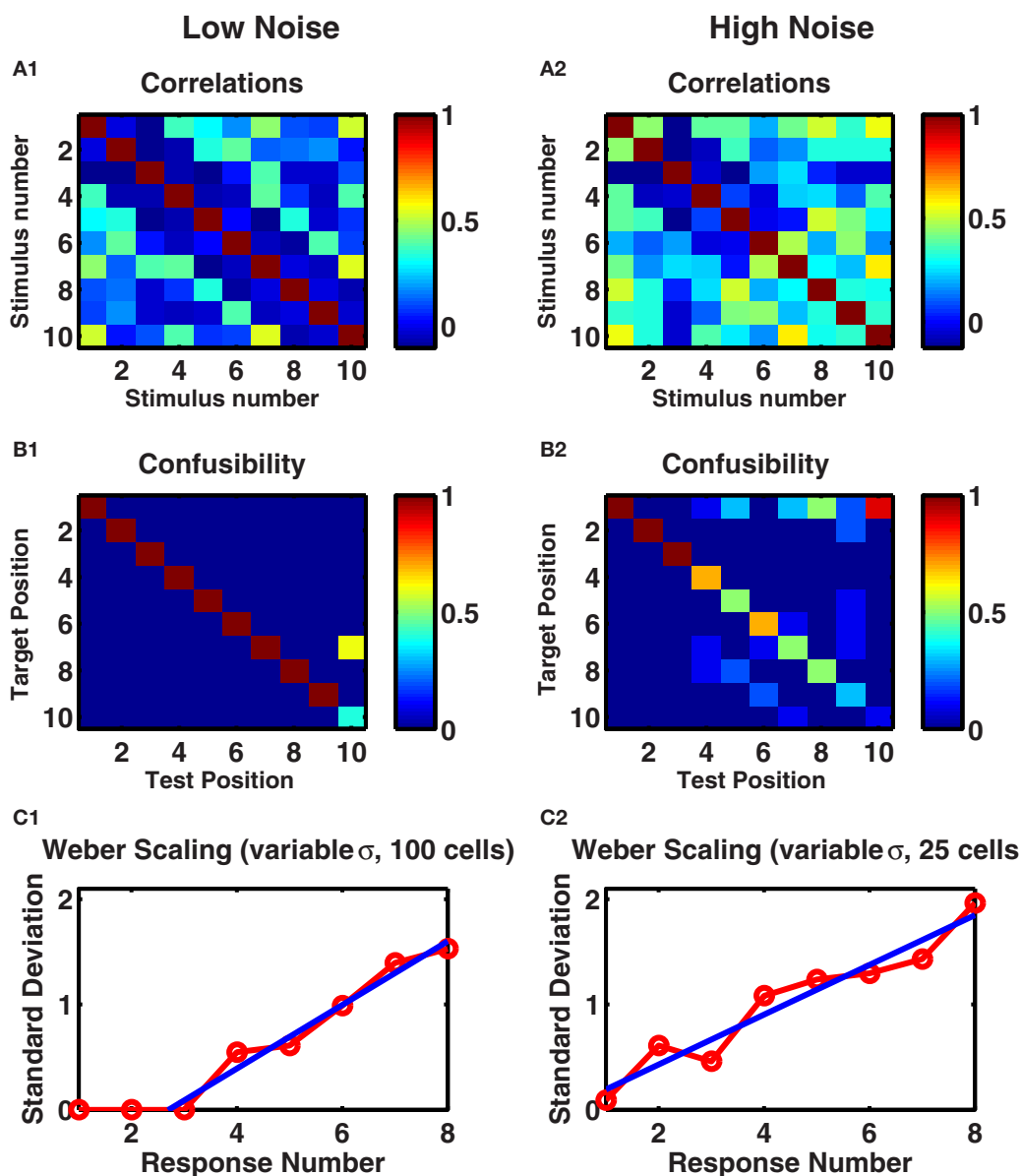
FIGURE 2 | A randomly connected recurrent network with depressing synapses counts identical stimuli. (A1,A2) Correlations between mean post-stimulus firing rates of all cell-groups as a function of stimulus number. **(B1,B2)** The confusability matrix indicates the probability of the network activity being most like a given target template following a given number of successive stimuli. Target position corresponds to each of the ten successive stimuli whose mean network activity was evaluated on ten preliminary trials. Recall position denotes each of ten successive stimuli on test trials—following each stimulus, the network activity was measured and compared with target stimuli. Mean of ten trials presented. Color scale: red = 100% correct, green = 50% correct, blue = 0% correct. **(A1,B1)** Internal noise, σ = 0.002. **(A2,B2)** Increased internal noise, σ = 0.005. **(C1,C2)** Standard deviations in the target position as a function of test position. Ten trials of each of ten networks, with different levels of noise, $0.1 < \sigma < 0.3$ in each trial. **(C1)** Network of 100 cell groups. Fitted line to points 3–8 is $y = 0.30x - 0.82$. **(C2)** Network of 25 cells. Fitted line to all points $y = 0.24x - 0.04$. **(C1,C2)** Straight line fits have higher adjusted r-square values than polynomial fits to $y(x)$ or $x(y)$, suggesting Weber's Law holds.

of $1/\sqrt{N_E}$. Similarly, near identical behavior was produced when the number of cell-groups was reduced, given the appropriate scaling of noise, so that a network with $N_E = 25$ and $\sigma = 0.001$ produced as reliable behavior as a network with $N_E = 100$ and $\sigma = 0.002$. However, when the number of excitatory cell-groups was reduced too much (for example, for $N_E < 15$) then, with current network parameters and random connections, the network would cycle through a small number of 2–4 discrete states so its ability to count inputs would be severely limited.

The effect of network size can be seen in **Figure 2C2**, in which we reproduce the analyses leading to **Figure 2C1**, but with the smaller network of 25 cell-groups. In this case, given the identical range of noise used, more errors occur at any stimulus number, so that even the response to the first stimulus is not completely

reliable. The standard deviation of the outputs of 10 such networks is statistically indistinguishable from a straight line through the origin, reproducing Weber's Law of scaling (see Discussion).

STIMULUS DURATION

Our network is not an integrator, but relies upon synaptic depression, which has a fixed time constant, to reduce the stability of active states. Therefore, it was not clear whether continuously applied stimuli of fixed durations could have the same effect on network activity as multiple, spaced individual stimuli. To test whether the same network could be responsive to stimulus duration, we reset the network following a range of stimuli of different durations then analyzed the resulting activity. The results in **Figure 3**, demonstrate the ability of the network to produce a response that is duration-dependent. Seven distinct states of activity are produced in the example network displayed (six if one excludes the unresponsive state following very short stimuli). Interestingly, the tuning curves of individual neurons differ from their tuning to numerosity—they are much broader and more of them are monotonic (**Figure 3B**).

STIMULUS INTENSITY

We assessed whether the same random network could produce resultant activity that depended on the strength of a fixed duration input current. Results of increasing stimulus strength are similar to those of increased duration in that tuning curves are broader and more monotonic. Interestingly, this is in line with electrophysiological recordings of activities of numerosity-tuned neurons in primates (Nieder and Merten, 2007). Given the broader tuning curves, many pairs of stable activity states were highly correlated (**Figure 4C**) but in the example shown, all 9 distinct stimulus amplitudes, ranging over a factor of five, were successfully encoded in distinct network states, with 100% reliability (**Figure 4D**).

DIFFERENTIATING NUMBER, DURATION AND INTENSITY OF STIMULI

A perfect integrator would produce a network state-dependent on the product of number, duration and intensity of stimuli. Indeed, one could argue that a drawback to the applicability of the perfect integrator to most sensory tasks is its inability, in the absence of other feedback mechanisms (Machens et al., 2005; Miller and Wang, 2006) to distinguish between number, duration and intensity of stimuli. Moreover, such integrators, as possessed by the head-direction system, or occulomotor system, typically require networks with highly specified architectures and often considerable fine-tuning of parameters. In our formalism, with randomly connected units, the network is robust, because groups of cells are individually bistable. In this manner the network resembles the discrete integrator (Koulakov et al., 2002; Goldman et al., 2003). However, since the connections are
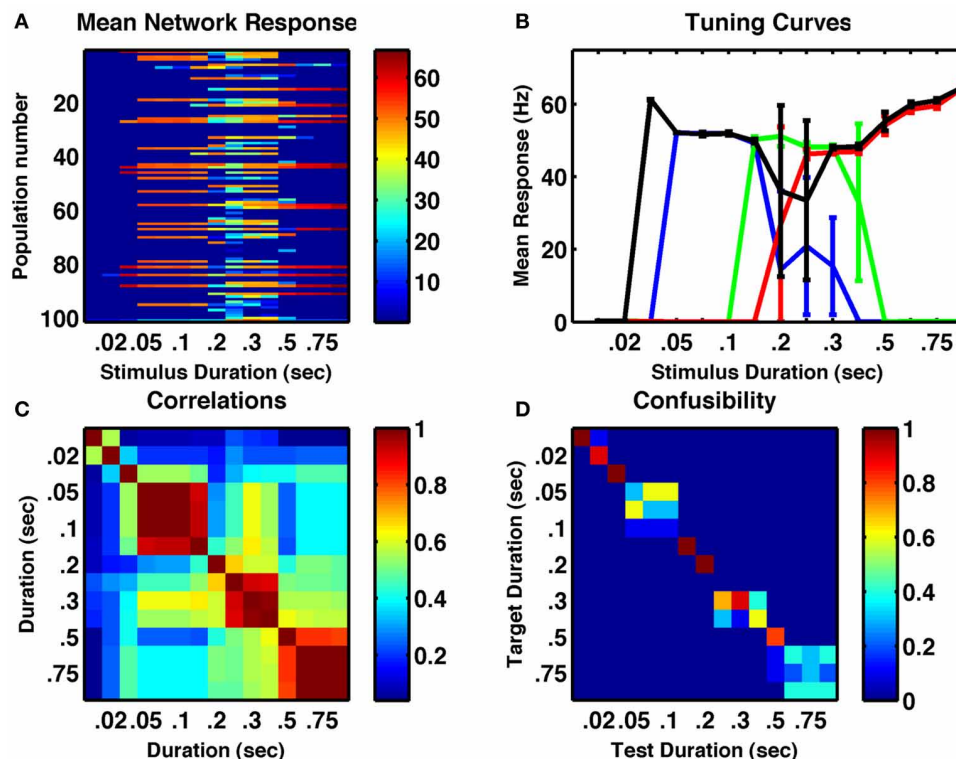
FIGURE 3 | A randomly connected network with depressing synapses can encode stimulus duration. (A) Mean response for all cell-groups following a single stimulus as a function of stimulus duration. Color indicates firing rate. **(B)** Responses of four example cell-groups indicate broad tuning.

(C) Correlation between network firing rates of cell-groups to different stimulus durations. **(D)** The confusability matrix (described in **Figure 1**) indicates the network can differentiate stimulus duration into seven completely distinct categories. Internal noise, $\sigma = 0.002$.
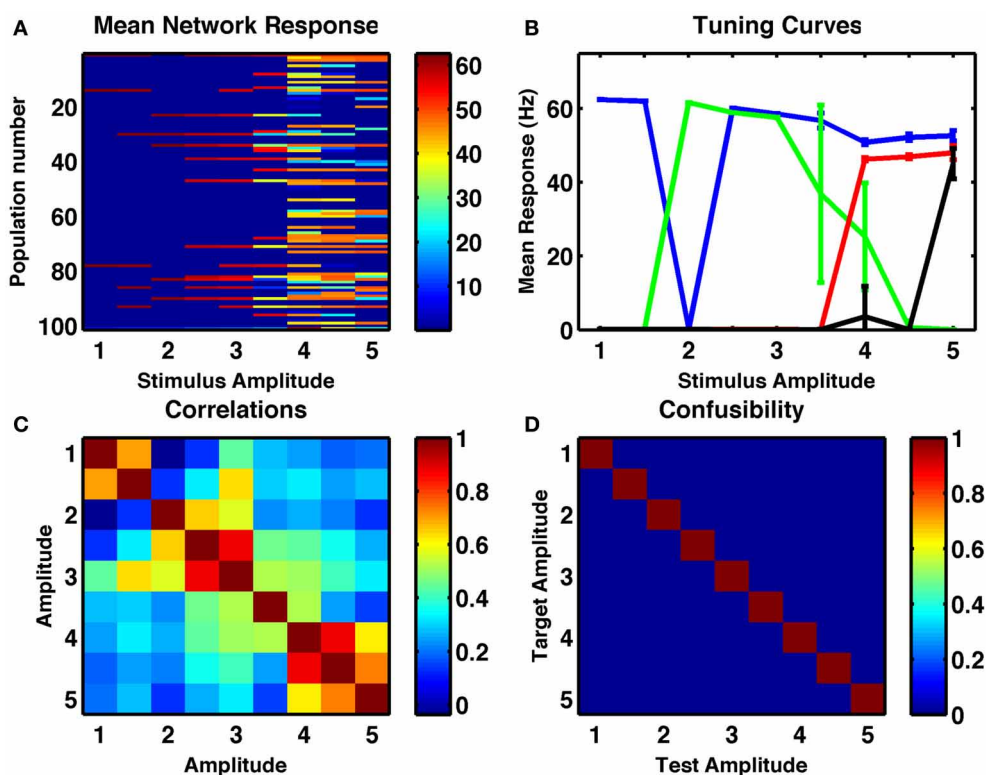
FIGURE 4 | A randomly connected network with depressing synapses can encode stimulus amplitude. (A) Mean response for all cell-groups following a single stimulus as a function of stimulus amplitude, ranging in steps of 0.5 from 1 to 5. Color indicates firing rate. **(B)** Responses of four example cell-groups indicate broad tuning to stimulus amplitude. **(C)** Correlation between network firing rates of cell-groups to different stimulus amplitudes. **(D)** The confusability matrix (described in **Figure 1**) indicates the network can differentiate stimulus amplitude into nine completely distinct categories. Internal noise, $\sigma = 0.002$.

random and not tuned to produce the one-dimensional line of stable points typical of an integrator, the network is unlikely to respond to changes in duration, amplitude and number of stimuli in qualitatively the same manner, as does an integrator. Rather, the stable activity on the randomly connected network appears to follow a high-dimensional, distributed representation—different bistable groups can switch on or off with different combinations of other bistable groups, without a systematic order to the switching. Therefore, it is plausible that multiple feature combinations of the stimulus could be separately encoded.

To test the ability of the network to represent multiple stimulus features, we first, within a single network, applied trains of transient stimuli of varying durations and constant amplitude. If the network were acting as an integrator, then it would respond to total stimulus time, such that a doubling of the duration combined with halving of the number of stimuli would result in the same network activity. However, we found this not to be the case (**Figures 5A,B**). Indeed, we analyzed the network's activity following sequences of up to 8 identical transient stimuli, with six different stimulus durations ranging from 0.05 to 0.3 s. We found for the intermediate stimulus duration of 0.15 s that not only was a unique, reliably different activity state produced following each of the eight successive stimuli, but also all 8 states were uniquely produced by that particular stimulus

duration and distinct from any states produced by any number of successive stimuli with either longer or shorter durations (**Figure 6A**).

An integrator would also respond to the product of amplitude and number of stimuli, or amplitude and duration of a single stimulus. However, the randomly coupled network produces distinct responses to trains of a few high-amplitude stimuli and many low-amplitude stimuli, as well as to intermediate combinations when all combinations have the same product of amplitude and number (**Figures 5C,D**). Moreover, when analyzing the network's activity following sequences of up to eight transient stimuli of constant duration, with seven different amplitudes (in the range 0.5–2.0) we found a very low likelihood for sequences with different amplitudes to be confused with each other and all 8 states following stimuli of intermediate amplitudes to be 90 or 100% correctly identified by both number and amplitude of stimuli (**Figure 6B**).

Figure 6C further indicates the distinctiveness of network response to stimuli of different amplitudes versus of different durations. Following a single transient stimulus, each of five different stimulus amplitudes in the range 1.0–2.0 produces either 3 or 4 different activity states that depend on stimulus duration. These states are both distinct from each other and distinct from any state produced by another stimulus amplitude (**Figure 6C**).
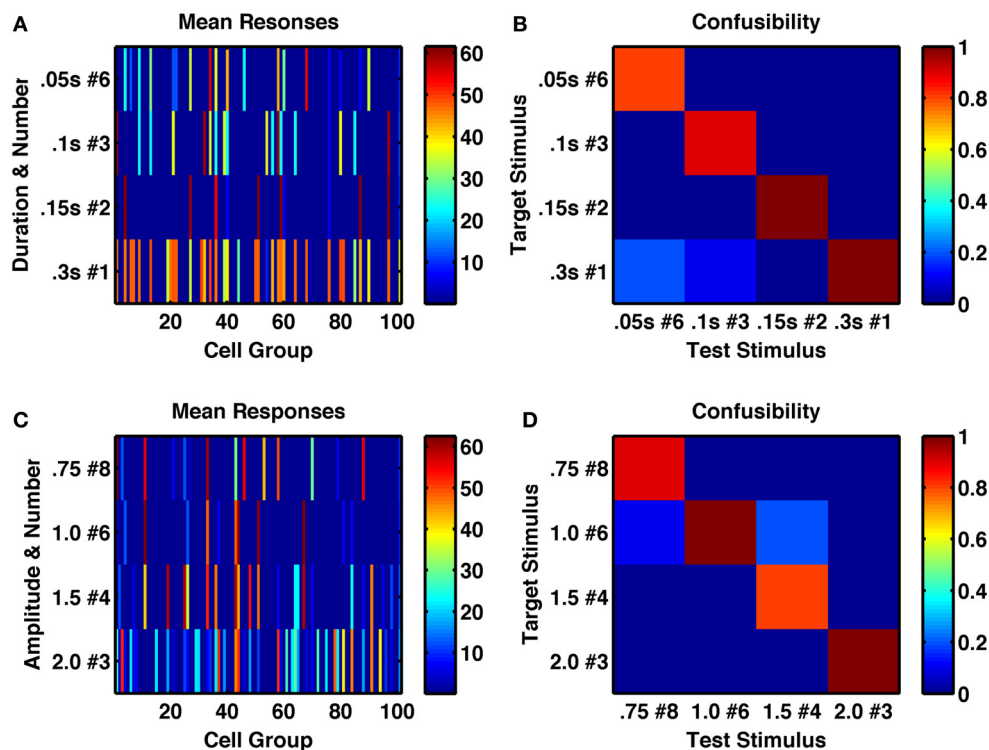
FIGURE 5 | A randomly connected network with depressing synapses produces distinctive responses to stimulus duration, amplitude and number of repetitions. (A) Mean network response, with color indicating firing rate of each cell-group, following the 6th of a series of 0.05 s stimulations (row 1), the 3rd of a series of 0.1 s stimulations (row 2), the 2nd of a series of 0.15 s stimulations (row 3) and a single 0.3 s stimulus, such that all stimulus combinations produce 0.3 s of total current (amplitude 1.5,

$\sigma = 0.002$). **(B)** Confusability matrix between the four types of stimulus combination of **(A)**, indicating the network's response is distinct to each stimulus combination. **(C)** Mean network response as in **(A)** to series of 8, 6 4, and 4 stimuli respectively of different amplitudes 0.75, 1.0, 1.5, and 2.0 (such that the product is constant at 6.0). Each stimulus has duration of 0.25 s, $\sigma = 0.002$. **(D)** Confusability matrix between responses to the four stimulus combinations of **(C)** demonstrates the responses are distinct.

We finally produced a $6 \times 3 \times 3$ array of stimuli with any combination of number ($N = 1 - 6$), duration ($T = 0.1$ s, 0.2 or 0.3 s) and intensity ($I = 1, 2,$ or 3) of applied current pulses. We assessed how network activity depended on these stimulus combinations. **Figure 6D** demonstrates that for a large number (27) of these stimulus combinations, the network activity is reliably propelled into a distinct state, unique to that single combination of duration, amplitude and number of stimuli. Since the stimuli are all constant, equal currents to all excitatory cell-groups in the network, the evolution of activity states depends entirely on the random cross-connections between cell-groups and the temporal dynamics of intra-group and inter-group synaptic transmission.

NETWORKS WITHOUT DEPRESSING SYNAPSES

When synaptic depression is removed from these networks—and static release probability is optimally tuned to allow for multiple stable activity states—the counting behavior of the network disappeared (**Figures 7A,C**). That is, successive stimuli simply reproduced the same state. The number of states produced by different durations and amplitudes of stimuli was reduced from 7–8 to 2–4 (**Figures 7B,D**). Also, under the same low-noise conditions as the networks shown in **Figures 1–6**, the reliability of responses to identical stimuli was greatly reduced. In fact,

with constant amplitude and varying duration, no states were distinctly produced by a single subset of stimuli.

In summary, it is short-term depression in the recurrent connections of bistable groups that produces itinerancy in the network states. Such itinerancy with consecutive stimuli enables the network to possess a counting behavior and to produce numerosity-tuned cells. The same synaptic depression imparts a preferred stimulus amplitude and duration for activation of a cell-group, increasing the number and reliability of amplitude-specific and duration-specific states.

DISCUSSION

Bistability relies upon positive feedback, which can arise from cell-intrinsic currents (Hounsgaard et al., 1984; Rinzel, 1985; Booth and Rinzel, 1995) or from network feedback (Kleinfeld et al., 1990; Camperi and Wang, 1998; Wang, 1999, 2001; Koulakov et al., 2002). Synaptic facilitation is a positive feedback mechanism in circuits of reciprocally connected excitatory cells, since the greater the mean firing rate, the greater the effective connection strength, further amplifying the excitatory input beyond that produced by the increased spike rate alone. This property of synaptic facilitation enhances the stability of memory states and renders them more robust to distractors (Itskov et al., 2011).
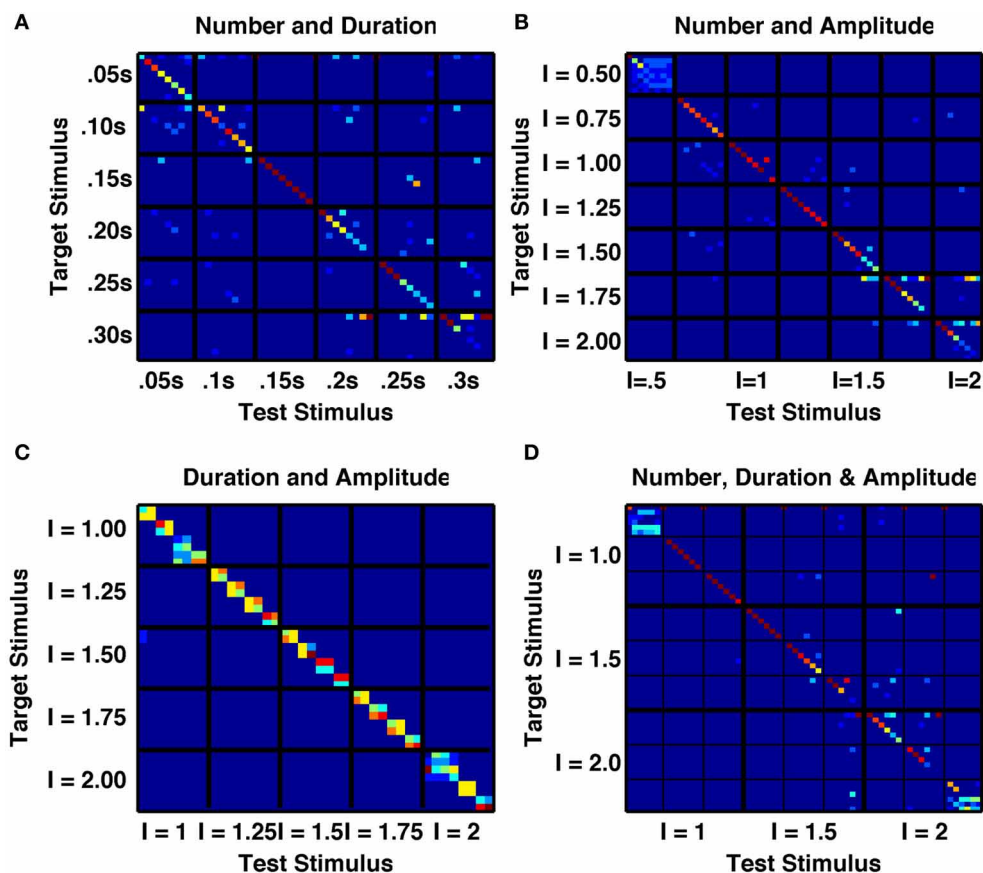
FIGURE 6 | A randomly connected network with depressing synapses produces distinct responses to multiple stimulus features.
(A) Confusability matrix between the network's activity states following 1–8 transient stimuli of one of six durations from 0.05 s to 0.30 s (48 stimulus combinations in total). 15 combinations are perfectly distinct.
(B) Confusability matrix between the network's activity states following 1–8 transient stimuli of one of seven amplitudes from 0.5 to 2.0 (56 stimulus combinations in total). 17 combinations are perfectly distinct.
(C) Confusability matrix between the network's activity states following a

single transient stimulus of one of eight durations from 0.05 s to 0.4 s and one of five amplitudes from 1.0 to 2.0 (32 stimulus combinations in total). Although no individual combination is perfectly distinguished from all others, 18 distinct states are apparent, with the majority of states responding to a single amplitude and two durations. **(D)** Confusability matrix between the network's activity states following 1–6 transient stimuli of one of three durations (0.1 s, 0.2 s, 0.3 s) and one of three amplitudes (1.0, 1.5, 2.0). 23 stimulus combinations are perfectly distinguished and over 25 distinct activity states are produced.

Other forms of positive feedback, such as depolarization-induced suppression of inhibition (DSI), which depends on activity in the post-synaptic cell, can similarly produce robustness in recurrent memory networks (Carter and Wang, 2007).

Conversely, depressing synapses in a self-exciting circuit produce negative feedback, by reducing the effective synaptic strength of the outputs of the most active cells. Such negative feedback reduces the stability of the attractor states produced by positive feedback. This effect has been demonstrated in a system known as the ring attractor, an example of a perfect integrator (Song and Wang, 2005), which in the absence of dynamic synapses can produce a "bump" of population activity in a marginal state. Once the bump has formed at a given location on the "ring" it can remain at that location so form the basis of a spatial memory. However, the stationary "bump" can be rendered unstable by synaptic depression and be replaced by one of two possible moving "bump" states with fixed velocity (York and van Rossum, 2009). Such an

effect is similar to that produced by intrinsic adaptation currents within the excitatory neurons of the ring attractor, which result in a pitchfork bifurcation as the single stationary state is replaced by two oppositely directed constant velocity states, whose absolute velocity increases as the underlying conductance increases (Ben-Yishai et al., 1997; Hansel and Sompolinsky, 1998; Laing and Longtin, 2001; Tegnèr et al., 2002).

In the randomly connected circuits that we simulate, synaptic depression in strong recurrent excitatory synapses also has the same effect on these excitatory cells as an adaptation current. Following the initial burst of excitatory input, the dynamic weakening of synaptic strength while vesicles need to be replaced causes a reduction in post-synaptic excitatory input, which affects the post-synaptic cell just as would an activity-dependent intrinsic inhibitory current. Thus, it is possible that synaptic depression could produce similar results to that of an adaptation current in successful models of binocular rivalry based on bistability
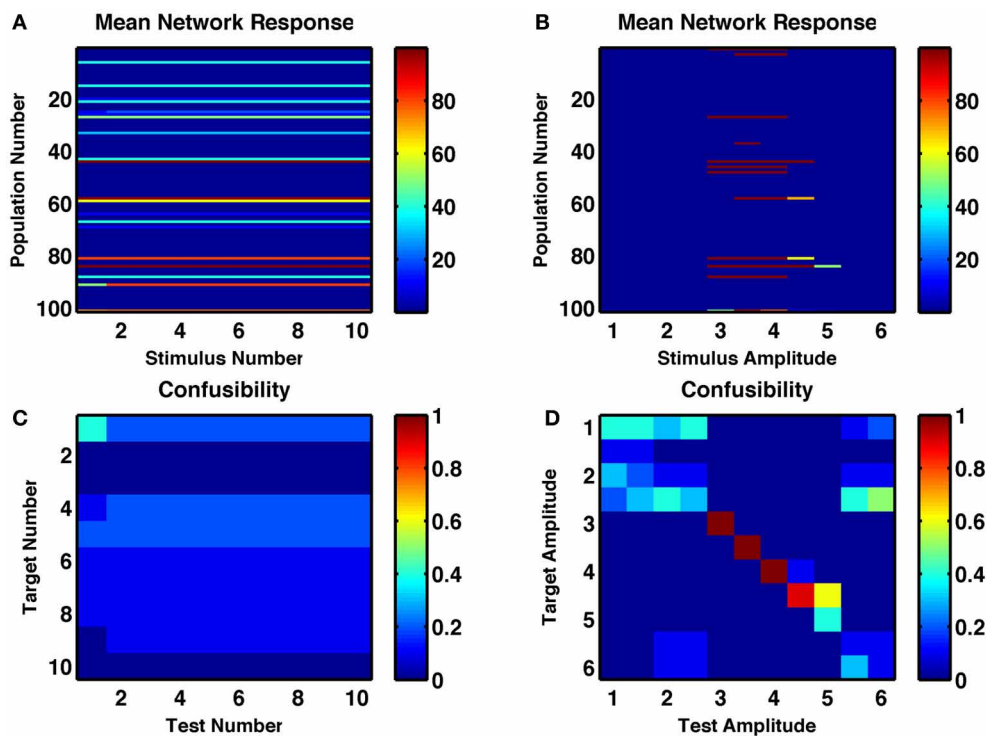
FIGURE 7 | A random recurrent network without depressing synapses shows no counting behavior and weakly tuned amplitude selectivity. (A,B) Post-stimulus firing rates of each group of neurons in response to **(A)** repetitions of an optimal-strength stimulus and **(B)** a single stimulus of varying amplitude. **(C,D)** Confusability matrix indicated how distinguishable the network responses are for the successive repetitions of the same stimulus **(C)** or for a single stimulus of different amplitudes **(D)**.

between groups of neurons (Moreno-Bote et al., 2007; Theodoni et al., 2011).

A randomly connected network of bistable neurons was shown to produce a diversity of neural responses (Rigotti et al., 2010) with neurons possessing mixed selectivity to conjunctions of stimulus features. In that work, different combinations of stimuli or inputs produced the different resulting distributions of stable network activity, allowing for appropriate responses in cognitive tasks. Here, we show that with the addition of depressing synapses, a similar network produces a diversity of responses to different dynamic features of a single stimulus of equal strength to all cells.

The randomly connected network responds differently from neural integrators, whether continuous (Seung, 1996; Miller et al., 2003; Song and Wang, 2005) or discrete (Koulakov et al., 2002; Goldman et al., 2003). For an integrator, increased signal amplitude affects the system in qualitatively the same manner as increased signal duration. The reason for the difference is that integrators are designed to have a one-dimensional sequence of stable fixed points—or a continuous line of fixed points representing a marginal phase (Ben-Yishai et al., 1995), sometimes called a line attractor (Seung, 1996)—whereas the randomly connected network is inherently of high dimensionality (Rigotti et al., 2010). Thus, even when an integrator either inherently (Compte et al., 2000; Song and Wang, 2005) or through its connections to a second output layer (Verguts and Fias, 2004), produces

non-monotonic, "peaked" tuning curves, the responses to number, duration and stimulus amplitude are not separable. That is, an integrator's activity following a given number of counts of one stimulus is identical to that following more counts of a weaker stimulus, or of a shorter duration stimulus—of course, in many situations other than counting, such integration is the desired network response (Zhang, 1996; Samsonovich and McNaughton, 1997; Romo et al., 1999; Seung et al., 2000; Song and Wang, 2005).

In many experiments analyzing numerosity coding, both behavioral (Merten and Nieder, 2009) and neural (Nieder and Miller, 2003) responses produce two features suggestive of logarithmic coding. First, errors are skewed, with a longer tail toward stimulus values higher than the stimulus producing peak response. Second, the standard deviation of number estimates—here calculated via the trial-to-trial variability in the network's estimate of stimulus number for each fixed actual number of stimuli—scales linearly with number of stimuli, a scaling known as Weber's Law (Weber, 1851). Our network does not exhibit the observed skew in neural responses, in particular because there is a tendency when errors are made, for the random attractor states visited to be more like the first attractor state (so an incorrect response of "one" is the most common). However, if we incorporate trial-to-trial variability in the level of noise (**Figure 2C**) then a Weber scaling is observed—errors become more likely, linearly with increasing number. Thus, the information pertaining to the encoded number, as contained within the distributed

representation of these networks, degrades in the expected manner, but it is likely a separate "readout" network of cells is needed to produce all the features observed in neural recordings. Such a "readout" network could also combine the different representations of number arising from stimuli of different properties into a single "pure number" representation—that is, it would produce pattern completion after this initial step of pattern separation.

Recent experiments have demonstrated associative forms of short-term plasticity (Brenowitz and Regehr, 2005; Erickson et al., 2010), which is more powerful, since it can be synapse-specific rather than cell-specific, so has greater information carrying capacity. Such associative-STP has been shown to be capable of temporarily coupling together specific pairs of bistable neural groups, so could form the basis for memory of sequences of discrete items (Botvinick and Watanabe, 2007; Miller and Wingfield, 2010).

In summary, we have shown that depression can destabilize discrete activity states and in so doing enables the network activity to change through repetitions of identical stimuli. Therefore, such networks could be of value in providing a basis for counting and for memory of sequences (Botvinick and Plaut, 2006; Botvinick and Watanabe, 2007). Indeed, our ongoing work suggests that memories of discrete sequences could be maintained in a network, which combines such effects of synaptic depression (**Figures 1–2**) with associative short-term plasticity (Erickson et al., 2010; Miller and Wingfield, 2010).

REFERENCES

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224.

Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 92, 3844–3848.

Ben-Yishai, R., Hansel, D., and Sompolinsky, H. (1997). Traveling waves and the processing of weakly tuned inputs in a cortical network module. *J. Comput. Neurosci.* 4, 57–77.

Booth, V., and Rinzel, J. (1995). A minimal, compartmental model for a dendritic origin of bistability of motoneuron firing patterns. *J. Comput. Neurosci.* 2, 299–312.

Botvinick, M., and Watanabe, T. (2007). From numerosity to ordinal rank: a gain-field model of serial order representation in cortical working memory. *J. Neurosci.* 27, 8636–8642.

Botvinick, M. M., and Plaut, D. C. (2006). Short-term memory for serial order: a recurrent neural network model. *Psychol. Rev.* 113, 201–233.

Brenowitz, S. D., and Regehr, W. G. (2005). Associative short-term synaptic plasticity mediated by endocannabinoids. *Neuron* 45, 419–431.

Brunel, N., and Nadal, J. P. (1998). Modeling memory: what do we learn from attractor neural networks? *C. R. Acad. Sci. III* 321, 249–252.

Camperi, M., and Wang, X. J. (1998). A model of visuospatial short-term memory in prefrontal cortex: recurrent network and cellular bistability. *J. Comput. Neurosci.* 5, 383–405.

Carter, E., and Wang, X. J. (2007). Cannabinoid-mediated disinhibition and working memory: dynamical interplay of multiple feedback mechanisms in a continuous attractor model of prefrontal cortex. *Cereb. Cortex* 17(Suppl. 1), i16–i26.

Compte, A., Brunel, N., Goldman-Rakic, P. S., and Wang, X. J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* 10, 910–923.

Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience.* Cambridge, MA: MIT Press.

Erickson, M. A., Maramara, L. A., and Lisman, J. (2010). A single brief burst induces GluR1-dependent associative short-term potentiation: a potential mechanism for short-term memory. *J. Cogn. Neurosci.* 22, 2530–2540.

Funahashi, S., Bruce, C. J., and Goldman-Rakic, P. S. (1991). Neuronal activity related to saccadic eye movements in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 65, 1464–1483.

Goldman, M. S., Levine, J. H., Major, G., Tank, D. W., and Seung, H. S. (2003). Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron. *Cereb. Cortex* 13, 1185–1195.

Hansel, D., and Sompolinsky, H. (1998). "Modeling feature selectivity in local cortical circuits," in *Methods in Neuronal Modeling: From Synapse to Networks*, eds C. Koch and I. Segev (Cambridge, MA: MIT press), 499–567.

Hebb, D. O. (1949). *Organization of Behavior.* New York, NY: Wiley.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. U.S.A.* 81, 3088–3092.

Hounsgaard, J., Hultborn, H., Jespersen, B., and Kiehn, O. (1984). Intrinsic membrane properties causing a bistable behaviour of alpha-motoneurones. *Exp. Brain Res.* 55, 391–394.

Itskov, V., Hansel, D., and Tsodyks, M. (2011). Short-term facilitation may stabilize parametric working memory trace. *Front. Comput. Neurosci.* 5:40. doi: 10.3389/fncom.2011.00040

Kleinfeld, D., Raccuia-Behling, F., and Chiel, H. J. (1990). Circuits constructed from identified Aplysia neurons exhibit multiple patterns of persistent activity. *Biophys. J.* 57, 697–715.

Koulakov, A. A., Raghavachari, S., Kepecs, A., and Lisman, J. E. (2002). Model for a robust neural integrator. *Nat. Neurosci.* 5, 775–782.

Laing, C. R., and Longtin, A. (2001). Noise-induced stabilization of bumps in systems with long-range spatial coupling. *Physica D.* 160, 149–172.

Lorente de Nó, R. (1933). Vestibulo-ocular reflex arc. *Arch. Neurol. Psychiatry* 30, 245–291.

Machens, C. K., Romo, R., and Brody, C. D. (2005). Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* 307, 1121–1124.

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382, 807–810.

Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5323–5328.

Merten, K., and Nieder, A. (2009). Compressed scaling of abstract numerosity representations in adult humans and monkeys. *J. Cogn. Neurosci.* 21, 333–346.

Miller, P., Brody, C. D., Romo, R., and Wang, X. J. (2003). A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cereb. Cortex* 13, 1208–1218.

Miller, P., and Wang, X. J. (2006). Discrimination of temporally separated stimuli by integral feedback control. *Proc. Natl. Acad. Sci. U.S.A.* 103, 201–206.

Miller, P., and Wingfield, A. (2010). Distinct effects of perceptual quality on auditory word recognition, memory formation and recall in a neural model of sequential memory. *Front. Syst. Neurosci.* 4:14. doi: 10.3389/fnsys.2010.00014

Moreno-Bote, R., Rinzel, J., and Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *J. Neurophysiol.* 98, 1125–1139.

Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11860–11865.

Nieder, A. (2013). Coding of abstract quantity by 'number neurons' of the primate brain. *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.* 199, 1–16.

Nieder, A., and Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *J. Neurosci.* 27, 5986–5993.

Nieder, A., and Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37, 149–157.

Nordlie, E., Gewaltig, M. O., and Plesser, H. E. (2009). Towards reproducible descriptions of neuronal network models. *PLoS Comput. Biol.* 5:e1000456. doi: 10.1371/journal.pcbi.1000456

Rigotti, M., Rubin, D. B., Wang, X. J., and Fusi, S. (2010). Internal

representation of task rules by recurrent dynamics: the importance of the diversity of neural responses. *Front. Comput. Neurosci.* 4:24. doi: 10.3389/fncom.2010.00024

Rinzel, J. (1985). Excitation dynamics: insights from simplified membrane models. *Fed. Proc.* 44, 2944–2946.

Romo, R., Brody, C. D., Hernández, A., and Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399, 470–474.

Samsonovich, A., and McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *J. Neurosci.* 17, 5900–5920.

Seung, H. S. (1996). How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. U.S.A.* 93, 13339–13344.

Seung, H. S., Lee, D. D., Reis, B. Y., and Tank, D. W. (2000). Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron* 26, 259–271.

Song, P., and Wang, X. J. (2005). Angular path integration by moving "hill of activity": a spiking neuron model without recurrent excitation of the head-direction system. *J. Neurosci.* 25, 1002–1014.

Tegnèr, J., Compte, A., and Wang, X. J. (2002). The dynamical stability of reverberatory circuits. *Biol. Cybern.* 87, 471–481.

Theodoni, P., Panagiotaropoulos, T. I., Kapoor, V., Logothetis, N. K., and Deco, G. (2011). Cortical microcircuit dynamics mediating binocular rivalry: the role of adaptation in inhibition. *Front. Hum. Neurosci.* 5:145. doi: 10.3389/fnhum.2011.00145

Tudusciuc, O., and Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14513–14518.

Tudusciuc, O., and Nieder, A. (2009). Contributions of primate prefrontal and posterior parietal cortices to length and numerosity representation. *J. Neurophysiol.* 101, 2984–2994.

Verguts, T., and Fias, W. (2004). Representation of number in animals and humans: a neural model. *J. Cogn. Neurosci.* 16, 1493–1504.

von Gersdorff, H., and Matthews, G. (1997). Depletion and replenishment of vesicle pools at a ribbon-type synaptic terminal. *J. Neurosci.* 17, 1919–1927.

Wang, X. J. (1999). Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. *J. Neurosci.* 19, 9587–9603.

Wang, X. J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* 24, 455–463.

Weber, E. H. (1851). *Annotationes Annatomicae et Physiologicae.* Leipsig: CF Koehler.

York, L. C., and van Rossum, M. C. (2009). Recurrent networks with short term synaptic depression. *J. Comput. Neurosci.* 27, 607–620.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensembles: a theory. *J. Neurosci.* 16, 2112–2126.

A model of microsaccade-related neural responses induced by short-term depression in thalamocortical synapses

Wu-Jie Yuan[1,2], Olaf Dimigen[3], Werner Sommer[3] and Changsong Zhou[1]*

[1] Department of Physics, Institute of Computational and Theoretical Studies, Centre for Non-linear Studies and the Beijing-Hong Kong-Singapore Joint Centre for Nonlinear and Complex Systems (Hong Kong), Hong Kong Baptist University, Kowloon Tong, Hong Kong, China

[2] College of Physics and Electronic Information, Huaibei Normal University, Huaibei, China

[3] Department of Psychology, Humboldt University at Berlin, Berlin, Germany

Microsaccades during fixation have been suggested to counteract visual fading. Recent experiments have also observed microsaccade-related neural responses from cellular record, scalp electroencephalogram (EEG), and functional magnetic resonance imaging (fMRI). The underlying mechanism, however, is not yet understood and highly debated. It has been proposed that the neural activity of primary visual cortex (V1) is a crucial component for counteracting visual adaptation. In this paper, we use computational modeling to investigate how short-term depression (STD) in thalamocortical synapses might affect the neural responses of V1 in the presence of microsaccades. Our model not only gives a possible synaptic explanation for microsaccades in counteracting visual fading, but also reproduces several features in experimental findings. These modeling results suggest that STD in thalamocortical synapses plays an important role in microsaccade-related neural responses and the model may be useful for further investigation of behavioral properties and functional roles of microsaccades.

Keywords: short-term depression, microsaccades, feedforward network, visual fading, fixation

1. INTRODUCTION

When the eyes fixate at a stationary object, they are never completely motionless, but perform involuntary, very small eye movements. These fixational eye movements are composed of three different types of movement: tremor, microsaccades, and drift. Tremor is an aperiodic, high-frequency fixational eye movement with the smallest amplitude of these three types of fixational eye movements. Microsaccades are involuntary jerk-like fixational eye movements. Drift is a typical fixational eye movement taking place between microsaccades with the slowest velocity of all the three types. Microsaccades are the largest and fastest fixational eye movements. It has been experimentally observed that microsaccades cause more variability in neuronal responses than both tremor and drift (Gur et al., 1997; Martinez-Conde, 2006). The most prominent contribution to fixational eye movements is generated by microsaccades (Rolfs, 2009). Therefore, both experimental and theoretical works have mainly focused on the role of microsaccades during fixation.

Over the past decade, the behavioral properties and functional roles of microsaccades have been widely investigated (for reviews, see Martinez-Conde et al., 2004, 2009; Rolfs, 2009). Importantly, it was found that the visual world quickly fades from view in the absence of fixational eye movements (Ditchburn and Ginsborg, 1952; Riggs and Ratliff, 1952). This suggests that microsaccades play an important functional role in counteracting visual fading during fixation (Ditchburn and Ginsborg, 1952; Martinez-Conde, 2006). Recently, the mechanism of microsaccades for

counteracting perceptual fading has received much research interest. Several studies have assumed that microsaccades refresh retinal images by moving the receptive fields of less adapted photoreceptors over stationary stimuli, thereby preventing perceptual fading (Ditchburn and Ginsborg, 1952; Martinez-Conde, 2006). However, the locus and properties of this retinal adaptation are not well known.

Therefore, the mechanism of microsaccades for counteracting visual fading is not well understood. This is largely because the neural correlates responsible for brain responses to microsaccades are unknown. So far, the brain responses due to microsaccades have been widely reported at different levels—from neuronal activities (Bair and O'keefe, 1998; Leopold and Logothetis, 1998; Martinez-Conde et al., 2002; Martinez-Conde, 2006) to electroencephalogram (EEG) (Yuval-Greenberg et al., 2008; Dimigen et al., 2009) and functional magnetic resonance imaging (fMRI) (Hsieh and Tse, 2009; Tse et al., 2010)—in a number of brain areas. It's mostly found that, microsaccades enhance neuronal firing and therefore raise excitatory response in early visual areas, such as lateral geniculate nucleus (LGN) and primary visual cortex (V1) (Martinez-Conde et al., 2002). Particularly, neural activity in V1 is crucial component for the understanding of visual information processing related to microsaccades.

Previous works (Riggs et al., 1953; Krauskopf, 1957; Sharpe, 1972; Engbert and Mergenthaler, 2006) have suggested that retinal adaptation might be responsible for visual fading in the absence of microsaccades during fixation. However, this

suggestion has not yet been verified directly in experiments. Moreover, the retinal adaptation in the absence of microsaccades has not been successfully described by using physiologically realistic model (Donner and Hemilä, 2007). Although some studies have found that microsaccades can increase the neural activity in the retina (Armington and Bloom, 1974; Greschner et al., 2002), the enhanced neural responses, which are the neural correlates of the perception of visibility during fading, have been only tested in LGN and V1 but not in the retina (Martinez-Conde et al., 2002; Martinez-Conde, 2006). While retinal adaptation cannot be excluded to contribute to visual fading in the absence of microsaccades, it is possible that the neural adaptation related to visual fading may take place at some stage between retina and early visual areas.

Over the past three decades, physiological studies have shown adaptation phenomena affecting neural activity in V1. Carandini et al. (2002) suggested three possible adaptation mechanisms: synaptic depression, intracortical inhibition and intrinsic cellular mechanisms. Of these three mechanisms, synaptic depression is well suited to explain the marked differences between the responses to transient and consecutive stimuli (Chance et al., 1998). Recently, a synaptic depression, short-term depression (STD), has been extensively found at thalamocortical synapses from LGN to V1 *in vitro* (Stratford et al., 1996; Bannister et al., 2002) and *in vivo* (Boudreau and Ferster, 2005) in the cat. Previously, network models of V1 neurons with the thalamocortical synaptic depression have been used to successfully explain some visual phenomena (Chance et al., 1998; Chance and Abbott, 2001; Carandini et al., 2002), including temporal phase shifts, spatial-phase adaptation, contrast saturation, cross-orientation suppression, and so on. However, the synaptic depression has not yet been used in a thalamocortical network to investigate the roles of microsaccades.

In this paper, we proposed an alternative explanation for visual fading by introducing STD in the thalamocortical system, without considering possible neural adaptation from retina. We used a computational model to investigate how microsaccades might induce neural responses in V1 by considering STD in thalamocortical synapses from LGN to V1. The adapted synapses subjected to STD can lead to response depression in V1, and induce visual fading because of sustained depression. Therefore, it is possible that the generation of microsaccades serves to counteract STD-induced depression of neuronal activity in order to counteract visual fading. Our model can reproduce several experimental findings of microsaccade-related neural responses (Martinez-Conde et al., 2002; Kagan et al., 2008). These results suggest that STD from LGN to V1 might play an important role in microsaccade-related neural responses, and provide theoretical insight into the understanding of more behavioral properties and functional roles of microsaccades.

2. MATERIALS AND METHODS

2.1. FEEDFORWARD MODEL

In sensory nervous system, substantial information processing can be performed by feedforward networks without considering recurrent connections, including perceptual learning (Tsodyks and Gilbert, 2004). A well-known model of feedforward networks

is that proposed by Poggio et al. (1992) on visual hyperacuity. In our work, extending this previous model, we constructed a simple feedforward network model consisting of two layers corresponding to LGN and V1 with STD in thalamocortical synapses, as shown in **Figure 1**. To focus on the effects of synaptic depression during fixation with microsaccades, we kept our model very simple, and did not consider corticocortical synaptic connections. In visual systems, a neuron sees only a small portion of the visual field. This small area is called the receptive field of the cell. This receptive field leads to a Gaussian tuning function of the mean firing rate of the neurons with respect to the orientation of fixated dot (Nelson et al., 1994; Ferster et al., 1996; Ferster and Miller, 2000; Seriès et al., 2004), which denotes the inputting orientation from fixated dot to the neurons by lateral synaptic inhibition connections (Amari, 1977; Pinault and Deschênes, 1998; Yuan et al., 2006, 2007). Since the input layer LGN consists of a number of Gaussian filters (receptive fields) as described by Poggio et al. (1992) and Tsodyks and Gilbert (2004), the afferent stimuli evoked by the fixated dot are transformed into firing trains in LGN neurons j, which can be described as Poisson spike trains with a time-independent rate R_j following Gaussian profile G_1 in space (shown in **Figure 1**). For the output layer, each V1 neuron i has connections coming from LGN relay neurons j (excitatory) with weights W_{ij} following Gaussian tuning curve G_2 (explained
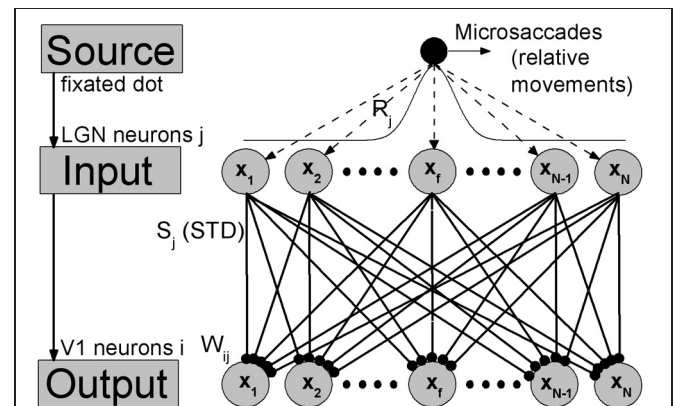
FIGURE 1 | The feedforward network model during fixation with microsaccades. Neurons in LGN and V1 are labeled and arranged by the center positions x_j and x_i of their receptive fields in the ranges from $-L$ to L, respectively. Gaussian filters (receptive fields) in LGN layer transform the afferent stimuli evoked by fixated dot into the inputs with Gaussian firing rate profile: $R_j = G_1(x_j - x_f) = A \exp^{-(x_j-x_f)^2/\sigma_1^2}$. A represents the amplitude of a visual input at fixated-dot position x_f and σ_1 is the width of the tuning curve, which denotes the width of the receptive fields. The output layer V1 is connected to input layer LGN by thalamocortical synapses with synaptic strengths S_j, which are subjected to the synaptic modification: STD. These connecting weights W_{ij} follow the Gaussian tuning curve: $W_{ij} = G_2(x_j - x_i) = \exp^{-(x_j-x_i)^2/\sigma_2^2}$, where $x_j - x_i$ denotes the position difference of receptive field centers between the input neuron j and the output neuron i. The microsaccades during fixation can be regarded as instantaneous relative movements of the fixated dot over LGN with microsaccadic magnitude Δ_M. In order to eliminate the effect of boundary due to limited network scale, the G_1 is extended to a period boundary function, i.e., $G_1(x_j - x_f) = A \exp^{-(x_j-x_f)^2/\sigma_1^2}$ for $|x_j - x_f| < L$; otherwise, $G_1(x_j - x_f) = A \exp^{-(2L-|x_j-x_f|)^2/\sigma_1^2}$.

in the caption of **Figure 1**) (Poggio et al., 1992; Tsodyks and Gilbert, 2004). The model is composed of Integrate-and-Fire neurons with chemical couplings of δ function. The dynamics of the membrane potential V_i of output neuron i in V1 is described by

$$\tau_m \frac{dV_i}{dt} = V_0 - V_i + \sum_{j=1}^{N} g W_{ij} S_j(t)(V_E - V_i)\delta(t - t_{sp}^j). \quad (1)$$

Here, we adopted the same parameter values as those in Abbott et al. (1997) and Chance et al. (1998), which model V1 cells according to empirical observations (Varela et al., 1997). The membrane time constant τ_m equals 30 ms, the resting potential V_0 is -70 mV, and the reversal potentials V_E for all the excitatory synapses are 0 mV. Each V1 neuron i integrates inputs coming from LGN neurons j at spike time t_{sp}^j distributed as Poisson spike trains. When the potential V_i reaches the threshold value -55 mV, the neuron i emits a spike, and then the membrane potential is reset to the relatively high value -58 mV (compared with the resting potential $V_0 = -70$ mV) in order to match experimental recordings (Varela et al., 1997). The parameter g represents the maximal synaptic conductance. The $S_j(t)$ in the thalamocortical synapses from LGN to V1 complies with STD plasticity, which will be described in the following.

　In simulations, N neurons in LGN and V1 are, respectively, spread uniformly in the ranges from $-L$ to L, which denote the physical positions of receptive field centers of these neurons. Compared with the responsive region of neurons induced by the fixated dot, L should be large enough that the new place of fixated dot after microsaccades is far from the boundary neurons. Here, in order to shorten the simulating time, the region from $-L$ to L is chosen as a narrow region with finite L. Meanwhile, to eliminate the effect of boundary due to the chosen narrow region, the input tuning curve G_1 is extended to a period boundary function (see the caption of **Figure 1**). In this way, the value of L does not change qualitatively the results. In our simulation, microsaccades are modeled by instantaneous relative displacement Δ_M of the tuning curve G_1. With suitable scale transformation, the size L and displacement Δ_M can be used to represent realistic range of microsaccades (Martinez-Conde et al., 2009). Here, we take $N = 1000$ neurons and $L = 10$. The main results, however, do not depend on these parameters.

2.2. SHORT-TERM DEPRESSION (STD)

Biophysically, synaptic depression can be regarded as the interaction between two processes, the activity-dependent depletion of the transmitter resources of synaptic vesicles and the slow replenishment of the resources. The depletion process means that the available transmitter is diminished immediately after the presynaptic spike time owing to the release of transmitter. Thus, each time a presynaptic spike arrives at synapse j, the synaptic strength S_j decreases immediately after the spike due to the use of transmitter resources. The depletion of a synapse is usually modeled by a multiplicative factor f (Abbott et al., 1997; Chance et al., 1998; Boudreau and Ferster, 2005):

$$S_j \to f S_j. \quad (2)$$

The parameter f $(0.0 < f < 1.0)$ denotes the ratio of the synaptic resources available immediately after release to those before release, and thereby determines the amount of depression at synapse j induced by each spike (the smaller the parameter f, the stronger the depression). The slow replenishment process can be modeled by exponential recovery from depression (Abbott et al., 1997; Chance et al., 1998):

$$\frac{dS_j}{dt} = \frac{1}{\tau_S}(1 - S_j). \quad (3)$$

The constant parameter τ_S determines the depression recovery time. Combining the Equations (2) and (3), the STD can be described by

$$\frac{dS_j}{dt} = \frac{1}{\tau_S}(1 - S_j) - (1 - f)S_j\delta(t - t_{sp}^j). \quad (4)$$

If the afferent neuron for the synapse j fires a Poisson spike train at rate R_j, the synaptic strength will quickly decrease to the approximate steady state (for a high rate) (Abbott et al., 1997):

$$S_j(ss) = \frac{1}{f + (1 - f)R_j\tau_S}, \quad (5)$$

when the depletion and replenishment processes reach a balance. According to the property of synaptic depression, it is obvious that the microsaccade can increase the activity in the nearby V1 neurons that have the receptive field of the landing position in our proposed feedforward network. This synaptic depression model gives a good fit of experimental data (Abbott et al., 1997). The two parameter values f and τ_S we used lie within the ranges indicated in the experimental data (Carandini et al., 2002; Boudreau and Ferster, 2005). In the following computations, we took $f = 0.75$ and $\tau_S = 200$ ms. Choosing different parameter values does not alter the qualitative results.

3. RESULTS

By using our feedforward model, we first describe the microsaccade-induced excitatory activity in V1 neurons that might contribute to counteract perceptual fading. Then, we will show that our model can reproduce experimental observations about V1 cortical responses after microsaccades as reported by Martinez-Conde et al. (2000, 2002). Moreover, our model can explain the saturation property of visual brain responses for large microsaccadic magnitude and velocity, which has been recently found by measuring scalp EEG (Dimigen et al., 2009).

　Since microsaccades are very fast movements (Martinez-Conde et al., 2009), for simplicity, we ignored the time course of microsaccades (Donner and Hemilä, 2007) in most of our simulations, i.e., the displacement by microsaccade of magnitude Δ_M happens immediately. However, we also studied the impact of velocity and showed that it also reproduces experimental findings. Here, we counted the total number of spikes N_{sp} of the V1 neurons in the model in a moving time bin ($T = 50$ ms) as a measure of the neural response.
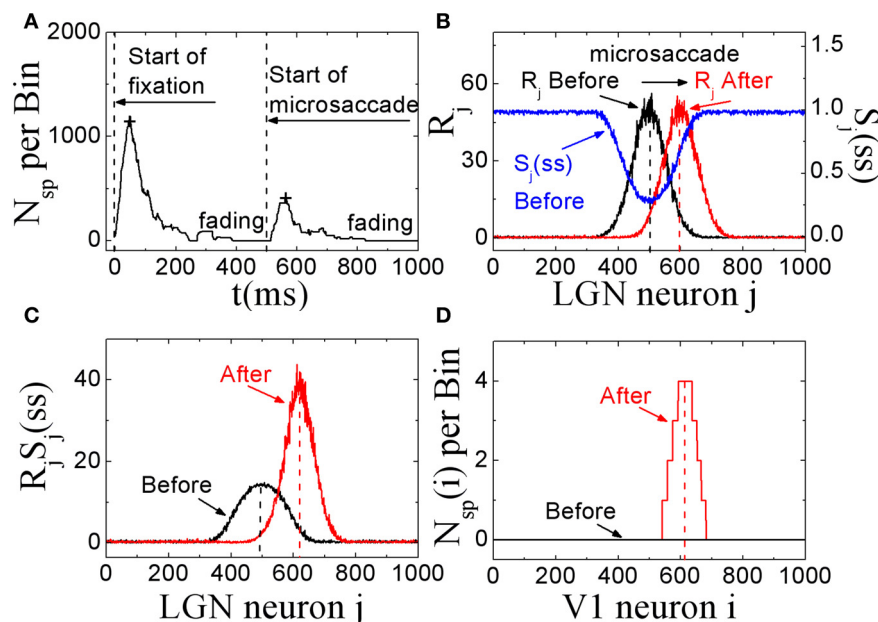
3.1. EXCITATORY RESPONSES TO MICROSACCADES AND A POSSIBLE EXPLANATION FOR COUNTERACTING VISUAL FADING

In these simulations, we showed that, STD in thalamocortical synapses can provide a possible explanation for microsaccades in counteracting visual fading. Here, we assumed that the fixation dot is the only relevant visual stimulus that generates the visual signal. As shown in **Figure 2A**, in this model with STD, neural activity in V1 begins to fade within several hundred milliseconds after the start of fixation in the absence of fixational eye movements (and head or body movements). If a microsaccade occurs, the neural excitation will return and persist for a few hundred milliseconds. If there are no more microsaccades, the neural activity in V1 will be fading completely in about 300 ms. In **Figures 2B–D**, we propose an explanation of the responses in terms of STD in thalamocortical synapses. During fixation in the absence of microsaccades, the spike trains evoked by the fixated dot with firing rates R_j in LGN persist in stimulating thalamocortical synapses (**Figure 2B**, black line). Due to depressing mechanism of STD in these synapses, the synaptic strengths will quickly decrease to steady state values. The strengths S_j with larger firing rate R_j will decrease to smaller steady state values $S_j(ss)$ (**Figure 2B**, blue line). The theoretical analysis (Abbott et al., 1997) showed that the steady state strengths $S_j(ss)$ are inversely proportional to R_j for high firing rates (see Equation 5). When there is a microsaccade, the network will generate a new neural input to stimulate V1 neurons by moving the fixated dot over the receptive fields of LGN neurons with less adapted thalamocortical synapses (in the sense of relative movement; **Figure 2B**,

red line). Before the microsaccade, the input of each thalamocortical synapse from LGN neuron j, which is proportional to $R_jS_j(ss)$ (Abbott et al., 1997), is rather small (**Figure 2C**, black line), and does not induce firing of V1 neurons for the parameter we used (**Figure 2D**, black line). However, immediately after the microsaccade, the new input of each thalamocortical synapse with less adaptation becomes much larger (**Figure 2C**, red line) due to the fast eye movement so that it can evoke spikes in V1 neurons (**Figure 2D**, red line). Afterwards, STD becomes effective to reduce the synaptic strengths and the response fades out again. These simulations indicate that, STD in thalamocortical synapses can give a potentially valid explanation for microsaccades in countering visual fading, which may suggest an important role of STD in microsaccade-related neural responses during fixation.

Next, we investigate in more detail the effect of microsaccadic frequency. As shown in **Figure 3A**, neural activity in V1 is sustained and does not fade away if there are microsaccades with high enough frequency. Here, we calculate the average neural activity related to microsaccades during fixation as a function of microsaccadic frequency F (**Figure 3B**). It is found that, the neural activity will start to increase obviously when microsaccadic frequency increases to 3–4 Hz. We also quantify the sensitivity of neuronal response to change of frequency F by an amount ΔF, which is the slope of average neural activity curve as the function of F in **Figure 3B**. As shown in **Figure 3C**, the sensitivity increases to high enough value when microsaccadic frequency arrives to 3–4 Hz. These results indicate that, the neural activity will be sustained and sensitive if the frequency of microsaccades

FIGURE 2 | The excitatory responses to microsaccades during fixation in the feedforward network model. (A) The change of neuronal activity in V1 neurons before and after microsaccade. The signs "+" reflect the response peaks. **(B)** The computer-generated input firing rates R_j in LGN neurons before and after microsaccade, and the STD-modified steady states $S_j(ss)$ of

synaptic strengthens before microsaccade. **(C)** The synaptic input $R_jS_j(ss)$ for each LGN neuron j before and after microsaccade. **(D)** The neuronal spike number $N_{sp}(i)$ per time bin (50 ms) of the output neuron i in V1 before and after microsaccade. Here the parameters are $g = 0.15$, $A = 50$, $\sigma_1 = \sigma_2 = 1.5$, and $\Delta_M = 2.0$.
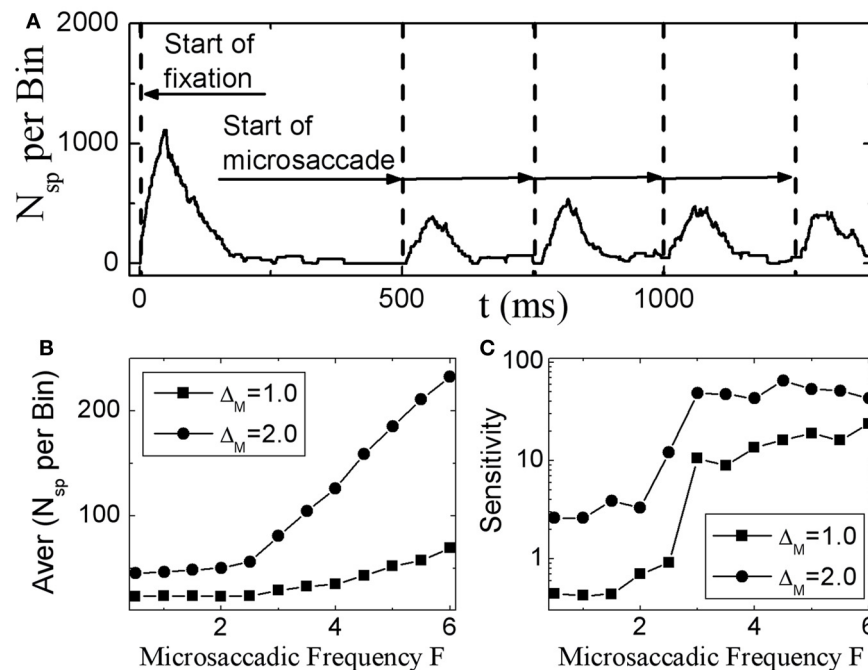
FIGURE 3 | Sustained responses in V1 neurons to microsaccades with high enough frequency. (A) The sustaining neuronal activity induced by periodic microsaccades with a frequency $F = 4$ Hz. The average neuronal activity during fixation (B), and its responsive sensitivity (C) correlated with random microsaccades (in Poisson trains) as a function of microsaccadic frequency F for different microsaccadic sizes Δ_M. Here the parameters are $g = 0.15$, $A = 50$, and $\sigma_1 = \sigma_2 = 1.5$. In (A), $\Delta_M = 2.0$ is chosen.

or macrosaccades is about 3–4 Hz, which is consistent with the fact that microsaccades occur 3–4 times per second (Otero-Millan et al., 2008; Martinez-Conde et al., 2009), though the real situation could be more complicated to involve other factors such as the variable sizes and speeds of the microsaccades.

3.2. REPRODUCING EXPERIMENTAL OBSERVATIONS

3.2.1. Different responses to microsaccades with flashing and stationary stimuli

Perceptual responses to flashed object have been experimentally studied in the presence of microsaccades (Martinez-Conde et al., 2002; Kagan et al., 2008) and saccade (Lappe et al., 2006). Particularly, Martinez-Conde et al. (2002) has experimentally compared neural activities in V1 induced by microsaccades with flashing and stationary (non-flashing) stimulus bars, in order to study how effective microsaccades are in generating neural activity by comparing them with previously characterized and well-known visual stimuli, flashing bars. In their experiment, the stimulus bars were in the receptive fields of the recorded V1 neuron both before and after the microsaccade. They used a white bar on a black background for on cells, and a black bar on a white background for off cells. Then, they calculated the spike probability of neurons in V1 to reflect neural response. As shown in **Figure 4**, the neural response after microsaccades is stronger when a rhythmically flashing bar is on during fixation, as compared to a condition in which the stimulus bar is always on (stationary). Here, our model can provide a possible understanding of this observation using STD. We considered
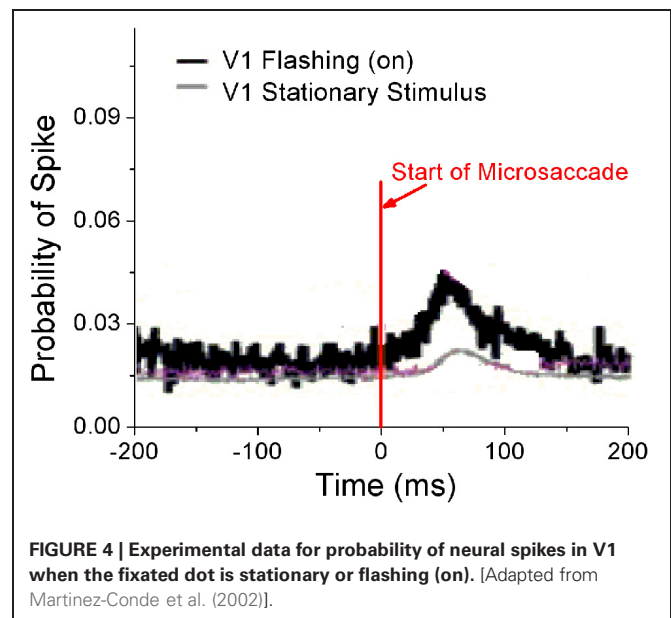
FIGURE 4 | Experimental data for probability of neural spikes in V1 when the fixated dot is stationary or flashing (on). [Adapted from Martinez-Conde et al. (2002)].

that the fixated dot as stimulus can be flashing (on–off) or stationary (**Figure 5A**). As shown in **Figure 5B**, the baseline before microsaccades and responsive peak after microsaccades when the flashing dot is on are both higher than those when the fixated dot is stationary, consistent with the experimental findings by Martinez-Conde et al. (2002). The higher responses are
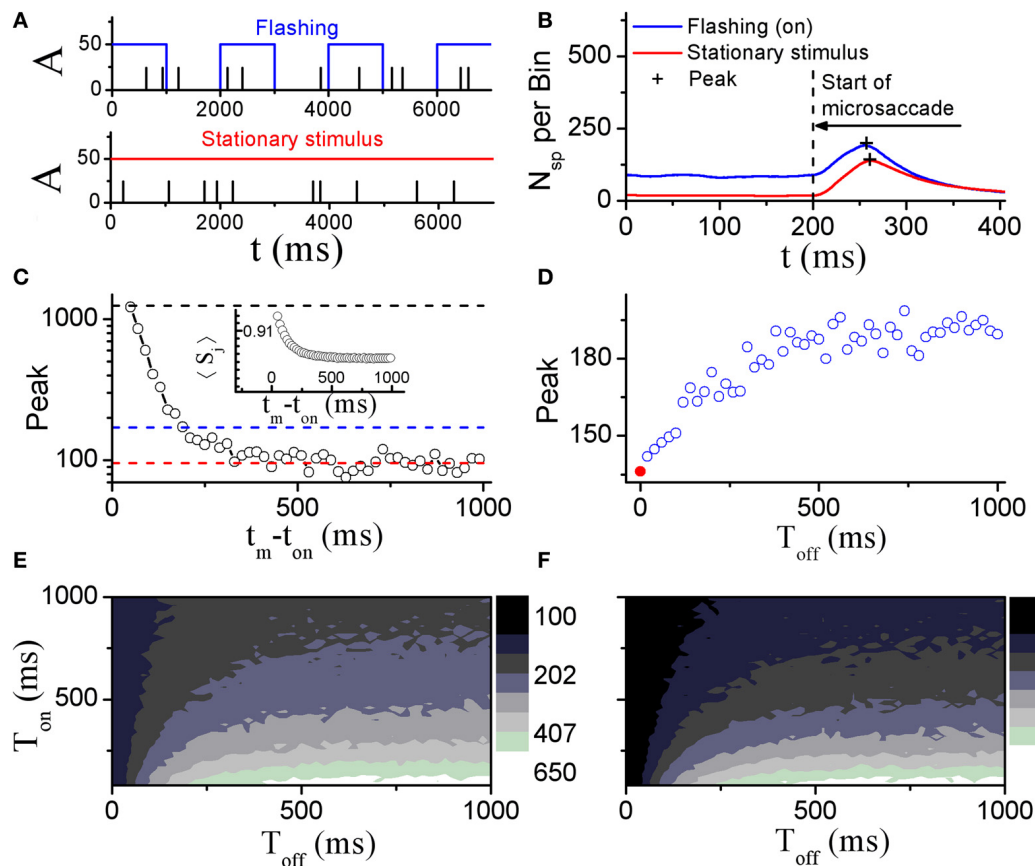
FIGURE 5 | Comparison of model neural responses to microsaccades in V1 when the fixated dot is stationary or flashing cyclically with T_{on} and T_{off} as indicated by the blue rectangles. (A) Stimulus brightness A (see **Figure 1**) of the fixated dot for a periodic flashing condition (blue; $T_{on} = 1$ s, $T_{off} = 1$ s) and stationary presentation (red; constant A) in the presence of microsaccades (black ticks, with fixed size Δ_M) in Poisson trains (1.5 Hz, here we choose the smaller frequency of microsaccades than the real microsaccadic frequency in order to avoid the correlated neural activity from one microsaccade to another due to the Poisson microsaccade trains in simulation). **(B)** Microsaccade increases neural activity in V1 when the fixated dot is stationary, and further increases the neural activity when it is flashing-on. "+"-signs denote the peaks of microsaccade-related neural activities. The results are obtained by averaging over all microsaccades in Poisson trains (1.5 Hz) during the on-state. **(C)** The response peak (i.e., the second response peak in

Figure 2A) evoked by a microsaccade as a function of the interval $t_m - t_{on}$ between onset of microsaccade and onset of flashing-on state. The mean values over all $t_m - t_{on}$ from 0 to T_{on} (blue dashed line) and the baseline of peak response (red dashed line) are approximately equal to the response peaks "+" in **(B)**. In addition, the response peak (the first response peak in **Figure 2A**) induced by onset of stimulus is plotted (black dashed line). Inset in **(C)**: network-averaged synaptic strength $\langle S_j \rangle$ as a function of $t_m - t_{on}$. **(D)** Response peaks "+" in **(B)** as a function of the off duration T_{off} of the flashing dot. The red ● at $T_{off} = 0$ corresponds to the stationary stimulus. **(E)** Phase diagram in T_{off}-T_{on} plane for the response peak "+" in the flashing condition in **(B)**. **(F)** As in **(E)**, but for the ratio of the two response peaks in **(B)** (flashing (on) to stationary stimulus). Here, data are obtained from simulation for 1000 s in **(B)**, **(D–F)** and averaged over 20 realizations for **(C)**. The other parameters are $g = 0.15$, $\sigma_1 = \sigma_2 = 1.5$, and $\Delta_M = 1.0$.

expected because of the additional onset response when the flashing bar is turned on. Namely, in our model, the observations are expected to be due to smaller synaptic depression during the shorter interval between the onset of the flashing-on and the onset of a microsaccade. To further understand the mechanism, we examined how the neural response **Figure 5C**) and network-averaged synaptic strength $\langle S(j) \rangle$ (inset of **Figure 5C**) depend on the time interval $t_m - t_{on}$ between the onsets of microsaccade and flashing-on. A smaller interval corresponds to a higher response activity due to a larger synaptic strength. When the interval $t_m - t_{on}$ is larger, the neural response will decrease to a relatively stable baseline (**Figure 5C**, red dashed line) due to the presence

of a large final stable synaptic depression after the larger interval (**Figure 5C**, inset). This baseline is the approximate response with the stationary stimulus since the synaptic strength for this case decreases to the same stable value as that for stationary stimuli. Because the time interval between the onsets of microsaccade and flashing-on is random (Martinez-Conde et al., 2002), the average response to all the microsaccades during flashing-on is approximately equal to the average activity over the whole possible intervals (i.e., $0 \leq t_m - t_{on} \leq T_{on}$, where T_{on} is the duration of flashing-on; **Figure 5C**, blue dashed line). Clearly, the interval-averaged response (**Figure 5C**, blue dashed line) is larger than the baseline, explaining that the microsaccade-related response with

flashing (on) stimulus is higher than that with stationary stimulus. Moreover, we compared the microsaccade-related responses to the response after a flashing bar turns on. The response after a flashing bar turns on (**Figure 5C**, black dashed line) is several times larger than the two microsaccade-related responses with flashing (**Figure 5C**, blue dashed line) and stationary stimuli (**Figure 5C**, red dashed line), consistent with the experimental observations in Martinez-Conde et al. (2002) and Kagan et al. (2008). Obviously, this is because the synaptic strengths with the thorough recovery from STD within $T_{off} = 1$ s are involved in the neural response after a flashing bar turns on, which is the same as the response at the start of fixation (the first response peak in **Figure 2A**).

To further study microsaccade-related neural responses due to STD with a flashing stimulus, we investigated effects of the flash-on duration T_{on} and the flash-off duration T_{off} (the time that passed since the last flash onset or offset, respectively) (**Figures 5D–F**). During the off state, the synaptic strengths will recover from the depression, reaching larger synaptic strengths with longer T_{off} till saturation. Thus, the microsaccade-related neural response increases with increasing T_{off} and then reaches saturation due to the thorough synaptic recovery for the larger T_{off} (**Figure 5D**). For the effect of the on-duration T_{on}, we can infer from **Figure 5C** that the neural response (**Figure 5C**, blue dashed line) will become larger with the decrease of T_{on}. Therefore, the ranges of the observed increase of microsaccade-related neural response and of the increased ratio of the response with flashing stimuli relative to stationary stimuli are in line with large T_{off} and small T_{on} (**Figures 5E** and **F**).

3.2.2. Saturation of activity for large microsaccadic magnitude and velocity

Dimigen et al. (2009) studied microsaccade-related brain activity in event-related brain potentials (ERP). ERP is the average of many epochs of EEG trials recorded from scalp for the same task, synchronized to the same event such as the stimulus onset or microsaccade onsets, yielding a clear pattern of brain response to the external signal when compared to the base line. (Picton et al., 2000; Handy, 2005; Ouyang et al., 2011). Dimigen et al. (2009) found that the tiny eye movements by microsaccades can generate sizable visual brain response in ERP comparable to usual saccadic eye movements and responses correlated with microsaccades tend to saturate for large microsaccades (**Figure 6A**, red line). Our model can provide a possible explanation for this phenomenon by the effect of microsaccade magnitude on neural activity, shown in **Figure 7**. A response peak appears soon after the microsaccade, and the value increases with the microsaccade magnitude Δ_M (**Figure 7A**). The increase is almost linear for small microsaccades, consistent with the finding by Dimigen et al. (2009). As the microsaccade magnitude increases further, the increasing response reaches saturation (**Figure 7B**). This saturation can be explained as follows. As shown in **Figures 7B** and **C**, the synaptic input $R_j S_j$ increases after a microsaccade by moving the fixated dot over the receptive fields of LGN neurons with less adapted thalamocortical synapses (in the sense of relative movement). But, when the moving distance due to large microsaccade exceeds the region with strong synapse-depression, the synaptic input will become
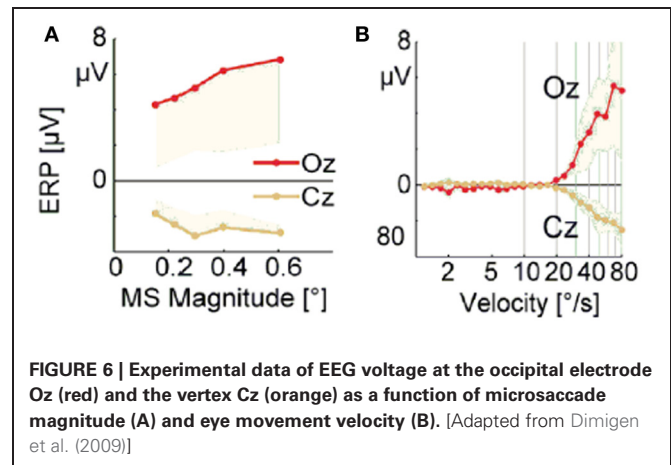
FIGURE 6 | Experimental data of EEG voltage at the occipital electrode Oz (red) and the vertex Cz (orange) as a function of microsaccade magnitude (A) and eye movement velocity (B). [Adapted from Dimigen et al. (2009)]

independent of the microsaccade magnitude, leading to saturated response.

So far, all the above simulations were done without considering the finite velocity of microsaccade. In **Figure 6B**, a relationship between instantaneous eye movement velocity (also including periods of drift) and the amplitude of occipital EEG response 100 ms later was observed by Dimigen et al. (2009). Our model can provide an understanding of this experimentally observed relationship when microsaccadic velocity is taken into consideration. In the simulation, we assume a constant velocity for microsaccadic movements, with a fixed duration of 15 ms for microsaccades of different sizes, following the experimental findings that there are approximately fixed microsaccadic durations (around 15 ms for human) for different microsaccadic velocities (Troncoso et al., 2008; Dimigen et al., 2009). The results shown in the inset of **Figure 7B** agree well with the pattern shown in **Figure 6B**, experimentally found in Dimigen et al. (2009). Though the analysis in Dimigen et al. (2009) included all samples of the eye movement trajectory, not only microsaccades, it is reasonable to believe that most of the medium-velocity samples belong to microsaccades (Martinez-Conde et al., 2004, 2009). From our model, we can understand that, at slow eye velocities, retinal displacements are small and the signal moves only slightly and slowly away from the strongly depressed region, without inducing strong neural response. This may be able to explain that tremor and drift do not induce significant neural responses (Gur et al., 1997; Martinez-Conde, 2006). With larger velocities, the signal quickly moves to a much less depressed region and induces sizable response.

4. DISCUSSION AND CONCLUSION

The prevailing theory suggests that visual fading in the absence of microsaccades is caused by retinal adaptation. However, retinal adaptation for visual fading has not been directly tested in experiments. On the other hand, STD in thalamocortical systems from LGN and V1 has been empirically confirmed and could play an important functional role. Based on these considerations, we have proposed an alternative potential biophysical foundation for the explanation how microsaccades counteract visual fading, thalamocortical STD. With a simple feedforward model, we showed that, without considering possible retinal adaptation, STD from
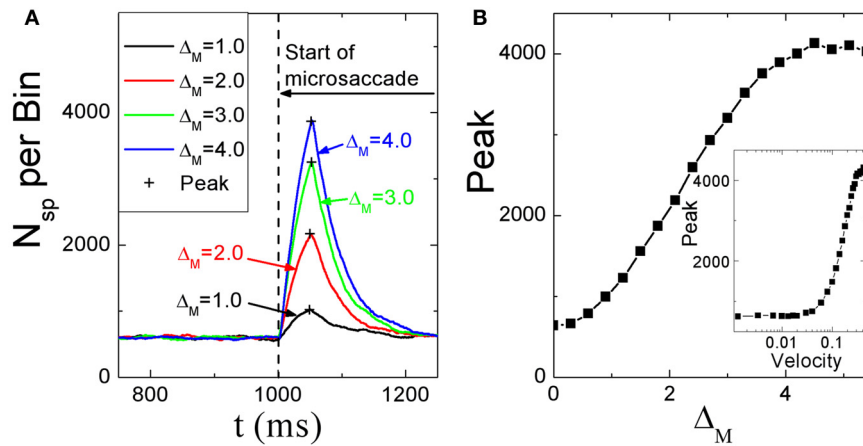
FIGURE 7 | (A) Microsaccade-induced neural activity for different microsaccade magnitudes. After microsaccade, there is a response peak. **(B)** Saturation of the response peak for large microsaccades. The inset in **(B)**: the effect of microsaccadic velocity on response peak in Log-Linear scale for comparison with **Figure 6B**, experimentally found in Dimigen et al. (2009). Here $g = 0.2$, $A = 100$, and $\sigma_1 = \sigma_2 = 1.5$. Data are averaged over 20 independent runs.

LGN to V1 alone can qualitatively reproduce several experimental observations about microsaccade-induced brain responses.

However, it is important to note that these two possible mechanisms of retinal adaptation and STD are not mutually exclusive. In fact, enhanced LGN activity by microsaccades as observed in experiments Martinez-Conde et al. (2000, 2002) could be an indication of possible retinal adaptation. In the real visual systems, the two possible mechanisms could yield different functional benefits for visual information processing, which are yet unknown. If retinal adaptation could be effectively described similar to STD and retinal neurons can be described similarly as in LGN and V1, and assuming that there is no STD in thalamocortical synapses, then from the viewpoint of simplified feedforward neural model, the response in V1 could be similar to what we described here for STD from LGN to V1. Possibly different functional/behavioral effects of the two mechanisms then would rely strongly on the biophysical details. Perhaps the most interesting possibility is that these two adaptation levels are actually arranged in a cascade. Such a cascading of adaptation is expected to enhance the sensitivity of adaptation, likely to sharpen the cortical neural responses to tiny and fast eye movements (or equivalently tiny and fast movement of the visual world). Further investigations are expected to reveal more behavioral properties and functional roles of microsaccades (for review, see Rolfs, 2009). The work presented in this paper will serve as a foundation for future studies.

To sum up, we proposed an alternative synaptic explanation for microsaccades in counteracting visual fading during fixation by introducing STD in the thalamocortical system. Moreover, the depression model can reproduce several experimental observations of microsaccade-related neural responses in V1. Our model and results are expected to provide quantitative method and theoretical insight into the study of microsaccades. Generally, our model may provide a useful tip for the understanding of visual information adaptation and transmission, and give a starting point for modeling visual process of microsaccades by considering more neurobiological ingredients, such as inhibitory connections within V1 and from LGN, and other types of synaptic plasticity and cascading with possible retinal adaptation.

REFERENCES

Abbott, L., Varela, J., Sen, K., and Nelson, S. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224.

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* 27, 77–87.

Armington, J., and Bloom, M. (1974). Relations between the amplitudes of spontaneous saccades and visual responses. *J. Opt. Soc. Am.* 64, 1263–1271.

Bair, W., and O'Keefe, L. (1998). The influence of fixational eye movements on the response of neurons in area mt of the macaque. *Vis. Neurosci.* 15, 779–786.

Bannister, N., Nelson, J., and Jack, J. (2002). Excitatory inputs to spiny cells in layers 4 and 6 of cat striate cortex. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357, 1793–1808.

Boudreau, C., and Ferster, D. (2005). Short-term depression in thalamocortical synapses of cat primary visual cortex. *J. Neurosci.* 25, 7179–7190.

Carandini, M., Heeger, D., and Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *J. Neurosci.* 22, 10053–10065.

Chance, F., and Abbott, L. (2001). Input-specific adaptation in complex cells through synaptic depression. *Neurocomputing* 38, 141–146.

Chance, F., Nelson, S., and Abbott, L. (1998). Synaptic depression and the temporal response characteristics of v1 cells. *J. Neurosci.* 18, 4785–4799.

Dimigen, O., Valsecchi, M., Sommer, W., and Kliegl, R. (2009). Human microsaccade-related visual brain

responses. *J. Neurosci.* 29, 12321–12331.

Ditchburn, R., and Ginsborg, B. (1952). Vision with a stabilized retinal image. *Nature* 170, 36–37.

Donner, K., and Hemilä, S. (2007). Modelling the effect of microsaccades on retinal responses to stationary contrast patterns. *Vision Res.* 47, 1166–1177.

Engbert, R., and Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7192–7197.

Ferster, D., Chung, S., and Wheat, H. (1996). Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature* 380, 249–252.

Ferster, D., and Miller, K. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annu. Rev. Neurosci.* 23, 441–471.

Greschner, M., Bongard, M., Rujan, P., and Ammermüller, J. (2002). Retinal ganglion cell synchronization by fixational eye movements improves feature estimation. *Nat. Neurosci.* 5, 341–347.

Gur, M., Beylin, A., and Snodderly, D. (1997). Response variability of neurons in primary visual cortex (v1) of alert monkeys. *J. Neurosci.* 17, 2914–2920.

Handy, T. (2005). *Event-related Potentials: A Methods Handbook (bradford books)*. Cambridge, MA: MIT Press.

Hsieh, P., and Tse, P. (2009). Microsaccade rate varies with subjective visibility during motion-induced blindness. *PLoS ONE* 4:e5163. doi: 10.1371/journal.pone.0005163

Kagan, I., Gur, M., and Snodderly, D. (2008). Saccades and drifts differentially modulate neuronal activity in v1: effects of retinal image motion, position, and extraretinal influences. *J. Vision* 8, 1–25.

Krauskopf, J. (1957). Effect of retinal image motion on contrast thresholds for maintained vision. *J. Opt. Soc. Am.* 47, 740–741.

Lappe, M., Kuhlmann, S., Oerke, B., and Kaiser, M. (2006). The fate of object features during perisaccadic mislocalization. *J. Vision* 6, 1282–1293.

Leopold, D., and Logothetis, N. (1998). Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Exp. Brain Res.* 123, 341–345.

Martinez-Conde, S. (2006). Fixational eye movements in normal and pathological vision. *Prog. Brain Res.* 154, 151–176.

Martinez-Conde, S., Macknik, S., and Hubel, D. (2000). Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nat. Neurosci.* 3, 251–258.

Martinez-Conde, S., Macknik, S., and Hubel, D. (2002). The function of bursts of spikes during visual fixation in the awake primate lateral geniculate nucleus and primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13920–13925.

Martinez-Conde, S., Macknik, S., and Hubel, D. (2004). The role of fixational eye movements in visual perception. *Nat. Rev. Neurosci.* 5, 229–240.

Martinez-Conde, S., Macknik, S., Troncoso, X., and Hubel, D. (2009). Microsaccades: a neurophysiological analysis. *Trends Neurosci.* 32, 463–475.

Nelson, S., Toth, L., Sheth, B., and Sur, M. (1994). Orientation selectivity of cortical neurons during intracellular blockade of inhibition. *Science* 265, 774–777.

Otero-Millan, J., Troncoso, X., Macknik, S., Serrano-Pedraza, I., and Martinez-Conde, S. (2008). Saccades and microsaccades during visual fixation, exploration, and search: foundations for a common saccadic generator. *J. Vision* 8, 1–18.

Ouyang, G., Herzmann, G., Zhou, C., and Sommer, W. (2011). Residue iteration decomposition (ride): a new method to separate erp components on the basis of latency variability in single trials. *Psychophysiology* 48, 1631–1647.

Picton, T., Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson, R. Jr., et al. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152.

Pinault, D., and Deschênes, M. (1998). Anatomical evidence for a mechanism of lateral inhibition in the rat thalamus. *Eur. J. Neurosci.* 10, 3462–3469.

Poggio, T., Fahle, M., and Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science* 256, 1018–1021.

Riggs, L., and Ratliff, F. (1952). The effects of counteracting the normal movements of the eye. *J. Opt. Soc. Am.* 42, 872–873.

Riggs, L., Ratliff, F., Cornsweet, J., and Cornsweet, T. (1953). The disappearance of steadily fixated visual test objects. *J. Opt. Soc. Am.* 43, 495–500.

Rolfs, M. (2009). Microsaccades: small steps on a long way. *Vision Res.* 49, 2415–2441.

Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in v1 neurons. *Nature* 412, 549–553.

Seriès, P., Latham, P., and Pouget, A. (2004). Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nat. Neurosci.* 7, 1129–1135.

Sharpe, C. (1972). The visibility and fading of thin lines visualized by their controlled movement across the retina. *J. Physiol.* 222, 113–114.

Stratford, K., Tarczy-Hornoch, K., Martin, K., Bannister, N., and Jack, J. (1996). Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. *Nature* 382, 258–261.

Troncoso, X., Macknik, S., and Martinez-Conde, S. (2008). Microsaccades counteract perceptual filling-in. *J. Vision* 8, 1–9.

Tse, P., Baumgartner, F., and Greenlee, M. (2010). Event-related functional mri of cortical activity evoked by microsaccades, small visually-guided saccades, and eyeblinks in human visual cortex. *Neuroimage* 49, 805–816.

Tsodyks, M., and Gilbert, C. (2004). Neural networks and perceptual learning. *Nature* 431, 775–781.

Varela, J., Sen, K., Gibson, J., Fost, J., Abbott, L., and Nelson, S. (1997). A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J. Neurosci.* 17, 7926–7940.

Yuan, W.-J., Luo, X.-S., Wang, B.-H., Wang, W.-X., Fang, J.-Q., and Jiang, P.-Q. (2006). Excitation properties of the biological neurons with side-inhibition mechanism in small-world networks. *Chin. Phys. Lett.* 23, 3115–3118.

Yuan, W.-J., Luo, X.-S., and Jiang, P.-Q. (2007). Study under AC Stimulation on excitement properties of weighted small-world biological neural networks with side-restrain mechanism. *Comm. Theor. Phys.* 47, 369–373.

Yuval-Greenberg, S., Tomer, O., Keren, A., Nelken, I., and Deouell, L. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* 58, 429–441.

Interaction of short-term depression and firing dynamics in shaping single neuron encoding

Ashutosh Mohan[1]*, Mark D. McDonnell[2] and Christian Stricker[1,3]

[1] The John Curtin School of Medical Research, Australian National University, Canberra, ACT, Australia
[2] Computational and Theoretical Neuroscience Laboratory, Institute for Telecommunications Research, University of South Australia, Mawson Lakes, SA, Australia
[3] Medical School, Australian National University, Canberra, ACT, Australia

We investigated how the two properties short-term synaptic depression of afferent input and postsynaptic firing dynamics combine to determine the operating mode of a neuron. While several computational roles have been ascribed to either, their interaction has not been studied. We considered two types of short-term synaptic dynamics (release-dependent and release-independent depression) and two classes of firing dynamics (regular firing and firing with spike-frequency adaptation). The input–output transformation of the four possible combinations of pre- and post-synaptic dynamics was characterized. Adapting neurons receiving input from release-dependent synapses functioned largely as coincidence detectors. The other three configurations showed properties consistent with integrators, each with distinct features. These results suggest that the operating mode of a neuron is determined by both the pre- and post-synaptic dynamics and that studying them together is necessary to understand emergent properties and their implications for neuronal coding.

Keywords: short-term depression, operating modes, emergent properties, firing properties, synaptic integration

INTRODUCTION

Synapses exhibit a range of activity-dependent plasticities at various timescales (Dobrunz et al., 1997; Dittman et al., 2000; Fuhrmann et al., 2004; Regehr, 2012). Short-term synaptic plasticity is the change in efficacy of the postsynaptic potential/current upon repeated stimulation lasting for a few to hundreds of milliseconds. Excitatory synapses in neocortex exhibit short-term depression and recover at a rate of about 1 s. Depression is dominant with minimal facilitation in layers 2/3, 4, and 5 of rat barrel cortex (Cowan and Stricker, 2004; Fuhrmann et al., 2004). However, the mechanisms underlying facilitation are much less clear. Hence, we restrict our investigation to synaptic depression and its role in encoding.

Functionally, these depressing synapses show two different types of dynamics, defined here as type 1 or 2. Type 1 synapses show depression due to vesicle-depletion (VDD) that reduces the probability of neurotransmitter release upon subsequent action potentials (Markram and Tsodyks, 1996; Markram et al., 1997; Matveev and Wang, 2000; Regehr, 2012). At these synapses, the recovery rate from depression is constant. Type 1 synapses are capable of signaling a stimulus rate change but not rate (Fuhrmann et al., 2004; Jedrzejewska-Szmek and Zygierewicz, 2010). Type 2 synapses on the other hand exhibit release-independent depression, i.e., they depress even when no neurotransmitter has been released (Dobrunz et al., 1997; Thomson, 1997; Brody and Yue, 2000; Cowan and Stricker, 2004; Fuhrmann et al., 2004; Muñoz-Cuevas et al., 2004; Regehr, 2012). Additionally, the recovery rate is frequency-dependent and increases with higher stimulus frequencies (Cowan and Stricker, 2004; Fuhrmann et al., 2004). Type 2 synapses are capable of

relaying both information about the stimulus rate and its rate change (Cowan and Stricker, 2004; Fuhrmann et al., 2004).

Previous work has largely focused on type 1 synapses that might endow single neurons and neuronal networks with specific capabilities. Type 1 synapses provide a gain control mechanism resulting in improved sensitivity of neurons to small changes in stimulus firing pattern (Abbott, 1997). Through simulations of networks in primary visual cortex, type 1 dynamics of thalamocortical synapses have been shown to precisely control the oscillatory response (Paik and Glaser, 2010). These properties also facilitate synchrony detection in a network (Senn et al., 1998). The functional implications of type 2 synapses have not been widely studied (but see Graham and Stricker, 2008; Scott et al., 2012). Previous studies of synaptic dynamics have primarily focused on its impact on information transfer in isolation, while neglecting the postsynaptic dynamics in detail (London et al., 2008; Fung et al., 2012).

As synaptic input is integrated at the postsynaptic side into a sequence of action potentials, the variations in firing dynamics also need consideration. The importance of studying both pre- and post-synaptic dynamics together for a holistic understanding of information processing has been recognized in the context of the dynamics of long-term plasticity and intrinsic plasticity of the postsynaptic membrane (Turrigiano et al., 1998; Xie et al., 2006; Triesch, 2007). To address this issue, we adopt the simple classification proposed by (Hodgkin, 1948)—class 1 and class 2 firing characteristics of a neuron (subsequently also called class 1 or 2 neuron). Class 1 firing is regular and there is a linear relationship between injected current and firing rate. Class 2 firing on the other hand shows spike-frequency adaption and

consequently a non-linear relationship between current and firing rate. From a dynamical systems point of view, class 1 and class 2 neurons exhibit saddle node on a limit cycle and Hopf bifurcations, respectively (Izhikevich, 2000). The rationale for adopting this classification is similar to that for adopting a phenomenological description for modeling synaptic dynamics—the focus is on functional dynamics without considering the physiological mechanisms that define them.

Here, we consider all four combinations between types and classes and study how pre- and post-synaptic properties together determine whether the neuron functions as an integrator of stimuli or a coincidence detector in the presence of synaptic background noise. That the cell is quiescent with a stimulus generating sparse firing is supported by several experimental studies (Shadlen and Newsome, 1998; Brecht and Sakmann, 2002). Further, we also study how each combination is affected by variations in noise properties and extent of depression exhibited by synapses. This investigation is especially relevant in the context of highly debated question of whether neurons use precise spike timings, thereby functioning as coincidence detectors or they work more broadly using spike rates, thereby functioning as integrators (Shadlen and Newsome, 1998; deCharms and Zador, 2000). This question is also highly relevant to whether neurons are capable of acting as integrators *in vivo* where there is an increase in background conductance due to synaptic activity (Rudolph and Destexhe, 2001).

METHODS
STIMULUS

Each stimulus consisted of N_{tot} number of presynaptic spikes delivered through N_{syn} number of synapses (either type 1 or 2) that relay excitatory postsynaptic potentials to the postsynaptic neuron with either class 1 or 2 firing characteristics. As shown in **Figure 1**, this stimulus was constructed as follows. N_{tot} Gaussian random numbers were generated with the specified parameters. Each of the generated numbers was assigned to a randomly picked synapse. The sum of all synaptic stimulations, thus, had a Gaussian distribution (in time). Simulations were performed by repeated iterations using a Gaussian stimulus, which was computed by distributing N_{tot} stimuli across N_{syn} number of synapses (see **Figure 1A2**). The timing of each presynaptic spike that comprises the stimulus was based on a Gaussian distribution with the following two parameters, μ_{stim} and σ_{stim} where the former is the mean of the stimulus distribution and the latter its standard deviation, subsequently also called dispersion. Specifically, since presynaptic spike times are generated based on a Gaussian distribution, this parameter signifies the time of stimulus peak. Small values of σ_{stim} imply tightly synchronized presynaptic spike arrivals while large values imply a less synchronized stimulus.

In order to facilitate comparison and interpretation of various values, σ_{stim} and, in general, all values capturing a time quantity were normalized by the membrane time constant τ_m. As an example, if $\sigma_{stim} = 0.1$, dispersion of the stimulus is 10% of the time constant. Since in a Gaussian distribution, 99.73% of all events occur within three times the standard deviation on either side of the mean, this implies that almost all presynaptic spikes arrive within 60% of τ_m.

Synapse model

The phenomenological model used is an extension of that proposed by Fuhrmann et al. (2004). Type 1 synapses show release-dependent depression with a constant rate of recovery. Type 2 synapses show release-independent depression and a frequency-dependent recovery rate. The model exhibits either type 1 or 2 dynamics depending on the parameter values.

The synaptic conductance (g_s) due to a single synapse is computed as:

$$g_s(t) = U_{SE}(t) \cdot P_v(t) \cdot A_{SE}$$

U_{SE} and P_V represent the maximal response when all synapses release their vesicles and probability of vesicle availability, respectively. Their product corresponds to the fraction of available vesicles that are released. A_{SE} is the maximal conductance. The variables in turn are governed by the following set of equations. The first is,

$$\frac{dP_V}{dt} = \frac{1 - P_V}{\tau_{VDD}} - U_{SE} \cdot P_V \cdot \sum_{N_{tot}} \delta(t - t_{AP}),$$

where τ_{VDD} is the time constant of the synaptic vesicle refilling process, δ is the Dirac delta function and t_{AP} is the time of arrival of an action potential. The formulation of release-independent depression is encapsulated with the variable U_{SE} being decremented from an initial availability of U_0 with a strength of S_{RID} followed by an exponential recovery with a characteristic time constant τ_{RID}, i.e.,

$$\frac{dU_{SE}}{dt} = \frac{U_0 - U_{SE}}{\tau_{RID}} - S_{RID} \cdot U_{SE} \cdot \sum_{N_{tot}} \delta(t - t_{AP}).$$

In analogy, the frequency-dependent recovery of type 2 synapses is captured by decrementing the recovery time constant with a strength of S_{FDR} upon the arrival of an action potential, i.e.,

$$\frac{d\tau_{RID}}{dt} = \frac{\tau_0 - \tau_{RID}}{\tau_{FDR}} - S_{FDR} \cdot \tau_{RID} \cdot \sum_{N_{tot}} \delta(t - t_{AP}).$$

In other words, the recovery rate becomes faster following which τ_{RID} approaches its original value with an exponential time course governed by τ_{FDR}.

For excitatory synapses, typical model parameter values of type 1 and 2 synapses were chosen based on parameter estimates using experimental data of Fuhrmann et al. (2004).
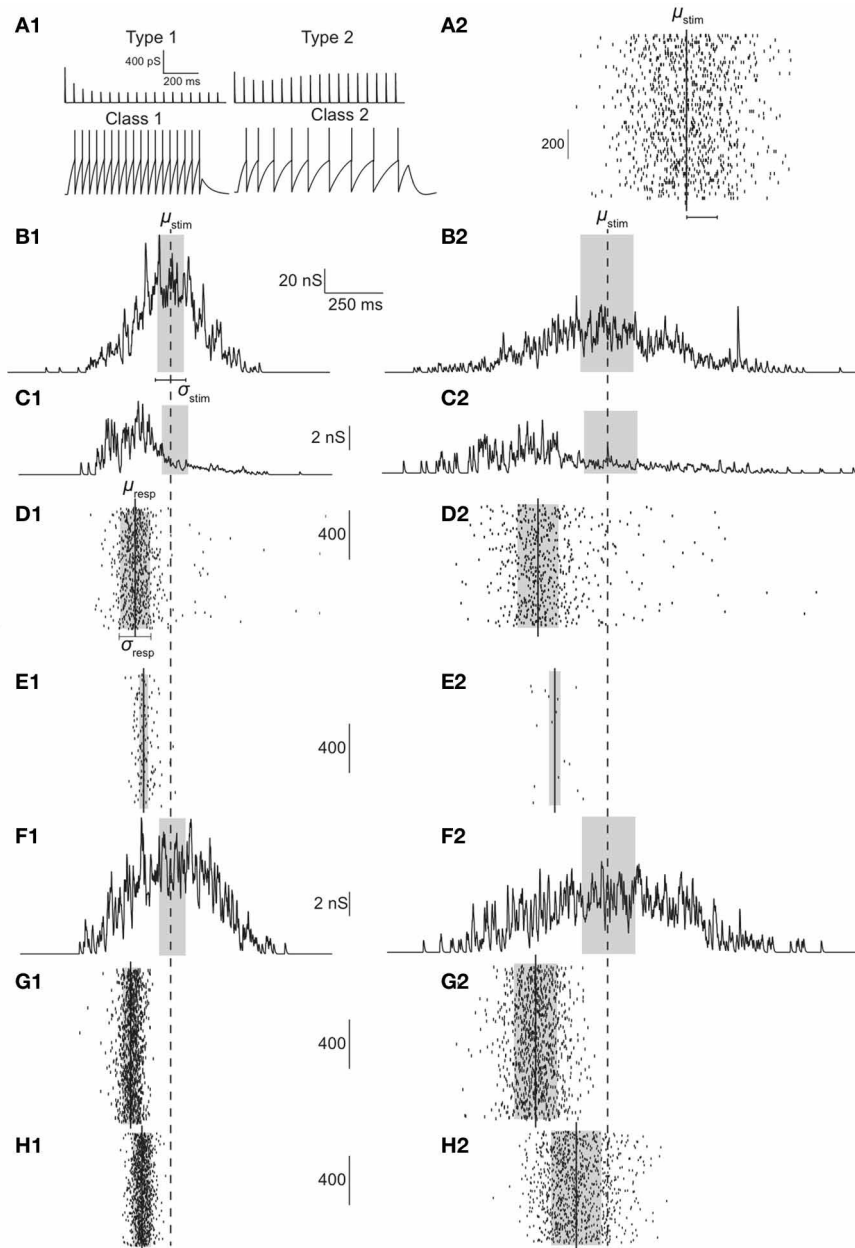
The model has six parameters with values as specified in **Table 1**.

Noise model

A noisy current I_N, was injected into neurons and modeled as an Ornstein–Uhlenbeck process (OUP) and approximated in discrete time simulations using the method proposed by Gillespie (1996), i.e.,

$$I_N(n) = \left(1 - \frac{\Delta t}{\tau_N}\right) \cdot I_N(n - 1) + \left(\sigma_N \sqrt{\frac{2\Delta t}{\tau_N}}\right) G(0, 1),$$

FIGURE 1 | (A1) Response of type 1 (top left) and type 2 (top right) synapses to 30 Hz stimulation. For a current injection of 0.5 nA response of class 1 (bottom left) and class 2 (bottom right) neurons are shown. **(A2)** Illustration of stimulus. In the example, 1000 Gaussian-distributed (in time) presynaptic spikes are relayed through 75 synapses. **(B)** Synaptic conductance if synapses were static; $\sigma_{stim} = 60$ **(B1)** or 120 ms **(B2)**. **(C,F)** Synaptic conductances with type 1 **(C)** synapses and type 2 **(F)** synapses. **(D,E)** With type 1 synapses, raster of spiking response over 1000 iterations for class 1 **(D)** and class 2 **(E)** neurons. **(G,H)** With type 2 synapses, raster of spiking response over 1000 iterations for class 1 **(G)** and class 2 **(H)** neurons. Dashed line signifies stimulus mean (μ_{stim}) while solid lines indicate response mean (μ_{resp}). Gray boxes indicate $\pm 0.5 \cdot \sigma_{stim}$ in **(B,C,** and **F)** and $\pm 0.5 \cdot \sigma_{resp}$ in **(D,E,G,** and **H)**.

where $G(0,1)$ is a zero mean, unit variance Gaussian distributed number. The sample time $\Delta\tau$ was set to 0.2 ms. This noise is characterized by the standard deviation (σ_N) and the correlation time (τ_N) which indicates the time window within which correlations in noise can be observed. As no two samples of white noise are correlated, an increase in the correlation time window results in

greater "coloring" of white noise. τ_N was varied to study how it interacted with short-term synaptic dynamics in shaping the neuronal response properties. The standard deviation of the process σ_N was set to a constant value of 50 pA and τ_N was varied in the simulations. Action potentials generated were almost always due to the stimulus and very rarely sole due to injected noise (<1%).

Table 1 | Synapse parameters.

Parameter/Type	Type 1	Type 2
U_0	0.6	0.25
τ_{VDD} [s]	0.5	0.005
τ_{FDR} [s]	0.9	0.9
τ_0 [s]	0.6	0.6
S_{RID}	0	0.25
S_{FDR}	0	0.30
A_{SE} [nS]	1	1

Table 2 | Neuron parameters.

C [pF]	1000
g_L [nS]	8
E_L [mV]	−70.6
V_T [mV]	−50.4
Δ_T [mV]	2
τ_W [ms]	144
a [nS]	1 (class 1) or 8 (class 2)
b [nA]	0.0805

RESPONSE

Each Gaussian stimulus comprising of several presynaptic spikes was relayed to the postsynaptic neuron through dynamic synapses. To explore the operating mode of the neuron, N_{syn} was varied between 75 and 125 in steps of 5 and the background noise correlation τ_n was varied between 50 and 100 in steps of 10. N_{tot} was set to 1000, unless mentioned otherwise. For each parameter set, individual Gaussian stimuli were repeated 5000 times and if the neuron spiked, the time of the first action potential was recorded. Resulting peri-stimulus time histograms (PSTHs) were characterized by a Gaussian distribution of width σ_{resp} and with respect to the stimulus distribution, shifted by a precession, t_{pre} (see **Figure 1**). Timing of only the first action potential was considered. While acknowledging the potential of spike trains to encode information, the focus of this study is on the encoding of stimulus information in the timing, reliability, and dispersion of the first action potential. Information encoded in repeated spiking is not considered.

Neuron model

We used an adaptive integrate-and-fire model formulated by Brette and Gerstner (2005); i.e.,

$$C\frac{dV}{dt} = f(V) - I_W(t) - I_N(t) - g_S(t) \cdot (V - E_e),$$

where V is membrane voltage, C is the membrane capacitance, $f(V)$ the function capturing the passive properties and the action potential generation dynamics, I_w the adaptation current, I_N the injected noise, g_S the synaptic conductance, and E_e the reversal potential for excitatory synapses. $f(V)$ is defined as:

$$f(V) = -g_L \cdot (V - E_L) + g_L \cdot \Delta_T \cdot \exp\left(\frac{V - V_T}{\Delta_T}\right),$$

where g_L is the leak conductance, E_L the leak reversal, Δ_T the slope factor, and V_T the spike threshold.

The adaptation current, I_W, is generated as follows:

$$\tau_W \frac{dI_W}{dt} = a \cdot (V - E_L) - I_W,$$

where τ_w is the time constant determining the rate of spike frequency adaptation. When an action potential is generated and the

membrane potential (V) goes over the threshold (V_T):

$$V \rightarrow E_L$$
$$I_w \rightarrow I_w + b$$

where b represents spike-triggered adaptation.

For class 1 neurons, the parameters were exactly those specified in Brette and Gerstner (2005), except that for class 1 and 2, a was set to 1 and 8, respectively. The variable that mainly determines the class is the subthreshold adaptation variable a with the spike-triggered adaptation variable b playing a more minor role in our simulations (Touboul and Brette, 2008). See **Table 2** for parameter values of the neuron models.

Response characteristics

We define the following variables that capture the characteristics of the spiking response, namely N_{iter} as the total number of iterations (set to 5000 in our simulations), N_{resp} as the number of spikes evoked over all iterations, R as the reliability of spike generation, defined as the ratio of number of spikes evoked across all iterations and the total number of iterations; i.e., $R = N_{resp}/N_{iter}$, t_{pre} as the precession of the mean of response Gaussian distribution with respect to the stimulus distribution, normalized by the membrane time constant τ_m, σ_{resp} as the width of the response Gaussian distribution, again normalized by τ_m and ζ as the sharpening of responses defined as the ratio between the stimulus and response dispersions ($\sigma_{stim}/\sigma_{resp}$).

Definition of operating modes

We considered the two operating modes coincidence detector and integrator. As an operational definition, we defined each mode in terms of one or more response parameters. Coincidence detectors were defined to be reliable ($R > 0.75$) only for tightly synchronized stimuli (defined as, $\sigma_{stim}/\tau_m < 0.4$) and otherwise unreliable ($R = 0.75$). Thus, a coincidence detector is selectively sensitive to synchronized inputs while failing to reliably relay dispersed inputs. Integrators were defined as being reliable over a range of stimulus synchronies ($0.2 < \sigma_{stim}/\tau_m < 1.2$) but requiring to exhibit a regular relationship between stimulus and response dispersion. Thus, an integrator relays stimulus information reliably with the response dispersion having a regular relationship with stimulus dispersion.

Simulation

All simulations were done in Igor Pro 6.2 (WaveMetrics Inc., Lake Oswego, OR, USA) on a Windows 7 workstation. For the

synapse model, the analytic solution was used instead of solving the differential equations (Scott et al., 2012). For the neuron model, the differential equations were solved numerically using a fourth order Runge–Kutta algorithm (Press et al., 2007). Five thousand trials took approximately 1 h with a time step $\Delta\tau = 0.2$ ms. All analysis was done using custom routines written in Igor Pro.

RESULTS

Type 1 synapses show release-dependent depression and constant recovery rate while type 2 synapses show release-independent depression with a faster recovery rate for higher presynaptic spike rates. **Figure 1A1** (top) shows these two examples stimulated at 25 Hz. Type 1 synapses depress and rapidly reach the steady state, the amplitude of which is inversely proportional to the stimulus frequency (Cowan and Stricker, 2004; Fuhrmann et al., 2004). Type 2 synapses on the other hand, depress but also recover rapidly and hence exhibit a larger steady state response the amplitude of which is more or less constant. Thus, it might be expected that type 1 synapses are effective to relay low frequency stimuli or high frequency stimuli that are highly synchronous. Type 2 synapses might be expected to be able to relay low and high frequency stimuli irrespective of the degree of synchronization.

On the postsynaptic side, class 1 neurons fire regularly and class 2 neurons show spike-frequency adaptation (see **Figure 1A1**; bottom). The ability to generate the first action potential is higher for class 2 neurons as the dynamics enables firing at arbitrarily low frequencies. Thus, class 1 neurons might be expected to be able to relay incoming stimuli irrespective of the degree of synchronization. Class 2 neurons cannot relay highly synchronized (i.e., not dispersed) inputs because the latter cannot depolarize the membrane sufficiently enough to counteract the hyperpolarizing current present in class 2 neurons.

Both synaptic and postsynaptic dynamics have implications in how presynaptic spike information is processed. This is illustrated in **Figures 1C–H**. As shown in **Figure 1C**, when a Gaussian stimulus is transmitted through type 1 synapses, the peak of the stimulus is shifted to the left (precession) in addition to a general decrease in amplitude due to depression. No such precession is observed with type 2 synapses (**Figure 1F**), which also depress less. As a result, even if the stimulus arriving from presynaptic neurons is the same, the response of class 1 differs depending on whether the stimulus is transmitted through type 1 (**Figure 1D**) or 2 synapses (**Figure 1G**). Similarly, the response of class 2 neurons differs based on whether the stimulus is transmitted via type 1 (**Figure 1E**) or 2 synapses (**Figure 1H**).

We systematically investigated the operating mode of a neuron for all possible combinations between synaptic types (T) and firing class (C); i.e., T1C1, T1C2, T2C1, and T2C2. In addition, the impact of the number of synapses comprising the stimulus and the injected background noise correlation was also studied. The number of synapses was chosen as a parameter because the extent of the number of synapses influences the amount of synaptic depression. The background noise correlation was included in order to study the interaction with the time constants of synaptic depression and recovery.

T1C2 ALLOWS FOR COINCIDENCE DETECTION

As predicted, the response of a neuron with class 2 firing receiving inputs through type 1 synapses is largely reliable for highly synchronous stimuli (smaller σ_{stim}). A reliable response is, by definition, when $R > 0.75$ (shaded regions in **Figures 2A2** and **2B2**). As the stimulus becomes more dispersed (increasing σ_{stim}), reliability decreases rapidly. This property is robust to variations in noise correlation and number of synapses. Dispersion of stimulus largely determines response precession (**Figures 2A1** and **2B1**). This property is also robust to variations in noise correlation and number of synapses for stimulus dispersion, $\sigma_{stim} < 0.8$.

Varying the synapse number while keeping τ_n to 50 ms reveals the extent to which presynaptic depression dynamics shape the response properties of the neuron. For example, if $N_{tot} = N_{syn}$, each synapse will, on average contribute only one event to the total stimulus. Since the first response of all synapses is identical and depression is apparent only from the second stimulus onwards, no effects of depression can be observed in this case. As the value of N_{syn} is decreased, each synapse receives a greater number of presynaptic spikes to the total stimulus and hence, the responses are subject to more depression. With changing synapse number, the ability for coincidence detection of the T1C2 configuration remains unaltered. Precession is largely determined by the stimulus dispersion (**Figure 2B1**). However, reliability is dependent on the number of synapses (**Figure 2B2**). A decrease in the number of synapses (increase in number of presynaptic spikes delivered to each synapse) results in greater overall depression and hence reduces reliability.

REMAINDER OF THE CONFIGURATIONS ARE LARGELY INTEGRATORS

Responses were reliable ($R > 0.75$) through out the range of simulated stimulus dispersions (0.1–1.4) for T1C1, T2C1, and T2C2 configurations. For T1C1, the reliability was primarily determined by the stimulus dispersion when the noise correlation was varied, keeping $N_{syn} = 1.0$ (**Figure 3A1**). Moreover, reliability did not decrease dramatically as demonstrated by T1C2 configuration, i.e., the coincidence detector. For varying number of synapses (with $\tau_N = 50$ ms), the reliability was determined by the stimulus dispersion and the number of synapses. As might be expected with an increasing number of synapses, reliability drops slightly (**Figure 3A2**) due to increased depression of type 1 synapses. Simultaneously increasing stimulus dispersion also improves reliability of response.

For integrators, an increase in stimulus dispersion must result in an increase in response jitter. We investigated this by computing the slope of this relation for various parameters. The relation between stimulus dispersion and response jitter was always more or less linear with varying slopes. For various values of noise correlation and synapse number, we computed the slope and plotted them against noise correlation (**Figure 3D1**) and number of synapses (**Figure 3D2**). T2C1 and T2C2 exhibited more or less similar slopes. Given that type 1 synapses depress rapidly, a surprising result was that the T1C1 configuration exhibited the steepest slope. This suggests that both pre- and postsynaptic dynamics together determine the operating mode of the neuron.
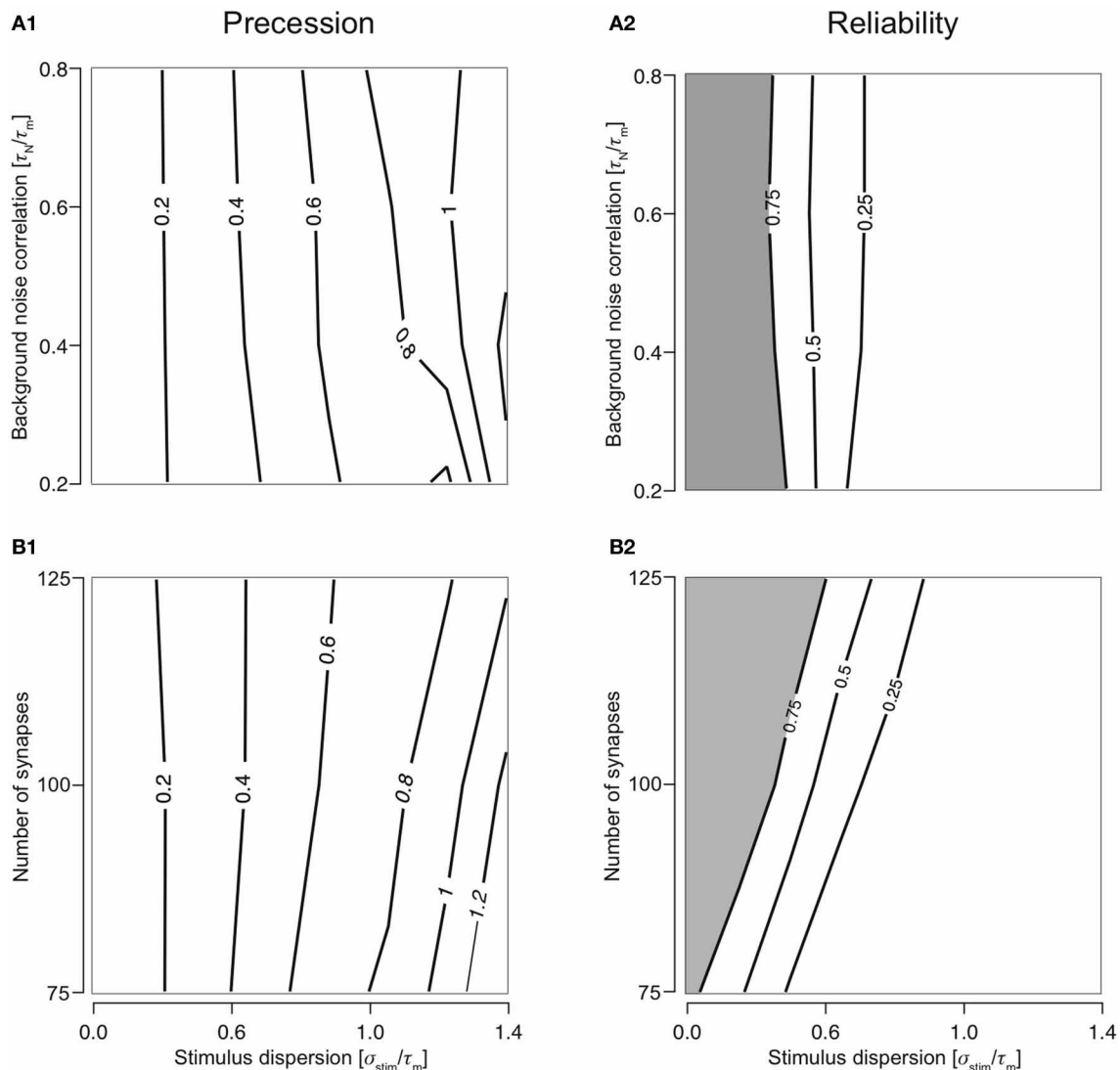
FIGURE 2 | Cellular response with varying number of synapses and background noise correlation. (A) Contour plots showing precession **(A1)** and reliability **(A2)** of response Gaussian distribution with respect to the stimulus distribution and changing background noise correlation (N_{syn} set to 100). Contour lines join points of equal value thus indicating regions in the two-dimensional parameter space (stimulus synchrony vs. noise correlation)

in which an response characteristic of the system is similar even while parameter values change. In addition, contour lines are useful in visualizing regions which are lesser or greater than a specified value. **(B)** Contour plots showing precession **(B1)** and reliability **(B2)** of response Gaussian distribution with respect to the stimulus distribution and synapse number (τ_N is set to 50 ms).

T1C1: PRESERVES SYNCHRONY MOST EFFECTIVELY

To study how the four configurations preserve stimulus synchrony in their response jitter, we investigated the behavior of response sharpening, $\xi = \sigma_{stim}/\sigma_{resp}$. Strictly speaking if $\xi < 1$, the response of the neuron does not preserve stimulus synchrony. Instead, the response jitter is more desynchronized than the stimulus. If $\xi = 1$, stimulus synchrony is preserved. If $\xi > 1$, response synchronization is greater than that of the stimulus; i.e., synchrony is enhanced. We define the region $0.5 < \xi < 1.5$ as preserving the stimulus synchrony in the response jitter. For T1C1 configuration, this region is larger (**Figures 4A1** and **4A2**) than for T2C1 (**Figures 4B1** and **4B2**) and T2C2 configurations

(**Figures 4C1** and **4C2**). For T2C1, the area is least compared to the other two configurations. T2C2 shows the highest sharpening, which is robust to variations in noise correlation (**Figure 4C1**) and number of synapses (**Figure 4C2**). This is consistent with previous work (Pinto et al., 1996; Marella and Ermentrout, 2008), which suggests that class 2 neurons show a greater tendency toward stochastic synchronization than class 1 neurons.

T1C1 neurons show the greatest preservation of stimulus synchrony, especially as dispersion of stimulus increases. An increase time constant of noise correlation results in an increase in the preservation of synchrony (ξ tends toward 1 or lower).
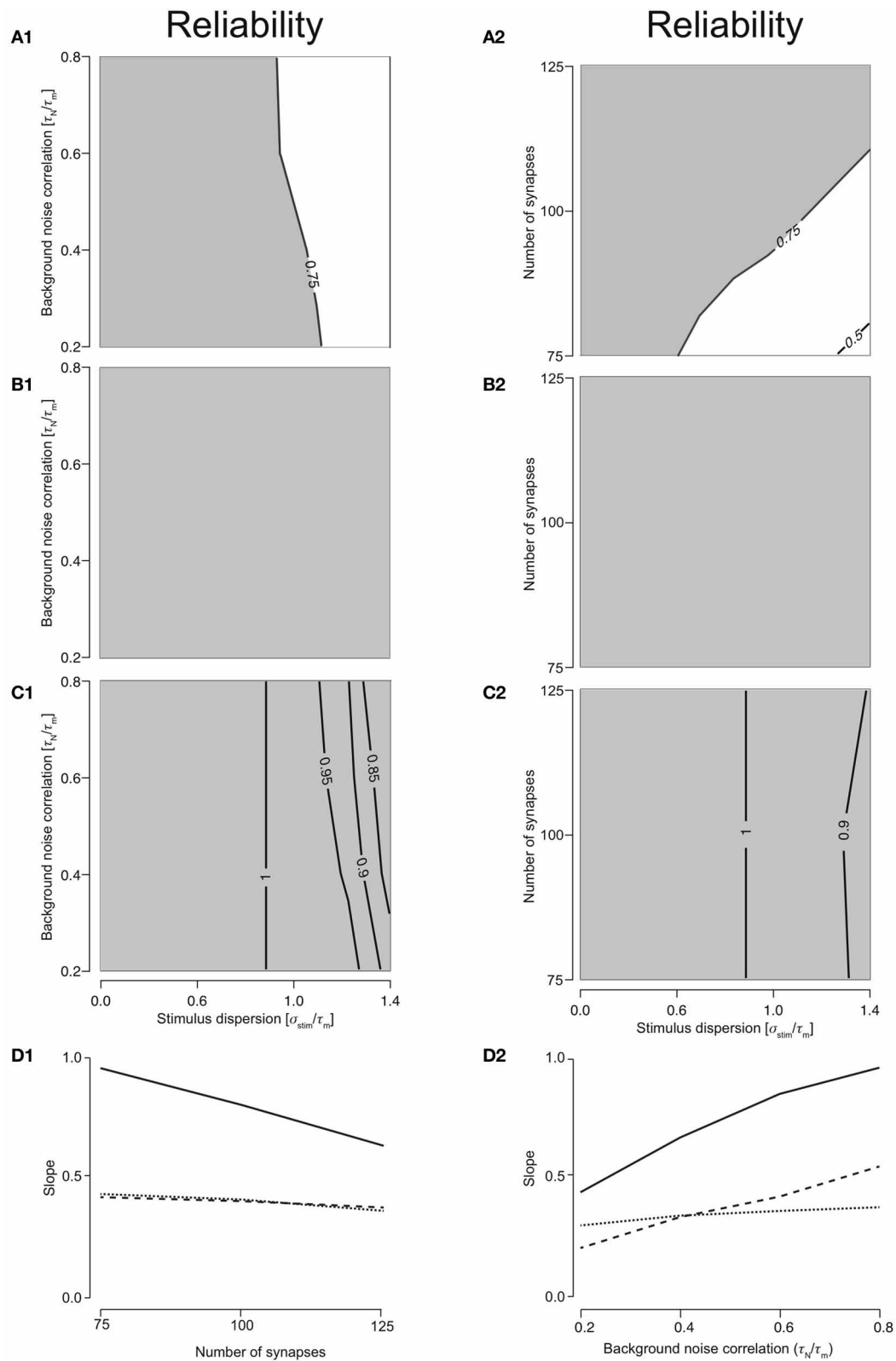
FIGURE 3 | Reliability of the cellular response with varying number of synapses and background noise correlation. (A) Contour plots reliability for varying noise correlation **(A1)** and number of synapses **(A2)** for class 1-type 1 configuration. **(B)** Contour plots showing reliability for varying

noise correlation **(B1)** and number of synapses **(B2)** for class 2-type 1 configuration. **(C)** Contour plots showing reliability for varying noise correlation **(C1)** and number of synapses **(C2)** for class 2-type 2

(Continued)

FIGURE 3 | Continued
configurations. In all of the above contour plots, when noise correlation is varied, N_{syn} is set to 1000 and when number of synapses is varied, τ_N is set to 50 ms. **(D)** Analysis of the slope of relationship between stimulus synchrony and response jitter. Graphs are plotted with the number of synapses **(D1)** or the noise correlation **(D2)** systematically changing along the abscissa with the corresponding slope plotted along the ordinate. **(D1)** For $N_{syn} = 100$ and $\tau_N = 50$, plot showing the relationship between number of synapses and ratio between response and stimulus dispersion. A straight line was fit and the slope computed. This was repeated for all parameter values to obtain relationship between number of synapses and the slope **(D2)** for T1C1 (solid), T2C1 (dashed), and T2C2 (dotted).

In order to explore the preservation of stimulus dispersion by T1C1, we studied the behavior of response sharpening (ξ) for three different total numbers of presynaptic spikes, $N_{tot} = 500$, 750, and 1000. As the total number of spikes increases, the area under the contour indicating synchrony preservation progressively decreases. For 500 stimuli, this area is largest (**Figures 5A1** and **5A2**) with the area decreasing for 750 (**Figures 5B1** and **5B2**) and even more for 1000 (**Figures 5C1** and **5C2**). For highly synchronized stimuli ($\sigma_{stim}/\tau_m < 0.5$), synchrony preservation was primarily determined by noise correlation and only to a much lesser extent by the number of synapses comprising the total stimulus. Type 1 synapses depress rapidly, especially when relaying highly synchronous stimuli at a high frequency. Thus, the response to a change in stimulus to the neuron after depression is minimal and hence it has little effect on synchrony preservation. But for a less synchronous stimulus, preservation of synchrony is dependent on the number of synapses. Type 1 synapses are in a less depressed state and hence small changes in synchrony are relayed to the postsynaptic neuron. Note that even though the area indicating synchrony preservation varies for different number of stimuli, the maximum sharpening for highly synchronous stimuli remains roughly the same (3.2–3.6). This suggests that for a small number of stimuli, synchrony preservation is more robust to variations in noise correlations and number of synapses.

T2C1: MOST RELIABLE INTEGRATOR

For T2C1, responses were always reliable ($R = 1$) when either noise correlation or number of synapses was varied (**Figures 3B1** and **3B2**). This is explained by the fact that type 2 depress less than type 1 synapses. Moreover, they undergo frequency-dependent recovery and hence are much more capable of reliably relaying presynaptic spikes to the neuron. But this property is not entirely dependent on synapse type alone. For T2C2, responses were reliable ($R > 0.75$) when noise correlation or number of synapses was varied (**Figures 3C1** and **3C2**). But reliability is not as perfect as with class 1 neurons. This is because class 2 neurons have a hyperpolarizing current, which reduces the firing an action potential; i.e., reliability. Thus, while synapses with smaller depression can influence a configuration to function as an integrator, synapse type alone does not govern operating mode. For example, class 2 neurons receiving type 1 synapses function as coincidence detectors (see above), but when class 1 neurons receive type 1 synapses, the operating mode is that of an integrator. Thus, operating mode of a configuration is set synergistically by both synaptic and neuronal dynamics.

T2C2: MAXIMUM RESPONSE SHARPENING

For T2C2, we studied the behavior of response sharpening (ξ) for three different total numbers of presynaptic spikes, N_{tot}—500,

750, and 1000 (see Methods). For 500 stimuli, this area is smallest (**Figures 6A1** and **6A2**) and increasing for 750 (**Figures 6B1** and **6B2**) and 1000 stimulus (**Figures 6C1** and **6C2**). For highly synchronized stimuli ($\sigma_{stim}/\tau_m < 0.5$), sharpening influenced by both variations noise correlation and the number of synapses comprising the total stimulus. This result is expected because with a greater number of spikes, the reliability of responses increases and resulting in a decrease in output dispersion.

DISCUSSION

In order to explore the interaction of short-term depression with neuronal firing dynamics in setting the operating mode of the neuron, we studied four canonical combinations of pre- and postsynaptic dynamics. Type 1 synapses show release-dependent depression and constant rate of recovery. They are capable of encoding the stimulus rate change in the response amplitude. Type 2 synapses, on the other hand show release-independent depression, and recover faster at higher rates. They are capable of maintaining substantial response amplitudes even at high stimulus rates. For the postsynaptic dynamics, we considered class 1 neurons that fire regularly and class 2 neurons, which exhibit spike-frequency adaptation. The first action potential response of all four possible combinations (T1C1, T1C2, T2C1 and T2C2) to a stimulus that was Gaussian distributed in time was characterized. We also investigated the sensitivity of these responses to correlations in background noise and to the number of synapses comprising the stimulus.

We found that the combination T1C2 can be characterized as a coincidence detector while the other three combinations were integrators each with specific features: T2C1 was an integrator with greatest reliability, T1C1 an integrator with greatest preservation of synchrony and T2C2 and integrator with greatest response sharpening. Specifically, the degree of reliability and preservation of synchrony varied across these integrators. The sensitivity to noise correlation and the extent of synaptic depression were different.

Though the results are based on simulations using models of dynamical synapses as well as neurons, we believe that our results capture the interactions realistically for the following reasons, Firstly, the synaptic dynamics are based on fitting the chosen model to EPSCs recorded in pairs of neurons *in vitro* (Scott et al., 2012). Individual EPSC peak conductances were set at 1 nS, a value that has been determined experimentally and modeled as alpha synapses with a decay time constant of 1 ms, which is similar to experimentally measured values (Stricker et al., 1996). In addition, varying the extents of type of classes did not systematically change the results in a qualitative sense (data not shown). We tested if it was indeed the adaptation current in class 2 neurons that produced the dynamics or whether an increased conductance of class 1 neurons might be sufficient to reproduce the effect.
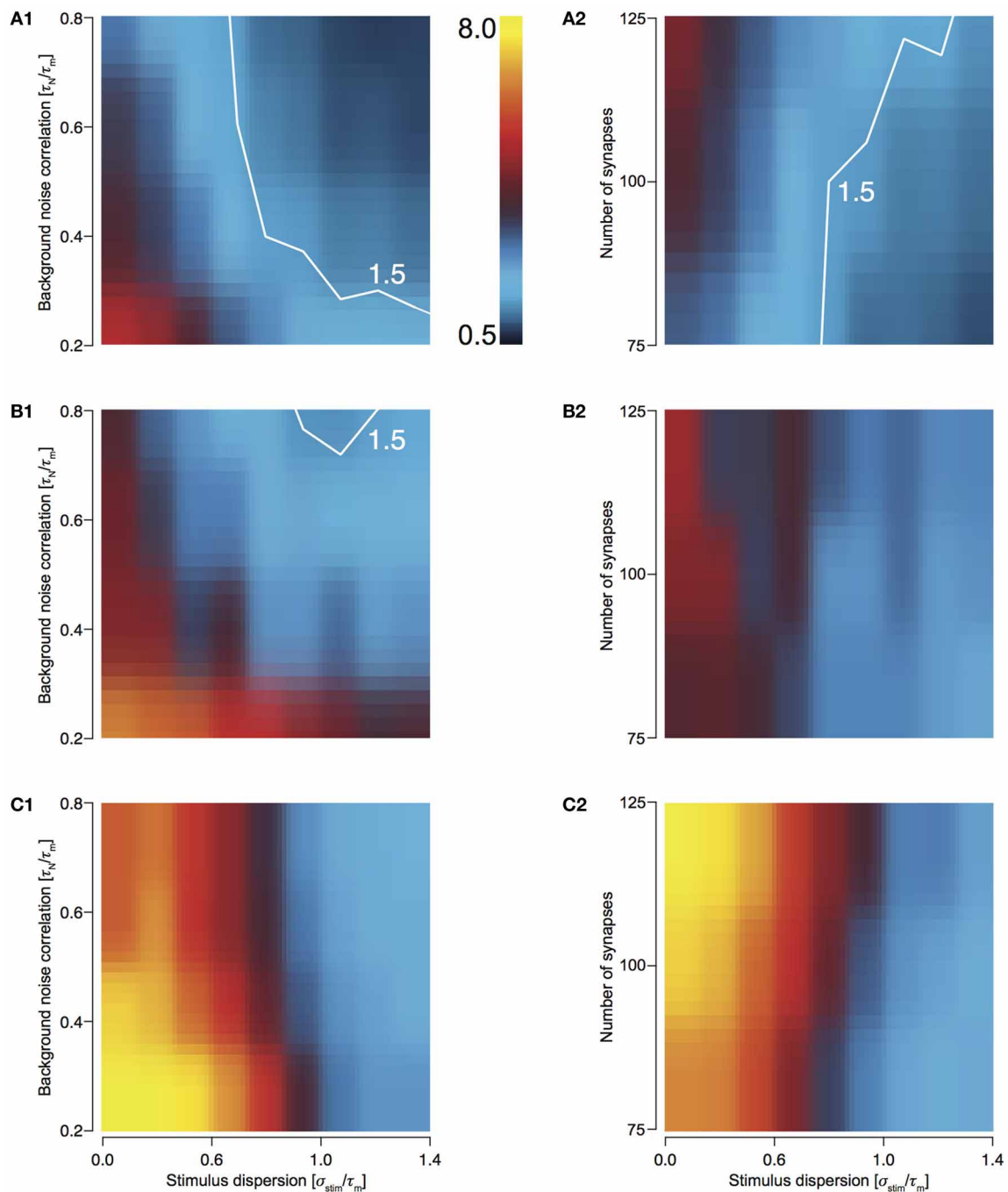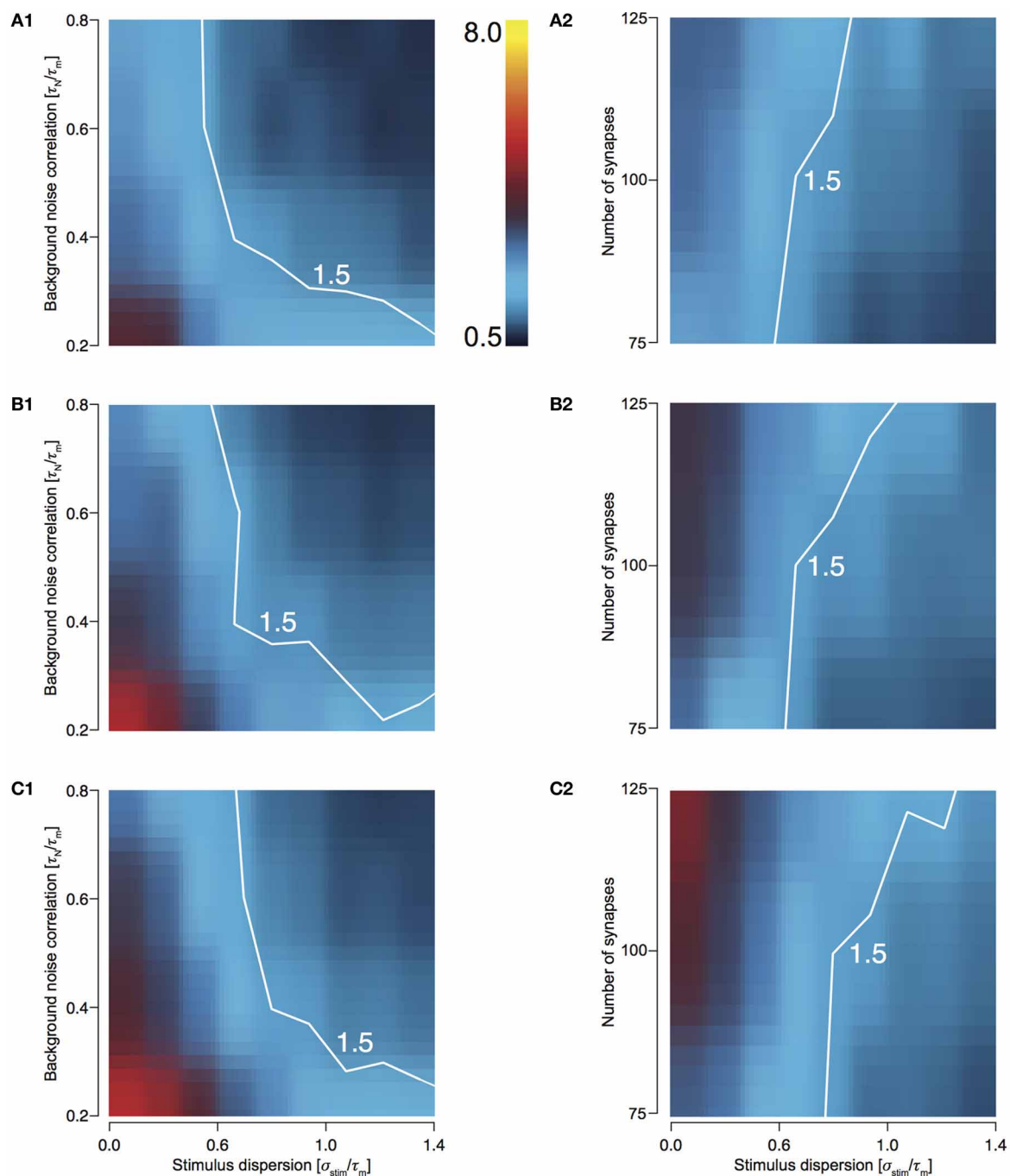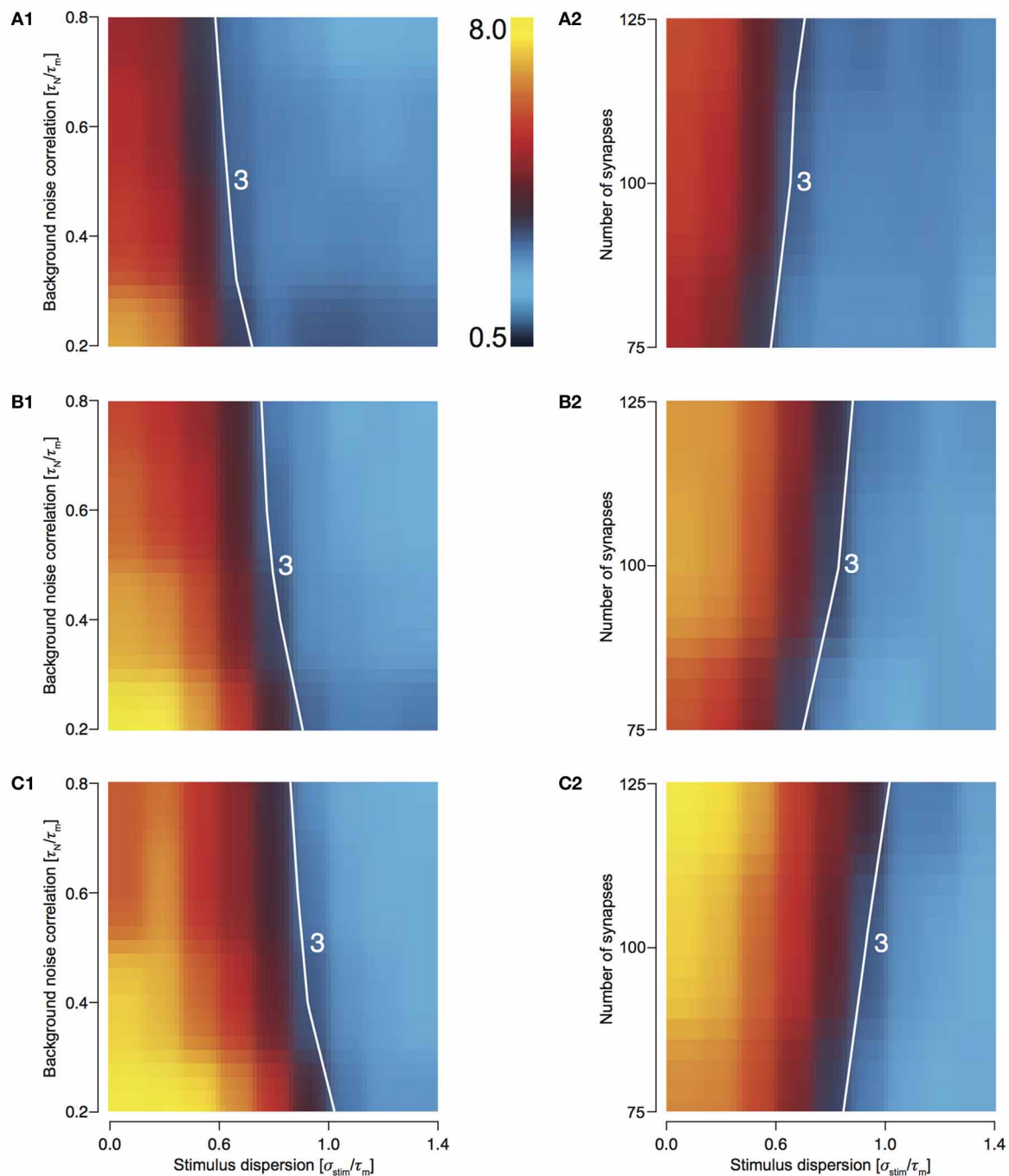
FIGURE 4 | Sharpening of the cellular response with varying number of synapses and background noise correlation. (A–C) Heat plots showing response sharpening for varying noise correlation **(A1)** and number of synapses **(A2)** for T1C1, T2C1, (**B1** and **B2**), and T2C2 (**C1** and **C2**) configurations. In all of the above heat plots, when noise correlation is varied,

N_{syn} is set to 100 and when number of synapses is varied, τ_N is set to 50 ms. Scaling of the heat plot is linear from values of 0.5 to 8. The white line in (**A1,A2**, and **B1**) is an isocline with a value of 1.5. The area circumscribed by this isocline encompasses sharpening values less than or equal to 1.5. The corresponding area in the other graphs is negligible.

Increasing the conductance of a class 1 neuron did not reproduce operating modes that were obtained with class 2 neurons but produced responses that were qualitatively similar to those with class 1 neurons. This is consistent with existing work that suggests that increase an in conductance converts class 2 into class 1 (Stiefel et al., 2008, 2009). Consequently, we think that our results robustly reflect the dynamics between type and class.

FIGURE 5 | Sharpening of the cellular response of T1C1 with varying number of synapses and background noise correlation. (A–C) Heat plots showing sharpening for varying noise correlation **(A1)** and number of synapses **(A2)** when total number of presynaptic spikes (N_{tot}; see Methods) was set to 500, 750 **(B1** and **B2)**, and 1000 **(C1** and **C2)**. In all of the above heat plots, when noise correlation is varied, N_{syn} is set to 100 and when number of synapses is varied, τ_N is set to 50 ms. Scaling of the heat plot is linear from values of 0.5 to 8. The white line in all the above graphs is an isocline with a value of 1.5. The area circumscribed by this isocline encompasses sharpening values less than or equal to 1.5.

Secondly, the postsynaptic neuron had an effective neuronal time constant of 60 ms (in the presence of synaptic background noise), which is similar to experimentally measured values both *in vitro and vivo* (Destexhe et al., 2003). For the cell to fire a

first action potential, typically about 45 synaptic events required to be activated within 10 ms. For class 2 neurons, the adapting current resembled a slow potassium conductance. There are two ways to interpret times of individual events that comprise

FIGURE 6 | Sharpening of the cellular response of T2C2 with varying number of synapses and background noise correlation. (A–C) Heat plots showing sharpening for varying noise correlation **(A1)** and number of synapses **(A2)** when total number of presynaptic spikes (N_{tot}; see Methods) was set to 500, 750, **(B1** and **B2)** and 1000 **(C1** and **C2)**. In all of the above heat plots, when noise correlation is varied, N_{syn} is set to 100 and when number of synapses is varied, τ_N is set to 50 ms. Scaling of the heat plot is linear from values of 0.5 to 8. The white line in all the above graphs is an isocline with a value of 3.0. The area circumscribed by this isocline encompasses sharpening values less than or equal to 3.0.

the stimulus. The first is to consider them presynaptic spike arrival times. The second is to consider them presynaptic spike times. Propagation delays are not considered and hence, if the second interpretation is followed, precessions reported might be

systematically overestimated. Timing of only the first spike was considered. Thus, our results are applicable in a context when the membrane potential of a class 1 or class 2 neuron is near threshold and presynaptic spikes are delivered through type 1 or

type 2 synapses. In this study, information encoded in repetitive spiking is not considered as it is affected not only by incoming signal but also back-propagating action potentials and steady state dynamics.

EMERGENT PROPERTIES THROUGH INTERACTION OF PRE- AND POST-SYNAPTIC DYNAMICS

An important question to answer is if the properties observed were largely the result of either pre- or post-synaptic dynamics alone or if these combinations gave rise to emerging characteristics. We think the latter is the case for the following reasons. Considering presynaptic dynamics separately, the prediction might be that T1C1 and T1C2 are coincidence detectors while T2C1 and T2C2 are integrators. In addition, combinations with type 1 synapses will have reliable responses only when inputs are sufficiently synchronized and combinations with type 2 synapses will have reliable responses over a much higher range of stimulus dispersion. In contrast, considering firing dynamics separately, the prediction might be that T1C1 and T2C1 are integrators and T1C2 and T2C2 are coincidence detectors. Furthermore, combinations with class 1 neurons exhibit reliable responses over a wide range of stimulus dispersion and those with class 2 neurons require synchronous inputs. Since class 2 neurons have a slow hyperpolarizing conductance, stimuli have to be sufficiently short and strong to evoke a response before the slow conductance is activated and decreases the probability of an action potential. However, only some of these predictions are correct. For instance, T1C2 is a coincidence detector, but T1C1 is an integrator with greatest synchrony preservation, even though presynaptic dynamics remain the same. All four configurations have unique properties and hence not considering the contribution of either result in an incomplete view of neuronal encoding. Intuitively, T2C1 is expected to be the most effective integrator and it is indeed from the standpoint of reliability. But T1C1 is a more effective integrator from the standpoint of the relation between stimulus dispersion and response jitter. Stimulus dispersion is more effectively captured by the response dispersion. This can be viewed as a tradeoff between synchrony preservation and reliability.

Both pre- and post-synaptic dynamics contribute for a specific operating mode to emerge. Our results suggest that a complete characterization of neuronal encoding can be obtained only by considering both pre- and post-synaptic dynamics together.

There is evidence for matching of synapse type with firing class in the literature. For example, synapses in layer IV show target-specificity with spiny stellates receiving predominantly type 1 synapses and star pyramids and pyramids receiving predominantly type 2 synapses (Cowan and Stricker, 2004). Such specificity has also been reported in the lobster pyloric network where a disruption of specificity results compromised function (Mamiya and Nadim, 2005). Since each combination performs specific stimulus to response transformations, a slight change in either synapse type or neuron class can cause significant changes in information processing of individual neurons and within the network.

IMPLICATIONS FOR SYNCHRONIZATION AND CODING

The background noise correlation was found to be a critical determinant of response sharpening (ξ) as preservation of stimulus synchrony or its enhancement would have important consequences for processing at the network level. When $\xi > 1$, stimulus synchrony is enhanced by postsynaptic neurons and, thus, the firing becomes more synchronized as excitation is transmitted through subsequent layers (Marsálek et al., 1997). The signal becomes temporally sharpened while losing information about the stimulus dispersion (Gerstein et al., 1989). From the perspective of single neuron oscillations, if ξ is taken to indicate the relation between successive cycles of oscillation, discharges of neurons might become more synchronized ($\xi > 1$), conserve synchrony ($\xi = 1$) or progressively lose synchrony ($\xi < 1$). While previous studies have considered either synaptic dynamics (Mamiya and Nadim, 2005; McDonnell et al., 2012) or neuronal dynamics (Ermentrout, 1996, 1998; Marella and Ermentrout, 2008) in shaping oscillatory dynamics in networks, there was virtually no study exploring how these properties might together determine synchronization of individual neurons and consequently the network. In fact, we show that the combination T1C1 is best suited for preserving input synchrony. In this context, T1C1 might aid in the preservation of asynchrony in a network and might aid in encoding of network information through desynchronization (Hanslmayr et al., 2012). But, in general, network effects of integrator configurations are much harder to speculate about without performing detailed simulations since the larger time window of summation (when compared with the integrators) allows for possible interactions with feedback connections of a recurrent network and the timing of the second action potential might be modulated by network effects. Even so, our results for integrators do have relevance for network processing since sharpening (see T1C1: Preserves Synchrony Most Effectively and T2C2: Maximum Response Sharpening) and delay to fire first action potential (data not shown) will influence the overall network encoding.

TYPE AND CLASS MIGHT ENHANCE INFORMATION PROCESSING

For the purpose of this paper, both synapse type and firing class were taken to be discrete properties. However, experimental evidence shows that type 1 and type 2 synapses exist along a continuum between release-dependence and release-independence and various experimental conditions can alter the extent of the release-dependence (Cowan and Stricker, 2004; Fuhrmann et al., 2004). Likewise, postsynaptic firing can vary smoothly between class 1 and class 2 properties (Stiefel et al., 2008, 2009). In addition, both the synapse type (unpublished data) and the firing class (Stiefel et al., 2008) can be altered concomitantly by neuromodulators like noradrenaline, and, thus, can be converted into each other. Further, there is intrinsic variability in firing dynamics among neurons of the same type (Schulz et al., 2006) that might be critical for maximizing information content (Padmanabhan and Urban, 2010). Our results suggest that variability in synapse type and firing class allows for specific neurons in the same network to capture and thereby encode different aspects of the stimulus. For instance, combinations with T1C2 properties would act as coincidence detectors. Upon exposure to a neuromodulator like

noradrenaline, both type and class are converted to become more T2C1-"like" and as a consequence, the same node in the network would act now as an integrator with greatest reliability. Any partial conversion along type and/or class would allow for other features about the stimulus to be encoded. For instance, the combination of T1C2 (coincidence detector) might be converted to a reliable integrator (T2C1) by concomitant conversion of type and class due to adrenergic modulation. For the same condition, T1C1 (integrator with greatest synchrony preservation) would be converted to T2C1, an integrator with improved reliability but loss of synchrony preservation. Thus, neurons in a network might be tuned to capture and encode various stimulus properties of interest.

REFERENCES

Abbott, L. F. (1997). Synaptic depression and cortical gain control. *Science* 275, 221–224.

Brecht, M., and Sakmann, B. (2002). Dynamic representation of whisker deflection by synaptic potentials in spiny stellate and pyramidal cells in the barrels and septa of layer 4 rat somatosensory cortex. *J. Physiol.* 543, 49–70.

Brette, R., and Gerstner, W. (2005). Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.* 94, 3637–3642.

Brody, D., and Yue, D. (2000). Release-independent short-term synaptic depression in cultured hippocampal neurons. *J. Neurosci.* 20, 2480–2494.

Cowan, A., and Stricker, C. (2004). Functional connectivity in layer IV local excitatory circuits of rat somatosensory cortex. *J. Neurophysiol.* 92, 2137–2150.

deCharms, R. C., and Zador, A. (2000). Neural representation and the cortical code. *Annu. Rev. Neurosci.* 23, 613–647.

Destexhe, A., Rudolph, M., and Paré, D. (2003). The high-conductance state of neocortical neurons *in vivo*. *Nat. Rev. Neurosci.* 4, 739–751.

Dittman, J. S., Kreitzer, A. C., and Regehr, W. G. (2000). Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. *J. Neurosci.* 20, 1374–1385.

Dobrunz, L. E., Huang, E. P., and Stevens, C. F. (1997). Very short-term plasticity in hippocampal synapses. *Proc. Natl. Acad. Sci. U.S.A.* 94, 14843–14847.

Ermentrout, B. (1996). Type I membranes, phase resetting curves, and synchrony. *Neural Comput.* 8, 979–1001.

Ermentrout, B. (1998). Linearization of F-I curves by adaptation. *Neural Comput.* 10, 1721–1729.

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

Fung, C. C. A., Wong, K. Y. M., Wang, H., and Wu, S. (2012). Dynamical synapses enhance neural information processing: gracefulness, accuracy, and mobility. *Neural Comput.* 24, 1147–1185.

Gerstein, G. L., Bedenbaugh, P., and Aertsen, M. H. (1989). Neuronal assemblies. *IEEE Trans. Biomed Eng.* 36, 4–14.

Gillespie, D. (1996). Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral. *Phys. Rev. E. Stat. Phys. Plasmas. Fluids. Relat. Interdiscip. Topics* 54, 2084–2091.

Graham, B., and Stricker, C. (2008). "Short term plasticity provides temporal filtering at chemical synapses," in *Proceedings of the 18th International Conference on Artificial Neural Networks, Part, II* (Prague), 268–276.

Hanslmayr, S., Staudigl, T., and Fellner, M.-C. (2012). Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Front. Hum. Neurosci.* 6:74. doi: 10.3389/fnhum.2012.00074

Hodgkin, A. L. (1948). The local electric changes associated with repetitive action in a non-medullated axon. *J. Physiol.* 107, 165–181.

Izhikevich, E. (2000). Neural excitability, spiking and bursting. *Int. J. Bifurcat. Chaos* 10, 1171–1266.

Jedrzejewska-Szmek, J., and Zygierewicz, J. (2010). Depressing synapse as a detector of frequency change. *J. Theor. Biol.* 266, 380–390.

London, M., Larkum, M. E., and Häusser, M. (2008). Predicting the synaptic information efficacy in cortical layer 5 pyramidal neurons using a minimal integrate-and-fire model. *Biol. Cybern.* 99, 393–401.

Mamiya, A., and Nadim, F. (2005). Target-specific short-term dynamics are important for the function of synapses in an oscillatory neural network. *J. Neurophysiol.* 94, 2590–2602.

Marella, S., and Ermentrout, G. B. (2008). Class-II neurons display a higher degree of stochastic synchronization than class-I neurons. *Phys. Rev. E* 77:041918. doi: 10.1103/PhysRevE.77.041918

Markram, H., Lubke, J., Frotscher, M., Roth, A., and Sakmann, B. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol.* 500, 409–440.

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382, 807–810.

Marsálek, P., Koch, C., and Maunsell, J. (1997). On the relationship between synaptic input and spike output jitter in individual neurons. *Proc. Natl. Acad. Sci. U.S.A.* 94, 735–740.

Matveev, V., and Wang, X. J. (2000). Differential short-term synaptic plasticity and transmission of complex spike trains: to depress or to facilitate? *Cereb. Cortex* 10, 1143–1153.

McDonnell, M. D., Mohan, A., Stricker, C., and Ward, L. M. (2012). Input-rate modulation of gamma oscillations is sensitive to network topology, delays and short-term plasticity. *Brain Res.* 1434, 162–177.

Muñoz-Cuevas, J., Vara, H., and Colino, A. (2004). Characterization of release-independent short-term depression in the juvenile rat hippocampus. *J. Physiol.* 558, 527–548.

Padmanabhan, K., and Urban, N. N. (2010). Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. *Nat. Neurosci.* 13, 1276.

Paik, S.-B., and Glaser, D. A. (2010). Synaptic plasticity controls sensory responses through frequency-dependent gamma oscillation resonance. *PLoS Comput. Biol.* 6:e1000927. doi: 10.1371/journal.pcbi.1000927

Pinto, D., Brumberg, J., Simons, D., and Ermentrout, G. (1996). A quantitative population model of whisker barrels: re-examining the Wilson-Cowan equations. *J. Comput. Neurosci.* 3, 247–264.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes*. 3rd Edn. Cambridge, UK: Cambridge University Press.

Regehr, W. G. (2012). Short-term presynaptic plasticity. *Cold Spring Harb. Perspect. Biol.* 4, a005702.

Rudolph, M., and Destexhe, A. (2001). Do neocortical pyramidal neurons display stochastic resonance? *J. Comput. Neurosci.* 11, 19–42.

Schulz, D. J., Goaillard, J.-M., and Marder, E. (2006). Variable channel expression in identified single and electrically coupled neurons in different animals. *Nat. Neurosci.* 9, 356–362.

Scott, P., Cowan, A. I., and Stricker, C. (2012). Quantifying impacts of short-term plasticity on neuronal information transfer. *Phys. Rev. E* 85:041921. doi: 10.1103/PhysRevE.85.041921

Senn, W., Segev, I., and Tsodyks, M. (1998). Reading neuronal synchrony with depressing synapses. *Neural Comput.* 10, 815–819.

Shadlen, M. N., and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* 18, 3870–3896.

Stiefel, K. M., Gutkin, B. S., and Sejnowski, T. J. (2008). Cholinergic neuromodulation changes phase response curve shape and type in cortical pyramidal neurons. *PLoS ONE* 3:e3947. doi: 10.1371/journal.pone.0003947

Stiefel, K. M., Gutkin, B. S., and Sejnowski, T. J. (2009). The effects of cholinergic neuromodulation on neuronal phase-response curves of modeled cortical neurons. *J. Comput. Neurosci.* 26, 289–301.

Stricker, C., Field, A. C., and Redman, S. J. (1996). Statistical analysis of amplitude fluctuations in EPSCs evoked in rat CA1 pyramidal neurones *in vitro*. *J. Physiol.* 490 (Pt 2), 419–441.

Thomson, A. (1997). Activity-dependent properties of synaptic transmission at two classes of connections made by rat neocortical pyramidal axons *in vitro*. *J. Physiol.* 502, 131–147.

Touboul, J., and Brette, R. (2008). Dynamics and bifurcations of the adaptive exponential integrate-and-fire model. *Biol. Cybern.* 99, 319–334.

Triesch, J. (2007). Synergies between intrinsic and synaptic plasticity

mechanisms. *Neural Comput.* 19, 885–909.

Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., and Nelson, S. B. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature* 391, 892–896.

Xie, X., Song, D., Wang, Z., Marmarelis, V. Z., and Berger, T. W. (2006). Interaction of short-term neuronal plasticity and synaptic plasticity revealed by nonlinear systems analysis in dentate granule cells. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 1, 5543–5546.

Adaptive neural information processing with dynamical electrical synapses

Lei Xiao[1], Dan-ke Zhang[2,3], Yuan-qing Li[2], Pei-ji Liang[1]* and Si Wu[3]*

[1] Department of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China
[2] School of Automation Science and Engineering, South China University of Technology, Guangzhou, China
[3] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

The present study investigates a potential computational role of dynamical electrical synapses in neural information process. Compared with chemical synapses, electrical synapses are more efficient in modulating the concerted activity of neurons. Based on the experimental data, we propose a phenomenological model for short-term facilitation of electrical synapses. The model satisfactorily reproduces the phenomenon that the neuronal correlation increases although the neuronal firing rates attenuate during the luminance adaptation. We explore how the stimulus information is encoded in parallel by firing rates and correlated activity of neurons, and find that dynamical electrical synapses mediate a transition from the firing rate code to the correlation one during the luminance adaptation. The latter encodes the stimulus information by using the concerted, but lower neuronal firing rate, and hence is economically more efficient.

Keywords: electrical synapses, short-term plasticity, information processing, adaptation, dynamical encoding

INTRODUCTION

In the central nervous system, neurons communicate with each other via two basic forms of synapse: chemical and electrical synapses (Kandel et al., 2000). A chemical synapse is asymmetric in structure, which passes information from a presynaptic neuron to a postsynaptic one through neurotransmitters release, and this occurs when the presynaptic neuron fires an action potential. An electrical synapse, on the other hand, is bidirectional, which allows signal to be transmitted in both ways. Compared to a chemical one, an electrical synapse is usually fast and underlies rapid communication among neighboring neurons of the same type.

It is well known that the strength of a chemical synapse can undergo a variety of short and long-term plasticity (Tsodyks and Markram, 1996; Bi and Poo, 1998; Dan and Poo, 2006). It has also been shown in experimental studies that the strength of an electrical synapse can be modulated similarly as a chemical one. For instances, it was found that titanic stimulation can lead to either long- or short-term potentiation of electrical synapses in goldfish (Yang et al., 1990; Pereda and Faber, 1996); in the rat thalamic reticular nucleus, titanic stimulation can cause long-term depression in the electrical synapses (Landisman and Connors, 2005; Haas et al., 2011); and in the vertebrate retina, electrical synapses can be dynamically regulated by either ambient illumination or circadian rhythms (Bloomfield and Volgyi, 2009).

Although a large volume of experimental data has revealed the abundant existence and the plasticity of electrical synapses in the neural system, their functional roles in neural information processing remain largely unclear (Connors and Long, 2004). In the thalamic reticular nucleus, electrical synapses may contribute to the shift between arousal states (Haas et al., 2011). In the retina, electrical synapses are sensitive to the background light conditions (Bloomfield and Volgyi, 2009), and the synchronous activity of electrically coupled ON direction-selective ganglion cells may encode the direction information of a moving stimulus (Ackert et al., 2006). It was also found that retinal ganglion cells (RGCs) coupled with electrical synapses exhibit stronger concerted activity than connected (indirectly) with chemical synapses in a circuit (Brivanlou et al., 1998; Jing et al., 2010).

In the present study, we investigate a potential role of electrical synapses in processing stimulus information during luminance adaptation. We first explore the effects of electrical and chemical synapses on generating neural correlation. We find that the neuronal correlation strength is much more sensitive to the plasticity of an electrical synapse than to the plasticity of a chemical one, indicating the potential importance of electrical synapses in modulating synchrony of neuronal activities. We then propose a phenomenological model for short-term facilitation of electrical synapses, based on the experimental finding that during the luminance adaptation, the neuronal correlation strength increases whereas the firing rates attenuate. The proposed model satisfactorily reproduces the experimental data. Finally, we explore the computational role of dynamical electrical synapses, and find that they contribute to generate a transition in encoding properties

during the adaptation. The implication of this transition is discussed.

MATERIALS AND METHODS

THE NEURON-PAIR MODELS

To investigate the effects of electrical and chemical synapses on generating correlated neuronal responses, we construct neuron-pair models coupled by either an electrical or a chemical synapse as shown in **Figures 1A,C**.

For neurons coupled with an electrical synapse, the dynamics of neurons are written as

$$C\frac{dV_i(t)}{dt} = -g_L[V_i(t) - V^{rest}] + g^{es}[V_j(t - d^{es})$$
$$-V_i(t)] + I_i^{ext}(t), \quad i, j = 1, 2 \tag{1}$$

where V_i is the membrane potential of the ith neuron, C the membrane capacitance, g_L the leaky conductance, $V^{rest} = -70$ mV the resting potential, and I_i^{ext} the external input current. g^{es} represents the conductance of the electrical synapse, which is a constant unless the synapse is undergoing plasticity. d^{es} denotes the transmission delay of the electrical synapse, which is in the range of 0.2–0.4 ms according to the experimental data (Brivanlou et al., 1998; Li et al., 2012). The neuron fires when its membrane potential reaches to a threshold $V^{th} = -50$ mV, and V_i is reset to be $V^{reset} = -70$ mV after firing.

For neurons connected by a chemical synapse, the dynamics of neurons are written as

$$C\frac{dV_i(t)}{dt} = -g_L[V_i(t) - V^{rest}] - g_{ij}^{cs}(t - d^{cs})[V_i(t)$$
$$-V^{rev}] + I_i^{ext}(i), \quad i, j = 1, 2 \tag{2}$$

where $V^{rev} = 0$ mV denotes the reversal potential. g_{ij}^{cs} is the conductance of the chemical synapse from the neuron j to i, whose dynamics is given by

$$\tau^s \frac{dg_{ij}^{cs}}{dt} = -g_{ij}^{cs} + u \sum_m \delta(t - t_j^m) \tag{3}$$

where τ^s is the synaptic time constant, t_j^m the moment when the mth spike of the jth neuron is generated, and u the increment of the chemical conductance due to a spike generation. d^{cs} denotes the transmission delay of the chemical synapse, which is in the range of 2–3 ms.

The external inputs to the neurons are given by (see **Figures 1A,C**)

$$I_i^{ext} = \mu(t) + \sigma[\sqrt{1-c}\xi_i(t) + \sqrt{c}\xi_c(t)], \quad i = 1, 2 \tag{4}$$

where $\mu(t)$ is the mean of the inputs. $\xi_i(t)$ is Gaussian white noise of zero mean and unit variance. Noise processes of the two neurons are independent to each other, i.e., $< \xi_i(t)\xi_j(t') > = \delta_{ij}\delta(t - t')$. $\xi_c(t)$ denotes the common noise to both

FIGURE 1 | Neuronal circuitry with electrical and chemical synapses.
(A) A model of electrically coupled neurons (left panel) and examples of simulated firing activities (right panel). Neurons receive external currents with both common and independent components. **(B)** The cross correlation function between electrically coupled neurons with bin size of 2 ms. Inset shows the cross correlation function with bin size of 0.1 ms. **(C)** A model of chemically connected neurons (left panel) and examples of simulated firing activities (right panel). Neurons receive external currents with both common and independent components. **(D)** The cross correlation function between chemically connected neurons with bin size of 2 ms. Inset shows the cross correlation function with bin size of 0.1 ms. The parameters values for the two conditions have been chosen to fit the two models to have similar behavior, which are: $C = 0.5$ nF, $g_L = 0.025\,\mu$S, $\tau^s = 5$ ms, $\mu = 0.62$ nA, $\sigma = 0.5$ nA, $g^{es} = 0.025\,\mu$S, $u = 0.05\,\mu$S, $c = 0$ for the electrical synapse, $c = 0.5$ for the chemical synapse, and simulation step size $= 0.1$ ms.

neurons. σ is the noise strength. The parameter $0 \leq c \leq 1$ determines the correlation strength between the inputs to the two neurons.

MEASURING THE CORRELATION STRENGTH

To quantify the characteristics of neural response, we divided time into small bins. A spike train is symbolized into "0" and "1" within a time bin, where $r_i(t) = 1$ means that the cell i fires in the tth time bin and "0" means that it does not fire. We use cross-correlation function (CCF) to measure the correlation strength between neurons. The value of CCF between two spike trains is calculated to be

$$CCF(\Delta t) = \frac{N \sum_{t=1+|\Delta t|}^{N-|\Delta t|} r_1(t) r_2(t + \Delta t)}{(N - 2|\Delta t|) \sqrt{\sum_{t=1}^{N} r_1(t)^2 \sum_{t=1}^{N} r_2(t)^2}} \quad (5)$$

where $r_i(t) = 0, 1$, for $i = 1, 2$, denotes the spike train generated by the ith neuron at the moment t and N indicates the length of spike train. The peak value around zero lag of CCF is used to represent the neuronal correlation strength.

THE EXPERIMENTAL DATA

In this study, we use two sets of experimental data. Both experiments were performed on isolated bullfrog retinas, and the experimental procedures and equipments have been described in detail in (Li et al., 2012; Xiao et al., 2012). The previous works did not study the model and the functional role of electrical synapses presented in this paper.

In the first experiment (Li et al., 2012), the bullfrog retina was exposed to flickering pseudo-random checker-boards for 100 s (frame refresh rate = 20 Hz), and a multi-electrode system was used to record the responses of RGCs simultaneously. **Figures 4A,B** present the experimental results. We use this set of data to fit the phenomenological model for short-term facilitation of electrical synapses during the luminance adaptation. The model is then applied to interpret the neural data in the second experiment.

In the second experiment (Xiao et al., 2012), the bullfrog retina was exposed to flicking pseudo-random checker-boards for 15 s followed by a sustained dark stimulation. The whole adaptation process to the dark stimulus lasted for about 5 s. **Figures 5A,B** present the experimental results. We use this set of data to explore the potential functional role of dynamical electrical synapses.

Both experiments were strictly conformed to the humane treatment and use of animals as prescribed by the Association for Research in Vision and Ophthalmology, and were approved by the Ethic Committee, School of Biomedical Engineering, Shanghai Jiao Tong University.

MEASURING THE STIMULUS INFORMATION CARRIED BY FIRING RATE AND CORRELATION

Denote $p(\mathbf{r}|s)$ the conditional probability of observing the neural response $\mathbf{r}$ given the stimulus s. We regard the dark stimulation and the random flicking check-boards as two stimuli, which occur with equal probability, i.e., $p(s) = 1/2$. For two neurons, $\mathbf{r} = \{r_1, r_2\}$. The bin size is 5 ms, unless it is stated specifically. The total amount of the stimulus information that can be extracted

from the neuronal data is given by the mutual information (Shannon, 1948),

$$I = - \int d\mathbf{r} p(\mathbf{r}) \log_2 p(\mathbf{r}) + \int d\mathbf{r} \sum_s p(s) p(\mathbf{r}|s) \log_2 p(\mathbf{r}|s) \quad (6)$$

To decompose the stimulus information into portions carried by different features of neuronal activities, we choose to use the information measure I^*, which is known to be directly linked to the decoding error of maximum likelihood inference based on a mismatched model (Wu et al., 2001; Oizumi et al., 2010). I^* quantifies the information gain when a mismatched neural encoding model $q(\mathbf{r}|s)$ is applied, and is calculated as (Merhav, 1994),

$$I^*(q) = \max_\beta \tilde{I}(q, \beta) \quad (7)$$

$$\tilde{I}(q, \beta) = - \int d\mathbf{r} p(\mathbf{r}) \log_2 \sum_s p(s) q(\mathbf{r}|s)^\beta$$

$$+ \int d\mathbf{r} \sum_s p(s) p(\mathbf{r}|s) \log_2 q(\mathbf{r}|s)^\beta \quad (8)$$

where β is a parameter to be optimized. This information measure has been applied recently for studying neural coding (Oizumi et al., 2010). By choosing the form of $q(\mathbf{r}|s)$ properly, the amount of the stimulus information contained in different features of neural responses can be obtained.

When two spike trains (binary variables) are considered, the joint probability of neural responses can be written as (Amari, 2001),

$$p(\mathbf{r}|s) = \frac{1}{Z} \exp\left(\sum_i \theta_i^1 r_i + \sum_{i<j} \theta_{ij}^2 r_i r_j \right) \quad (9)$$

where Z is the normalization factor, and the parameters θ_i^1 is related to the firing rate of the ith neuron, θ_{ij}^2 is related to the correlation between the ith and jth neurons. The values of θ^1 and θ^2 can be uniquely determined by matching $p(\mathbf{r}|s)$ with the real distribution of the data.

Suppose we choose $q(\mathbf{r}|s)$ to be the probability distribution which has the same firing rates as $p(\mathbf{r}|s)$ but with vanishing correlation between neurons, i.e.,

$$q(\mathbf{r}|s) = \frac{1}{Z_1} \exp\left(\sum_i \theta_i^1 r_i \right) \quad (10)$$

where Z_1 is the normalization factor. The parameter θ_i^1 is determined by the requirement that the firing rates remain the same for both distributions p and q. Thus, the value $I^*(q)$, refer to as I_1 hereafter, is the amount of the stimulus information contained in the firing rates of neurons. Its discrepancy to the mutual information, denoted as $I_2 = I - I_1$ hereafter, is the amount of the stimulus information contained in the correlation. The relative contributions of firing rate and correlation are measured by the ratios, $R_i = I_i/I$, for $i = 1, 2$.

RESULTS

NEURAL CORRELATIONS GENERATED BY ELECTRICAL AND CHEMICAL SYNAPSES

Neurons can be connected by either electrical or chemical synapses. We investigate how different forms of synapse affect neuronal correlation. The neuron-pair models coupled by either an electrical or a chemical synapse as shown in **Figures 1A,C** are used (see Materials and Methods). The correlation strength is measured by the *CCF* between the spike trains generated by two neurons.

Figure 1B shows the *CCF* for the electrically coupled neurons. We see that the *CCF* exhibits a narrow peak for the bin size of 2 ms, indicating that the two neurons' responses are largely synchronized. If the bin size is 0.1 ms, the *CCF* has dual peaks around $\Delta t = 0$ due to the transmission delay of the electrical synapse (**Figure 1B** inset). These results agree with the experimental data for electrically coupled RGCs in the bullfrog retina (Li et al., 2012).

Figure 1D shows the *CCF* for the neurons connected via a chemical synapse. We see that the *CCF* has a much broader distribution than for the electrical synapse. This property is general and reflects that a chemical synapse is slow and that the correlation it generates is usually small.

Figure 2 displays how the synaptic strength affects the neuronal correlation strength. For the electrical synapse, the correlation strength varies significantly for different conductance values of g^{es} (**Figure 2A**). On the other hand, for the chemical synapse, the correlation strength is rather insensitive to the coupling parameter u (**Figure 2B**, the chemical conductance g^{cs} increases with u). In the case of electrical synapse, the neuronal correlation can be very strong even when the input correlation is very small (for very small *c*-values); whereas, in the case of chemical synapse, the neuronal correlation can only be strong when the input correlation is sufficiently large (for very large *c*-values). An intuitive justification for this is that a chemical synapse is slow and its effect on coordinating neuronal activities is diminished by input noises

and the resetting of the membrane potential after neural firing.

We have only presented the result for the case that there is a single excitatory chemical synapse from the neuron 1 to 2 (see **Figure 1B**). For the case that there exists a reciprocal chemical synapse from the neuron 2 to 1, the property about correlation strength shown in **Figure 2B** still holds (data not shown).

In the present study, the membrane potential of a neuron after firing was reset to be the resting value, i.e., $V^{reset} = V^{rest} = -70$ mV. Alternatively, we could reset the neuron to be hyperpolarized after firing, e.g., $V^{reset} = -85$ mV. We found that this did not change our results qualitatively, and that hyperpolarization tended to increase the robustness of neural correlation mediated by gap-junction to noises.

We further check for fixed synaptic strength, how the correlation strength changes with the neuronal firing rates. As expected, the correlation strength increases with the firing rates (**Figure 3**). This is understandable, since larger firing rates enlarge the effects of neuronal interaction via both forms of synapse.
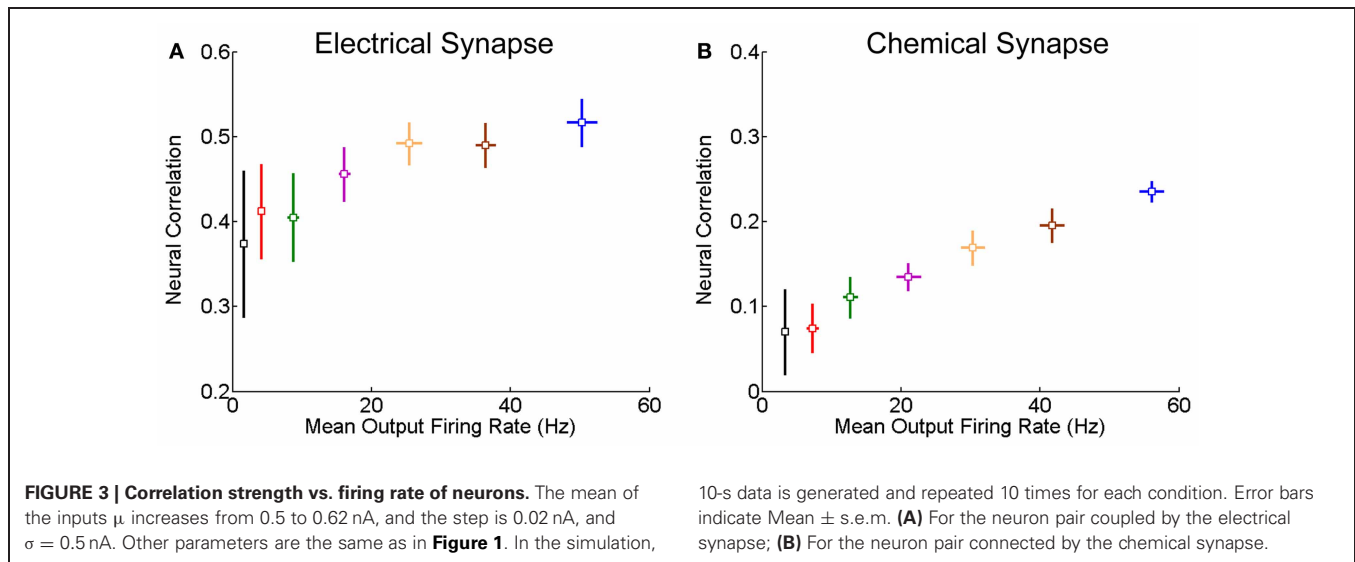
A PHENOMENOLOGICAL MODEL FOR SHORT-TERM FACILITATION OF ELECTRICAL SYNAPSE

We explore how an electrical synapse may vary with time during the adaptation of neuronal responses. The experiment was performed on an isolated bullfrog retina, which was exposed to flickering pseudo-random checker-boards for 100 s (Li et al., 2012; see Materials and Methods). A multi-electrode system was used to record the responses of RGCs simultaneously.

As shown in **Figure 4A**, the responses of RGCs exhibits a clear adaptive behavior, in terms of that the firing rates of RGCs first increase quickly at the onset of the stimulation and then they decrease gradually to a much lower value. We measure during this adaptation process, how the correlation strength between neurons coupled by an electrical synapse changes with time. The result is presented in **Figure 4B**, which shows that the correlation strength first increases with time in the first 20 s (**Figure 4B** inset) and then decreases gradually.

FIGURE 2 | The effect of coupling strength on the correlation strength between neurons. In the simulation, 10-s data is generated and repeated for 10 times. Bin size is 2 ms, and error bars indicate

Mean ± s.e.m. **(A)** For the electrical synapse. **(B)** For the chemical synapse. The parameters are the same as in **Figure 1**, but varying parameter *c*.

FIGURE 3 | Correlation strength vs. firing rate of neurons. The mean of the inputs μ increases from 0.5 to 0.62 nA, and the step is 0.02 nA, and σ = 0.5 nA. Other parameters are the same as in **Figure 1**. In the simulation,

10-s data is generated and repeated 10 times for each condition. Error bars indicate Mean ± s.e.m. **(A)** For the neuron pair coupled by the electrical synapse; **(B)** For the neuron pair connected by the chemical synapse.

The fact that the neuronal correlation increases whereas the firing rates attenuate at the initial stage of the adaptation is not a trivial property. According to the result in **Figure 3**, for fixed synapse strength, the neuronal correlation should decrease with the attenuation of firing rates. We therefore, suspect that an enhancement of the neuronal interaction efficacy is going on during the adaptation (see Discussion for alternative mechanisms). Furthermore, it has been shown that the plasticity of a chemical synapse is insufficient to induce large change in the correlation strength (**Figure 2B**). Thus, we propose that it is the short-term facilitation of the electrical synapse leading to this paradox phenomenon.

To describe the experimental data, we propose the following phenomenological model for short-term facilitation of an electrical synapse, which is given by

$$\tau^f \frac{dg^{es}}{dt} = -\left(g^{es} - g_0^{es}\right) + u^f \left(g_{max}^{es} - g^{es}\right) \exp\left(\frac{-|\Delta T|}{\tau^l}\right) \quad (11)$$

where τ^f is the time constant of short-term facilitation. g_0^{es} and g_{max}^{es} are the static and the maximum values of g^{es}, respectively. ΔT denotes the time difference between two adjacent spikes generated by the two neurons. τ^l determines the time window for plasticity and u^f the rate of facilitation. This plasticity rule states that if two neurons fire strongly and synchronously in a short-time window, their electrical synapse is temporally enhanced. We fix the parameters in the model by the experimental data in **Figures 4A,B**. Once their values are determined, the model will be used to explain the results from another experimental data shown in **Figure 5**.

To mimic the luminance adaptation condition, we set the mean of the inputs to be $\mu(t) = 0.8e^{-t/a}$ with $t = 0$ being the moment of the stimulation onset. $\mu(t)$ decreases with time, reflecting that the current from bipolar/amacrine cells to a RGC attenuates during luminance adaptation (Baccus and Meister, 2002). We choose the parameter $a = 20$ s, so that the simulation results match the experimental data.

Combining Equations (1) and (11), we simulate the neuronal responses during the adaptation. **Figure 4C** displays how the firing rate of a neuron changes over time, which reproduces the adaption behavior observed in the experiment (**Figure 4A**). **Figure 4D** displays how the correlation strength changes over time, which reproduces the experimental observations shown in **Figure 4B**, namely, the correlation strength increases in the first 20 s and then decreases gradually to a stable value. This increment is due to the short-term facilitation of the electrical synapse in the first 20 s, as shown in **Figure 4E**. As a comparison, we also simulate the change of correlation strength between neurons when they are connected with constant electrical synapse strength (**Figure 4F**). In this case, the correlation strength linearly decreased with time during the adaptation process (the first 20 s; inset of **Figure 4F**), and is unable to explain the experimental observation.

COMPUTATIONAL ROLE OF DYNAMICAL ELECTRICAL SYNAPSES
In the above we have demonstrated that short-term facilitation of electrical synapses can well justify the neuronal response properties during the adaptation. But, what is the functional meaning of this short-term plasticity?

To answer this question, we analyzed another set of experimental data in which the RGCs of a bullfrog retina were exposed to a sustained dark stimulation after having responded to flicking pseudo-random checker-boards for 15 s (see Materials and Methods). The whole adaptation process to the dark stimulus lasted for about 5 s. **Figures 5A,B** present the experimental results, which show that the firing rates of RGCs attenuated over time and that the neuronal correlation via electrical synapses increased over time. Similar to the analysis in section A Phenomenological Model for Short-Term Facilitation of Electrical Synapse by considering short-term facilitation of electrical synapses, our model, i.e., Equations (1 and 11), successfully reproduces the experimental data (**Figures 5C,D**).

To ascertain the computational contribution of the enhanced correlation and consequently the functional role of short-term
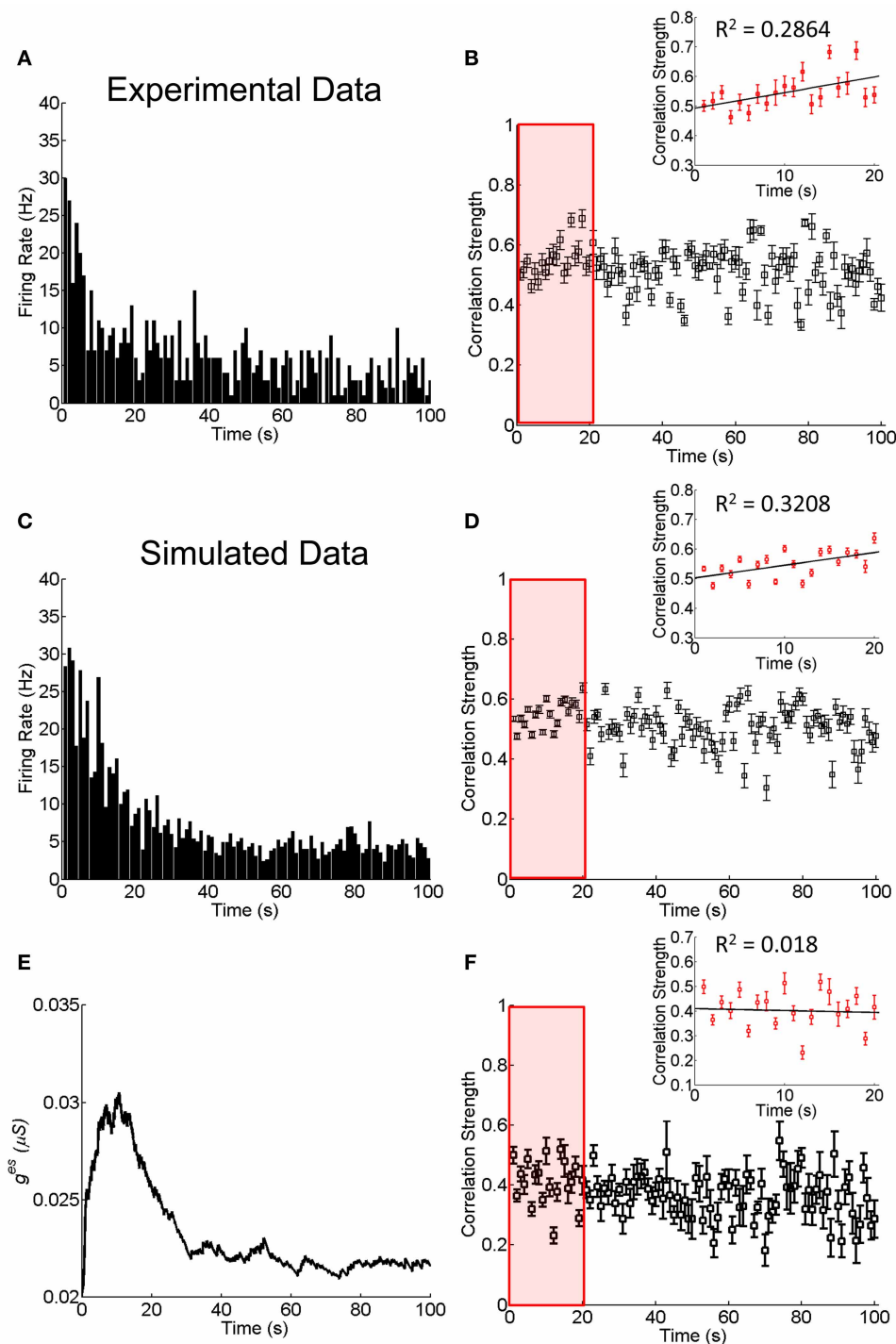
FIGURE 4 | Experimental and simulated results about the change of the correlation strength between two electrically coupled neurons during the adaptation to 100-s flicking pseudo-random checker-boards. (A) An example of retinal ganglion cell's response to 100-s flicking pseudo-random checker-boards. The time window for calculating firing rate is 1 s. **(B)** The change of the correlation strength measured in the experiment. Inset highlights the change of the correlation strength in the first 20-s (red box), fitted by a straight line. $n = 11$ neuron pairs are used and error bars indicate Mean ± s.e.m. **(C)** The simulated neural responses to 100-s flicking pseudo-random checker-boards. **(D)** The change of the correlation strength based on the short-term facilitation in Equation 6. Inset shows the change of correlation strength in the first 20-s (red box). **(E)** The change of electrical coupling strength due to the short-term facilitation during the adaptation. **(F)** The change of the correlation strength during the adaptation with a constant electrical coupling strength ($g^{es} = 0.025\,\mu S$). Inset shows the change of correlation strength in the first 20-s (red box). The simulation is repeated 10 times and error bars indicate Mean ± s.e.m. The parameters $\tau^f = 10\,s$, $\tau^l = 50\,ms$, $u^f = 50\,\mu S$, and other parameters are the same as in **Figure 1**.
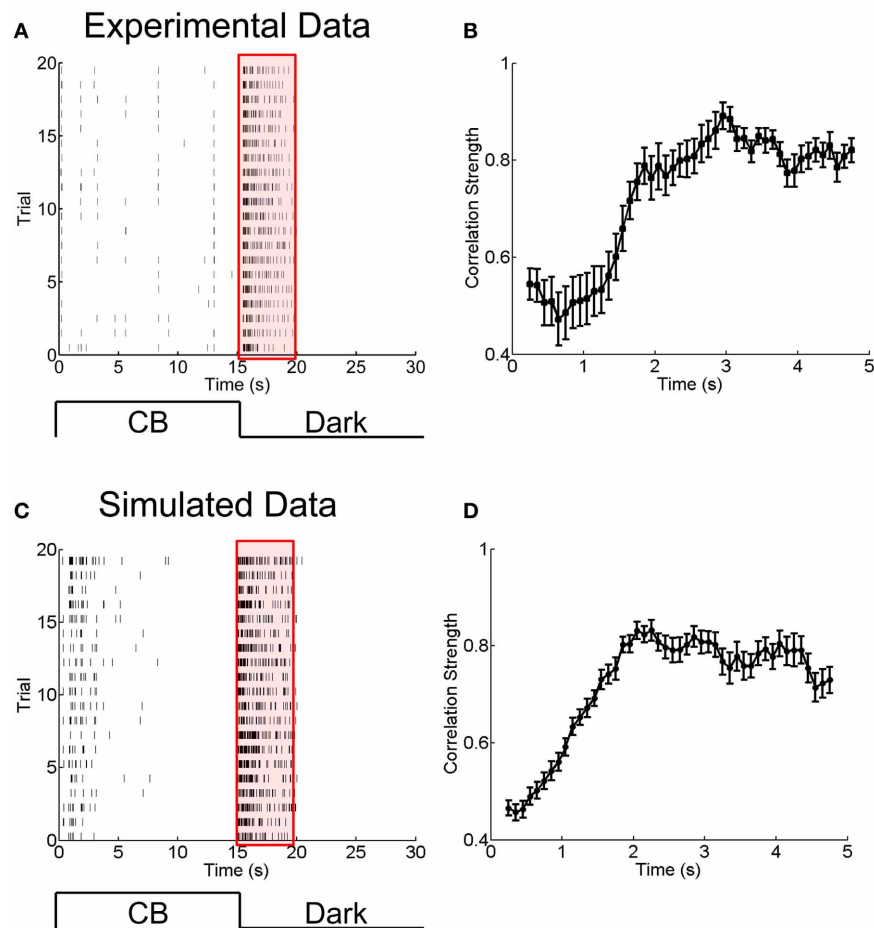
FIGURE 5 | Experimental and simulated results for the change of the correlation strength between two electrically coupled neurons during the adaptation to a sustained dark stimulation. (A) An example of raster plots of a neuron's response to 15-s flicking pseudo-random checker-boards followed by 15-s darkness stimulation, repeated 20 times. Bin size is 2 ms. **(B)** The change of the correlation strength measured in the experiment during the adaptation to darkness (red box in A). $n = 20$ neuron pairs are used. Error bars Mean ± s.e.m. **(C)** and **(D)** Simulated results for the neuronal adaptive behavior and the change of the correlation strength when the short-term facilitation of electrical synapses is considered. The parameters are the same as in **Figure 4**.

facilitation of electrical synapses, we analyze how the stimulation information is encoded separately in the firing rates and the neural correlation during the adaptation. The information analysis approach is introduced in Materials and Methods.

Figure 6 shows the results calculated by Equations (6–10) based on the experimental data shown in **Figure 5**. We see that during the adaptation, the stimulus information contained in the firing rates decays dramatically with time, whereas, the stimulus information contained in the correlation of electrical coupled neurons tend to increase with time (**Figure 6A**). Their relative contributions exhibit a very interesting behavior: at the beginning of neuronal response to dark stimulation, more than 90% of the stimulus information is encoded by the firing rates; whereas after about 2 s, more than 50% of the stimulus information is encoded in the correlation (**Figure 6B**). This result implies that during the adaptation, there exists a transition in the encoding strategies of the neural system, namely, from the firing rate code to the correlation one, and a computational role of short-term

facilitation of electrical synapse is to implement this transition operation.

DISCUSSIONS

In the present study we have investigated the potential computational roles of dynamical electrical synapses in neural information processing. We find that electrical synapses are more efficient than chemical synapses in modulating the concerted activity of neurons. That is because an electrical synapse tends to equate the sub-threshold membrane potentials of connected neurons and hence is more efficient in controlling synchronous firing of neurons. On the other hand, a chemical synapse only conveys signal when a neuron fires, and its effect in coordinating synchronous firing can be easily disturbed by fluctuations in sub-threshold potentials of the neurons due to input noises.

Based on the experimental data, we propose a phenomenological model for short-term facilitation of electrical synapses, which successfully reproduces the seemingly paradox phenomenon that

FIGURE 6 | Information entropy and information ratios carried by firing rate and correlated activity during the luminance adaptation. (A) Information carried by firing rate and correlation of electrically coupled neurons during the adaptation to the dark stimulation. Bin size is 5 ms. **(B)** Information ratios carried by firing rate and correlation during the adaptation. $n = 20$ neuron pairs are used. Error bars Mean ± s.e.m.

the increment of neural correlation is associated with the attenuation of firing rates. In a recent work, Cortes et al. (2012). found that chemical synapses with neuronal spike-frequency adaptation can also generate this paradoxical behavior. Nevertheless, for the particular neural data considered in this study, namely, the responses of RGCs, we believe that short-term facilitation of electrical synapses is a more plausible mechanism for two reasons. First, RGCs are abundantly connected by electrical synapses, and their interaction through chemical synapses is indirect (mediated by bipolar and amacrine cells); and secondly, the *CCF* between RGCs measured in the experiment has a very narrow peak and it exhibits dual peaks when the bin size is sufficiently small, which are the typical syndromes of electrical synapses.

We investigated how the stimulus information is encoded separately in the firing rates and the correlations of RGCs during the luminance adaptation. We find that there exists a transition from the firing rate code to the correlation one at the late stage of the adaptation. The latter encodes the stimulus information by using the concerted, but less active, firings of neurons, and hence is economically more efficient. Our finding suggests that dynamical electrical synapses can play profound roles in neural information processing.

AUTHOR CONTRIBUTIONS

Lei Xiao, Si Wu, and Pei-ji Liang designed experiments. Lei Xiao and Pei-ji Liang performed experiments; Lei Xiao, Dan-ke Zhang, Yuanqing Li, and Si Wu implemented simulations, and performed model analysis; and Lei Xiao, Si Wu, and Pei-ji Liang wrote the paper.

ACKNOWLEDGMENTS

REFERENCES

Ackert, J. M., Wu, S. H., Lee, J. C., Abrams, J., Hu, E. H., Perlman, I., et al. (2006). Light-induced changes in spike synchronization between coupled ON direction selective ganglion cells in the mammalian retina. *J. Neurosci.* 26, 4206–4215.

Amari, S. (2001). Information geometry on hierarchy of probability distributions. *IEEE Trans. Inf. Theory* 47, 1701–1711.

Baccus, S. A., and Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron* 36, 909–919.

Bi, G., and Poo, M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464–10472.

Bloomfield, S. A., and Volgyi, B. (2009). The diverse functional roles and regulation of neuronal gap junctions in the retina. *Nat. Rev. Neurosci.* 10, 495–506.

Brivanlou, I. H., Warland, D. K., and Meister, M. (1998). Mechanisms of concerted firing among retinal ganglion cells. *Neuron* 20, 527–539.

Connors, B. W., and Long, M. A. (2004). Electrical synapses in the mammalian brain. *Annu. Rev. Neurosci.* 27, 393–418.

Cortes, J. M., Marinazzo, D., Series, P., Oram, M. W., Sejnowski, T. J., and Van Rossum, M. C. (2012). The effect of neural adaptation on population coding accuracy. *J. Comput. Neurosci.* 32, 387–402.

Dan, Y., and Poo, M. M. (2006). Spike timing-dependent plasticity: from synapse to perception. *Physiol. Rev.* 86, 1033–1048.

Haas, J. S., Zavala, B., and Landisman, C. E. (2011). Activity-dependent long-term depression of electrical synapses. *Science* 334, 389–393.

Jing, W., Liu, W. Z., Gong, X. W., Gong, H. Q., and Liang, P. J. (2010). Influence of GABAergic inhibition on concerted activity between the coupled ON direction selective ganglion cells. *Neuroreport* 21, 797–801.

Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (2000). *Principles of Neural Science.* New York, NY: McGraw-Hill.

Landisman, C. E., and Connors, B. W. (2005). Long-term modulation of electrical synapses in the mammalian thalamus. *Science* 310, 1809–1813.

Li, H., Liu, W. Z., and Liang, P. J. (2012). Adaptation-dependent synchronous activity contributes to receptive field size change of bullfrog retinal ganglion cell. *PLoS ONE* 7:e34336. doi: 10.1371/journal.pone.0034336

Merhav, N. (1994). On information rates for mismatched decoders. *IEEE Trans. Inf. Theory* 40, 1953–1967.

Oizumi, M., Ishii, T., Ishibashi, K., Hosoya, T., and Okada, M. (2010). Mismatched decoding in the brain. *J. Neurosci.* 30, 4815–4826.

Pereda, A. E., and Faber, D. S. (1996). Activity-dependent short-term enhancement of intercellular coupling. *J. Neurosci.* 16, 983–992.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423, 623–656.

Tsodyks, M., and Markram, H. (1996). Plasticity of neocortical synapses

enables transitions between rate and temporal coding. *Artif. Neural Netw.* 96, 445–450.

Wu, S., Nakahara, H., and Amari, S. (2001). Population coding with correlation and an unfaithful model. *Neural Comput.* 13, 775–797.

Xiao, L., Wu, S., and Liang, P. J. (2012). "Population coding strategy during retinal adaptation," in *8th International Meeting on Substrate-Integrated Microelectrode Arrays* (Reutlingen: NMI Natural and Medical Sciences Institute at the University of Tuebingen), 123–124.

Yang, X. D., Korn, H., and Faber, D. S. (1990). Long-term potentiation of electrotonic coupling at mixed synapses. *Nature* 348, 542–545.

Reduction in LFP cross-frequency coupling between theta and gamma rhythms associated with impaired STP and LTP in a rat model of brain ischemia

Xiaxia Xu , Chenguang Zheng and Tao Zhang *

Computational Neuroscience Lab, The College of Life Sciences, Nankai University, Tianjin, China

The theta-gamma cross-frequency coupling (CFC) in hippocampus was reported to reflect memory process. In this study, we measured the CFC of hippocampal local field potentials (LFPs) in a two-vessel occlusion (2VO) rat model, combined with both amplitude and phase properties and associated with short and long-term plasticity indicating the memory function. Male Wistar rats were used and a 2VO model was established. STP and LTP were recorded in hippocampal CA3-CA1 pathway after LFPs were collected in both CA3 and CA1. Based on the data of relative power spectra and phase synchronization, it suggested that both the amplitude and phase coupling of either theta or gamma rhythm were involved in modulating the neural network in 2VO rats. In order to determine whether the CFC was also implicated in neural impairment in 2VO rats, the coupling of CA3 theta–CA1 gamma was measured by both phase-phase coupling (*n:m* phase synchronization) and phase-amplitude coupling. The attenuated CFC strength in 2VO rats implied the impaired neural communication in the coordination of theta-gamma entraining process. Moreover, compared with modulation index (MI) a novel algorithm named cross frequency conditional mutual information (CF-CMI), was developed to focus on the coupling between theta phase and the phase of gamma amplitude. The results suggest that the reduced CFC strength probably attributed to the disruption of the phase of CA1 gamma envelop. In conclusion, it implied that the phase coupling and CFC of hippocampal theta and gamma played an important role in supporting functions of neural network. Furthermore, synaptic plasticity on CA3-CA1 pathway was reduced in line with the decreased CFC strength from CA3 to CA1. It partly supported our hypothesis that directional CFC indicator might probably be used as a measure of synaptic plasticity.

Keywords: two-vessel occlusion, cross frequency conditional mutual information (CF-CMI), synaptic plasticity, hippocampus, neural information flow (NIF)

INTRODUCTION

Hippocampus is known to be one of the most important brain regions closely related to the learning and memory processes with synaptic plasticity as the accepted cellular basis (Howland and Wang, 2008; Shang et al., 2010; Sydow et al., 2011; Foster, 2012). One of the functional indices of synaptic plasticity is long term potentiation (LTP) (Quan et al., 2010), which is a long lasting enhancement of synaptic strength induced by high-frequency stimulating presynaptic neurons (Bliss and Lomo, 1973). In addition, the early transient potentiation phase of LTP lasting 10 min or less is termed short-term potentiation (STP) and is considered to be one candidate mechanism for short term memory (STM) (Erickson et al., 2010).

Synchronized neural oscillations were supposed to facilitate simultaneous firing of neural population and may be related to cognitive processes (Basar et al., 2001; Ward, 2003; Zhang, 2011). Conventionally, neural oscillation is classified into five frequency bands e.g., delta 1–4 Hz, theta 4–8 Hz, alpha 8–13 Hz, beta 13–30 Hz, and gamma 30–150 Hz (Buzsaki and Draguhn, 2004), which are possibly associated with different brain status. Among these rhythms, both theta and gamma rhythms in hippocampus, modulated during perception and memory tasks, are supposed to be most relevant to cognition (Kahana et al., 2001; Behrendt, 2010). We previously utilized an approach of general partial directed coherence (gPDC), which was one of directional algorithms, to determine the directionality of neural information flow (NIF) between CA3 and CA1 (Xu et al., 2012). It was found that coupling directional index was significantly reduced at either theta or gamma frequency bands between hippocampal CA3 and CA1 regions in brain ischemic rats, which might be associated with the alteration of LTP (Xu et al., 2012). In addition, a previous study showed that the coupling direction indices from thalamus to medial prefrontal cortex were considerably

Abbreviations: CFC, cross frequency coupling; CF-CMI, cross frequency conditional mutual information; fEPSP, field excitatory postsynaptic potential; CMI, conditional mutual information; LFP, local field potential; LTP, long term potentiation; MWM, Morris water maze; MI, modulation index; NIF, neural information flow; PAC, phase-amplitude coupling; PLV, phase locking value; STM, short term memory; STP, short term potentiation; 2VO, two vessel occlusion.

decreased at the theta rhythm in the rat model of depression, and increased after memantine treatment, which might be also associated with the LTP alterations and cognitive impairment (Zhang et al., 2011). However, so far the above NIF measurements of directional index have only been performed in a same frequency band rather than cross frequency bands. Accordingly, a question has been raised as to whether there is a causality relationship between rhythms, such as theta and gamma rhythms, between two brain regions.

Recently, several studies reported that there were two forms of cross frequency coupling (CFC) between theta and gamma rhythms, namely *n:m* phase-phase coupling (Belluscio et al., 2012) and phase-amplitude coupling(Canolty et al., 2006). It suggested that the alterations of CFC were possibly involved in the changes of cognitive function (Chrobak et al., 2000; Lisman, 2005; Sauseng et al., 2009). Modulation index approach (Canolty et al., 2006) can be employed to measure phase-amplitude coupling (PAC) between hippocampal CA3 and CA1. However, the measurement of modulation index is affected by both the amplitude and phase signals. Therefore, a novel measurement is needed, which focuses on the coupling between theta phase and the phase of gamma amplitude. In the present study, a novel approach, named cross frequency mutual information (CF-CMI), was developed based on conditional mutual information (Palus et al., 2001; Palus and Stefanovska, 2003). In contrast to an approach of MI, which transiently combines the amplitude envelope of high-frequency with the phase of low frequency rhythm into analytic signals, the approach of CF-CMI focuses on the phase–phase coupling between two different rhythms. This novel coupling measurement may provide an underlying indication of the coupling strength possibly corresponding to the information coding in hippocampus.

In this study, Male Wistar rats were used and the two vessel occlusion (2VO) (Xu et al., 2012) model was successfully established. Local field potentials were collected before STP and LTP performed on hippocampal CA3 and CA1 pathway. The phase locking value (PLV) measurement was used to measure the phase synchronization between CA3 and CA1 regions over a particular rhythm, such as theta or gamma rhythm. In order to determine whether the CFC was also implicated in neural impairment in 2VO rats, we examined the theta-gamma coupling between CA3 and CA1 in hippocampus, which were done by both phase-phase coupling (*n:m* phase synchronization) and PAC. Furthermore, the CF-CMI was used to measure the coupling strength between theta phase and the phase of gamma amplitude. An issue was addressed as to whether such a directional index of NIF between cross-frequency bands is able to reveal the variations of hippocampal synaptic plasticity in brain ischemia, combining with the alterations of STP and LTP on CA3-to-CA1 neural pathway.

MATERIALS AND METHODS
EXPERIMENTAL ANIMALS
Experiments were performed on male Wistar rats (280–300 g, around 8-week old), which were provided from the Laboratory Animal Center; Academy of Military Medical Science of People's Liberation Army, and reared in the animal house of Medical School, Nankai University. Animals were housed in a 12 h light/dark cycle with freely feed and water and randomly divided into two groups (*n* = 12), namely Con group (*n* = 6) and 2VO group (*n* = 6). A rat model of 2VO was established, which was as same that in our previous reports (Li et al., 2011; Xu et al., 2012). Rats were reared for 3 weeks since operation. All procedures were carried out in accordance with the Ethical Commission at Nankai University, China.

ELECTROPHYSIOLOGICAL EXPERIMENT
Rats was placed in a stereotaxic frame (Narishige, Japan) under 30% urethane anesthesia (4 ml/kg, i.p., Sigma-Aldrich, St. Louis, MO, USA). The skull was opened and a small hole (2 mm in diameter) in its left side was drilled. Two Stainless steel electrodes were slowly implanted into CA3 and CA1 sites (CA3: 4.2 mm posterior to the bregma, 3.5 mm lateral to midline, 2.5 mm ventral below the dura; CA1: 3.5 mm posterior to the bregma, 2.5 mm lateral to midline, 2.0 mm ventral below the dura), respectively. Ground and reference electrodes were placed symmetrically over the two hemispheres of the cerebellum. The signals of local field potential were collected concurrently from the regions of CA3 and CA1 at a sampling rate of 1000 Hz.

After LFPs were collected, STP and LTP recordings were performed in the same brain regions. First, low-frequency stimulations (0.05 Hz) for 20 min were delivered to Schaffer collateral evoking a response of 50% of its maximum. And then tetanic stimulation (10 pulses at 100 Hz for 2 s repeated 10 times) was delivered and field excitatory postsynaptic potentials (fEPSPs) were recorded at 20 kHz sampling rate every 20 s for 60 min. fEP-SPs slope was used to measure synaptic efficacy (Li et al., 2011). As the average responses, STP and LTP were measured at the first 10 min and between 50 min and 60 min after induction, respectively. The initial data was analyzed by Clampfit 9.0 (Molecular Devices, Sunnyvale, CA, USA).

PHASE LOCKING VALUE (PLV)
PLV is a widely used method to measure the strength of phase synchronization within rhythms between brain regions (Rosenblum et al., 1996). ϕ_a and ϕ_b signed the phase of the two signals and PLV is defined as

$$PLV = \left| \frac{1}{N} \sum_{j=1}^{N} \exp(i[\phi_a(j\Delta t) - \phi_b(j\Delta t)]) \right|$$

N stands for the length of the signal and $\frac{1}{\Delta t}$ is the sampling frequency. The value of PLV is within [0, 1] with 1 indicates fully synch and 0 no syncing at all.

n:m PHASE SYNCHRONIZATION
Cross frequency phase-phase coupling between theta and gamma rhythms was determined by *n:m* phase synchronization, where the ration of *n:m* stood for stable *n* cycles of the gamma oscillator for every *m* theta oscillator.

The radial distance (r) values, determined as: $r_{n:m} = \left| \frac{1}{N} \sum_{t=1}^{N} e^{i[m \times \phi_{theta}(t) - n \times \phi_{gamma}(t)]} \right|$ were used to determine the strength of cross frequency phase-phase coupling.

The distribution of $r_{n:m}$ for different rations, e.g., 1:1, 1:2, ..., 1:10, etc. was calculated. A Larger value of r indicated a more unimodal distribution of $\Delta\varphi_{n:m}(t) = m \times \phi_{\text{theta}}(t) - n \times \phi_{\text{gamma}}(t)$, i.e., stronger phase coupling (Rayleigh test for uniformity) (Tass et al., 1998; Belluscio et al., 2012).

PHASE AMPLITUDE COUPLING (PAC)

Modulation index (MI) was used to evaluate the cross frequency PAC between CA3 and CA1 regions. The main idea of MI measure was to create a composite signal with amplitude envelope of the high frequency ($A_{\text{famp}}(t)$) as its amplitude and instantaneous phase of the low frequency ($\phi_{\text{fph}}(t)$) as its phase.

$$Z_{\text{fph, famp}}(t) = A_{\text{famp}}(t) \times \exp(i \times \phi_{\text{fph}}(t))$$

This composite signal created a joint probability density function on the complex plane. The initial value of MI is calculated as the absolute value of the average of the composite signal:

$$MI_{\text{raw}} = abs\,(\text{mean}(Z_{\text{fph,fam}}(t)))$$

For further processing, surrogate data need to be generated by bringing a random time lag τ between $\phi_{\text{fph}}(t)$ and $A_{\text{famp}}(t)$: $Z_{\text{surr}}(t, \tau) = A_{\text{famp}}(t + \tau) \times \exp(i \times \phi_{\text{fph}}(t))$.

Finally, MI is defined as $MI = (MI_{\text{raw}} - \mu)/\sigma$, where μ is the mean of the surrogate lengths and σ is a standard deviation.

In this case, Morlet wavelets of the depth 7 were applied to generate analytic representations with a frequency range of 1–20 Hz in CA3 and 30–80 Hz in CA1. And then Hilbert transform was used to obtain CA3 $\phi_{\text{fph}}(t)$ and CA1 $A_{\text{famp}}(t)$, respectively. Finally, a window length of 40 s with 50% overlap and 100 trials of surrogate data were employed in the study.

PHASE-AMPLITUDE COUPLING BASED ON CONDITIONAL MUTUAL INFORMATION

In order to measure the strength of directional CFC between CA3 and CA1 regions, an improved algorithm named cross frequency conditional mutual information (CF-CMI) was made. Specifically, we firstly extracted the phase of broadband-filtered theta rhythm (from 4 Hz to 8 Hz) in CA3 region (ϕ_{theta}) and the amplitude of the narrowband-filtered gamma rhythm (from 30 Hz to 80 Hz, step = 1 Hz) in CA1 region (amp$_{\text{gamma}}$) by Hilbert transformation. Since amp$_{\text{gamma}}$ did not vary very fast, we band-filtered it from 1 Hz to 10 Hz. And then the phase of amp$_{\text{gamma}}$ was extracted by a second Hilbert transformation signed as ϕ_{amp_γ}. Finally, CMI (Palus et al., 2001; Palus and Stefanovska, 2003) was applied to measure the directional coupling between ϕ_{theta} and ϕ_{amp_γ}.

Briefly, supposing two processes $\{X_{CA3}\}$ and $\{X_{CA1}\}$ (from the amplitude envelope of signals in CA1), their instantaneous phases $\{\varphi_{\text{theta}}\}$ and $\{\phi_{\text{amp}_r}\}$ can be estimated by application of the discrete Hilbert transform. Accordingly, the "net" information about the τ − future of the process $\{\phi_{\text{amp}_r}\}$ contained in process $\{\phi_{\text{theta}}\}$ using $C = I(\phi_{\text{theta}}; \Delta_\tau\phi_{\text{amp}_\gamma} | \phi_{\text{amp}_\gamma})$.

To establish possible causality relations, we consider phase increments,

$$\Delta_\tau\phi_{\text{amp}_\gamma} = |\phi_{\text{amp}_\gamma}(t + \tau) - \phi_{\text{amp}_\gamma}(t)|$$

Then the conditional mutual information is defined as, $I(\phi_{\text{theta}}; \Delta_\tau\phi_{\text{amp}_\gamma} | \phi_{\text{amp}_\gamma}) = H(\phi_{\text{theta}} | \phi_{\text{amp}_\gamma}) + H(\Delta_\tau\phi_{\text{amp}_\gamma} | \phi_{\text{amp}_\gamma}) - H(\phi_{\text{theta}}, \Delta_\tau\phi_{\text{amp}_\gamma} | \phi_{\text{amp}_\gamma})$.

DATA AND STATISTICAL ANALYSIS

All data were presented as mean ± SEM. Of the STP and LTP test, field excitatory postsynaptic potentials (fEPSPs) slopes were expressed as the percentage change of the baseline. Statistical comparisons were made using the Wilcoxon rank sum test. The analyses were performed using SPSS 17.0 software with the significant level setting at $P < 0.05$.

RESULTS

Traces show representative sections of original neurograms obtained from recordings of LFPs made one normal Wistar rat at hippocampal CA1 region (black line in upper panel of **Figure 1A**) and CA3 area (black line in upper panel of **Figure 1C**) as well as a 2VO rat at CA1 (gray line in upper panel of **Figure 1A**) and CA3 (gray line in upper panel of **Figure 1C**). The signals were obtained at 1000 Hz sampling frequency and a 5 s sampling period.

POWER SPECTRUM OF LFP

Digitized LFPs signals were subjected off-line to a fast Fourier transformation to produce a power spectrum. Based on Wilcoxon rank sum test, it shows that there is no significant difference of total power between Con group and 2VO group in either theta frequency band (4–8 Hz) or slow gamma frequency band (30–50 Hz) in CA1 region (**Figure 1B**). In addition, there are significant decreases of total power in both theta and slow gamma frequency bands in 2VO group compared to that in Con group in hippocampal CA3 region (theta, $F = -2.882$, $p = 0.004$; gamma, $F = -2.882$, $p = 0.004$, **Figure 1D**).

PHASE SYNCHRONIZATION

Figure 2A showed the phase synchronization analysis at theta and slow gamma frequency bands for control and 2VO groups. The original signals were filtered into 1–50 Hz range (bandwidth = 1 Hz, step = 1 Hz). Based on the Hilbert transform, the phases of the filtered signals were generated and then used to compute the PLV. It was found that PLVs at both theta and gamma frequency bands were much lower in 2VO group compared to that in Con group (theta: $F = -2.882$, $p = 0.004$; gamma: $F = -2.562$, $p = 0.010$, **Figure 2A**).
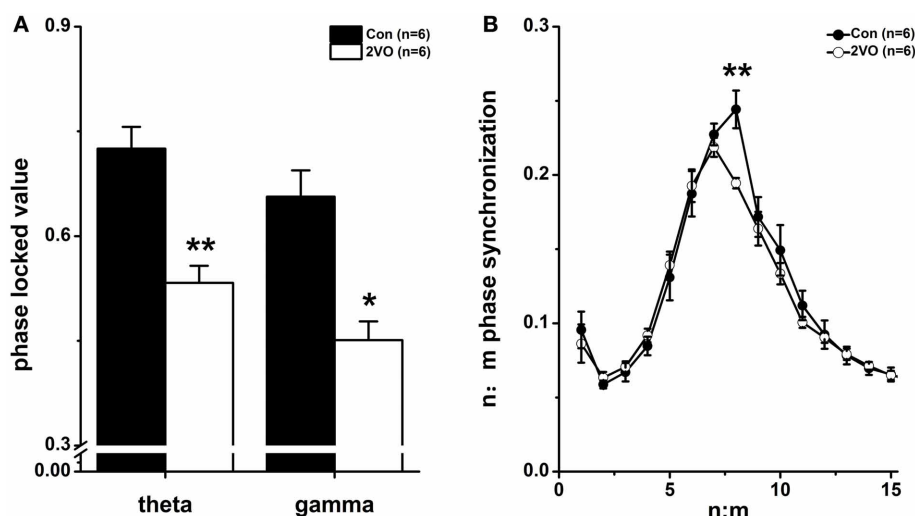
CROSS FREQUENCY PHASE–PHASE COUPLING

With the purpose of investigating the cross frequency theta-gamma phase coupling quantitatively, the radial distance values (r) of the circular distribution from the phase differences between $m \times$ theta (CA3) and $n \times$ low gamma (CA1) phases for 15 ratios were calculated (**Figure 2B**). Rayleigh test showed that there were a distinct peak at $n:m = 1:8$ ratio ($p < 0.05$) in Con group and another peak at $n:m = 1.7$ ($p < 0.05$) in 2VO group. Furthermore, Wilcoxon rank sum test showed that there was a significant difference of 1:8 phase synchronization values between these two groups ($F = -2.882$, $p = 0.004$). It implied that cross
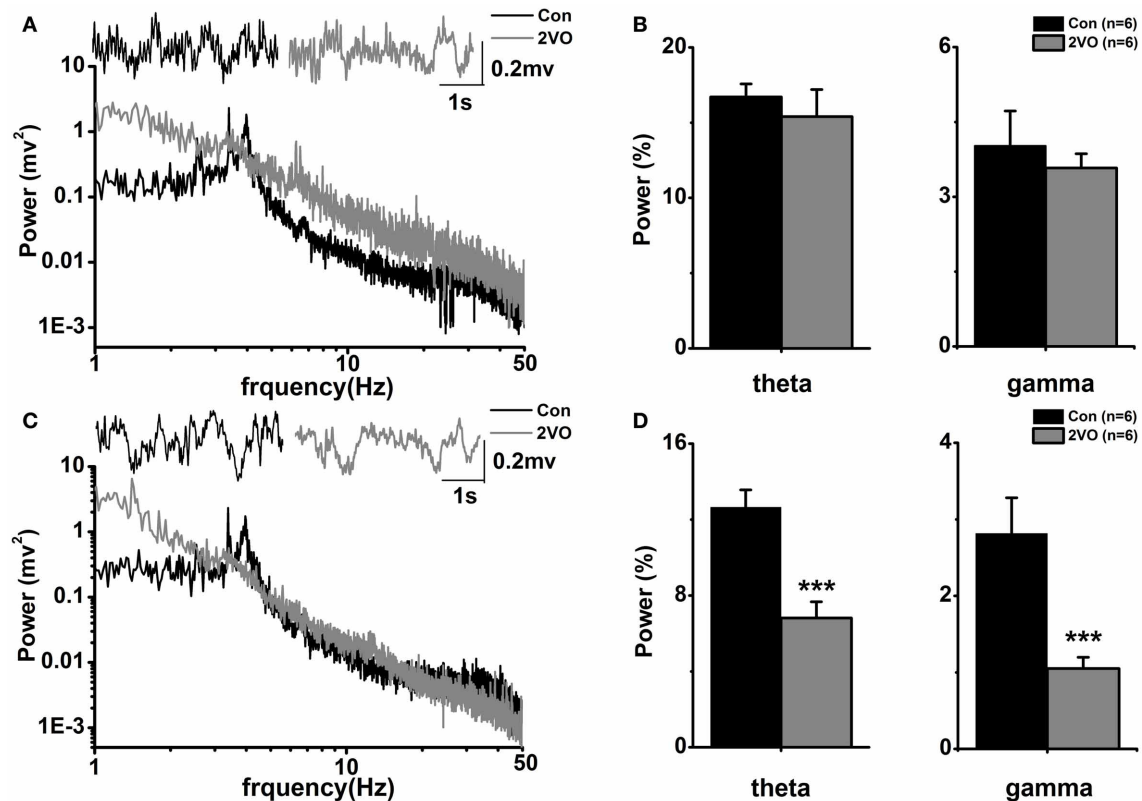
FIGURE 1 | Power spectral analysis in the two groups. (A) Representative local field potential traces and corresponding power spectra in hippocampal CA1 regions in one normal rat (black line) and one 2VO rat (gray line). (B) Statistical analysis of relative theta and gamma power spectra in CA1 region in the two groups. (C) Same display as (A) in CA3 region. (D) Same display as (B) in CA3 region. ***$p < 0.001$ comparison between Con and 2VO groups.

FIGURE 2 | Phase synchronization index. (A) Phase locking value (PLV) of LFPs between CA3 and CA1 at theta and gamma frequency bands in Con and 2VO groups ($n = 6$). (B) Phase–phase ($n{:}m$) coupling between theta and gamma oscillations. Mean radial distance values (r values) from the distribution of the difference between $m \times$ theta and $n \times$ gamma phases calculated for different ratios in these two groups. *$p < 0.05$ and **$p < 0.01$ comparison between Con group and 2VO group.

frequency phase–phase coupling might be weakened in brain ischemia rats.

CROSS FREQUENCY PHASE-AMPLITUDE COUPLING

Figure 3 showed the mean modulation indices in both Con and 2VO groups, which reflected cross frequency PAC between CA3 phase sequences (1–20 Hz, step = 1 Hz) and CA1 amplitude sequences (30–80 Hz, step = 1 Hz). Larger values of MI indicate stronger cross frequency coupling. In normal animals, the maximal coupling was found at both 40 Hz of CA1 amplitude and 6 Hz of CA3 phase (**Figure 3A**), while the strong PAC between CA3 and CA1 existed at slow gamma band (30–50Hz). However, this cross frequency PAC was almost disappeared in brain ischemic rats (**Figure 3B**).

REDUCED PHASE-AMPLITUDE DIRECTIONAL COUPLING ASSOCIATED WITH IMPAIRED STP AND LTP

Stimulating Schaffer collateral evokes basal field excitatory post-synaptic potentials (fEPSPs) in the hippocampal CA1 region. **Figure 4A** shows the time courses of fEPSPs slopes normalized to the 20 min baseline period. It can be seen that the fEPSPs slopes are increased immediately after the high-frequency stimulation and then stabilized to a level above the baseline period. The mean fEPSP slopes of the first 10 min after HFS were examined as STP results. Based on Wilcoxon rank sum test, it was found that the mean fEPSPs slope was lower in 2VO group than that in control group ($113 \pm 3.42\%$ vs. $126 \pm 1.51\%$, $p < 0.001$, **Figure 4B**-left). Furthermore, LTP was measured as the mean fEPSP slopes in 45–60 min after HFS. It could be seen that the mean fEPSPs slope was much lower in 2VO group than that in control group ($103 \pm 2.65\%$ vs. $118 \pm 0.50\%$, $p < 0.001$, **Figure 4B**-right).

Figures 4C–E shows the data of statistical CFC analysis. It was found that the value of MI was enormously lower in 2VO

group compared to that in control one ($F = -2.882$, $p = 0.004$, **Figure 4C**). In order to measure the directional cross-frequency coupling (CFC) between theta rhythm in CA3 and gamma rhythm in CA1, LFP signals were filtered over 1–50 Hz with 1 Hz bandwidth, using FIR band filter with hamming window (filter order = 512). Two types of phase sequence were extracted by means of Hilbert transform, one from original LFP signals within theta frequency band and another from the amplitude of LFP signals within gamma frequency band. And the novel algorithm of CF-CMI was applied to determine the directionality of NIL between these two areas. It can be seen that the value of CF-CMI measurement is much lower in 2VO rats compared to that in control animals ($F = -2.882$, $p = 0.004$, **Figure 4D**). There was no statistical difference of gamma power spectra in one theta circle between these two groups (**Figure 4D**).
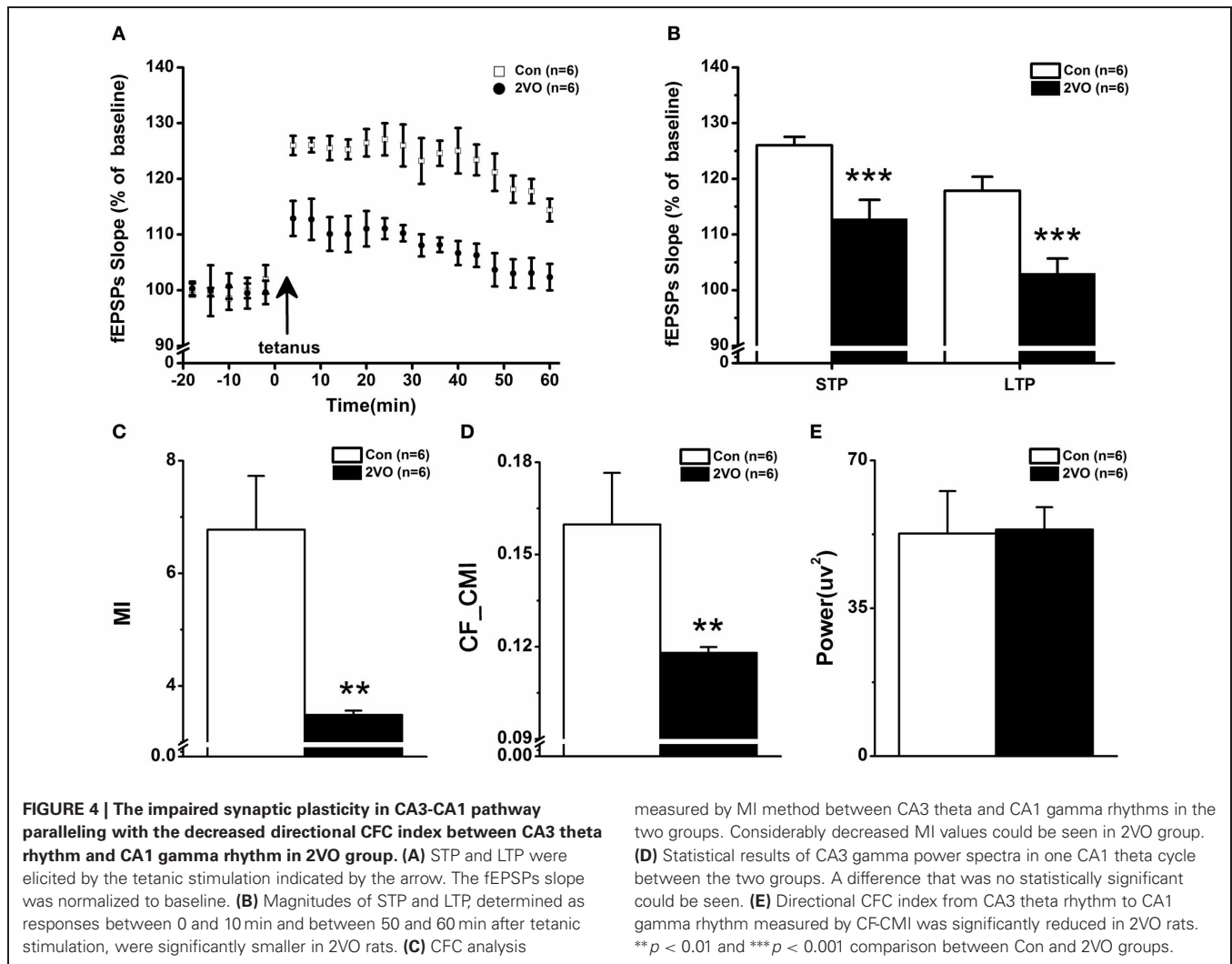
DISCUSSION

In this study, a 2VO rat model was employed with impairments cognition functions (Li et al., 2011). In addition, a novel algorithm was developed to measure the CFC directionality between CA3 and CA1 regions in hippocampus. It was found that the CFC directional index from CA3 theta rhythm to CA1 gamma rhythm was significantly reduced, which was interestingly in line with the alteration of STP and LTP in CA3-CA1 pathway in brain ischemic state. The above result shows great promise for our hypothesis that the CFC directionality could be an indicator of the synaptic plasticity in hippocampal CA3-CA1 pathway.

Phase synchronization within both theta and gamma rhythms was believed to be crucial to the cognitive behaviors (Basar-Eroglu et al., 1992; Gallinat et al., 2006), while cognitive impairment usually accompanied with reduced phase synchronization (Yener et al., 2007; Ford et al., 2008). In the present study, it was found that both theta and gamma synchronizations were
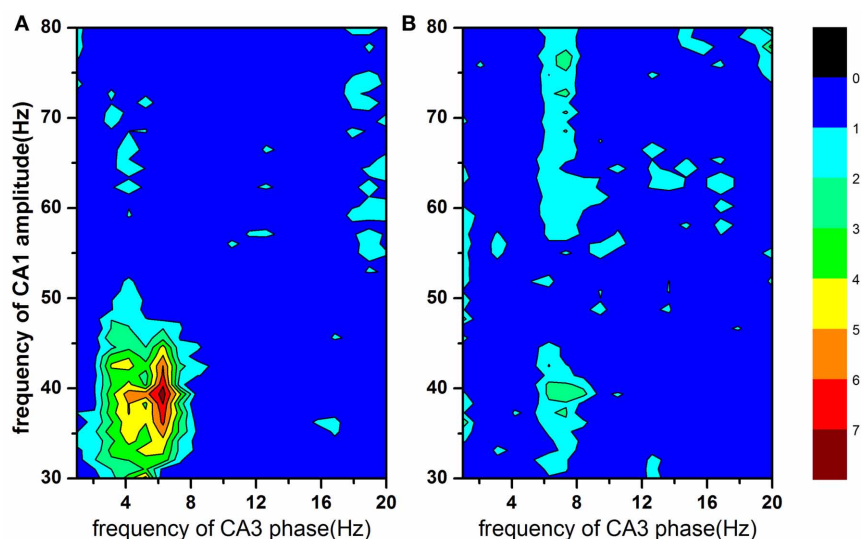
FIGURE 3 | The modulation index as a function of analytic amplitude (30–80 Hz) in CA1 and analytic phase (1–20 Hz) in CA3. Larger MI value indicates stronger cross frequency coupling. Strong phase-amplitude coupling between CA3 and CA1 existed at CA1 slow gamma band (30–50 Hz) in normal rats (panel **A**), however, almost disappeared in brain ischemia rats (panel **B**).

FIGURE 4 | The impaired synaptic plasticity in CA3-CA1 pathway paralleling with the decreased directional CFC index between CA3 theta rhythm and CA1 gamma rhythm in 2VO group. (A) STP and LTP were elicited by the tetanic stimulation indicated by the arrow. The fEPSPs slope was normalized to baseline. **(B)** Magnitudes of STP and LTP, determined as responses between 0 and 10 min and between 50 and 60 min after tetanic stimulation, were significantly smaller in 2VO rats. **(C)** CFC analysis measured by MI method between CA3 theta and CA1 gamma rhythms in the two groups. Considerably decreased MI values could be seen in 2VO group. **(D)** Statistical results of CA3 gamma power spectra in one CA1 theta cycle between the two groups. A difference that was no statistically significant could be seen. **(E)** Directional CFC index from CA3 theta rhythm to CA1 gamma rhythm measured by CF-CMI was significantly reduced in 2VO rats. **p < 0.01 and ***p < 0.001 comparison between Con and 2VO groups.

considerably decreased in 2VO group compared to that in Con group (**Figure 2A**), implying that there was a disturbance of neural synchronized coordination in brain ischemic state. The fact that the reduction of phase synchronization was associated with cognitive deficits was in line with the findings in Schizophrenia and Alzheimer subjects (Yener et al., 2007; Ford et al., 2008). Moreover, the analysis of cross frequency phase coupling (Belluscio et al., 2012) showed that the *n:m* (1:8) theta-gamma rhythm coding in Con group was changed to the *n:m* (1:7) in 2VO group (**Figure 2B**). Previous studies indicated that in computational models, identical gamma cycles with an equal number of spikes in each cycle were distributed across the entire theta cycle to support a multi-item working memory buffer (Lisman and Idiart, 1995; Jensen and Lisman, 1996). Each gamma cycle contains a discrete item (or position in space), and approximately seven gamma cycles could store 7 ± 2 sequential items. Thus, the reduction of the ratio might imply the impairment of memory capacity. However, the underlying physiological mechanism is still under further investigation. Our result of reduced ratios between theta and gamma rhythms (from 8:1 to 7:1,

Figure 2B) in 2VO rats might indicate the impaired memory capacity (Sauseng et al., 2009) induced by 2VO operation.

Another form of CFC is the amplitude of gamma rhythm nesting in theta cycles, measured by modulation index (Bragin et al., 1995; Lakatos et al., 2005; Mormann et al., 2005; Canolty et al., 2006). One speculation of this coupling was that because of relative long conduction delays, theta rhythm was well suited to synchronize the networks over long distances while gamma rhythm nested in the theta cycle to coordinate cell assemblies involved in information dissemination process (von Stein and Sarnthein, 2000). In this study, it was found that CA1 low gamma rhythm, however not the high gamma rhythm, significantly nested in CA3 theta rhythm in Con rats (**Figure 3A**). Theta-gamma coupling was supposed to be relevant to cognitive function (Palva et al., 2005, 2010; Sauseng et al., 2008). In addition, it was reported that the low gamma rhythm was coherent between CA3 and CA1 in hippocampus, entrained by theta phase (Colgin et al., 2009). Therefore, we focused on the alteration of theta-gamma coupling in CA3-CA1 pathway associated with the cognitive disorder by 2VO. interestingly,

such coupling phenomenon disappeared in brain ischemic state (**Figure 3B**), suggesting that the impaired cognitive function in 2VO rats was relevant to decreased theta-gamma coupling in CA3-CA1 pathway. Meanwhile, we did not pay attention to other neural pathways, such as cortico-hippocampal interactions and/or hippocampal DG-CA3 interactions, at the present study.

It is well known that there is a directional alteration of neural information flow, so as to measure directional CFC is more important to explore the relationship between the patterns of neural oscillation and cognitive functions. In our previous study, the algorithm of general partial directed coherence was utilized to determine the directionality of NIF between hippocampal CA3 and CA1 over either theta or gamma frequency band (Xu et al., 2012). We found that the coupling directional index was considerably decreased in the above two frequency bands in brain ischemic state, respectively. It was indicating that the strength of CA3 driving CA1 was significantly reduced. Subsequently, a hypothesis was raised that there was causality relationship in cross-frequency between hippocampal CA3 and CA1. MI algorithm has been used to measure CFC. From its formula, it can be seen that there are two factors affecting MI measurement. One is the cross phase coupling between these two frequency bands, and another is the amplitude of the high frequency band. Obviously, it will be better if we can distinguish between these two factors during the measurement of CFC.

In the present study, a novel algorithm CF-CMI is focused on measuring the coupling between theta phase and phase of gamma amplitude. Given that conditional mutual information is a directional algorithm over an identical frequency band, the developed CF-CMI should be a unidirectional coupling measurement across different frequency bands between two brain regions. Our data showed that there were no significant differences of the gamma power spectra in one theta circle between the two groups (**Figure 4E**). However, CF-CMI measurement presented that the value of directional CFC was much lower in 2VO group than that in Con group (**Figure 4D**), indicating that it was the phase information of signals rather than the amplitude of signal, which played an essential role in changing STP and LTP on CA3-CA1 pathway in brain ischemic rats (**Figure 4B**). The data further implied that the decreased information transmission along the CA3-CA1 pathway in cross-frequency of theta and slow gamma rhythms might be related to the impairment of STP and LTP in 2VO rats.

Taken together, our findings suggest that cognitive deficits caused by brain ischemia, such as learning and memory dysfunction, are implicated in the alteration of phase-phase coupling strength in theta and gamma oscillations. Moreover, the CA3-CA1 synaptic plasticity is impaired, which is in line with the decreased directional CFC from CA3 theta rhythm to CA1 gamma rhythm. It suggests that the modifications of diverse brain rhythms and their interaction, such as theta and gamma, are involved in regulating the behavioral functions. In addition, combining the impaired synaptic plasticity and reduced values of directional CFC, we would be able to understand that the directional CFC is likely to be another indicator of synaptic plasticity compared to that of NIF directionality obtained from same oscillatory rhythm. However, studying the relationship between the directional CFC and synaptic plasticity is still at an early stage of development. It remains open issues as to if there are other brain rhythms involved, which may indicate an alteration of cognitive functions.

REFERENCES

Basar, E., Basar-Eroglu, C., Karakas, S., and Schurmann, M. (2001). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *Int. J. Psychophysiol.* 39, 241–248.

Basar-Eroglu, C., Basar, E., Demiralp, T., and Schurmann, M. (1992). P300-response: possible psychophysiological correlates in delta and theta frequency channels. A review. *Int. J. Psychophysiol.* 13, 161–179.

Behrendt, R. P. (2010). Contribution of hippocampal region CA3 to consciousness and schizophrenic hallucinations. *Neurosci. Biobehav. Rev.* 34, 1121–1136.

Belluscio, M. A., Mizuseki, K., Schmidt, R., Kempter, R., and Buzsaki, G. (2012). Cross-frequency phase-phase coupling between theta and gamma oscillations in the hippocampus. *J. Neurosci.* 32, 423–435.

Bliss, T. V., and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.* 232, 331–356.

Bragin, A., Jando, G., Nadasdy, Z., Hetke, J., Wise, K., and Buzsaki, G. (1995). Gamma (40-100 Hz) oscillation in the hippocampus of the behaving rat. *J. Neurosci.* 15, 47–60.

Buzsaki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., et al. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313, 1626–1628.

Chrobak, J. J., Lorincz, A., and Buzsaki, G. (2000). Physiological patterns in the hippocampo-entorhinal cortex system. *Hippocampus* 10, 457–465.

Colgin, L. L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., et al. (2009). Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature* 462, 353–357.

Erickson, M. A., Maramara, L. A., and Lisman, J. (2010). A single brief burst induces GluR1-dependent associative short-term potentiation: a potential mechanism for short-term memory. *J. Cogn. Neurosci.* 22, 2530–2540.

Ford, J. M., Roach, B. J., Hoffman, R. S., and Mathalon, D. H. (2008). The dependence of P300 amplitude on gamma synchrony breaks down in schizophrenia. *Brain Res.* 1235, 133–142.

Foster, T. C. (2012). Dissecting the age-related decline on spatial learning and memory tasks in rodent models: N-methyl-D-aspartate receptors and voltage-dependent Ca(2)(+) channels in senescent synaptic plasticity. *Prog. Neurobiol.* 96, 283–303.

Gallinat, J., Kunz, D., Senkowski, D., Kienast, T., Seifert, F., Schubert, F., et al. (2006). Hippocampal glutamate concentration predicts cerebral theta oscillations during cognitive processing. *Psychopharmacology (Berl.)* 187, 103–111.

Howland, J. G., and Wang, Y. T. (2008). Synaptic plasticity in learning and memory: stress effects in the hippocampus. *Prog. Brain Res.* 169, 145–158.

Jensen, O., and Lisman, J. E. (1996). Novel lists of 7 +/− 2 known items can be reliably stored in an oscillatory short-term memory network: interaction with long-term memory. *Learn. Mem.* 3, 257–263.

Kahana, M. J., Seelig, D., and Madsen, J. R. (2001). Theta returns. *Curr. Opin. Neurobiol.* 11, 739–744.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., and Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J. Neurophysiol.* 94, 1904–1911.

Li, Z., Wang, Y., Xie, Y., Yang, Z., and Zhang, T. (2011). Protective effects of exogenous hydrogen sulfide on neurons of hippocampus in a rat model of brain ischemia. *Neurochem. Res.* 36, 1840–1849.

Lisman, J. (2005). The theta/gamma discrete phase code occuring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus* 15, 913–922.

Lisman, J. E., and Idiart, M. A. (1995). Storage of $7+/-2$ short-term memories in oscillatory subcycles. *Science* 267, 1512–1515.

Mormann, F., Fell, J., Axmacher, N., Weber, B., Lehnertz, K., Elger, C. E., et al. (2005). Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. *Hippocampus* 15, 890–900.

Palus, M., Komarek, V., Hrncir, Z., and Sterbova, K. (2001). Synchronization as adjustment of information rates: detection from bivariate time series. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 63, 046211.

Palus, M., and Stefanovska, A. (2003). Direction of coupling from phases of interacting oscillators: an information-theoretic approach. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 67, 055201.

Palva, J. M., Monto, S., Kulashekhar, S., and Palva, S. (2010). Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7580–7585.

Palva, J. M., Palva, S., and Kaila, K. (2005). Phase synchrony among neuronal oscillations in the human cortex. *J. Neurosci.* 25, 3962–3972.

Quan, M. N., Tian, Y. T., Xu, K. H., Zhang, T., and Yang, Z. (2010). Post weaning social isolation influences spatial cognition, prefrontal cortical synaptic plasticity and hippocampal potassium ion channels in Wistar rats. *Neuroscience* 169, 214–222.

Rosenblum, M. G., Pikovsky, A. S., and Kurths, J. (1996). Phase synchronization of chaotic oscillators. *Phys. Rev. Lett.* 76, 1804–1807.

Sauseng, P., Klimesch, W., Gruber, W. R., and Birbaumer, N. (2008). Cross-frequency phase synchronization: a brain mechanism of memory matching and attention. *NeuroImage* 40, 308–317.

Sauseng, P., Klimesch, W., Heise, K. F., Gruber, W. R., Holz, E., Karim, A. A., et al. (2009). Brain oscillatory substrates of visual short-term memory capacity. *Curr. Biol.* 19, 1846–1852.

Shang, X. L., Zhao, J. H., Cao, Y. P., and Xue, Y. X. (2010). Effects of synaptic plasticity regulated by 17beta-estradiol on learning and memory in rats with Alzheimer's disease. *Neurosci. Bull.* 26, 133–139.

Sydow, A., Van der Jeugd, A., Zheng, F., Ahmed, T., Balschun, D., Petrova, O., et al. (2011). Tau-induced defects in synaptic plasticity, learning, and memory are reversible in transgenic mice after switching off the toxic Tau mutant. *J. Neurosci.* 31, 2511–2525.

Tass, P., Rosenblum, M. G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., et al. (1998). Detection of n:m phase locking from noisy data: application to magnetoencephalography. *Phys. Rev. Lett.* 81, 3291–3294.

von Stein, A., and Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int. J. Psychophysiol.* 38, 301–313.

Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. *Trends Cogn. Sci.* 7, 553–559.

Xu, X., Li, Z., Yang, Z., and Zhang, T. (2012). Decrease of synaptic plasticity associated with alteration of information flow in a rat model of vascular dementia. *Neuroscience* 206, 136–143.

Yener, G. G., Guntekin, B., Oniz, A., and Basar, E. (2007). Increased frontal phase-locking of event-related theta oscillations in Alzheimer patients treated with cholinesterase inhibitors. *Int. J. Psychophysiol.* 64, 46–52.

Zhang, T. (2011). Neural oscillations and information flow associated with synaptic plasticity. *Sheng Li Xue Bao* 63, 412–422.

Zhang, M., Zheng, C., Quan, M., An, L., Yang, Z., and Zhang, T. (2011). Directional indicator on neural oscillations as a measure of synaptic plasticity in chronic unpredictable stress rats. *Neurosignals* 19, 189–197.

Stability analysis of associative memory network composed of stochastic neurons and dynamic synapses

Yuichi Katori[1,2]*, Yosuke Otsubo[3], Masato Okada[3,4] and Kazuyuki Aihara[2]

[1] FIRST, Aihara Innovative Mathematical Modelling Project, Japan Science and Technology Agency, Tokyo, Japan
[2] Institute of Industrial Science, The University of Tokyo, Tokyo, Japan
[3] Graduate School of Frontier Science, The University of Tokyo, Chiba, Japan
[4] RIKEN Brain Science Institute, Saitama, Japan

We investigate the dynamical properties of an associative memory network consisting of stochastic neurons and dynamic synapses that show short-term depression and facilitation. In the stochastic neuron model used in this study, the efficacy of the synaptic transmission changes according to the short-term depression or facilitation mechanism. We derive a macroscopic mean field model that captures the overall dynamical properties of the stochastic model. We analyze the stability and bifurcation structure of the mean field model, and show the dependence of the memory retrieval performance on the noise intensity and parameters that determine the properties of the dynamic synapses, i.e., time constants for depressing and facilitating processes. The associative memory network exhibits a variety of dynamical states, including the memory and pseudo-memory states, as well as oscillatory states among memory patterns. This study provides comprehensive insight into the dynamical properties of the associative memory network with dynamic synapses.

Keywords: dynamic synapse, short-term plasticity, neural network, associative memory network, mean field model, bifurcation analysis

1. INTRODUCTION

Dynamic synapses change their transmission efficacy depending on the activity of the presynaptic neuron, and the postsynaptic response can be decreased (short-term depression) or increased (short-term facilitation) (Markram and Tsodyks, 1996; Tsodyks and Markram, 1997; Markram et al., 1998; Thomson, 2000; Wang et al., 2006). Synaptic transmission is carried out by the flow and diffusion of chemical components. Activation of a presynaptic neuron and generation of an action potential causes influx of calcium ions into the presynaptic membrane. A chemical reaction with the calcium ions triggers the release of the neurotransmitters and induces the post synaptic current. If many action potentials are generated in a short period of time, the calcium concentration and the fraction of releasable neurotransmitters change, and the transmission efficacy increases or decreases transiently. Change in the transmission efficacy is modeled by variables that represent the releasable neurotransmitters and the utilization parameter that defines the fraction of the neurotransmitter release by each action potential, reflecting the calcium concentration.

Stochastic neuron models with dynamic synapses and the corresponding mean field models have been proposed in previous studies, and their dynamical properties and possible roles of the dynamic synapses have been intensively investigated (Igarashi et al., 2010; Otsubo et al., 2010; Katori et al., 2012). Synaptic depression is known to enable neuronal gain control (Abbott et al., 1997), and to contribute to the destabilization of the network activity and generation of an oscillatory state (Pantic et al., 2002; Melamed et al., 2008; Otsubo et al., 2010). Synaptic facilitation is believed to enhance the working memory function in the prefrontal cortex (Mongillo et al., 2008). Furthermore, in a network with both depression and facilitation synapses, changes in the efficacy of dynamic synapses are suggested to reorganize the effective network structure, thereby contributing to flexible information processing in the prefrontal cortex (Katori et al., 2011).

An associative memory network retrieves a memory pattern according to their network dynamics in which the memory patterns are stored in their synaptic connections. Associative memory networks have also been well investigated (Anderson and Bower, 1972; Nakano, 1972; Amari, 1977; Hopfield, 1982; Adachi and Aihara, 1997). Dynamics of memory retrieval can be characterized as the convergence of the state of the network to a fixed-point attractor that corresponds to a stored memory pattern (Hopfield, 1982). In this type of conventional model of an associative memory network, the state of the network usually remains in the attractor. In contrast to this, in an associative memory network with the depression synapses, the memory retrieved state can be destabilized and the state of the network can move to another attractor that corresponds to another memory pattern. Such transitive dynamics among several memory patterns has been also investigated (Tsuda et al., 1987; Adachi and Aihara, 1997; Kanamaru et al., 2013). Although stochastic neural networks with depression and facilitation synapses have been studied (Torres et al., 2007; Mejias and Torres, 2009), a comprehensive understanding of the dynamics of associative memory networks with dynamic synapses has not yet been achieved.

In the present study, we focus on the associative memory network with stochastic neurons and dynamic synapses.

In particular, we target stability analysis on the associative memory network with correlated memory patterns. The properties of the dynamic synapses are characterized by parameters that specify the time constants of recovery from an active state to a resting state of synapses. In the models of short-term plasticity the difference between depression and facilitation can be specified using theses parameters. We investigate how the dynamics of the associative memory network depends on these parameters.

In the following sections, first, we explain the model of a stochastic neural network with dynamic synapses. Next we derive the corresponding macroscopic mean field models that approximate the dynamical properties of the stochastic model. Furthermore, we analyze structural details of the dynamical system in the macroscopic mean field model, and we show how the network behavior and the memory-retrieval performance are influenced by noise intensity and the properties of the dynamic synapses. Finally, we discuss the results of the analyses from a viewpoint of neuroscience as well as possible future studies.

2. MATERIALS AND METHODS
2.1. ASSOCIATIVE MEMORY NETWORK WITH STOCHASTIC NEURONS AND DYNAMIC SYNAPSES
In this study, we use an associative memory network comprising N binary neurons. The state of the neuron is determined stochastically depending on inputs to the neuron. The state of the ith binary neuron at time t is denoted by the variable $s_i(t)$, which represents a resting state [$s_i(t) = 0$] or an active state [$s_i(t) = 1$] of the neuron. The state of the neuron changes according to the following probabilistic dynamics (Amit et al., 1985; Mejias and Torres, 2009):

$$\text{Prob}[s_i(t + 1) = 1] = g_\beta(h_i(t)), \tag{1}$$

$$g_\beta(h_i(t)) = \frac{1}{2}\left(1 + \tanh[\beta h_i(t)]\right), \tag{2}$$

where $g_\beta(h)$ is a neural response function with the noise intensity $1/\beta = T$. The noise intensity T determines the smoothness of the response function; for $T \rightarrow +0$ the model becomes deterministic. Note that we use $\{0, 1\}$ to represent the neural activity in $s_i(t)$, whereas we use $\{-1, 1\}$ to represent the memory patterns as we describe later. The equation

$$h_i(t) = \sum_{j \neq i}^{N} J_{ij}[2s_j(t)x_j(t)u_j(t)/U_{se} - 1] \tag{3}$$

represents the total input to the ith neuron. The quantity J_{ij} represents the absolute strength of the synaptic connection from the jth to ith neuron. U_{se} represents the fraction of released neurotransmitters in absence of depression and facilitation, and is the steady state value of the variable $u_i(t)$.

The properties of dynamic synapses activated by the jth neuron are modeled using the variables x_j and u_j, which represent the fraction of releasable neurotransmitters and the utilization parameter, respectively (Tsodyks et al., 1998). The releasable neurotransmitters x_j decreases with activation of the synapse, which is triggered by the presynaptic neural activation. If there is no

presynaptic activation, x_j recovers its steady state $x_j = 1$ with time constant τ_R. The utilization parameter u_j increases with the activation of the synapse and recovers its steady state $u_j = U_{se}$ with time constant τ_F. This dynamics can be described by the following equations (Tsodyks and Markram, 1997; Tsodyks et al., 1998):

$$x_j(t + 1) = x_j(t) + \frac{1 - x_j(t)}{\tau_R} - s_j(t)x_j(t)u_j(t), \tag{4}$$

$$u_j(t + 1) = u_j(t) + \frac{U_{se} - u_j(t)}{\tau_F} + U_{se}(1 - u_j(t))s_j(t). \tag{5}$$

The efficacy of synaptic transmission is determined by the product of $x_j(t)$ and $u_j(t)$; the efficacy decreases (short-term depression) or increases (short-term facilitation) according to the parameters τ_R, τ_F, and U_{se}.

Associative memory networks work well if the memory patterns are mutually orthogonal, but otherwise it does not necessarily work well. Moreover, in the associative memory network with depression synapses, the appearance of the oscillatory states is influenced by the similarity among the memory patterns (Otsubo et al., 2010). To evaluate the influence of the similarity among memory patterns in the network with both depression and facilitation synapses, we construct the associative memory network with correlated memory patterns by considering a parent memory pattern $\boldsymbol{\xi}$ and p child patterns $\boldsymbol{\xi}^\mu$ (Amari, 1977; Toya et al., 2000) as follows:

$$\boldsymbol{\xi} = (\xi_1, \ldots, \xi_N), \tag{6}$$

$$\boldsymbol{\xi}^\mu = (\xi_1^\mu, \ldots, \xi_N^\mu), \mu = 1, \ldots, p. \tag{7}$$

Note that here we use the $\{-1, 1\}$ to represents the memory patterns. A schematic of the relationship between these patterns for $p = 3$ is shown in **Figure 1**. Elements of the memory patterns are randomly generated according to the probability

$$\text{Prob}[\xi_i = \pm 1] = \frac{1}{2}, \tag{8}$$

$$\text{Prob}[\xi_i^\mu = \pm 1] = \frac{1 \pm b\xi_i}{2}, \tag{9}$$

where b is the correlation level among memory patterns and takes values in the interval $[0, 1]$. For $b = 0$, child patterns are mutually orthogonal for $N \rightarrow \infty$; for $b = 1$, the child patterns are the same as the parent pattern. Here we use the child patterns as the memory patterns. The direction cosine between memory patterns can described as $\cos \theta_0 = \frac{1}{N} \sum_{i=1}^{N} \xi_i \xi_i^\mu = b$ and $\cos \theta = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu \xi_i^\nu = b^2$, where $\mu \neq \nu$ (Otsubo et al., 2010).

According to the Hebbian rule, we use the following absolute strength of synaptic connection J_{ij}:

$$J_{ij} = \frac{1}{N} \sum_{\mu}^{p} \xi_i^\mu \xi_j^\mu, \tag{10}$$

where the self-recurrent connection does not exist (i.e., $J_{ii} = 0$). The absolute strength represents the synaptic response on the connected neurons when the synapses do not undergo any

FIGURE 1 | Schematic of correlated memory patterns for $p = 3$. The direction cosine between the parent pattern $\boldsymbol{\xi}$ and a memory pattern $\boldsymbol{\xi}^{\mu}$ is $\cos \theta_0 = b$. The direction cosine between two memory patterns is $\cos \theta = b^2$.
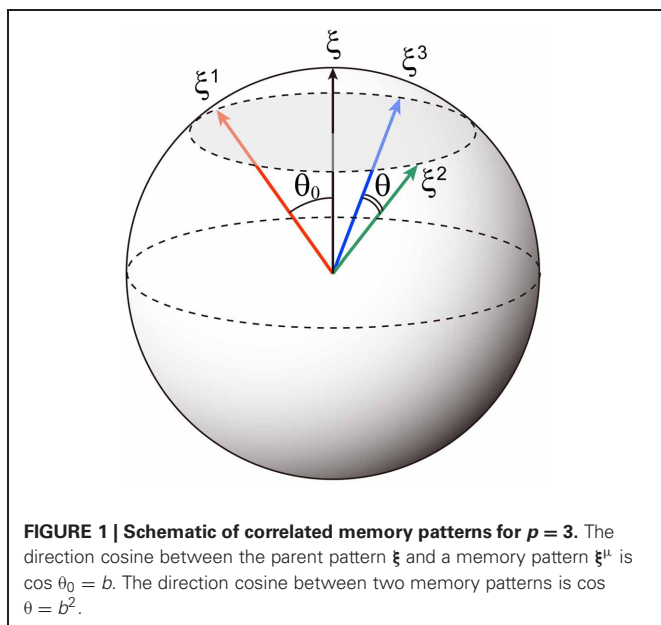
depression and facilitation. The connections with positive or negative values of the absolute strength correspond to excitatory or inhibitory synaptic connections, respectively.

2.2. MEAN FIELD THEORY

To analyze the macroscopic properties of the associative memory network with stochastic neurons, we consider dynamical mean field theory with the sublattice method (Coolen, 2001; Otsubo et al., 2010), which allows us to analyze the mean field model with the non-homogeneous network structure of the associative memory network.

First, we derive the microscopic mean field model by taking the noise average of each variable in the stochastic neural network model. We get the following equations from Equations (1) to (3):

$$\langle s_i(t+1) \rangle = g_{\beta}(\langle h_i(t) \rangle), \tag{11}$$

$$\langle h_i(t) \rangle = \sum_{j \neq i}^{N} J_{ij}[2\langle s_j(t)x_j(t)u_j(j) \rangle/U_{se} - 1]. \tag{12}$$

Because of the non-convexity of the response function g_{β} and the excitatory feedback connection, the network can stabilize the self-sustained active states (Barbieri and Brunel, 2007). Similarly to Equation (11), we obtain the following equations corresponding to Equations (4) and (5):

$$\langle x_j(t+1) \rangle = \langle x_j(t) \rangle + \frac{1 - \langle x_j(t) \rangle}{\tau_R} - \langle s_j(t)x_j(t)u_j(t) \rangle, \tag{13}$$

$$\langle u_j(t+1) \rangle = \langle u_j(t) \rangle + \frac{U_{se} - \langle u_j(t) \rangle}{\tau_F} + U_{se}\langle(1 - u_j(t))s_j(t) \rangle. \tag{14}$$

Here, we assume that the correlations among variables $s_j(t)$, $x_j(t)$, and $u_j(t)$ are negligible on the basis of the following considerations. The correlations among the variables $s_j(t)$, $x_j(t)$, and $u_j(t)$

can be separated into three pairs. The state of $s_j(t)$ is determined by the state of other neurons in the previous time step, and the state of $x_j(t)$ and $u_j(t)$ are determined by the state of each variable in the previous time step. Thus, the correlation between $s_j(t)$ and $x_j(t)$ is of the order $1/N$, and this correlation disappears as $N \to \infty$ (Igarashi et al., 2010). Similarly, the correlation between $s_j(t)$ and $u_j(t)$ also disappears as $N \to \infty$. Accordingly, we assume the following independent relations between variables:

$$\langle s_j(t)x_j(t)u_j(t) \rangle = \langle s_j(t) \rangle \langle x_j(t) \rangle \langle u_j(t) \rangle, \tag{15}$$

$$\langle s_j(t)u_j(t) \rangle = \langle s_j(t) \rangle \langle u_j(t) \rangle. \tag{16}$$

Note that the independency between $x_j(t)$ and $u_j(t)$ is reported to hold if there is no facilitation (Tsodyks et al., 1998). Thus, we evaluate the validity of this assumption by comparison between the simulation and the mean field model derived by this assumption. As we show in "Results" section, the mean field model shows good approximations. By using these relations (15) and (16), the microscopic mean field model is derived as

$$m_i(t+1) = g_{\beta}\left[\sum_{j \neq i}^{N} J_{ij}\left(2m_j(t)X_j(t)U_j(t)/U_{se} - 1\right)\right], \tag{17}$$

$$X_i(t+1) = X_i(t) + \frac{1 - X_i(t)}{\tau_R} - m_i(t)X_i(t)U_i(t), \tag{18}$$

$$U_i(t+1) = U_i(t) + \frac{U_{se} - U_i(t)}{\tau_F} + U_{se}(1 - U_i(t))m_i(t), \tag{19}$$

where $m_i(t) \equiv \langle s_i(t) \rangle$, $X_i(t) \equiv \langle x_i(t) \rangle$, and $U_i(t) \equiv \langle u_i(t) \rangle$.

We now derive the mean field model that describes the macroscopic dynamical properties of the associative memory network. Here we use the sublattice method (Coolen, 2001) with p-dimensional pattern vectors $\boldsymbol{\eta} = (\eta^1, \ldots, \eta^p)^T \in \{-1, 1\}^p$. A set of neurons $\{1, \ldots, N\}$ is divided into 2^p groups on the basis of these pattern vectors. Suppose that $\bar{\boldsymbol{\xi}}_i = (\xi_i^1, \ldots, \xi_i^p)^T \in \{-1, 1\}^p$, a sublattice is defined as a set of neurons belonging to a given pattern vector. The sublattice belonging to the pattern vector $\boldsymbol{\eta}$ is defined as

$$\mathcal{I}_{\boldsymbol{\eta}} = \{i | \bar{\boldsymbol{\xi}}_i = \boldsymbol{\eta}\}, \tag{20}$$

$$\{1, \ldots, N\} = \bigcup_{\boldsymbol{\eta}} \mathcal{I}_{\boldsymbol{\eta}}, \tag{21}$$

where $\mathcal{I}_{\boldsymbol{\eta}}$ is called a sublattice.

The absolute strength of synaptic connection (Equation 10) can be rewritten with the expression of the sublattice as follows:

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{P} \eta^{\mu}\eta'^{\mu} = \frac{1}{N}\boldsymbol{\eta} \cdot \boldsymbol{\eta}', \tag{22}$$

for $i \in \mathcal{I}_{\boldsymbol{\eta}}$, and $j \in \mathcal{I}_{\boldsymbol{\eta}'}$.

We assumed that neurons within the same sublattice $\mathcal{I}_{\boldsymbol{\eta}}$ follow the same dynamics and that the variables in the microscopic mean

field model (Equations 17–19) can be described as

$$m_i(t) = m_\eta(t), X_i(t) = X_\eta, \text{ and } U_i(t) = U_\eta \text{ for } i \in \mathcal{I}_\eta. \quad (23)$$

With these assumptions, we obtain the following macroscopic mean field model of the associative memory network:

$$m_\eta(t+1) = F_{m_\eta}(\{m_\eta(t)\}, \{X_\eta(t)\}, \{U_\eta(t)\}), \quad (24)$$

$$X_\eta(t+1) = F_{X_\eta}(\{m_\eta(t)\}, \{X_\eta(t)\}, \{U_\eta(t)\}), \quad (25)$$

$$U_\eta(t+1) = F_{U_\eta}(\{m_\eta(t)\}, \{X_\eta(t)\}, \{U_\eta(t)\}), \quad (26)$$

where

$$F_{m_\eta}(\{m_\eta(t)\}, \{X_\eta(t)\}, \{U_\eta(t)\})$$

$$= g_\beta \left(\sum_{\eta'} p_{\eta'} \eta \cdot \eta' \left(2m_{\eta'}(t) X_{\eta'}(t) U_{\eta'}(t)/U_{se} - 1 \right) \right), \quad (27)$$

$$F_{X_\eta}(\{m_\eta(t)\}, \{X_\eta(t)\}, \{U_\eta(t)\})$$

$$= X_\eta(t) + \frac{1 - X_\eta(t)}{\tau_R} - m_\eta(t) X_\eta(t) U_\eta(t), \quad (28)$$

$$F_{U_\eta}(\{m_\eta(t)\}, \{X_\eta(t)\}, \{U_\eta(t)\})$$

$$= U_\eta(t) + \frac{U_{se} - U_\eta(t)}{\tau_F} + U_{se}(1 - U_\eta(t))m_\eta(t), \quad (29)$$

where $p_\eta = |\mathcal{I}_\eta|/N$ denotes the relative sublattice size.

We represent the steady state for the macroscopic mean field model by $\bar{m}_\eta$, $\bar{X}_\eta$, and $\bar{U}_\eta$. The steady state for the Equations (24–26) with $(t \to \infty)$ is given by the following self-consistent equations:

$$\bar{m}_\eta = g_\beta \left(\sum_{\eta'} p_{\eta'} \eta \cdot \eta' \left(\frac{2\bar{m}_{\eta'}(1 + \tau_F \bar{m}_{\eta'})}{1 + (\tau_F + \tau_R) U_{se} \bar{m}_{\eta'} + U_{se}\tau_F\tau_R \bar{m}_{\eta'}^2} - 1 \right) \right), \quad (30)$$

$$\bar{X}_\eta = \frac{1}{1 + \tau_R \bar{U}_\eta \bar{m}_\eta}, \quad (31)$$

$$\bar{U}_\eta = \frac{U_{se}(1 + \tau_F \bar{m}_\eta)}{1 + \tau_F U_{se} \bar{m}_\eta}. \quad (32)$$

To investigate the stability of the system given by Equations (24–26) around the steady state given by Equations (30–32), we consider the locally linearized equations with small perturbations $\delta m_\eta(t)$, $\delta X_\eta(t)$, and $\delta U_\eta(t)$ around the steady state as follows:

$$m_\eta(t) = \bar{m}_\eta + \delta m_\eta(t), \quad (33)$$

$$X_\eta(t) = \bar{X}_\eta + \delta X_\eta(t), \quad (34)$$

$$U_\eta(t) = \bar{U}_\eta + \delta U_\eta(t). \quad (35)$$

We obtain the following locally linearized equations on the small perturbations around the steady state with Jacobian matrix K.

$$\begin{pmatrix} \delta m_\eta(t+1) \\ \delta X_\eta(t+1) \\ \delta U_\eta(t+1) \end{pmatrix} = K \begin{pmatrix} \delta m_\eta(t) \\ \delta X_\eta(t) \\ \delta U_\eta(t) \end{pmatrix}. \quad (36)$$

The stability of the system can be determined by the eigenvalues of the Jacobian matrix on the steady state; the stability is distinguished by the absolute value of the eigenvalues. Elements on the Jacobian matrix K are given as

$$\frac{\partial F_{m_\eta}}{\partial m_{\eta'}} = g_\beta'(h) p_{\eta'} \eta \cdot \eta' (2X_{\eta'}(t) U_{\eta'}(t)/U_{se}), \quad (37)$$

$$\frac{\partial F_{m_\eta}}{\partial X_{\eta'}} = g_\beta'(h) p_{\eta'} \eta \cdot \eta' (2m_{\eta'}(t) U_{\eta'}(t)/U_{se}), \quad (38)$$

$$\frac{\partial F_{U_\eta}}{\partial U_{\eta'}} = g_\beta'(h) p_{\eta'} \eta \cdot \eta' (2m_{\eta'}(t) X_{\eta'}(t)/U_{se}), \quad (39)$$

where

$$g_\beta'(h) = \frac{\beta}{2} \left(1 - \tanh^2(\beta h) \right), \quad (40)$$

$$h = \sum_{\eta'} p_{\eta'} \eta \cdot \eta' \left(2m_{\eta'} X_{\eta'} U_{\eta'}/U_{se} - 1 \right). \quad (41)$$

Furthermore, the remaining matrix elements are given by

$$\frac{\partial F_{X_\eta}}{\partial m_{\eta'}} = -U_\eta X_\eta \delta_{\eta,\eta'}, \quad (42)$$

$$\frac{\partial F_{X_\eta}}{\partial X_{\eta'}} = \left(\left(1 - \frac{1}{\tau_R} \right) - m_\eta U_\eta \right) \delta_{\eta,\eta'}, \quad (43)$$

$$\frac{\partial F_{X_\eta}}{\partial U_{\eta'}} = -m_\eta X_\eta \delta_{\eta,\eta'}, \quad (44)$$

$$\frac{\partial F_{U_\eta}}{\partial m_{\eta'}} = U_{se}(1 - U_\eta) \delta_{\eta,\eta'}, \quad (45)$$

$$\frac{\partial F_{U_\eta}}{\partial X_{\eta'}} = 0, \quad (46)$$

$$\frac{\partial F_{U_\eta}}{\partial U_{\eta'}} = \left(\left(1 - \frac{1}{\tau_F} \right) - U_{se} m_\eta \right) \delta_{\eta,\eta'}, \quad (47)$$

where $\delta_{\eta,\eta'}$ is Kronecker's delta, namely, $\delta_{\eta,\eta'}$ is 1 if the $\eta = \eta'$, and 0 otherwise. By using this Jacobian matrix, we analyze the stability of the steady states in the following section.

In the following analysis, we fix the number of stored pattern to be $p = 3$. In this case, neurons can be divided into eight sublattices with the following combination of η:

$$\eta \in \{(1,1,1)^T, (1,1,-1)^T, (1,-1,1)^T, (1,-1,-1)^T,$$
$$(-1,1,1)^T, (-1,1,-1)^T, (-1,-1,1)^T, (-1,-1,-1)^T\}. \quad (48)$$

Since the memory patterns are provided by Equations (8) and (9), the number of neurons $|\mathcal{I}_\eta|$ in the sublattice $\mathcal{I}_\eta$ is given as follows (Otsubo et al., 2010):

$$|\mathcal{I}_\eta| = \begin{cases} N(1 + 3b^2)/8, & \text{if } \eta = (1, 1, 1)^T, (-1, -1, -1)^T, \\ N(1 - b^2)/8, & \text{otherwise.} \end{cases} \quad (49)$$

The model with $p = 3$ is composed of 24 variables in total.

3. RESULTS

In this section, we present the results of simulation in the stochastic model and of analyses of the macroscopic behavior in the associative memory model with dynamic synapses. In particular, we analyze the changes in the structure of the dynamics, depending on the parameters T, τ_F, τ_R, and U_{se}.

To quantify the similarity between the state of the network $s(t)$ and the μth memory pattern ξ^μ, we use an overlap given by

$$M^\mu(t) = \frac{1}{N} \sum_{i=1}^{N} \xi_i^\mu [2s_i(t) - 1]. \quad (50)$$

In the Equation (50), if $2s_i(t) - 1$ is equal to $\xi_i^\mu, \forall i$, then $M^\mu(t) = 1$. This means that if the state of neurons completely matches the μth memory pattern, the overlap becomes unity. In the formulation of the macroscopic mean field model, the above equation can be rewritten as follows:

$$M^\mu(t) = \sum_{\eta'} p_{\eta'} \eta'^\mu [2m_{\eta'}(t) - 1]. \quad (51)$$

Furthermore, the state of the network is classified according to the symmetry of the overlaps by using the effective dimension (ED) which is defined in the following. We consider only the case with $p = 3$. If the values of three overlaps at time t are equal or nearly equal, namely, if they satisfy $|M^\mu(t) - M^\nu(t)| < \epsilon, \forall (\mu, \nu) \in \{(1, 2), (2, 3), (3, 1)\}$, then $ED(t) = 1$, where $\epsilon = 10^{-5}$. If the values of all the overlaps are different i.e., if they satisfy $|M^\mu(t) - M^\nu(t)| > \epsilon, \forall (\mu, \nu) \in \{(1, 2), (2, 3), (3, 1)\}$, then $ED(t) = 3$. Otherwise, namely, if the values of two of three overlaps are equivalent, $ED(t) = 2$. The mean effective dimension (MED) is defined as $MED = \sum_{t=1}^{L} ED(t)/L$, where L is the length of a given time course.

We classified the state of the network according to the overlaps and the ED. There are four different types of steady state (fixed point), described as follows. In the memory state (MEM), one of the memory patterns or inverted memory patterns is retrieved. In the symmetric (asymmetric) mixed state [SMIX(AMIX)], one of the symmetric (asymmetric) mixture of the memory patterns is retrieved. In the paramagnetic state (PARA), the network does not retrieve any patterns and the state of each neuron is random. The oscillatory states have been classified according to the ED of the macroscopic mean field model giving rise to three oscillatory regimes: OS1, OS2, and OS3 states, which satisfy $MED = 1$, $1 < MED \leq 2$, and $2 < MED \leq 3$, respectively.

Figure 2 shows typical time courses indicating that the state of the network converges to the steady states. The top panels in

each subfigure in **Figure 2** show a raster plot; the dots indicate the active state of the neuron with $s_i(t) = 1$. The initial states of the simulation in the stochastic model are $x_i(t) = 1$, $u_i(t) = U_{se}$, and $s_i(t)$ are set to be 0 or 1 randomly so that the overlaps are almost zero in the initial state. We used $N = 10^4$ neurons in the simulation. The bottom panels show overlaps $M^1(t)$, $M^2(t)$, and $M^3(t)$ of the stochastic model (solid curves) and its corresponding steady states in the macroscopic mean field model (dashed lines). Appearance of the steady states of the stochastic model is consistent with the corresponding macroscopic mean field model.

In the MEM state (**Figure 2A**), one of the memory patterns or inverted memory patterns is retrieved. The state of the network converges to a steady state, which corresponds to a stable fixed point in the macroscopic mean field model. The steady state can be represented with the overlaps as e.g., $(M^1, M^2, M^3) = (M, M^*, M^*)$, where M and M^* satisfy $M > M^* > 0$, and the corresponding memory pattern is ξ^1. There are six possible MEM states: the states obtained by the permutations on the three memory patterns and its inversion. **Figure 2A** shows a typical time course of the process of convergence to the MEM state (to the memory pattern ξ^3 in the **Figure 2A**) in the stochastic model.

In the SMIX state (**Figure 2B**), the mixture of the memory patterns or the inverted memory patterns is retrieved. There are two possible SMIX states; the SMIX states are represented as $(M^1, M^2, M^3) = (\bar{M}, \bar{M}, \bar{M})$ for the mixture of the stored patterns and $(M^1, M^2, M^3) = (-\bar{M}, -\bar{M}, -\bar{M})$ for its inverse, where $\bar{M} > 0$. The corresponding memory patterns are $\text{sgn}(\xi^1 + \xi^2 + \xi^3)$ and $-\text{sgn}(\xi^1 + \xi^2 + \xi^3)$, respectively. **Figure 2B** shows a typical time course that the network converges to the SMIX state [to the mixture of the stored patterns $\text{sgn}(\xi^1 + \xi^2 + \xi^3)$ in the **Figure 2B**].
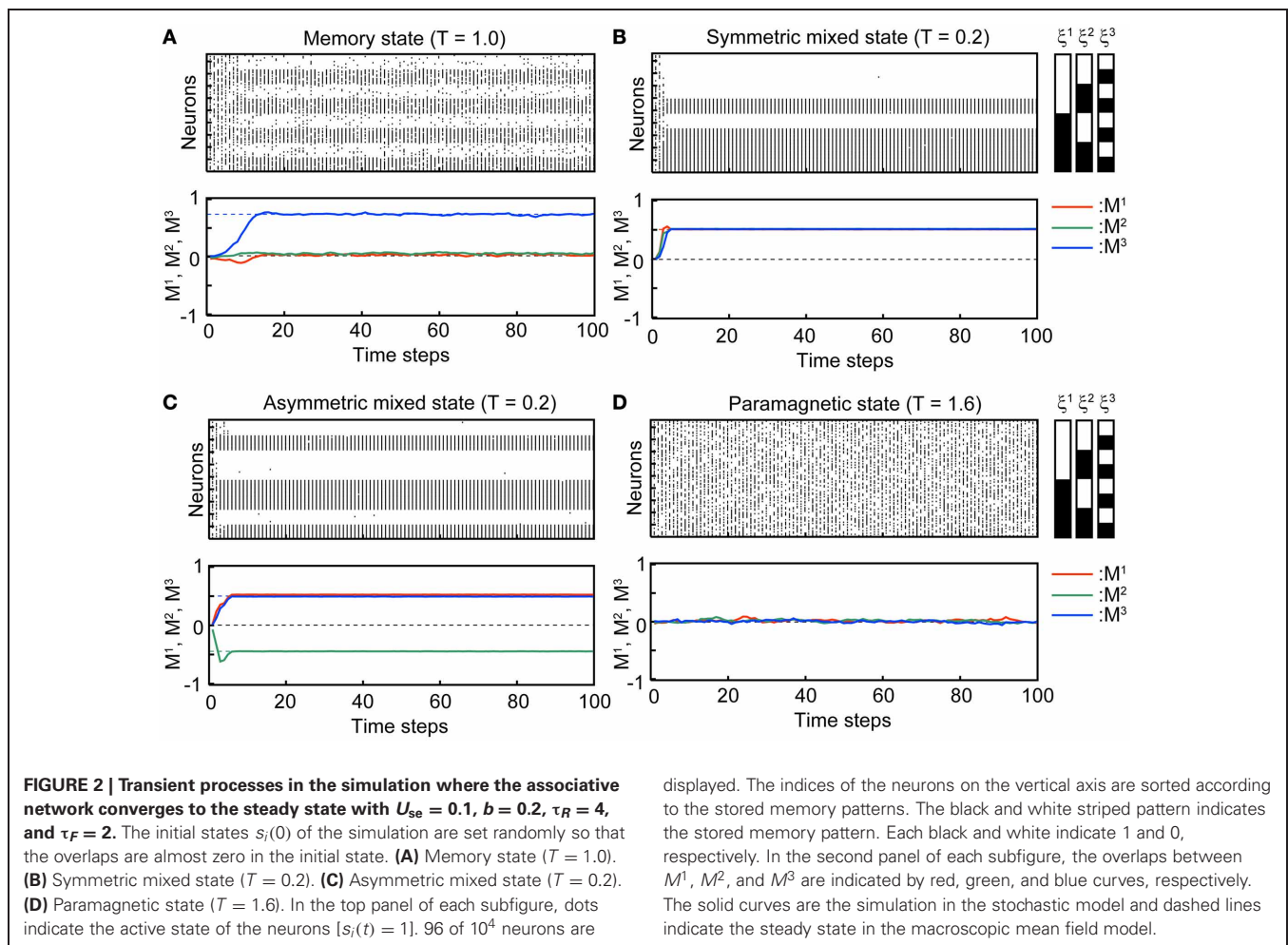
In the AMIX state (**Figure 2C**), one of the asymmetric mixture of the memory patterns is retrieved. The AMIX state can be represented as e.g., $(M^1, M^2, M^3) = (-M'', M', M')$, where $M' > M'' > 0$, and the corresponding memory pattern is $\text{sgn}(-\xi^1 + \xi^2 + \xi^3)$. There are six possible AMIX states: the states obtained by the permutations on the three memory patterns and its inversion. **Figure 2C** shows a typical time course of the state of the network when the state converges to the AMIX state that corresponds to the pattern $\text{sgn}(\xi^1 - \xi^2 + \xi^3)$.

In the PARA state, the state of each neuron is random. Thus, the PARA state is represented as $(M^1, M^2, M^3) = (0, 0, 0)$. **Figure 2D** shows that the network stays on the PARA state.

Figure 3 shows typical time courses of the oscillatory states in the stochastic model with $N = 10^4$ and the corresponding macroscopic mean field model. Dynamics of the mean field model is shown in the third panel in each subfigure and is consistent with that of the corresponding stochastic model.

In the OS1 state shown in **Figure 3A**, the network oscillates between the mixed state and the inverse of the mixed state; thus, the overlaps M^1, M^2, and M^3 oscillate in phase and the $ED = 1$. The time course of overlaps in the macroscopic mean field model is shown in the third panel in **Figure 3A**.

In the OS2 state shown in **Figure 3B**, the network oscillates between one of the memory patterns and its inverse pattern; one of the overlap (M^1 in the **Figure 3B**) oscillates with larger amplitude than others. The remaining two overlaps oscillate in phase.

FIGURE 2 | Transient processes in the simulation where the associative network converges to the steady state with $U_{se} = 0.1$, $b = 0.2$, $\tau_R = 4$, and $\tau_F = 2$. The initial states $s_i(0)$ of the simulation are set randomly so that the overlaps are almost zero in the initial state. **(A)** Memory state ($T = 1.0$). **(B)** Symmetric mixed state ($T = 0.2$). **(C)** Asymmetric mixed state ($T = 0.2$). **(D)** Paramagnetic state ($T = 1.6$). In the top panel of each subfigure, dots indicate the active state of the neurons [$s_i(t) = 1$]. 96 of 10^4 neurons are displayed. The indices of the neurons on the vertical axis are sorted according to the stored memory patterns. The black and white striped pattern indicates the stored memory pattern. Each black and white indicate 1 and 0, respectively. In the second panel of each subfigure, the overlaps between M^1, M^2, and M^3 are indicated by red, green, and blue curves, respectively. The solid curves are the simulation in the stochastic model and dashed lines indicate the steady state in the macroscopic mean field model.

Because the model is symmetric, three possible patterns of oscillation exist and the realization of the oscillatory pattern depends on the initial state of the network.

In the OS3 state shown in **Figures 3C,D**, there are two sub-modes of oscillatory states. The first mode oscillates symmetrically between one of memory patterns and its inverted patterns, and appearance of the oscillation circulate among the three memory patterns (see **Figure 3C**). The order of the three memory pattern randomly changes in the stochastic model. In the macroscopic mean field model, the oscillatory pattern with the orders $M^1 \rightarrow M^2 \rightarrow M^3$ and $M^3 \rightarrow M^2 \rightarrow M^1$ coexist (the oscillatory pattern with the order $M^1 \rightarrow M^2 \rightarrow M^3$ is shown in the third panel of **Figure 3C**). The second mode shows asymmetric oscillation among three memory patterns (see **Figure 3D**) or among three inverted patterns. The order of circulation in the three memory (or inverted-memory) patterns is random in the simulation.

Figure 4 shows the qualitative difference in the bifurcation diagrams with respect to the noise intensity T in three different parameter regions: the pseudo-constant region ($\tau_R = 4$ and $\tau_F = 2$), the depression-dominant region ($\tau_R = 10$ and $\tau_F = 2$), and the facilitation-dominant region ($\tau_R = 4$ and $\tau_F = 24$). Here, we set $b = 0.2$ and $U_{se} = 0.1$.

In the pseudo-constant region (**Figure 4A**), the time constants τ_R and τ_F are relatively small, then the effect of the short-term plasticity quickly disappears, and the transmission efficacy of the dynamic synapses remains nearby its steady state. **Figure 4A** shows the bifurcation diagram with respect to the noise intensity T in the pseudo-constant region with $\tau_R = 4$ and $\tau_F = 2$. In the relatively low noise range with $T < 0.4$, AMIX, SMIX, and MEM states coexist as the stable fixed points. The absolute values of the overlaps decreased with T. As T increases, the fixed points that correspond to the AMIX states are destabilized via the saddle-node (SN) bifurcation at $T = 0.429$. Each of two SMIX states intersects with three unstable fixed points and becomes unstable at $T = 0.781$ via the transcritical (TC) bifurcation, which is stabilized again at $T = 1.161$ via another TC bifurcation. The two SMIX states disappear by coalescing with an unstable fixed point at $T = 1.488$ via the pitchfork (PF) bifurcation, and the stable fixed point that corresponds to the PARA state emerges. All six MEM states disappear at $T = 1.248$ via the SN bifurcation.

In the depression-dominant region (**Figure 4B**), τ_R is relatively large, and the effect of decreases in the releasable neurotransmitters remains long. In this region, the position of the fixed point shrink to the low-noise side and quasi-periodic circles that correspond to oscillatory states appear. As T increases, AMIX, SMIX,
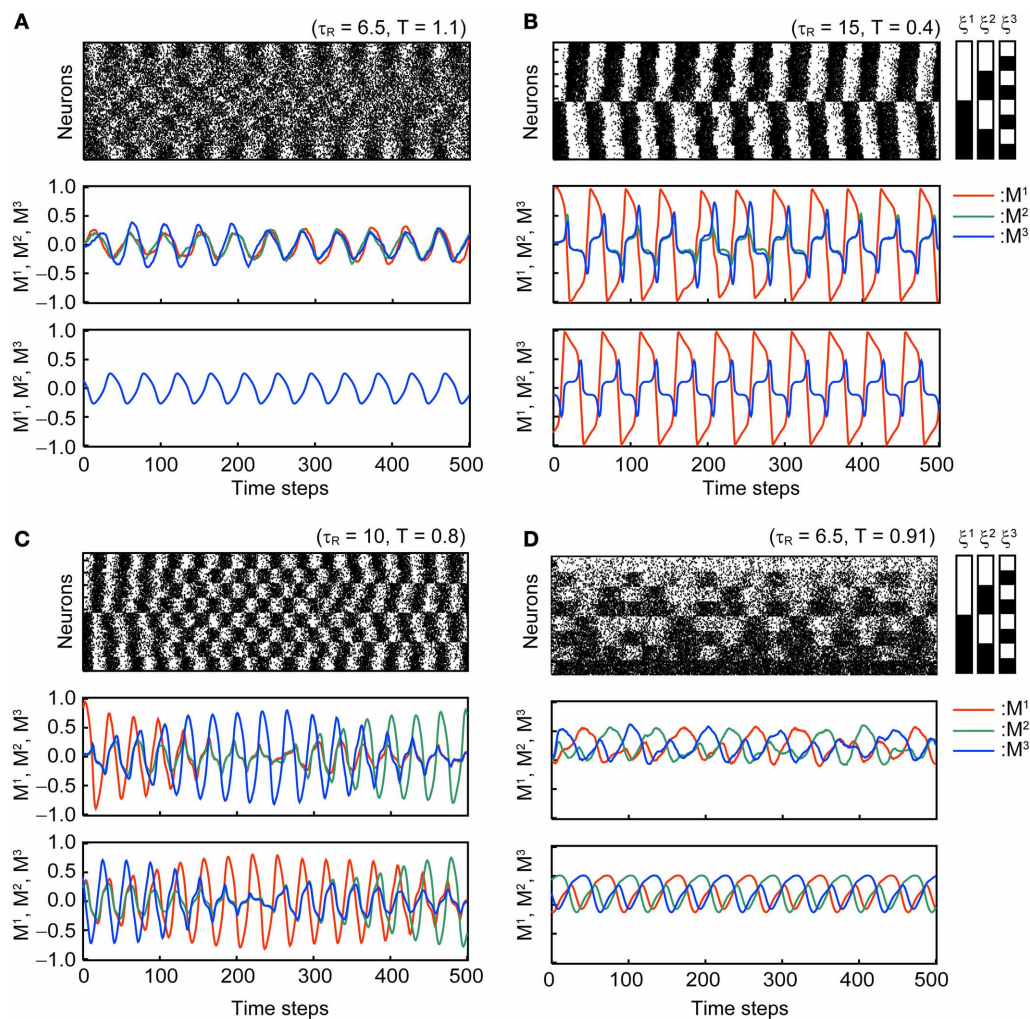
FIGURE 3 | Oscillatory states of the associative network with
$U_{se} = 0.1$, $b = 0.2$, **and** $\tau_F = 2$. The first and second panels of each
subfigure are the simulation results in the same format as in **Figure 2**.
The third panel of each subfigure shows the time course of the overlaps
in the macroscopic mean field model. Subfigures **(A–D)** display the

different modes of oscillation. **(A)** Oscillatory state with MED = 1 (OS1).
(B) Oscillatory state with $1 <$ MED ≤ 2 (OS2). **(C)** Oscillatory state
showing the symmetric oscillatory pattern with $2 <$ MED ≤ 3 (OS3).
(D) Oscillatory state showing the asymmetric oscillatory pattern with
$2 <$ MED ≤ 3 (OS3).

and MEM states are destabilized via the Neimark-Sacker (NS)
bifurcations at $T = 0.212$, $T = 0.311$, and $T = 0.576$, respectively. The oscillatory states appear at $T = 0.569$ and exhibit
quasi-periodic oscillation on an invariant circle. There exists a
multi-stable state of the stable fixed point and quasi-periodic
states on the range from $T = 0.569$ to $T = 0.576$. As T increases,
OS2, OS3, and OS1 appear in this order. The oscillatory states
disappear via the NS bifurcation at $T = 1.180$.

In the facilitation-dominant region (**Figure 4C**), τ_F is relatively large, and the effect of increase in the utilization parameter
remains long. In this region, the range of the fixed points that
correspond to the MEM, SMIX, and AMIX is expanded. The overall bifurcation structure is similar to that of the pseudo-constant
region, but the SMIX state is destabilized at $T = 1.845$ via the NS
bifurcation. Furthermore, the OS1 state appear at $T = 1.811$ and
disappear at $T = 1.964$ via the NS bifurcation.

Figure 5A shows a bifurcation diagram for comparison
between the macroscopic mean field model and the simulation
when we set $U_{se} = 0.1$, $\tau_R = 10$, $\tau_F = 2$, $b = 0.2$, and $N = 10^4$
with several initial values. The simulation shows good agreement with the corresponding macroscopic mean field model.
Figure 5B shows an orbit of an OS3 state for $U_{se} = 0.1$, $\tau_R = 6.5$,
$\tau_F = 2$, $b = 0.2$, and $T = 0.91$ in the simulation with $N = 10^4$
(red dots) and in the macroscopic mean field model (the blue
solid curve). The quasi-periodic orbit in the macroscopic mean
field model also shows good agreement with the simulation. We
have confirmed that the simulation result becomes closer to the
macroscopic mean field model when N is increased.

The phase diagrams in **Figures 6**, **7** show sets of bifurcation
points that switch the stability of the fixed points and the distribution of the oscillatory states obtained by the brute-force methods.
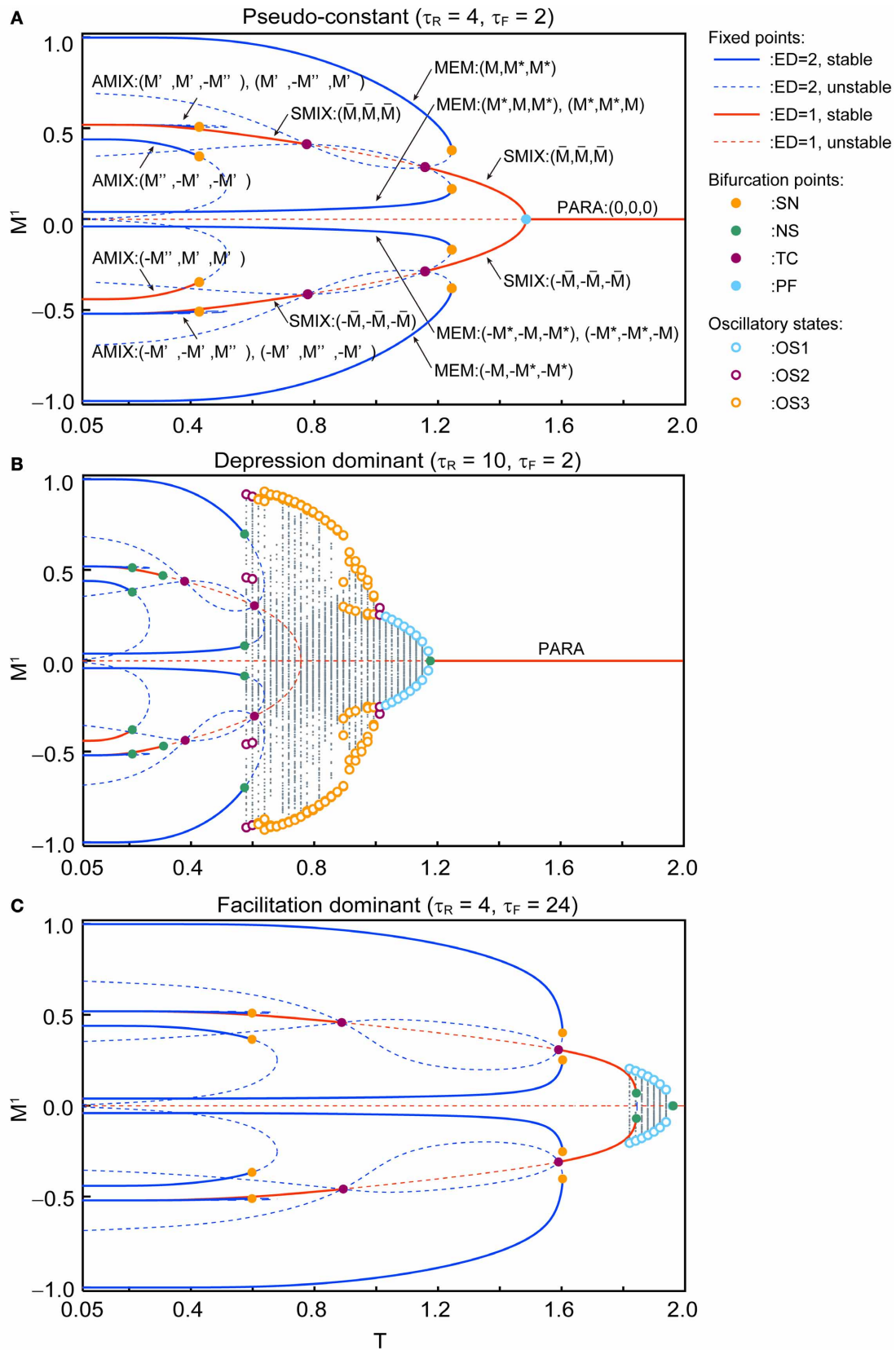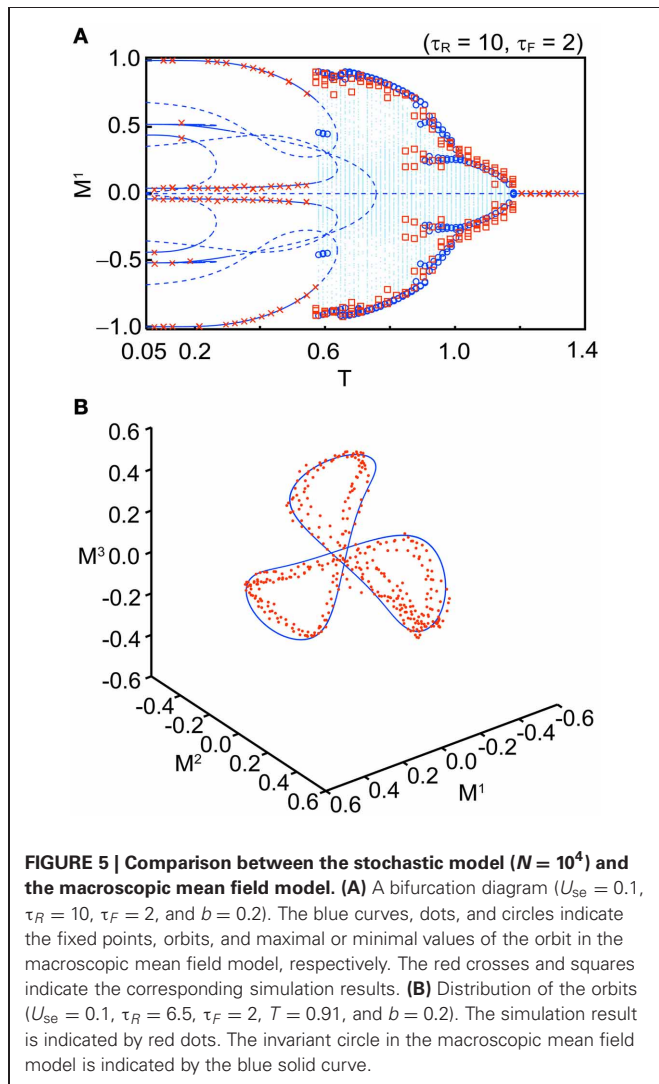We calculated the time evolution of the macroscopic mean field

FIGURE 4 | Continued

FIGURE 5 | Comparison between the stochastic model ($N = 10^4$) and the macroscopic mean field model. (A) A bifurcation diagram ($U_{se} = 0.1$, $\tau_R = 10$, $\tau_F = 2$, and $b = 0.2$). The blue curves, dots, and circles indicate the fixed points, orbits, and maximal or minimal values of the orbit in the macroscopic mean field model, respectively. The red crosses and squares indicate the corresponding simulation results. **(B)** Distribution of the orbits ($U_{se} = 0.1$, $\tau_R = 6.5$, $\tau_F = 2$, $T = 0.91$, and $b = 0.2$). The simulation result is indicated by red dots. The invariant circle in the macroscopic mean field model is indicated by the blue solid curve.

model on each parameter points; the parameter points where the orbit converges to the oscillatory states are indicated by colored dots in **Figures 6**, **7**. In the higher-noise boundary of the oscillatory state, the oscillatory states are separated by the supercritical type of the NS bifurcation; the region of the oscillatory states is well separated by the sets of the NS bifurcation. On the other hand, the oscillatory states appear with the subcritical type of NS bifurcation in the lower-noise boundary. Thus, the oscillatory states and the steady states coexist as multi-stable states in this region. Similar bifurcation structure is found in the uniformly connected network (Katori et al., 2012).
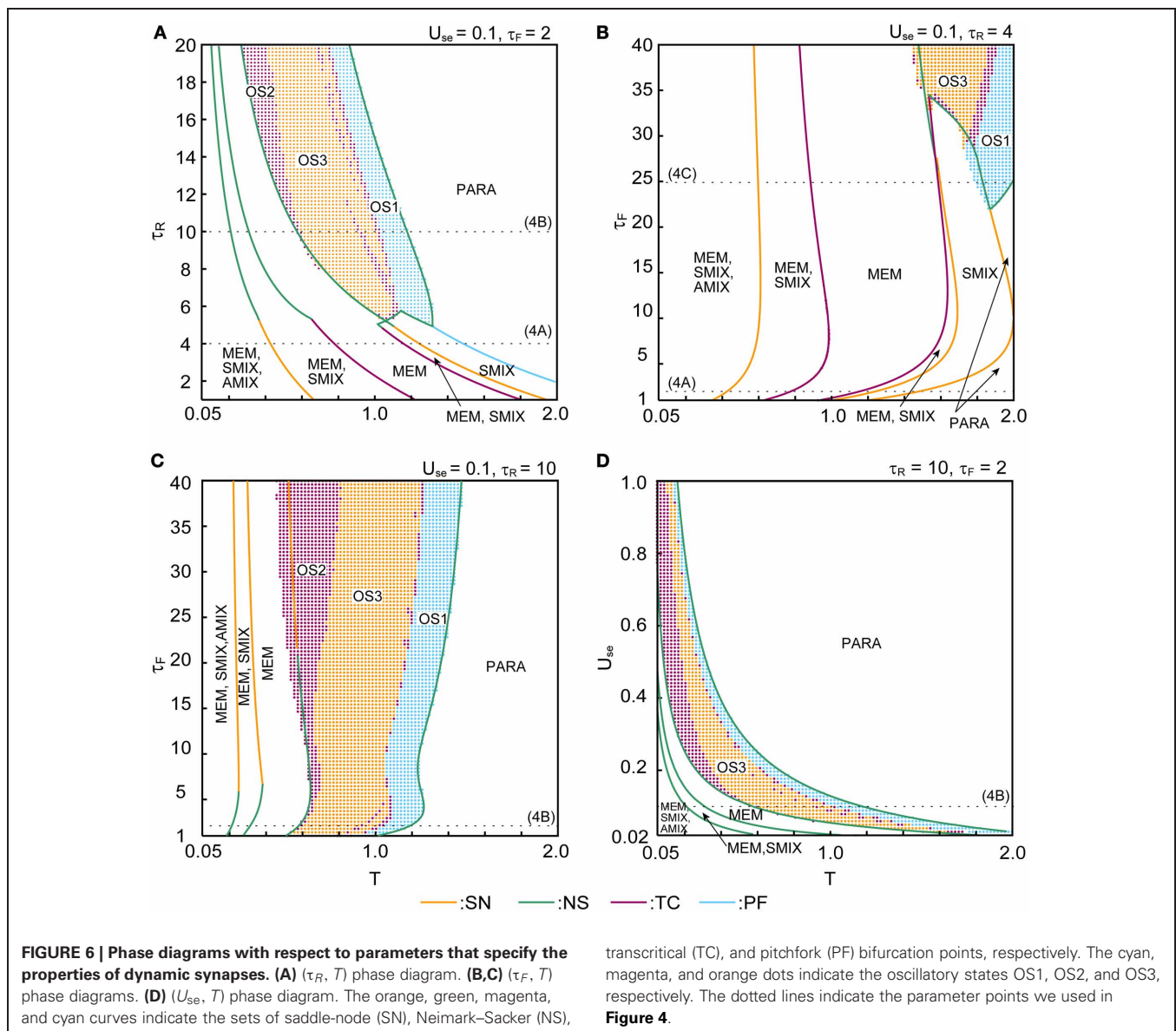
The (T, τ_R) phase diagram in **Figure 6A** shows changes in the dynamical properties of the network from the pseudo-constant region to the depression-dominant region. As τ_R increases, the regions of the stable fixed point of MEM, SMIX, and AMIX shrink, while the regions of the PARA state and the oscillatory states expand. The (T, τ_F) phase diagram shown in **Figure 6B** illustrates the dynamical properties from the pseudo-constant region to the facilitation-dominant region. As τ_F increase, the regions of MEM, SMIX, and AMIX expand, while the region of the PARA state shrinks. Furthermore, the oscillatory states appear. As τ_F increases from the depression-dominant region (**Figure 6C**), the regions of the oscillatory states expand. As U_{se} increases, the region of the PARA state expands, while regions of other states shrink.

The (T, b) phase diagrams in **Figure 7** show that the dynamical properties of the network depend on the correlation level between the memory patterns. As b increases, the region of the SMIX state expands, while regions of the other states shrink. In the depression-dominant range (**Figure 7B**), as the correlation level b increases, the region of the OS3 state shrinks but that of OS1 state remain, which corresponds to the oscillatory state between SMIX states. In the facilitation-dominant range (**Figure 7C**), the overall bifurcation structure is similar to that of the pseudo-constant range, but the region of MEM states expands.

4. DISCUSSION

In this study, we investigated the dynamical properties of an associative memory network composed of a stochastic neural network with both short-term depression and facilitation synapses on the basis of the macroscopic mean field model. We analyzed the behavior of the network in broad ranges of parameters that specify the noise intensity and the properties of the dynamic synapses. We found that the associative memory network exhibits the variety of dynamics, including the memory state, SMIX and AMIX, and several modes of the oscillatory states, and that its properties change with various types of bifurcations.

The performance of the memory retrieval can be characterized by the appearance of the MEM state in which the state of the network successfully converges to one of the memory patterns. In addition to the MEM state, in the relatively-low-noise range, there exists SMIX and AMIX states that correspond to pseudo-memory patterns. In this parameter range, the retrieval of the memory pattern is not assured and depends on the initial state of the network. In the high-noise range, the network tends to the PARA state, which corresponds to the state in which the pattern of neural activity is disrupted and randomized because of the noise. We classified the oscillatory states into three modes according to the ED. The OS1

FIGURE 6 | Phase diagrams with respect to parameters that specify the properties of dynamic synapses. (A) (τ_R, T) phase diagram. **(B,C)** (τ_F, T) phase diagrams. **(D)** (U_{se}, T) phase diagram. The orange, green, magenta, and cyan curves indicate the sets of saddle-node (SN), Neimark–Sacker (NS), transcritical (TC), and pitchfork (PF) bifurcation points, respectively. The cyan, magenta, and orange dots indicate the oscillatory states OS1, OS2, and OS3, respectively. The dotted lines indicate the parameter points we used in **Figure 4**.

state corresponds to oscillation between the pseudo-memory patterns, and it appears in the relatively high noise range. The OS2 state is the oscillation between one of the memory patterns and its inverse pattern, and it appears next to the MEM state. The OS3 state is the transitive state between memory patterns and their inverse patterns. Such transitive dynamics is related to the itinerant dynamics in terms of chaotic dynamics (Tsuda et al., 1987; Adachi and Aihara, 1997; Kanamaru et al., 2013).

The appearance of the above mentioned states of the network depends on the properties of the dynamic synapses (**Figure 6**) and on the correlation level between memory patterns (**Figure 7**). In the pseudo-constant region (**Figure 4A**), the state of the network converges to one of the fixed points like the conventional associative memory model (Anderson and Bower, 1972;

Nakano, 1972; Hopfield, 1982). In the depression-dominant region, which is archived by increasing the recovery time constant τ_R from the pseudo-constant region, the area of successful memory retrieval shrinks, whereas the oscillatory states appear as shown in **Figure 6A**. Increase in the fraction of neurotransmitter-release U_{se} intensifies the influence of the depression. As U_{se} increase, the area of the PARA state expands, whereas the areas of other states shrink (**Figure 6D**). In the facilitation-dominant region, which is archived by increasing the time constant τ_F from the pseudo-constant region, the area of the memory retrieval expands (**Figures 6B, 7C**), which suggests that the facilitation synapses contribute to the memory retrieval (Mongillo et al., 2008). As the correlation level among memory patterns increases (**Figure 7**), the network loses the ability to retrieve the memory pattern, and the state of the network tends to become the
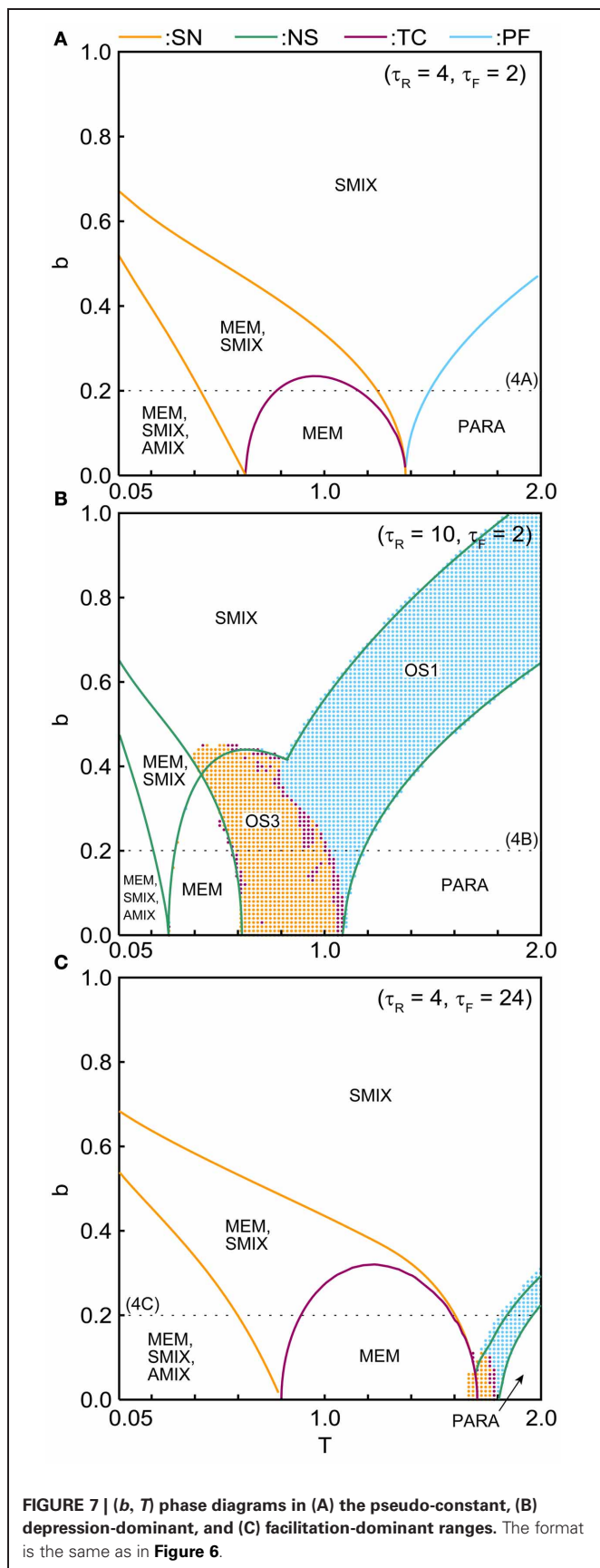
FIGURE 7 | (b, T) phase diagrams in (A) the pseudo-constant, (B) depression-dominant, and (C) facilitation-dominant ranges. The format is the same as in **Figure 6**.

pseudo-memory pattern. In the region of the oscillatory states, the oscillatory state among the memory patterns shrinks, whereas the oscillatory state between pseudo-memory patterns remains (**Figure 7B**).

These results have implications regarding brain functions. The distribution of facilitation and depression synapses in the brain varies according to the region of the brain. Many facilitation synapses exist in the prefrontal lobe and, whereas many depression synapses appear in the parietal lobe (Wang et al., 2006). The facilitation synapses may form a synaptic working memory and contribute to the prefrontal function, which requires a flexible executive function. Conversely, the depression synapses might be involved in memory search or mental rotation, which requires to imagine to handle an object in the parietal cortex (Tagaris et al., 1996). The oscillatory states OS3 observed in the present model correspond to the states that the neural network sequentially retrieves stored memory patterns. The oscillatory state appears with the incorporation of depression synapses. Furthermore, the area of the oscillatory state expands with increase in the time constant of the facilitation process. These findings imply that the depression and facilitation synapses contribute to various brain functions e.g., a generation of sequential actions or the flexible information representation (Katori et al., 2011).

The main findings of this work are consistent with previously reported studies on associative memory networks, and we revealed further details of the network dynamics. In the previous study on the associative memory network with both depression and facilitation synapses by Torres et al. (2007), the mean activities with active and inert neurons are considered to construct the mean field model, in which the number of the variables in the model is on the order of p. On the other hand, in our present study, we constructed the mean field model formulated with the sublattice method that enables to analyze the non-homogeneous network structure of the associative memory network; the number of the variables is on the order of 2^p. In the case with $p = 1$, these two mean field models are equivalent, whereas these are differences in cases with $p \geq 2$. Here we discussed the case with $p = 3$ and reported that the associative memory network exhibits a variety of dynamical states, including the memory and pseudo-memory states, as well as several oscillatory states among memory patterns. Furthermore, we reported the dependency of these states on the noise level and the parameters that specify the properties of the dynamic synapses, including details of bifurcation structure.

Although, we have considered the properties of the steady state and the oscillatory state as the attractors in the present study, properties of a transient process of memory retrieval should be evaluated. The relation between the stability of the memory retrieved states and irregularity of the neural activity (Mongillo et al., 2012) remains to be further investigated. In the present study, we used a simple neuron model, namely the discrete-time and binary neuron model. Meanwhile, the behavior observed in the present model should be qualitatively and/or quantitatively evaluated in more realistic neuron models e.g., integrate-and-fire or Hodgkin–Huxley model in the future.

REFERENCES

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science* 275, 220–224.

Adachi, M., and Aihara, K. (1997). Associative dynamics in a chaotic neural network. *Neural Netw.* 10, 83–98.

Amari, S. (1977). Neural theory of association and concept-formation. *Biol. Cybern.* 26, 175–185.

Amit, D., Gutfreund, H., and Sompolinsky, H. (1985). Spin-glass models of neural networks. *Phys. Rev. A* 32, 1007–1018.

Anderson, J. R., and Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychol. Rev.* 79, 97–123.

Barbieri, F., and Brunel, N. (2007). Irregular persistent activity induced by synaptic excitatory feedback. *Front. Comput. Neurosci.* 1:5. doi: 10.3389/neuro.10.005.2007

Coolen, A. (2001). Statistical mechanics of recurrent neural networks IIdynamics. *Handb. Biol. Phys.* 4, 619–684.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.

Igarashi, Y., Oizumi, M., and Okada, M. (2010). Mean field analysis of stochastic neural network models with synaptic depression. *J. Phys. Soc. Jpn.* 79, 84001.

Kanamaru, T., Fujii, H., and Aihara, K. (2013). Deformation of attractor landscape via cholinergic presynaptic modulations: a computational study using a phase neuron model. *PLoS ONE* 8:e53854. doi: 10.1371/journal.pone.0053854

Katori, Y., Igarashi, Y., Okada, M., and Aihara, K. (2012). Stability analysis of stochastic neural network with depression and facilitation synapses. *J. Phys. Soc. Jpn.* 81, 114007.

Katori, Y., Sakamoto, K., Saito, N., Tanji, J., Mushiake, H., and Aihara, K. (2011). Representational switching by dynamical reorganization of attractor structure in a network model of the prefrontal cortex. *PLoS Comput. Biol.* 7:e1002266. doi: 10.1371/journal.pcbi.1002266

Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382, 807–810.

Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. U.S.A.* 95, 5323–5328.

Mejias, J., and Torres, J. (2009). Maximum memory capacity on neural networks with short-term synaptic depression and facilitation. *Neural Comput.* 21, 851–871.

Melamed, O., Barak, O., Silberberg, G., Markram, H., and Tsodyks, M. (2008). Slow oscillations in neural networks with facilitating synapses. *J. Comput. Neurosci.* 25, 308–316.

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science* 319, 1543–1546.

Mongillo, G., Hansel, D., and van Vreeswijk, C. (2012). Bistability and spatiotemporal irregularity in neuronal networks with nonlinear synaptic transmission. *Phys. Rev. Lett.* 108, 158101.

Nakano, K. (1972). Associatron-a model of associative memory. *IEEE Trans. Syst. Man. Cybern.* SMC-2, 380–388.

Otsubo, Y., Nagata, K., Oizumi, M., and Okada, M. (2010). Instabilities in associative memory model with synaptic depression and switching phenomena among attractors. *J. Phys. Soc. Jpn.* 79, 84002.

Pantic, L., Torres, J., Kappen, H., and Gielen, S. (2002). Associative memory with dynamic synapses. *Neural Comput.* 14, 2903–2923.

Tagaris, G., Kim, S., Strupp, J., Andersen, P., Ugurbil, K., and Georgopoulos, A. (1996). Quantitative relations between parietal activation and performance in mental rotation. *Neuroreport* 7, 773–776.

Thomson, A. M. (2000). Facilitation, augmentation and potentiation at central synapses. *Trends Neurosci.* 23, 305–312.

Torres, J., Cortes, J., Marro, J., and Kappen, H. (2007). Competition between synaptic depression and facilitation in attractor neural networks. *Neural Comput.* 19, 2739–2755.

Toya, K., Fukushima, K., Kabashima, Y., and Okada, M. (2000). Bistability of mixed states in a neural network storing hierarchical patterns. *J. Phys. A Math. Gen.* 33, 2725–2737.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835.

Tsodyks, M. V., and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. U.S.A.* 94, 719–723.

Tsuda, I., Koerner, E., and Shimizu, H. (1987). Memory dynamics in asynchronous neural networks. *Prog. Theor. Phys.* 78, 51–71.

Wang, Y., Markram, H., Goodman, P. H., Berger, T. K., Ma, J., and Goldman-Rakic, P. S. (2006). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat. Neurosci.* 9, 534–542.