Behavioral/Systems/Cognitive

Gating of Signal Propagation in Spiking Neural Networks by Balanced and Correlated Excitation and Inhibition

Jens Kremkow, Ad Aertsen, and Arvind Kumar
Bernstein Center Freiburg, Neurobiology and Biophysics, Faculty of Biology, University of Freiburg, 79104 Freiburg, Germany

Both ongoing and natural stimulus driven neuronal activity are dominated by transients. Selective gating of these transients is mandatory for proper brain function and may, in fact, form the basis of millisecond-fast decision making and action selection. Here we propose that neuronal networks may exploit timing differences between correlated excitation and inhibition (temporal gating) to control the propagation of spiking activity transients. When combined with excitation–inhibition balance, temporal gating constitutes a powerful mechanism to control the propagation of mixtures of transient and tonic neural activity components.

Introduction

Appropriate action selection is one of the fundamental problems that the brain needs to solve in its interaction with the environment. This problem is even enhanced by the intrinsically parallel architecture of the brain. A mechanism based on winner-take-all dynamics is often suggested for this purpose (Ermentrout, 1992). However, winner-take-all dynamics require strong recurrent inhibition, which may not be available in every brain region (Tepper et al., 2004). Moreover, winner-take-all dynamics are typically slow if the competing signals are of comparable strength, and may even result in unpredictable responses in the presence of transient synchrony within and among the competing inputs (Ermentrout, 1992). Another approach to action selection could be based on the gating of signals, such that only one (or few) effectively propagate to the next processing stage. Thus, recently Vogels and Abbott (2009) demonstrated that detailed balance of excitation and inhibition (amplitude gating) can efficiently gate the propagation of firing rates. However, while this mechanism allows for a rapid selection among competing rate inputs, it fails to control activity transients on a timescale of up to a few tens of milliseconds [Vogels and Abbott (2009), their Fig. 3*f*]. In a natural stimulus environment, such activity transients frequently occur as a consequence of stimulus-driven changes in firing rates or by the emergence of correlations in the network (Buracas et al., 1998; Bar-Yosef and Nelken, 2007; Butts et al., 2007; Hromádka et al., 2008; Haider et al., 2010). Therefore, reliable gating of activity transients is a prerequisite of proper brain function.

Here we propose an alternative gating mechanism that exploits temporal differences between correlated excitation and inhibition (Okun and Lampl, 2008; Gentet et al., 2010). We show that small differences between the timing of excitation and inhibition (referred to temporal gating from hereon) can act as a modulator of neuronal integration time and, thereby, serve as a highly selective gate for signal transients. The proposed mechanism allows for both gating at a millisecond timescale and reliable control of onset transients and correlated rate inputs. When combined with amplitude gating, temporal gating constitutes a powerful mechanism to control the propagation of mixtures of transient and tonic neural activity components.

Received July 26, 2010; revised Sept. 22, 2010; accepted Sept. 23, 2010.

This work was supported by the German Federal Ministry of Education and Research (BMBF Grant 01GQ0420 to Bernstein Center for Computational Neuroscience Freiburg) and the 6th Research Framework Programme of the European Union (Grant 15879-FACETS). We thank Tim Vogels, Benjamin Staude, Ioannis Vlachos, Guillaume S. Masson, and Laurent U. Perrinet for helpful discussions.

Correspondence should be addressed to Arvind Kumar at the above addresses. E-mail: arvind.kumar@biologie.uni-freiburg.de.

DOI:10.1523/JNEUROSCI.3874-10.2010

 0270-6474/10/3015760-09$15.00/0

Materials and Methods

Models

Neurons

The neurons in the network were modeled as leaky integrate-and-fire neurons, with the subthreshold dynamics of the membrane potential $V^i(t)$ in neuron i described by the following equation:

$$C\frac{d}{dt}V^i(t) + G_{\text{rest}}[V^i(t) - V_{\text{rest}}] = I^i_{syn},$$

where I^i_{syn} is the total synaptic input current into neuron i, and C and G_{rest} denote the passive electrical properties of its membrane at rest (V_{rest}). When the membrane potential reached a fixed spike threshold V_{th} above rest, a spike was emitted, the membrane potential was reset to its resting value, and synaptic integration was halted for 2 ms, mimicking the refractory period in real neurons. The parameters used in the simulations were as follows: $C = 290$ pF, $G_{\text{rest}} = 29$ nS, $V_{\text{rest}} = -70$ mV, and $V_{\text{th}} = -57$ mV (Destexhe et al., 1998; Muller et al., 2007).

Network

We simulated a network of 28,125 neurons, with 22,500 excitatory and 5625 inhibitory neurons (ratio 4:1). The excitatory neurons were arranged on a Cartesian grid of 150×150, the inhibitory neurons on a grid of 75×75. Both grids represented the same cortical space of 1×1 mm^2. To avoid boundary effects, the network was folded as a torus. The recurrent network was sparsely connected with a connection probability of ~5% ($K_{\text{exc}\to\text{exc}} = 1120$, $K_{\text{exc}\to\text{inh}} = 280$, $K_{\text{inh}\to\text{exc}} = 1120$, $K_{\text{inh}\to\text{inh}} = 280$). As an exception to this rule, excitatory neurons that participated in the signal path received less than 5% connection from the background neurons, to maintain the same in-degree. The local connection probabil-

ity was described as a function of their cell body distance using a Gaussian profile (Hellwig, 2000). The width of the Gaussian describing the connection probability σ_e was set to 0.6 mm for the excitatory and $\sigma_i = 0.1$ mm for the inhibitory connections (Stepanyants et al., 2008) (Fig. 1*a*). In addition to the local connections from within the network, each neuron received excitatory connections from outside, mimicking both long-range connections within the same cortical area and input from other areas in cortex (Braitenberg and Schüz, 1991; Kumar et al., 2008a). The external inputs were modeled as uncorrelated Poisson-type spike trains. The rate and strength of external inputs (1500 excitatory synapses per neuron, each at ~2 Hz) were adjusted to obtain a low-firing-rate (~3 Hz), asynchronous (pairwise correlation ≈ 0.01), and irregular activity state in the network (Brunel, 2000; Kumar et al., 2008b).

Synapses

Synaptic inputs were modeled as transient conductance changes, using exponential functions with $\tau_{exc} = 1.5$ ms and $\tau_{inh} = 10$ ms (Kuhn et al., 2004; Muller et al., 2007). Excitatory and inhibitory synaptic delays in the embedding network were set to 2 ms. The weights (peak conductance) of the excitatory synapses targeting excitatory neurons (J_{EE}) were set to 0.5 nS, such that a postsynaptic potential of ~0.15 mV was elicited at rest (Matsumura et al., 1996; Bruno and Sakmann, 2006; Kumar et al., 2008a). The weight of excitatory synapses targeting inhibitory neurons (J_{EI}) was set to 1 nS (Cruikshank et al., 2007; Kremkow et al., 2010). The synaptic weights of all inhibitory synapses (J_{II}, J_{IE}) were set to a 0.5 nS conductance value. This set of synaptic strengths and external input (see above) generated both a realistic effective membrane time constant of ~5 ms (Anderson et al., 2000; Destexhe et al., 2003; Léger et al., 2005) and a realistic low-firing-rate, asynchronous-irregular network activity regime (Brunel, 2000; Kumar et al., 2008b; Ecker et al., 2010).

Embedding of a feedforward network as signal path

To construct a signal path, we embedded a feedforward network, containing three successive groups of neurons (sender, gate, and receiver groups), into the recurrent network. The sender group was composed solely of 100 excitatory neurons, whereas the gate and receiver group contained 100 excitatory and 25 inhibitory neurons each (Fig. 1*b*). The inhibitory neurons provided the balancing inhibition in the signal path and controlled the effective integration time of the excitatory neurons. The three groups were spatially separated to avoid cross talk between groups (Kumar et al., 2008a) (Fig. 1*b*). The neurons constituting a group were randomly chosen from a local pool of 300 excitatory and 75 inhibitory neurons. The center of the neuron pool defined the physical location of the group in the network. Each neuron in the signal path received convergent inputs from 60 randomly selected excitatory neurons of the preceding group. The neurons in the gate and receiver groups received excitatory synapses from 1060 neurons (<5%) of the background network to maintain uniform in-degree in the network. The delay of excitatory synapses in the signal path was set to 5 ms (Aertsen et al., 1996; Diesmann et al., 1999). This choice of the excitatory delays of 5 ms is somewhat arbitrary. These delays are supposed to mimic the latency between different neighboring brain regions, and such delays can be in the order of up to 10 ms. Further, we point out that these delays do not influence the dynamics of the embedding network and the embedded feedforward network, and can be reduced to any arbitrary value without influencing the gating. In the default state of the signal path, the delay of inhibitory synapses onto the excitatory neurons within a group was set to 2 ms (Cruikshank et al., 2007; Kremkow et al., 2010).

Embedding multiple signal paths

To construct two signal paths that merged in the gate group, we embedded the first signal path as described above. The sender group of the second signal path was then added. It was spatially separated from the first sender group to avoid cross talk. The excitatory neurons in the gate and receiver groups were common in both signal paths. The inhibitory neurons of the gate group in both signal paths were randomly chosen from a local pool of 75 inhibitory neurons. Thus, both signal paths might share a small fraction of inhibitory neurons in the gate group (see Fig. 5*a*).

Stimuli

We used three different stimulus classes to study signal propagation and gating in the signal path.

Pulse packet input. The input to the sender group was a volley of spikes with α spike times drawn from a Gaussian distribution of width σ (Aertsen et al., 1996; Diesmann et al., 1999). The number of synapses from the presynaptic neurons to the sender group was set to 60, the same as within the signal path.

Uncorrelated rate input. We used a Poisson process to stimulate the sender group with uncorrelated firing rate input (Vogels and Abbott, 2009). The onset time of the stimulus was set to 500 ms to mimic a sudden activation of the signal path (Butts et al., 2007). The rate of the Poison process was set to 200 Hz to induce a strong onset transient.

Correlated rate input. To introduce correlation in the firing rate input, we used the multiple-interaction process (MIP) (Kuhn et al., 2003). This process introduces correlations by copying spikes from a Poisson "mother process" into a desired number of "children" processes. The copy probability (c) defines the correlation among the children processes. As in the uncorrelated rate input, the onset time constant of the correlated rate stimulus was set to 500 ms. The rate of the MIP population was set to 20 Hz to illustrate that even at low firing rates, spike correlations [$c = 0.5$, e.g. (see Fig. 3*d*)] could induce strong transients.

Data analysis

Event characterization

To characterize the spiking activity of each group and, thus, the α and σ of the pulse packet sent to the next group, we measured "events," i.e., distinct, short bursts of spikes in the peristimulus time histogram (PSTH) of all excitatory neurons in a group. The spike count of all excitatory neurons in an event described the response strength α. Likewise, σ described the temporal spread of the spike times, that is, the time difference between the start and end of the "event." Thus, the propagation of activity in the signal path was determined by the response of the three groups of the signal path, quantified by α and σ.

Signal propagation

To determine whether a signal, composed of multiple pulse packets, could successfully propagate from the sender to the receiver, we measured the similarity between the spiking activities in both groups. The similarity was calculated by cross-correlating the PSTHs of the excitatory neurons in both groups. In case of successful propagation, the signal in the receiver group should have a similar temporal sequence of peaks in the PSTH. However, due to the synaptic delay in the signal path, this sequence would be shifted in time. Therefore, when calculating the cross-correlation between sender and receiver, we accounted for this time shift.

Simulation tools

All network simulations were written in Python (http://www.python.org) using PyNN (Davison et al., 2008) (http://neuralensemble.org/trac/PyNN) as interface to the simulation environment NEST (Morrison et al., 2005; Gewaltig and Diesmann, 2007; Eppler et al., 2008) (http://www.nest-initiative.org). The dynamical equations were integrated at a fixed temporal resolution of 0.1 ms. Simulation management was performed using the Python package NeuroTools (http://neuralensemble.org/trac/NeuroTools). Simulation data were analyzed in Python using the scientific libraries SciPy (http://www.scipy.org) and NumPy (http://numpy.scipy.org) and visualized using the plotting library Matplotlib (http://matplotlib.sourceforge.net).

Results

To systematically study the concept of temporal gating, we embedded a signal path in a large-scale recurrent cortical network model (Fig. 1*a*,*b*). The signal path was a feedforward network consisting of three successive groups of neurons. The first group, composed of excitatory neurons activated by the stimulus, is referred to as the sender. The second and third groups contained both excitatory and inhibitory neurons, with the inhibitory neurons providing disynaptic inhibition to the excitatory neurons

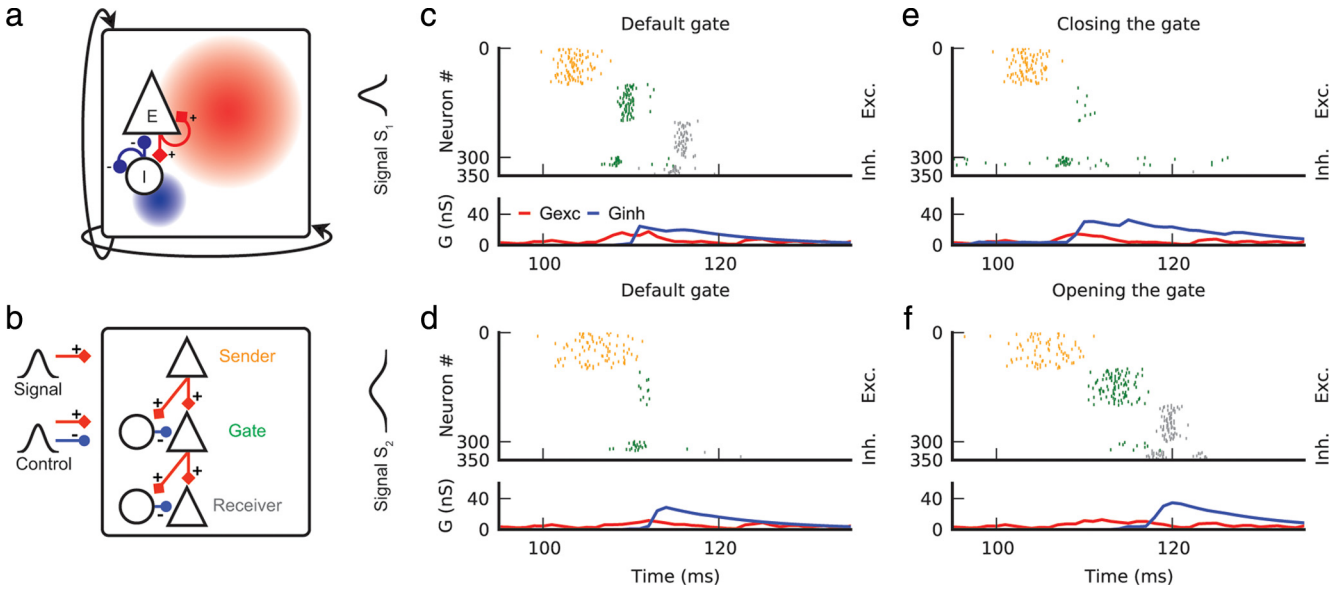

Figure 1. Temporal gating of spiking activity by delayed correlation between excitation and inhibition. ***a***, Scheme of cortical network model. Twenty-eight thousand, one hundred twenty-five neurons [80% excitatory (△) and 20% inhibitory (○)] were arranged as a two-dimensional sheet. Space constants of excitatory and inhibitor neurons are shown as red and blue shaded regions, respectively. ***b***, A signal path (sender, gate, receiver) was embedded in the network. The sender neuron group transmitted an input signal via the gate group toward the receiver group. A depolarizing or hyperpolarizing input to the gate was used as control signal. ***c***, ***d***, Top, Spiking activity of the sender, gate, and receiver groups. Bottom, Excitatory and inhibitory synaptic conductances (Gexc, Ginh) for a representative excitatory neuron in the gate group. ***c***, A strong ($\alpha = 60$), synchronous ($\sigma = 3.5$ ms) signal S_1 propagated through the gating group and successfully activated the receiver group. ***d***, By contrast, a less synchronous ($\sigma = 7$ ms) signal S_2 failed to propagate. ***e***, Propagation of S_1 could be blocked by advancing inhibitory spiking in the gate group, thereby reducing Δt of the excitatory neurons. ***f***, Propagation of S_2 could be rescued by delaying (or abolishing) inhibitory spiking in the gate group, thereby increasing Δt of the excitatory neurons.

within the same group. The third group is referred to as receiver, whereas the second group, which is in a strategic position to influence propagation from sender to receiver, is referred to as the gate (Fig. 1*b*). However, the gating mechanism is flexible and in a feedforward network with more than three groups, any group can be considered as a gate group between its preceding and succeeding partner groups. The embedding network was tuned into an asynchronous-irregular state to provide realistic background activity (Kumar et al., 2008a) (cf. Materials and Methods).

Temporal gating of signal propagation

To illustrate the mechanism of temporal gating, we stimulated the sender group with a volley of spikes (pulse packet; cf. Materials and Methods), with α spike times drawn from a Gaussian distribution of width σ.

We first considered the propagation of a synchronous stimulus S_1 (small σ). In the default state of the signal path (cf. Materials and Methods), S_1 generated a transient response in the sender, which successfully propagated through the gate to the receiver (Fig. 1*c*). This propagation can be described in terms of the synfire attractor, a defining feature of feedforward network dynamics (Diesmann et al., 1999; Kumar et al., 2008a, 2010; Kremkow et al., 2010), with the fate of signal propagation being determined by the properties of the pulse packet stimulus. The (α, σ) state space is subdivided by a separatrix (schematically shown in Fig. 2*a*, blue line). If the pulse packet starts above the separatrix (class *P* stimulus), the signal converges into a fixed point (FP) and propagates in a stable manner (Figs. 1*c*, 2*a*). If, by contrast, the pulse packet starts below the separatrix (class *F* stimulus), the signal progressively dies out, indicating failure to propagate (Figs. 1*d*, 2*a*). Thus, the location of the separatrix is the critical parameter that controls the propagation of a stimulus (Diesmann et al., 1999; Kumar et al., 2008a, 2010; Kremkow et al., 2010). Consequently, altering the location of the separatrix could change a class *P* stimulus into a class *F* stimulus (or vice versa), thereby providing a simple mechanism for gating the propagation of more or less transient spiking activity (Fig. 2*a*). Various properties of the feedforward network, such as group size, connection probability, and weights determine the location of the separatrix in the (α, σ) space, which, however, can also be modulated by the dynamical properties of the feedforward network, such as noise level of the embedding background activity (Diesmann et al., 1999; Gewaltig et al., 2001). Furthermore, we recently showed that the delay between excitation and inhibition within one group changes the slope of the separatrix by effectively altering the integration time (Δt) of the neurons (Kremkow et al., 2010). Interestingly, recent data from *in vivo* measurements of the delay between excitation and inhibition (Okun and Lampl, 2008) are in the right range to alter the slope of the separatrix in an appreciable way. Several biologically plausible mechanisms could lead to dynamic variations in the excitation/inhibition delay and, thereby, to associated modulations of Δt. For instance, neuromodulators could change Δt by differentially altering neuronal excitability or synaptic strengths (Kruglikov and Rudy, 2008; Antal et al., 2010). Similarly, stimulus properties could introduce variations in Δt (Zhang et al., 2003; Wilent and Contreras, 2005).

We exploited the control over the delay between excitation/inhibition to close the gate, i.e., to block the propagation of the S_1-type stimulus as shown in Figure 1*c*. Thus, we reduced the temporal delay between excitation and inhibition in the gating group by injecting a small depolarizing control pulse into the inhibitory neurons. As a result, the gate closed: the response in the gate group fell below the separatrix, causing activity in the receiver group to fade (Figs. 1*e*, 2*a*). To further explore the temporal gating mechanism, we systematically varied Δt by altering the inhibitory synaptic delay, while keeping the excitatory delay fixed. For different values of Δt, Figure 2*b* shows α and σ for

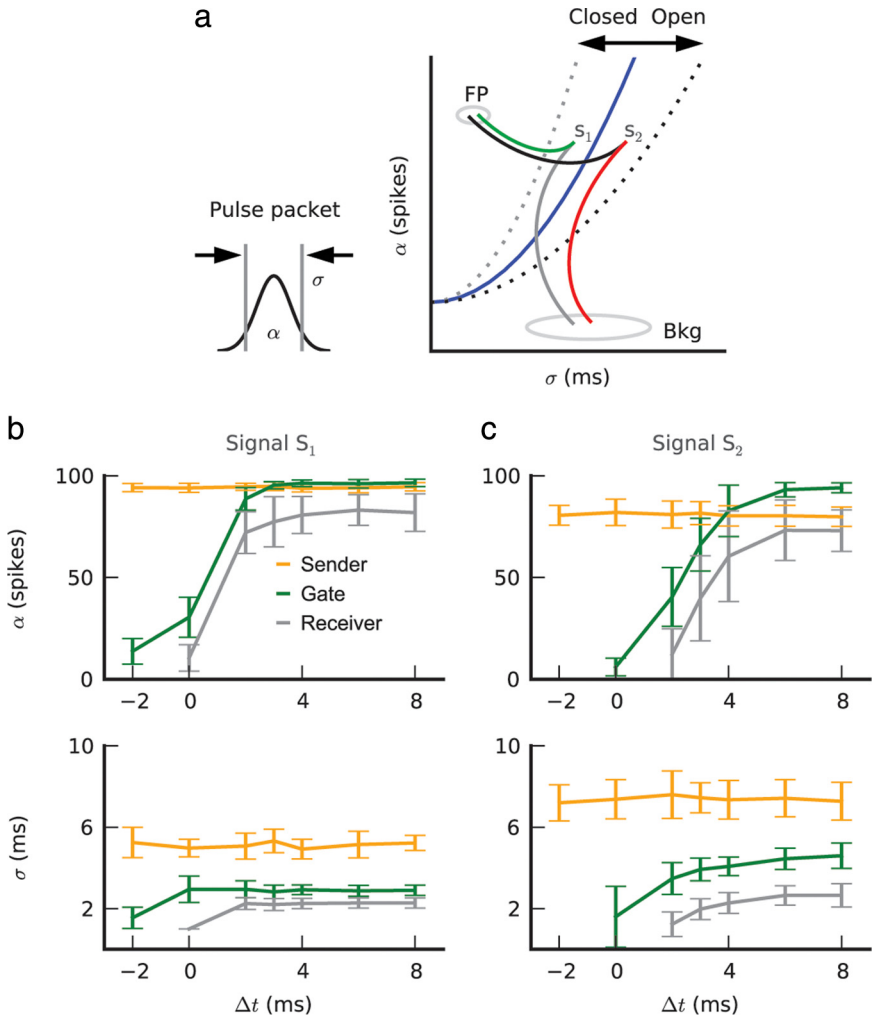

Figure 2. Temporal delay between excitation and inhibition controls signal gating. ***a***, Scheme of the pulse packet propagation state space, which is subdivided by a separatrix (Diesmann et al., 1999) (blue line). If the pulse packet starts above the separatrix (S_1), the signal converges into a fixed point (FP) and propagates in a stable manner (green trajectory). If, by contrast, the pulse packet starts below the separatrix (S_2), the signal fails to propagate and fades into the background (Bkg.) (red trajectory). When the effective integration time Δt is reduced, the separatrix moves up (gray dotted line), thus blocking the propagation of S_1 (gray trajectory). By contrast, when the integration time is increased, the separatrix moves down (black dotted line), thus allowing the propagation of S_2 (black trajectory). ***b***, α and σ in sender, gate, and receiver groups for the S_1 stimulus (class *P*) as a function of effective integration time (Δt). Error bars represent SD ($n = 20$). In this example, S_1 successfully propagates in the default state ($\Delta t = 2$ ms). Further increase in Δt does not influence signal propagation. However, a small decrease in Δt (~1 ms) suffices to completely block signal propagation. ***c***, Same as ***b*** for stimulus S_2 (class *F*). In this example, S_2 does not propagate in the default state ($\Delta t = 2$ ms). Propagation can be rescued by increasing Δt. Note that, in contrast to ***b***, the prolonged opening of the gate causes activity in the gate to be higher than in the receiver. Also, for increasing integration time (larger Δt), signal propagation becomes more reliable, which is expressed as a reduction of trial-by-trial variability in the receiver group.

the three excitatory neuron groups in the signal path. For small integration time ($\Delta t \leq 1$ ms in this example), the gate response was weak and fell below the separatrix. Consequently, the response in the receiver group was indistinguishable from the background network activity. However, increasing the integration time ($\Delta t \approx 2$ ms in this example) shifted the gate activity above the separatrix, inducing a strong response (high α, small σ) in the receiver group. The same mechanism can be used to open the gate, i.e., to facilitate propagation of class *F* stimuli, which would otherwise fail to propagate (Fig. 1*f*). To test this, we applied a hyperpolarizing (rather than depolarizing) control pulse to the inhibitory neurons, thereby delaying or even suppressing their spiking activity. As a result, the integration time was increased, such that more excitatory neurons in the gate group responded. Its activity thereby rose above the separatrix and, hence, ensured a strong response (high α, small σ) in the receiver group (Fig. 1*f*). A systematic increase in the integration time Δt of the gate neurons resulted in a monotonic increase in both α and σ of the gate activity (Fig. 2*c*). Due to the synfire attractor, the activity of the receiver group was further synchronized. Note that, in contrast to Figure 2*b*, the prolonged opening of the gate caused activity in the gate to be higher than in the receiver. Also, it is noteworthy that for increasing integration time (higher Δt), signal propagation became more reliable, as expressed in a reduction in trial-by-trial variability in the receiver group (Fig. 2*c*, top panel).

In summary, the sigmoidal relationship between Δt and the activity in the gate and receiver groups demonstrates the effectiveness of the temporal gating (Fig. 2*b*,*c*).

As there is an ongoing debate (Kuhn et al., 2004; Vogels and Abbott, 2005; Kumar et al., 2008a,b) on whether synapses in a point neuron model are best described by conductance transients (as done here) or current transients [the more common choice, e.g., Brunel (2000)], we resimulated the network model with current-based synapses. The sigmoidal relationship between Δt and the activity in the gate and receiver groups remained unchanged when synapses were modeled as current transients (compare supplemental Fig. S1, available at www.jneurosci.org as supplemental material, and Fig. 2*b*,*c*), demonstrating that the "temporal gating" does not depend on the choice of synapse or input activity model.

Gating of signals with strong transients

Next we tested whether temporal gating is also effective in controlling the propagation of signal onset transients. Such signal onset transients arise when the input signal is abruptly changed (Deger et al., 2009). To mimic sharp changes in the input, the sender group, we stimulated the sender group with Poisson-type spiking activity, starting at time $t = 500$ ms (Fig. 3*a*). In the default state of the gate group, with balanced excitation and inhibition, the onset transient propagated to the receiver group, whereas the tonic part of the input was blocked (Fig. 3*b*). The control of such transients, however, is critical because in a dynamic natural environment, such transients are likely to occur frequently (Buracas et al., 1998; Haider et al., 2010). In the default state of the gate, the effective integration time is large enough for the transients to elicit a response in the downstream neurons (gate group). Thus, an effective way to control the flow of transients is to reduce the integration time of the downstream neurons (e.g., by temporal gating) (see Figs. 1, 2). Indeed, when the delay between excitation and inhibition in the gate group was reduced, while still keeping the overall excitation and inhibition in balance, it was possible to control the propagation of both the

stimulus transient and its tonic component (Fig. 3*c*). Note that in this particular example, when the gate is closed (Fig. 3*c*), only the inhibitory population in the gate group is active, while the activity of the excitatory neurons is completely suppressed by the balancing inhibitory input. Reducing the gain of the inhibitory neurons would result in a biologically more realistic residual spiking activity in the excitatory neurons in the gate group. Transients in the spiking activity in neural systems may also arise due to spike time correlations among neurons (Alonso et al., 1996; Dan et al., 1998; Chen et al., 2006). The emergence of such correlations could be due to large-scale fluctuations in ongoing network activity (Arieli et al., 1996; Kenet et al., 2003) or stimulus-induced network activity (Bar-Yosef and Nelken, 2007; Alonso et al., 2008; Hromádka et al., 2008). To generate a signal composed of tonic and transient components, we used the MIP model (Kuhn et al., 2003) (Fig. 3*d*; cf. Materials and Methods). Once again, when the sender group was stimulated by a MIP-type input signal, the control of the activity was imprecise and signal transients reflecting the spike correlations propagated uninhibited (Fig. 3*e*). However, reducing the delay between excitation and inhibition at the gate group provided an effective control over the propagation of signal transients (Fig. 3*f*). In summary, Figure 3, *b* and *e*, illustrates the limitations of the Vogels and Abbott mechanism in gating of both onset transients (Fig. 3*b*) and transients induced by correlations (Fig. 3*e*). Both types of transients were fast enough to pass before the delayed feedforward inhibition could quench them. By contrast, Figure 3, *c* and *f*, shows the improved effectiveness of the temporal gating mechanism in controlling the propagation of these activity transients.

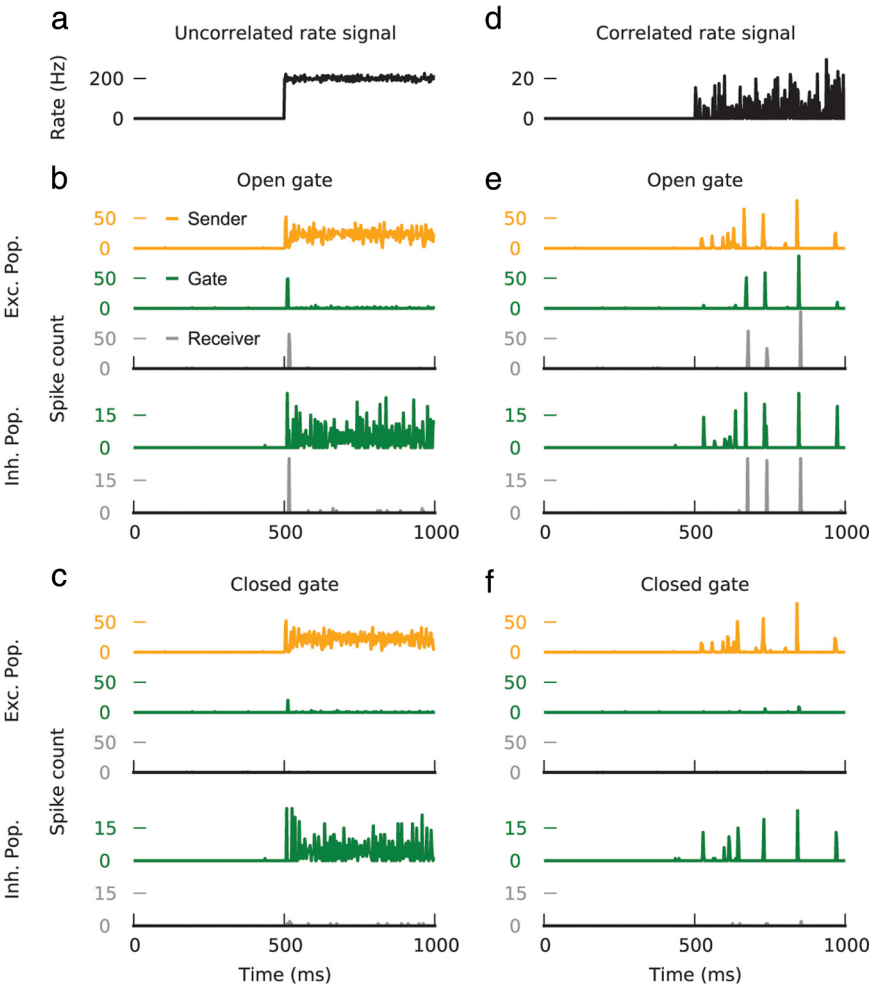

Figure 3. Effectiveness of gating tonic activity and activity transients. ***a***, Uncorrelated rate stimulus as input to the signal path. ***b***, Population activity in the excitatory (top) and inhibitory (bottom) neurons in the signal path. In the default state ($\Delta t = 2$ ms), gating by balancing excitation and inhibition successfully blocks propagation of the tonic component of the stimulus, but fails to block the stimulus onset transient. ***c***, By reducing the effective integration time of the excitatory neurons ($\Delta t \leq 1$ ms), temporal gating effectively blocks both the onset transient and the tonic stimulus component. ***d***, Firing rate stimulus with transient input correlations. ***e***, As in ***b***, gating by balancing excitation and inhibition fails to block activity transients, here induced by transient input correlations. ***f***, As in ***c***, temporal gating effectively blocks activity transients.

Combining temporal and amplitude gating

The effectiveness of temporal gating in controlling the propagation of activity transients has its price: it precludes the propagation of sustained (tonic) activity devoid of transients (Fig. 3*c*,*f*). While the importance of such sustained average neural activity can be argued, in view of the ubiquitous fluctuations in the ongoing activity and dynamic natural environment, it is of general interest to know whether it is possible to do both, to control the flow of activity transients, while allowing tonic activity to propagate. Because temporal gating is most effective in controlling the propagation of activity transients, whereas detailed balance (or amplitude gating) can be used to control the propagation of tonic activity, a combination of the two might provide a general mechanism to control the propagation of arbitrary neural activity patterns. Therefore, we examined whether a combination of temporal gating and amplitude gating would be possible within the same network.

For this we stimulated the signal path with uncorrelated Poisson activity, starting at time $t = 500$ ms. This type of stimulus contains both a transient onset and a tonic phase, which makes it optimal for probing the effectiveness of both gating mechanisms. The amplitude gating mechanism was implemented by varying the strength of the inhibition in the gate group, thereby controlling the balance of excitation and inhibition in the excitatory neurons of the gate group (Vogels and Abbott, 2009). To this end, we varied the synaptic strength from the excitatory neurons in the sender group onto the inhibitory population (J_{EI}), while keeping the synaptic strength from the excitatory neurons onto the excitatory neurons (J_{EE}) fixed. Thus, we define the inhibitory gain factor as J_{EI}/J_{EE}. In all examples shown so far (i.e., Figs. 1–3), the inhibitory gain factor was 2. This gain factor ensured an effective feedforward inhibition in the neurons of the gate group (cf. Materials and Methods) (Cruikshank et al., 2007; Kremkow et al., 2010). Figure 4 shows the results of systematically varying both the degree of temporal gating and amplitude gating. Simultaneous closing of both the temporal ($\Delta t = -2$ ms) and amplitude (inhibitory gain = 2) gates blocked both the transient and the tonic component in the gate group (Fig. 4*a*–*c*). Opening the tem-

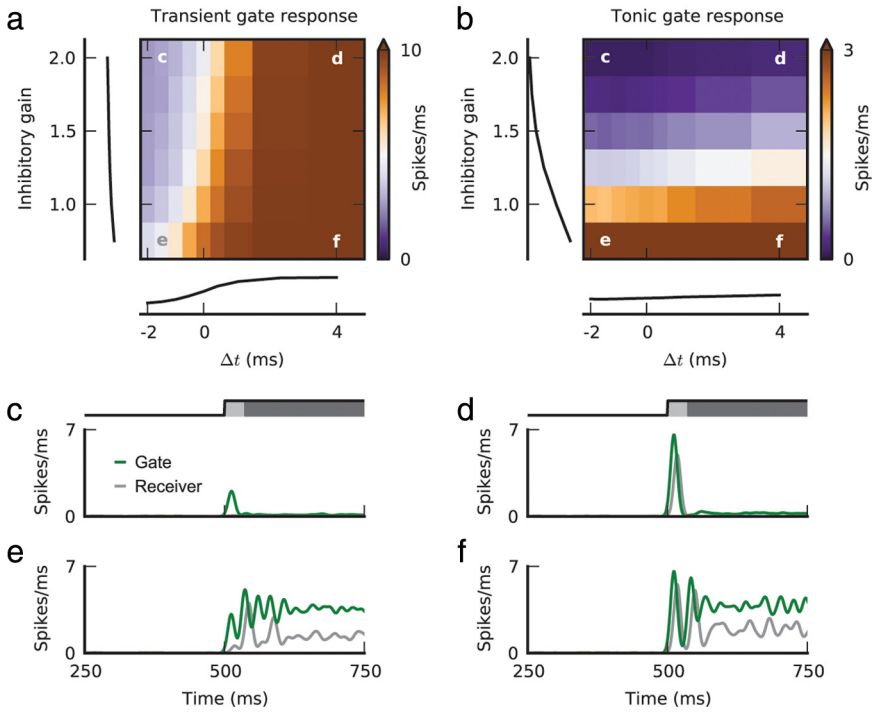

Figure 4. Combining temporal and amplitude gating. ***a***, Transient response (first 10 ms, marked light gray in ***c*** and ***d***) in the gate as a function of effective integration time Δt (temporal gating) and inhibitory gain (amplitude gating) with uncorrelated rate stimulus as input to the signal path. Mean transient response in the gate does not change as a function of inhibitory gain for all values of Δt, indicating that inhibitory gain cannot gate the propagation of transient component of the activity (left). By contrast, mean transient response in the gate as a function of Δt for all values of inhibitory gain shows a sigmoidal shape, indicating that Δt can control the propagation of transient component of the activity (bottom). ***b***, Tonic response (excluding first 10 ms, marked dark gray in ***c*** and ***d***) in the gate as a function of effective integration time Δt and inhibitory gain. Mean tonic response in the gate as a function of inhibitory gain for all values of Δt, varies in a threshold linear manner, indicating that inhibitory gain can control the propagation of tonic component of the activity (left). By contrast, mean tonic response in the gate does not change as a function of Δt for all values of inhibitory gain, indicating that Δt cannot gate the tonic component of the activity (bottom). ***c–f***, Population activity of the excitatory neurons in the gate and receiver group during four distinct gating modes, marked with corresponding letters in ***a*** and ***b***. ***c***, Simultaneously closing of the temporal and the amplitude gates blocks both the onset transient and the tonic component, resulting in a very small activity in the receiver group. ***d***, Opening the temporal gate, while keeping the amplitude gate closed, allows the onset transient to propagate. ***e***, Opening the amplitude gate, while keeping the temporal gate closed, allows the tonic component to propagate. ***f***, Both the transient and the tonic component are propagated when both gates are open.

poral gate ($\Delta t = 4$ ms), while keeping the amplitude gate closed (inhibitory gain = 2), allowed only the onset transient to induce activity in the gate group (Fig. 4*a*,*b*,*d*). These two examples, thus, show the effectivity of temporal gating at high inhibitory gain (compare Figs. 3 and 4*c*,*d*). By contrast, opening the amplitude gate (e.g., inhibitory gain = 0.75) resulted in tonic activity in the gate group (Fig. 4*a*,*b*,*e*,*f*). Here, the weaker feedforward inhibition could not balance the direct excitation during the tonic phase (Kremkow et al., 2010). However, during the transient phase the inputs from the sender group arrived synchronously, activating the inhibitory neurons in the gate group simultaneously, thereby providing sufficient balancing inhibition for the temporal gating to function and block the transient [compare closed temporal gate (Fig. 4*e*) and open temporal gate (Fig. 4*f*)]. Note that using inhibitory gain as a gating variable may result in an attenuated response in the receiver group. However, because a high signal-to-noise ratio is maintained, this reduced activity could easily be amplified in a subsequent stage.

In summary, Figure 4 shows that amplitude gating is effective in gating tonic component of the activity (Fig. 4*b*), while temporal gating is effective in gating the activity transients (Fig. 4*a*). When the two gating mechanisms are implemented within the same network, together they provide a powerful mechanism to control the propagation of a wide variety of temporal activity profiles.

Gating of multiple signals

So far we considered how isolated transients or those that form part of stimulus-evoked or ongoing activity modulations propagate along a single signal path. However, in a modular system like the brain, action selection can be thought of in terms of directing one of multiple signals to a next processing stage or, alternatively, as directing a particular signal to one of multiple subsequent processing stages (Bienenstock, 1995). In both these scenarios, the brain needs a mechanism to control the propagation of activity over multiple (and interacting) signal pathways. Here, we test the effectiveness of temporal gating in selecting between two neuronal activity signals that are statistically identical but uncorrelated, and dominated by transients.

Figure 5*a* shows the scheme of two signal paths, embedded in a recurrent random network (cf. Materials and Methods). This arrangement of sender, gate, and receiver groups is similar to that used in Figures 1–4, except that here two spatially separated sender groups (Sender_X and Sender_Y) can send signals (S_X, S_Y) to the gate. To control the propagation of either signal S_X or S_Y, we systematically varied the delay between excitation and inhibition from the respective sender to the gate (Δt_X and Δt_Y, respectively). To quantify the propagation of the two signals, we measured the cross-correlations (ρ_X and ρ_Y) between the receiver group output and each of the two input signals (cf. Materials and Methods). Figure 5*b* shows the difference between the two: $\rho = \rho_X - \rho_Y$ as a function of both Δt_X and Δt_Y. Positive values of ρ indicate that the receiver group output is more similar to the input signal S_X, and negative values reflect a higher similarity with input signal S_Y. When ρ is close to zero, the receiver output is equally correlated (including uncorrelated) with both input signals; in this case, the scenario is evidently not suitable for action selection.

When the two signals were presented simultaneously, selective transmission of signal S_X (Fig. 5*b*,*c*) could be obtained by increasing Δt_X and decreasing Δt_Y (effectively closing signal path S_Y) (compare Fig. 2). Similarly, selective transmission of signal S_Y (Fig. 5*b*,*f*) could be obtained by increasing Δt_Y and decreasing Δt_X (closing signal path S_X). The exact value of Δt_X to gate signal S_X depended on the value of Δt_Y (Fig. 5*b*) and vice versa. When the gate was closed for both signal paths (Fig. 4*e*), neither of the two input signals propagated to the receiver and, hence, ρ was close to zero. On the other hand, when the gate was opened for both signal paths (Fig. 5*d*), both signals propagated and the response of the receiver group was equally correlated to both input signals, again resulting in ρ being close to zero.

Thus, the temporal gating mechanism appears well suited for selective gating of multiple signals to a next processing stage.

Similarly, temporal gating can also be used to selectively direct a particular signal to one of several next processing stages (not shown here).

Discussion

In a modular system like the brain, appropriate gating or control over activity propagation is critical for information processing and action selection (Bienenstock, 1995; Kumar et al., 2010). In fact, for various biological reasons, gating of neural activity emerges as a superior strategy over winner-take-all dynamics for action selection (Ermentrout, 1992; Tepper et al., 2004). In any case, a key requirement for an appropriate gating mechanism is that it should be able to control neural activity with arbitrary statistical properties.

In particular, it is important to be able to gate neuronal activity that is dominated by strong transients, which are ubiquitous in a dynamic sensory environment, as has been reported under various conditions *in vivo* in all sensory pathways. For example, it has been demonstrated in the rodent whisker-barrel system that cortex is primarily driven by transient synchronous activity of thalamic neurons (Bruno and Sakmann, 2006). Similarly, neurons with overlapping receptive fields of the same polarity (on- or off-center) in the cat visual thalamus show spike correlation in the millisecond range and, consequently, drive postsynaptic neurons in the visual cortex much more effectively (Alonso et al., 1996). Furthermore, transient subthreshold dynamics have been measured intracellularly in the cat visual cortex *in vivo* during presentation of natural stimuli (Haider et al., 2010; Y. Fregnac, personal communication). Likewise, subthreshold membrane potential *in vivo* in the rat auditory cortex showed large transients, induced by synchronous synaptic input, occurring irregularly in an otherwise quiet background (DeWeese and Zador, 2006). Together, these observations in different sensory pathways in several species corroborate the importance of activity transients. Consequently, a proper gating mechanism should be able to cope with these transients.

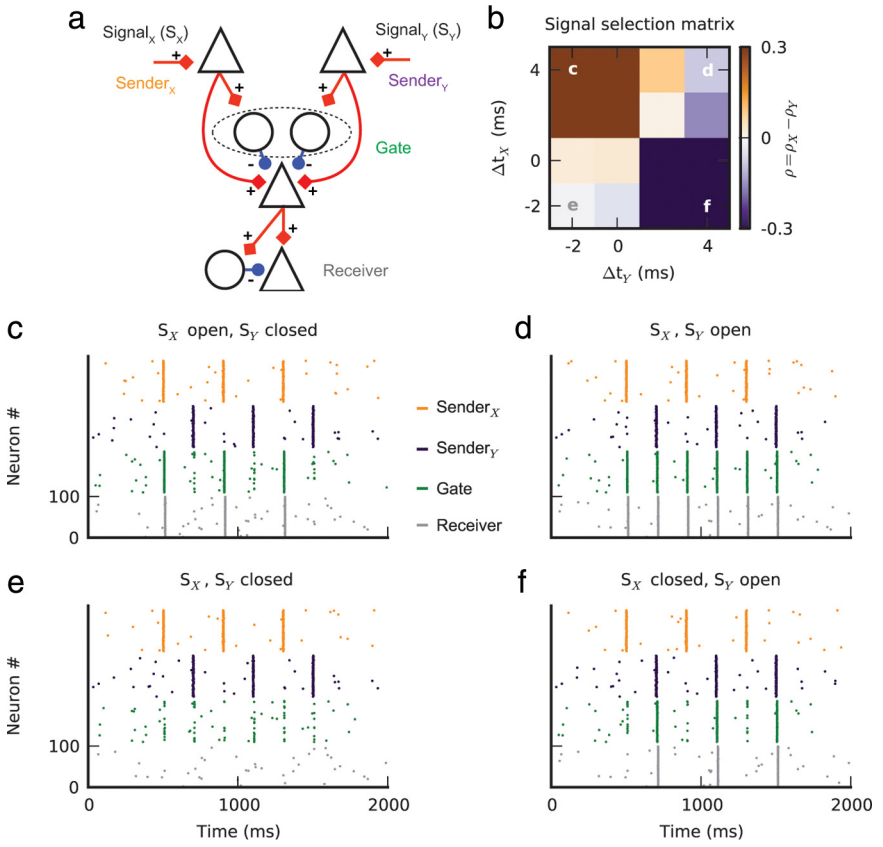

Figure 5. Selective gating of multiple signals. ***a***, Scheme of two interacting signal paths. ***b***, Selectivity of signal propagation in two interacting signal paths. The color map shows the selectivity measure $\rho = \rho_X - \rho_Y$ as a function of effective integration time Δt_X and Δt_Y. Positive values (brown color) of ρ indicate that the receiver group output is more similar to signal S_X, and negative values (blue color) reflect a high similarity with signal S_Y (i.e., selective gating of either S_X or S_Y). At $\rho \approx 0$, the receiver output is equally correlated with both signals (i.e., no selective gating). ***c–f***, Spiking activity of the sender, gate, and receiver groups during four distinct gating modes, marked with corresponding letters in ***b***. ***c***, With the gate open for S_X and closed for S_Y, the activity in the receiver group resembles the signal S_X. ***d***, With both gates open, the activity in the receiver group is a superposition of S_X and S_Y, thereby reducing ρ to zero. ***e***, Closing both gates blocks the propagation of both signals. ***f***, With the gate open for S_Y and closed for S_X, the activity in the receiver group resembles the signal S_Y.

Recently, Vogels and Abbott (2009) proposed a gating mechanism based on the detailed balance of excitation and inhibition to control the propagation of neuronal activity in feedforward networks. This gating strategy, however, is only effective in controlling the propagation of tonic components of neuronal activity modulation—transients "escape" from such gating and propagate to the next stage in any case. This is because the mechanism fails in gating transients that are fast enough to pass the small time window, induced by the lagging inhibition, i.e., in the range of some tens of milliseconds [Vogels and Abbott (2009), their Fig. 3*f*]. As a consequence, gating based only on detailed balance of excitation and inhibition is truly limited in its applicability.

The temporal gating mechanism proposed here extends this idea by changing the delay between excitation and inhibition, and thereby obtaining the means to control the temporal scale of transients that can pass or, alternatively, are blocked. The delay between excitation and inhibition affects the neurons integration time (Pouille and Scanziani, 2001; Cruikshank et al., 2007). Thereby, it can be used to block or facilitate activity propagation. Thus, this gating mechanism presents an example of exploiting a single neuron-level property to modulate network behavior—in this case, the location of the state space separatrix—such as to control the propagation of spiking activity within the network. As our study showed, the temporal gating mechanism is a simple yet powerful mechanism to control the propagation of activity transients. Thus it is complementary to the amplitude gating mechanism proposed by Vogels and Abbott (2009). When the two mechanisms are combined, they can control neural activity with different temporal profiles.

The key parameter of temporal gating is the delay between excitation and inhibition. The delayed correlated inhibition in our model was induced by disynaptic inhibition, a connectivity pattern observed ubiquitously in the CNS [hippocampus (Buzsáki, 1984), thalamocortical connection (Swadlow, 2003; Cruikshank et al., 2007), local cortical network (Kapfer et al., 2007; Silberberg and Markram, 2007), and long-range horizontal cortical network (Hirsch and Gilbert, 1991; Tucker and Katz, 2003a,b)]. In addition, neuromodulators (Kruglikov and Rudy,

2008; Antal et al., 2010) and stimulus properties (Zhang et al., 2003; Wilent and Contreras, 2005) could change the latency between excitation and inhibition in a neuron specific manner, by differentially altering neuronal excitability or synaptic strengths. Thus, the required neuronal hardware for the proposed temporal gating mechanism is quite general and not restricted to any specific brain area. Furthermore, the gating mechanism does not depend on the choice of the synapse model (current based vs conductance based) (Fig. 2 and supplemental Fig. S1, available at www.jneurosci.org as supplemental material). Moreover, the delay values used here were well within the range measured experimentally *in vivo* (Okun and Lampl, 2008). The effectivity of temporal gating in controlling the propagation of activity transients, however, precludes propagation of sustained (tonic) neural activity devoid of transients (Fig. 3*c*,*f*). Because detailed balance (or amplitude gating) is suitable in controlling the propagation of tonic activity, a combination of the two mechanisms provides a general solution to control the propagation of arbitrary neural activity patterns.

Here we presented a possibility of combining the two mechanisms and showed that, indeed, by comodulating excitation–inhibition balance and their latencies, it is possible to control the propagation of neural activity for combinations of transient and tonic components (Fig. 4). In addition, more sophisticated mechanisms might be invoked. For instance, dynamic modulation of excitation–inhibition balance can be achieved with depressing or facilitating synapses. Thus, introducing synapses with short-term plasticity (Klyachko and Stevens, 2006) into our temporal gating mechanism might be an alternative way of combining the strengths of the two gating mechanisms.

Gating of transients is most effective when the activity transients are close to the separatrix (Fig. 2*a*). If, by contrast, the activity transients are close to the stable fixed point (FP) (Fig. 2*a*) or too similar to the background fluctuations (Bkg) (Fig. 2*a*), temporal gating may not be as effective. Thus, if temporal coding is indeed used in neuronal systems, it should be possible to predict the spectrum of neural activity transients, given the properties of synapses and the connection probabilities between processing stages. In a similar vein, knowledge of the neural activity transients may provide important hints concerning the connectivity between subsequent processing stages.

Preliminary experimental evidence already points toward the involvement of a temporal gating mechanism in decision making and action selection. For instance, animals can make decisions on a millisecond timescale (Yang et al., 2008). Mechanisms based on winner-take-all dynamics would not support such rapid decision making, while temporal gating allows for signal gating and, thereby, action selection (decision making) on a millisecond timescale. Moreover, precisely timed perturbations of the cortical network can systematically alter the behavior of the animal (Seidemann et al., 1998). However, more direct tests of this type of gating mechanism have not been performed thus far. As noted above, the temporal gating method exploits circuit level behavior, which makes this mechanism particularly interesting for *in vivo* experiments. Recent developments in optogenetic techniques, which allow selective activation/inactivation of specific groups of neurons (Zhang et al., 2007), could be used to test our proposal of temporal gating *in vivo* in a highly controlled manner.

References

Aertsen A, Diesmann M, Gewaltig MO (1996) Propagation of synchronous spiking activity in feedforward neural networks. J Physiol Paris 90:243–247.

Alonso JM, Usrey WM, Reid RC (1996) Precisely correlated firing in cells of the lateral geniculate nucleus. Nature 383:815–819.

Alonso JM, Yeh CI, Stoelzel CR (2008) Visual stimuli modulate precise synchronous firing within the thalamus. Thalamus Relat Syst 4:21–34.

Anderson JS, Carandini M, Ferster D (2000) Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. J Neurophysiol 84:909–926.

Antal M, Acuna-Goycolea C, Pressler RT, Blitz DM, Regehr WG (2010) Cholinergic activation of m2 receptors leads to context-dependent modulation of feedforward inhibition in the visual thalamus. PLoS Biol 8:e1000348.

Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273:1868–1871.

Bar-Yosef O, Nelken I (2007) The effects of background noise on the neural responses to natural sounds in cat primary auditory cortex. Front Comput Neurosci 1:3.

Bienenstock E (1995) A model of neocortex. Netw Comput Neural Syst 6:179–224.

Braitenberg V, Schüz A (1991) Anatomy of the cortex: statistics and geometry. Berlin: Springer.

Brunel N (2000) Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. J Comput Neurosci 8:183–208.

Bruno RM, Sakmann B (2006) Cortex is driven by weak but synchronously active thalamocortical synapses. Science 312:1622–1627.

Buracas GT, Zador AM, DeWeese MR, Albright TD (1998) Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. Neuron 20:959–969.

Butts DA, Weng C, Jin J, Yeh CI, Lesica NA, Alonso JM, Stanley GB (2007) Temporal precision in the neural code and the timescales of natural vision. Nature 449:92–95.

Buzsáki G (1984) Feed-forward inhibition in the hippocampal formation. Prog Neurobiol 22:131–153.

Chen Y, Geisler WS, Seidemann E (2006) Optimal decoding of correlated neural population responses in the primate visual cortex. Nat Neurosci 9:1412–1420.

Cruikshank SJ, Lewis TJ, Connors BW (2007) Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex. Nat Neurosci 10:462–468.

Dan Y, Alonso JM, Usrey WM, Reid RC (1998) Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. Nat Neurosci 1:501–507.

Davison AP, Brüderle D, Eppler J, Kremkow J, Muller E, Pecevski D, Perrinet L, Yger P (2008) PyNN: A common interface for neuronal network simulators. Front Neuroinformatics 2:11.

Deger M, Cardanobile S, Helias M, Rotter S (2009) The Poisson process with dead time captures important statistical features of neural activity. BMC Neuroscience 10:P110.

Destexhe A, Contreras D, Steriade M (1998) Mechanisms underlying the synchronizing action of corticothalamic feedback through inhibition of thalamic relay cells. J Neurophysiol 79:999–1016.

Destexhe A, Rudolph M, Paré D (2003) The high-conductance state of neocortical neurons in vivo. Nat Rev Neurosci 4:739–751.

DeWeese MR, Zador AM (2006) Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. J Neurosci 26:12206–12218.

Diesmann M, Gewaltig MO, Aertsen A (1999) Stable propagation of synchronous spiking in cortical neural networks. Nature 402:529–533.

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

Eppler JM, Helias M, Muller E, Diesmann M, Gewaltig MO (2008) PyNEST: a convenient interface to the NEST simulator. Front Neuroinformatics 2:12.

Ermentrout B (1992) Complex dynamics in winner-take-all neural nets with slow inhibition. Neural Netw 5:415–431.

Gentet LJ, Avermann M, Matyas F, Staiger JF, Petersen CCH (2010) Membrane potential dynamics of gabaergic neurons in the barrel cortex of behaving mice. Neuron 65:422–435.

Gewaltig MO, Diesmann M (2007) Nest (neural simulation tool). Scholarpedia 2:1430.

Gewaltig MO, Diesmann M, Aertsen A (2001) Propagation of cortical synfire activity: survival probability in single trials and stability in the mean. Neural Netw 14:657–673.

Haider B, Krause MR, Duque A, Yu Y, Touryan J, Mazer JA, McCormick DA (2010) Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. Neuron 65:107–121.

Hellwig B (2000) A quantitative analysis of the local connectivity between pyramidal neurons in layers 2/3 of the rat visual cortex. Biol Cybern 82:111–121.

Hirsch JA, Gilbert CD (1991) Synaptic physiology of horizontal connections in the cat's visual cortex. J Neurosci 11:1800–1809.

Hromádka T, DeWeese MR, Zador AM (2008) Sparse representation of sounds in the unanesthetized auditory cortex. PLoS Biol 6:e16.

Kapfer C, Glickfeld LL, Atallah BV, Scanziani M (2007) Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. Nat Neurosci 10:743–753.

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

Klyachko VA, Stevens CF (2006) Excitatory and feed-forward inhibitory hippocampal synapses work synergistically as an adaptive filter of natural spike trains. PLoS Biol 4:e207.

Kremkow J, Perrinet LU, Masson GS, Aertsen A (2010) Functional consequences of correlated excitatory and inhibitory conductances in cortical networks. J Comput Neurosci 28:579–594.

Kruglikov I, Rudy B (2008) Perisomatic gaba release and thalamocortical integration onto neocortical excitatory cells are regulated by neuromodulators. Neuron 58:911–924.

Kuhn A, Aertsen A, Rotter S (2003) Higher-order statistics of input ensembles and the response of simple model neurons. Neural Comput 15:67–101.

Kuhn A, Aertsen A, Rotter S (2004) Neuronal integration of synaptic input in the fluctuation-driven regime. J Neurosci 24:2345–2356.

Kumar A, Rotter S, Aertsen A (2008a) Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. J Neurosci 28:5268–5280.

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

Kumar A, Rotter S, Aertsen A (2010) Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. Nat Rev Neurosci 11:615–627.

Léger JF, Stern EA, Aertsen A, Heck D (2005) Synaptic integration in rat frontal cortex shaped by network activity. J Neurophysiol 93:281–293.

Matsumura M, Chen D, Sawaguchi T, Kubota K, Fetz EE (1996) Synaptic interactions between primate precentral cortex neurons revealed by spike-triggered averaging of intracellular membrane potentials *in vivo*. J Neurosci 16:7757–7767.

Morrison A, Mehring C, Geisel T, Aertsen AD, Diesmann M (2005) Advancing the boundaries of high-connectivity network simulation with distributed computing. Neural Comput 17:1776–1801.

Muller E, Buesing L, Schemmel J, Meier K (2007) Spike-frequency adapting neural ensembles: beyond mean adaptation and renewal theories. Neural Comput 19:2958–3010.

Okun M, Lampl I (2008) Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nat Neurosci 11:535–537.

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

Seidemann E, Zohary E, Newsome WT (1998) Temporal gating of neural signals during performance of a visual discrimination task. Nature 394:72–75.

Silberberg G, Markram H (2007) Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. Neuron 53:735–746.

Stepanyants A, Hirsch JA, Martinez LM, Kisvárday ZF, Ferecskó AS, Chklovskii DB (2008) Local potential connectivity in cat primary visual cortex. Cereb Cortex 18:13–28.

Swadlow HA (2003) Fast-spike interneurons and feedforward inhibition in awake sensory neocortex. Cereb Cortex 13:25–32.

Tepper JM, Koós T, Wilson CJ (2004) Gabaergic microcircuits in the neostriatum. Trends Neurosci 27:662–669.

Tucker TR, Katz LC (2003a) Recruitment of local inhibitory networks by horizontal connections in layer 2/3 of ferret visual cortex. J Neurophysiol 89:501–512.

Tucker TR, Katz LC (2003b) Spatiotemporal patterns of excitation and inhibition evoked by the horizontal network in layer 2/3 of ferret visual cortex. J Neurophysiol 89:488–500.

Vogels TP, Abbott LF (2005) Signal propagation and logic gating in networks of integrate-and-fire neurons. J Neurosci 25:10786–10795.

Vogels TP, Abbott LF (2009) Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. Nat Neurosci 12:483–491.

Wilent WB, Contreras D (2005) Dynamics of excitation and inhibition underlying stimulus selectivity in rat somatosensory cortex. Nat Neurosci 8:1364–1370.

Yang Y, DeWeese MR, Otazu GH, Zador AM (2008) Millisecond-scale differences in neural activity in auditory cortex can drive decisions. Nat Neurosci 11:1262–1263.

Zhang F, Wang LP, Brauner M, Liewald JF, Kay K, Watzke N, Wood PG, Bamberg E, Nagel G, Gottschalk A, Deisseroth K (2007) Multimodal fast optical interrogation of neural circuitry. Nature 446:633–639.

Zhang LI, Tan AYY, Schreiner CE, Merzenich MM (2003) Topography and synaptic shaping of direction selectivity in primary auditory cortex. Nature 424:201–205.

frontiers in
SYSTEMS NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 24 October 2011
doi: 10.3389/fnsys.2011.00086

The role of inhibition in generating and controlling Parkinson's disease oscillations in the basal ganglia

Arvind Kumar*[1,2], *Stefano Cardanobile*[1,3], *Stefan Rotter*[1,3] *and Ad Aertsen*[1,2]

[1] *Bernstein Center Freiburg, University of Freiburg, Germany*
[2] *Department of Neurobiology and Biophysics, University of Freiburg, Germany*
[3] *Computational Neuroscience, University of Freiburg, Germany*

Edited by:
Raphael Pinaud, University of Oklahoma Health Sciences Center, USA

Reviewed by:
Mark D. Humphries, Ecole Normale Superieure, France
Rafal Bogacz, University of Bristol, UK
Alan Dorval, University of Utah, USA

***Correspondence:**
Ad Aertsen, Department of Neurobiology and Biophysics, University of Freiburg, D-79104 Freiburg, Germany.
e-mail: aertsen@biologie.uni-freiburg.de

Movement disorders in Parkinson's disease (PD) are commonly associated with slow oscillations and increased synchrony of neuronal activity in the basal ganglia. The neural mechanisms underlying this dynamic network dysfunction, however, are only poorly understood. Here, we show that the strength of inhibitory inputs from striatum to globus pallidus external (GPe) is a key parameter controlling oscillations in the basal ganglia. Specifically, the increase in striatal activity observed in PD is sufficient to unleash the oscillations in the basal ganglia. This finding allows us to propose a unified explanation for different phenomena: absence of oscillation in the healthy state of the basal ganglia, oscillations in dopamine-depleted state and quenching of oscillations under deep-brain-stimulation (DBS). These novel insights help us to better understand and optimize the function of DBS protocols. Furthermore, studying the model behavior under transient increase of activity of the striatal neurons projecting to the indirect pathway, we are able to account for both motor impairment in PD patients and for reduced response inhibition in DBS implanted patients.

Keywords: basal ganglia, Parkinson's disease, oscillations, deep-brain-stimulation, spiking neural networks

INTRODUCTION

Parkinson's disease (PD) is a prominent brain disorder, characterized by a host of motor and cognitive dysfunctions, caused by dopamine depletion in the basal ganglia (BG). A distinctive feature of PD is the presence of aberrant oscillations (12–30 Hz, β-band) of the local field potential in the subthalamo-pallidal (STN-GPe) network (Brown and Williams, 2005; Hammond et al., 2007; Eusebio et al., 2008). These oscillations are causally linked to deficits in movement initiation and execution, resting state tremor, and other related symptoms (Tass et al., 2010). While consensus has emerged on the existence of STN-GPe oscillations and their relevance for behavioral pathologies, the mechanisms underlying these oscillations, however, are still debated. Nevertheless, it is an empirical fact that high-frequency deep-brain-stimulation (DBS) of the STN effectively alleviates the hypokinetic disorders associated with PD (Benabid, 2003; Coffey, 2009). However, despite the success of DBS as a therapeutic protocol, also in other brain disorders (Krack et al., 2010), the mechanisms by which it quenches the β-band oscillations in PD are still only poorly understood (Benabid, 2003; McIntyre et al., 2004; Kringelbach et al., 2007; Nambu, 2008).

Here, we propose a unified explanation for the absence of oscillations in the normal state, for the emergence of oscillations in the dopamine-depleted state, and for the efficacy of DBS in quenching these oscillations in PD. Using a biologically realistic, large-scale spiking neural network model of the STN-GPe loop, we show how the strength of inhibitory input to the GPe neurons controls its oscillatory activity. Similar suggestions have been previously made using abstract firing rate based models of the basal ganglia (Gillies et al., 2002; van Albada and Robinson, 2009; van Albada et al., 2009). We would like to stress that our explanation does not rely on the prevailing assumption of potentiated synaptic coupling between the two populations (Magill et al., 2001). However, increased functional coupling between STN and GPe could be both a cause and consequence of altered activity in these networks. Here, we explore the possibility of obtaining oscillatory activity in the STN-GPe network without explicitly scaling up the synaptic coupling between the two neuron populations. Specifically, we show that increased inhibition to GPe neurons from upstream brain areas, as observed in the dopamine-depleted striatum in PD, unleashes oscillations. By contrast, increased inhibition to STN neurons restrains them. This novel insight into basal ganglia function and dysfunction also provides new understanding of DBS functioning, and it is used here to suggest innovative and more effective DBS protocols.

MATERIALS AND METHODS

LARGE-SCALE SPIKING NETWORK MODEL OF GPe AND STN

We simulated a network of 3,000 neurons, of which 1,000 excitatory neurons were assigned to the STN population and the remaining 2,000 inhibitory neurons to the GPe population. Neurons were implemented as leaky-integrate-and-fire (LIF) neurons. Passive properties of the neurons were drawn from a distribution to introduce heterogeneity into the network (cf. **Tables 1** and **2**). Neurons in the GPe population received excitatory synaptic input from the STN (connection probability 5%; Kita and Kitai, 1991) and inhibitory synaptic inputs from other GPe neurons (connection probability 2%; Kita and Kitai, 1994; Sadek et al., 2007). STN neurons received inhibitory connections from the

Table 1 | Tabular description of network model.

NEURAL POPULATIONS			
Sub-thalamic nucleus (STN)		N_{STN}	Leaky IaF neurons (excitatory)
Globus pallidus (GPe)		N_{GPe}	Leaky IaF neurons (inhibitory)
CONNECTIVITY			
Name	**Source**	**Target**	**Pattern**
GPe–GPe	GPe	GPe	Random convergent; connection probability $C_{GPe \to GPe}$, weight $J_{GPe–GPe}$, delay d_{intra}
GPe–STN	GPe	STN	Random convergent; connection probability $C_{GPe \to STN}$, weight $J_{GPe–STN}$, delay d_{inter}
STN–STN	STN	STN	Random convergent; connection probability $C_{STN \to STN}$, weight $J_{STN–STN}$, delay d_{intra}
STN–GPe	STN	GPe	Random convergent; connection probability $C_{STN \to GPe}$,weight $J_{STN–GPe}$, delay d_{inter}
NEURON MODEL			
Name	IaF neuron		
Type	Leaky integrate-and-fire		
Subthreshold dynamics	if $(t > t^* + \tau_{ref})\ \tau_m(dV/dt) = -V + I(t)/C_m$ else $V(t) = V_{reset}$		
Spiking	If $V(t) \geq \Theta$ 1. set $t^* = t$ 2. emit spike with time stamp t^* and set $V(t) = V_{reset}$		
SYNAPSE MODEL			
Type	Conductance bases synapse		
Synaptic current	$I_{syn}(t) = G_{syn}(t)(V_m(t) - E_{syn})$ $V_m(t)$ is membrane potential and E_{syn} is the reversal potential of the synapse		
Synaptic conductance dynamics	$G_{syn}(t) = J \frac{t}{\tau_{syn}} e^{-\frac{t}{\tau_{syn}}}$ for $t \geq 0$ where t is the time of spike and τ_{syn} is synaptic time constant		
BACKGROUND INPUT			
Type	**Target**	**Connection**	**Description**
Poisson generator	N_{STN}	Excitatory	Independent for each neuron, rate ν_{STN}, weight J
Poisson generator	N_{GPe}	Excitatory	Independent for each neuron, rate $\nu_{GPe,}$ weight J
STRIATUM INPUT			
Type	**Target**	**Connection**	**Description**
Poisson generator	N_{GPe}	Inhibitory	Independent for each neuron, rate $\nu_{Str,}$ weight $J_{Str \to GPe}$
DBS INPUT			
DBS type	**Target**	**Figure**	**Description**
Poisson type inhibition of STN	N_{STN}	**Figures 2E–H**	Independent inhibitory input for each STN neuron, rate ν_{DBS}
STN lesion	N_{STN}	**Figure 5A–E**	Permanent silencing of a fraction of STN neurons
Periodic blanking of axons in STN	N_{STN}	**Figure 5F–J**	Periodic blanking of the incoming axons in the STN with a frequency f_{stim}
Periodic inhibition of STN	N_{STN}	**Figure 5K–O**	Periodic inhibitory input to the STN neurons with a frequency f_{stim}
MEASUREMENTS			
Spikes from all GPe and STN neurons			

GPe (connection probability 5%; Bevan et al., 1997; Baufreton et al., 2009; cf. **Tables 1** and **2**). Published literature suggest that there is only very little recurrent excitatory connections within the STN (Hammond and Yelnik, 1983; Sato et al., 2000), thus, in our model STN received excitatory synaptic inputs from other STN neurons with a connection probability of 2%. Synaptic input was modeled as transient conductance changes, using exponential functions. The synaptic weights are provided in the **Table 2**. Note that the results of the model are not critically dependent on the exact values of synaptic strengths and other network parameters.

All neurons in the STN received external excitatory input from the cortex, modeled as uncorrelated Poisson spike trains. Similarly, all the GPe neurons received uncorrelated Poisson spike trains as input so as to achieve observed baseline firing rates in the GPe. Thalamus could be the source of some of these excitatory connections. This input and synaptic strengths were tuned to obtain realistic baseline firing rates in STN (∼15 Hz) and GPe (∼45 Hz; Bergman et al., 1994; Raz et al., 2000), observed in healthy animals. STN and GPe neurons are reported to act as pacemakers *in vitro*, i.e. these neurons spike in a near-periodic manner without any external input (Surmeier et al., 2005). However, in the healthy state, *in vivo*, the spiking of STN and GPe neurons is not periodic (Raz et al., 2000), indicating that the pacemaker behavior is overridden by cortical inputs. This aperiodic spiking behavior of STN and GPe neurons *in vivo* was well captured by modeling these neurons as integrate-and-fire neurons stimulated with Poisson type external input.

Table 2 | Simulation parameters.

Name	Value	Description
POPULATIONS		
N_{stn}	1,000	Size of STN population
N_{GPe}	2,000	Size of GPe population
CONNECTIVITY		
$C_{GPe \to GPe}$	0.05	Connection probability from GPe to GPe (Kita and Kitai, 1994; Sadek et al., 2007)
$C_{GPe \to STN}$	0.02	Connection probability from GPe to STN (Parent and Hazrati, 1995; Bevan et al., 1997; Baufreton et al., 2009)
$C_{STN \to STN}$	0.02	Connection probability from STN to STN (Hammond and Yelnik, 1983; Sato et al., 2000)
$C_{STN \to GPe}$	0.05	Connection probability from STN to GPe (Kita and Kitai, 1991)
NEURON MODEL		
g_{leak}	15 nS	Membrane leak conductance
C_m	300 pF	Membrane capacitance
τ_m	20 ms	Resting membrane time constant
Θ	-54 ± 5 mV (uniformly distributed)	Fixed firing threshold
V_0	-70 mV	Resting potential
V_{reset}	V_0	Reset potential
τ_{ref}	2 ms	Absolute refractory period
SYNAPSE MODEL		
τ_{exc}	1.0 ms	Rise time of excitatory conductance
τ_{inh}	10 ms	Rise time of inhibitory conductance
E_{exc}	0 mV	Reversal potential of excitatory synapses
E_{inh}	-80 mV	Reversal potential of inhibitory synapses
$J_{STN–STN}$	1.3 mV	At a holding potential of -70 mV
$J_{STN–GPe}$	1.3 mV	At a holding potential of -70 mV
$J_{GPe–GPe}$	-0.45 mV	At a holding potential of -55 mV
$J_{GPe–STN}$	-0.7 mV	At a holding potential of -55 mV
SYNAPTIC DELAYS		
d_{intra}	2 ms	delay for GPe to GPe (inhibitory) and STN to STN (excitatory) synapses
d_{inter}	5 ms	delay for GPe to STN (inhibitory; Fujimoto and Kita, 1993) and STN to GPe (excitatory; Kita et al., 2005) synapses
INPUT		
ν_{STN}	1500–3250 Hz	Total rate of external Poisson type excitatory input to the STN
ν_{GPe}	2000–3250 Hz	Total rate of external Poisson type excitatory input to the GPe
ν_{Str}	0–60 Hz	Firing rate of individual striatal neurons. Each GPe neuron received input from 500 striatum neurons
ν_{DBS}	0–60 Hz	Rate of Poisson type inhibition to the STN. Each STN neuron received only one Poisson input for this purpose

To obtain networks with slightly different ongoing activity states (corresponding to the healthy state), shown in **Figure 2**, we simulated the STN-GPe network with different combinations of external input (ν_{STN} and ν_{GPe}; cf. **Table 2**). Similarly, to obtain network activities with different degree of synchrony corresponding to the PD state, shown in **Figure 5**, we simulated the STN-GPe network with different combinations of ν_{STN}, ν_{GPe}, and ν_{Str} (cf. **Table 2**).

Details of the spiking network model are provided in tabular form (cf. **Tables 1** and **2**) to facilitate reproduction of the simulation results.

APERIODIC STIMULATION PROTOCOL

For the aperiodic stimulation we first chose a minimal interval Δt between the delivery of pulses. After each pulse, we chose a uniformly distributed random integer γ from the possible values $\{1, 2, \ldots, n\}$ and set the next pulse to be delivered after $\gamma \Delta t$. Both Δt and n should be chosen small enough, to have sufficiently many short intervals occur to quench the oscillations (HFS), while avoiding large intervals that might lead to entrainment (LFS). Here we chose $n = 3$ and systematically varied Δt between 5 and 15 ms.

TRANSIENT STIMULATION OF THE GPe

D2 type dopamine receptor expressing neurons in the striatum project to the GPe and are thought to represent the inhibition of the task (NoGO behavior; Frank and O'Reilly, 2006; Kravitz et al., 2010). To understand the effect of a transient NoGo task related inhibitory input to the GPe from the striatum we injected Poisson types inhibitory input to some of the GPe neurons for 20 ms. The fraction of GPe neuron stimulated and the strength of inhibition were systematically varied.

ANALYSIS OF NETWORK ACTIVITY

We used the following descriptors to quantify the network activity states:

Mean firing rate

The firing rate of individual neurons was estimated as the average spike count over the full simulation period, excluding the first 500 ms of initial network transients. The mean network firing rate was then obtained by averaging the firing rates of all neurons in the network. Here, we distinguish between the units of the rate of the periodic processes and the point processes. For periodic processes we use the standard unit "Hertz" (abbreviated as Hz), while for the point processes we use the unit "Becquerel" (abbreviated as Bq).

Synchrony index

The population synchrony in the network was estimated by the Fano Factor (FF) of the population spike count (Kumar et al., 2008)

$$\mathrm{FF[pop]} = \mathrm{Var[pop]} / \mathrm{E[pop]}. \quad (1)$$

where E[pop] and Var[pop] denote the mean and variance of the spike counts of the neural population, respectively. To obtain a reliable estimate of the population activity, we recorded the spike trains of all neurons in the network and binned their cumulative activity (bin width = 5 ms). A population of independent Poisson processes yields FF[pop] = 1, whereas positive correlations in the population activity result in an increase of Var[pop] and, hence, of FF[pop].

Oscillation index

To estimate the strength of oscillations we used the fact that oscillations introduce peaks in the power spectral density of the population activity. Therefore, we estimated the spectrum ($S(f)_{\text{pop}}$) of the population activity. Oscillations in the STN-GPe network were typically restricted to a narrow frequency band (15–25 Hz; **Figure 1C**). Thus, we defined the oscillation index as the relative power in this frequency band

$$\text{OI[pop]} = \frac{\int_{15}^{25} S(f)_{\text{pop}}\, df}{\int_{0}^{Fs/2} S(f)_{\text{pop}}\, df} \tag{2}$$

where Fs refers to the sampling frequency. To estimate the spectrum S_{pop}, we used Fs = 1 kHz. When the network was oscillating strongly, most of the power was contained in the 15–25 Hz band (**Figure 2C**) and, hence, OI was close to unity.

SIMULATION AND DATA ANALYSIS TOOLS

All network simulations were written in python[1] using PyNN[2] as an interface to the simulation environment NEST[3].

RESULTS

The first clue to understand the emergence of oscillations in the basal ganglia system comes from the architecture of the STN-GPe network (**Figure 1A**). The STN-GPe network is a sparsely connected, recurrent network, with the STN being composed of excitatory neurons and the GPe, by contrast, of inhibitory neurons. Furthermore, both neural populations are mutually coupled via inhibitory and excitatory synapses, respectively. Previous theoretical work on mutually coupled excitatory-inhibitory networks has shown that strong recurrent inhibition can induce "fast" oscillations, caused by delayed feedback from the inhibitory population (Brunel and Wang, 2003) or, alternatively, "slow" oscillations, which additionally require inhibitory input to the inhibitory neurons (Terman et al., 2002; Holgado et al., 2010). In the latter case, the oscillation frequency is determined by an interplay of effective membrane time constants and synaptic delay, and the resulting frequency is usually in the β-range.

Theoretical and computational work thus far has focused on the role of increased efficacy of the mutual coupling between STN and GPe to explain PD related oscillations in the basal ganglia. In PD, however, there is only indirect and weak evidence for potentiated projections emanating from GPe neurons (Magill et al., 2001; Shen and Johnson, 2005). Moreover, oscillations in a recurrent network could be both cause and effect of altered coupling. By contrast, recent experimental data suggest that, in fact, striatal inhibitory input to the GPe is increased (Mallet et al., 2006; Liang et al., 2008), possibly due to potentiated glutamatergic synapses in the striatum (Smith et al., 2009), emerging in the dopamine-depleted state in PD, or due to increased cortical input to the striatum (Tseng et al., 2001). In addition, in the dopamine-depleted state, the medium-spiny neurons in the striatum are more excitable (Tseng et al., 2001) and recurrent inhibition is weakened, due to an increased firing rate of cholinergic neurons (Raz et al., 1996); both findings provide additional indirect evidence for increased activity in the striatum in PD. Therefore, we studied the dynamics of the STN-GPe network under the influence of increased inhibitory input

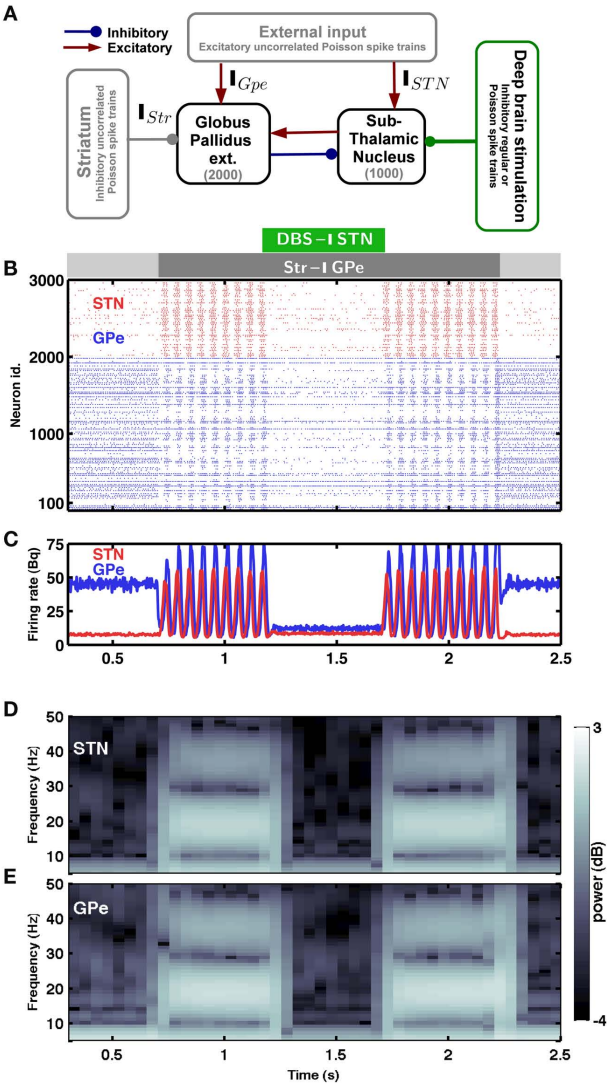

FIGURE 1 | Spiking activity in the sub-thalamic nucleus – globus pallidus network. (A) Scheme of the STN-GPe spiking neural network model. **(B)** Raster plot of the spiking activity of STN (red) and GPe (blue) neurons. The strength of striatal inhibition to the GPe is indicated by the light gray (weak) and dark gray (strong) horizontal bars. The green horizontal bar indicates Poisson type inhibitory stimulation of the STN. **(C)** Time-resolved population firing rates of the STN (red) and GPe (blue) networks. **(D,E)** Spectrogram of the STN **(D)** and GPe **(E)** population activity. In the presence of weak striatal inhibition to the GPe, the network exhibited asynchronous and aperiodic activity (Hauptmann et al., 2009). An increase of the striatal inhibition to GPe (dark gray horizontal bar) induced synchrony and β-band oscillations in the STN-GPe network. Such oscillations could be quenched by inhibiting the STN network with Poisson type input (green horizontal bar). The oscillations reappeared again as soon as the inhibitory stimulation of the STN was turned off.

[1]http://www.python.org
[2]http://neuralensemble.org/trac/PyNN
[3]http://www.nest-initiative.org

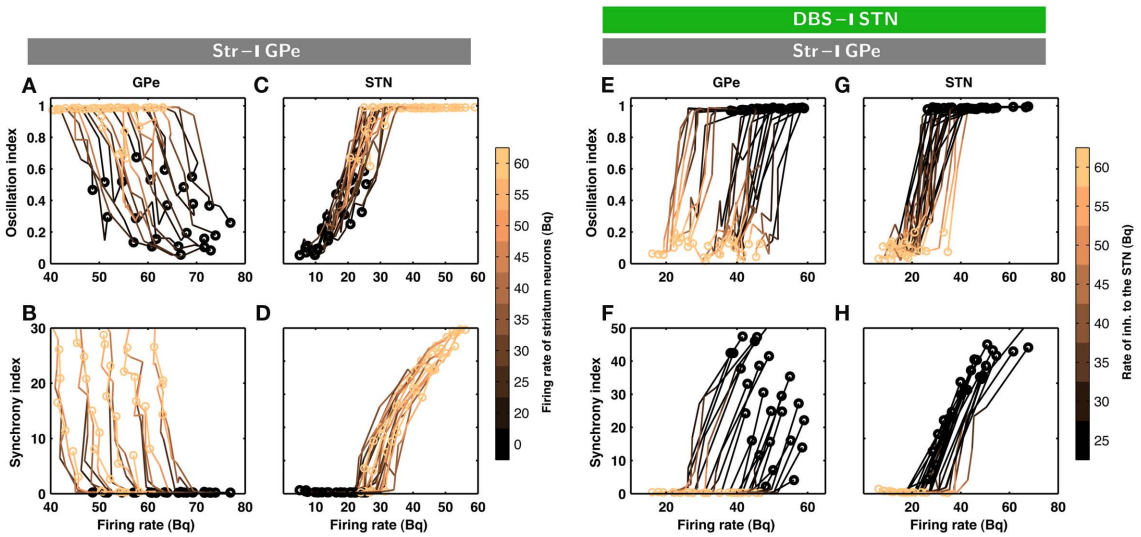

FIGURE 2 | Evolution of inhibition-induced unleashing of oscillations. Each trace corresponds to a different set of network parameters (see **Table 2**); a change from dark to bright colors represents an increase in striatal inhibition of the GPe neurons. **(A)** Oscillation index (cf. Materials and Methods) and firing rate in the GPe network for different levels of inhibitory input to the GPe. Note that even a small increase in striatal firing rate was sufficient to unleash oscillations. The dark circles refer to the state of STN-GPe network in absence of striatal inhibition. **(B)** Synchrony index (cf. Materials and Methods) and firing rate in the GPe network for different levels of inhibitory input to the GPe. **(C)** Same as in A for STN network. **(D)** Same as in B for STN network. Observe that both synchrony and oscillations increased in the STN-GPe network with increasing inhibitory input, whereas firing rate decreased in the GPe and increased in the STN. **(E–H)** Evolution of inhibition-induced quenching of oscillations. Every trace corresponds to a different set of network parameters (cf. **Table 2**); a change from dark to bright colors represents an increase in the rate of Poisson type inhibition to the STN neurons. **(E)** Oscillation index and firing rate in the GPe network for different levels of inhibitory input to the STN. The yellow circles refer to the state of STN-GPe network in absence of striatal inhibition. The dark circles refer to the state of STN-GPe network with strong striatal inhibition. **(F)** Synchrony index and firing rate in the GPe network for different levels of inhibitory input to the STN. **(G)** Same as in **(E)** for STN network. **(H)** Same as in **(F)** for STN network. Observe that firing rate, synchrony and oscillations decreased with increasing inhibitory input in both sub-networks of the STN-GPe network.

from the striatum. Specifically, we explore the possibility whether the STN-GPe network can exhibit oscillations without scaling up the synaptic coupling between the two neuron populations.

STRIATUM ACTIVITY UNLEASHES OSCILLATIONS IN THE BASAL GANGLIA NETWORK

The minimal anatomy of the neural circuitry of the basal ganglia already hints toward a crucial role for synaptic inputs in shaping the oscillatory activity in the network. Specifically, the strength of oscillatory modes should increase if inhibitory input to an inhibitory population (Str to GPe) exceeds a certain level. Because increased inhibition of the inhibitory population (GPe) would increase activity in the excitatory population (STN) which can reverberate in the coupled excitatory-inhibitory network. Similarly, an increase of excitatory input to an excitatory population (e.g. Ctx to STN) can also induce oscillations. Results from large-scale numerical simulations of a STN-GPe spiking network model confirm our predictions from these qualitative considerations (**Figures 1B–E**). Immediately after an increase in the ν_{Str} oscillations appeared in the STN-GPe network (**Figures 1B–E**; region marked by gray horizontal bar). We tested the effect of the inhibition of GPe neurons by increased striatal activity (ν_{Str}) on STN-GPe networks with different average firing rates and oscillations. Indeed a progressive increase in ν_{Str} strengthened the amplitude of oscillations in both STN and GPe neurons for all the network studied here (**Figures 2A–D**). Consistent with experimental data, in all the networks, increase on the oscillations was associated with an increase in firing rate of STN neurons and a decrease in firing rate of GPe neurons, respectively (Bergman et al., 1994; Raz et al., 2000). Finally, our suggestion that increased striatal firing rates of are sufficient to unleash oscillations and, hence, Parkinsonian symptoms (oscillations in the case of our model), in the basal ganglia network is in full accordance with recent experimental findings (Kravitz et al., 2010), which showed that selective increase in the firing rate of D2 type striatum neurons, which specifically project to the GPe, can induce PD symptoms in mice.

MECHANISM OF OSCILLATIONS

When increased striatal activity inhibits GPe neurons the STN neurons are released from inhibition, resulting in a slow increase in the activity in the STN network. Increased excitation originating from STN causes an increase of activity in GPe after a certain delay, the duration of which is determined by the combination of axonal transmission, synaptic transduction and dendritic integration. Increase GPe activity is eventually able to stop STN firing, thereby resetting the system to the initial state. Thus, in this mechanism STN activity leads the GPe activity.

This mechanism implies that oscillations onset is accompanied with a decrease in the activity of the GPe neurons. Alternatively, it is possible that strong cortical excitatory input to the STN could increase firing rates of the STN neurons, initiating the process of oscillations as described above.

In both the scenarios of increased striatal to the GPe or cortical input to the basal ganglia network, oscillations appear because activity in the excitatory population (STN) builds up on a slow time-scale governed by the membrane time constants of the STN neurons. In summary, oscillations in the STN-GPe network are, in fact, alternating reverberations of increased excitation and inhibition, maintained by an effective increase in excitatory input to STN neurons. In this sense, they are analogous to the "slow" oscillations observed in networks of integrate-and-fire neurons (Brunel and Wang, 2003).

Because STN firing rates are increased due to high striatal output, one could speculate that this may lead to an increase in the effective coupling from GPe to STN, and thus cause of oscillations. However, according to Brunel's theory (Brunel, 2000), this increase could result in fast as well as slow oscillations, and also in asynchronous firing. Moreover, the strength of GPe to STN coupling in the absence of STN to GPe feedback is not increased as STN firing rate increases in the high striatal output regime (**Figure 3**). In fact, the coupling strength decreases for high STN rate, thus it cannot be the mechanism underlying the oscillations.

SPECTRUM OF CORRELATIONS WITHIN STN AND GPe NETWORKS

To further characterize the activity of the STN-GPe network in non-oscillatory and oscillatory states we calculated the oscillation index (OI_{cc}) of cross-correlograms of the STN and GPe neuron pairs in non-oscillatory ($OI = 0.15$) and oscillatory ($OI = 0.97$) states. In both states, GPe neurons showed a wide distribution of pairwise correlations ($PwC_{GPe} = 0.27 \pm 0.17$ and 0.39 ± 0.19, respectively; **Figure 4A**). By comparison, STN neurons exhibited a narrow distribution of pairwise correlation coefficients, with small correlation in the non-oscillatory state ($PwC_{STN} = 0.12 \pm 0.08$) and high pairwise correlation in the oscillatory state ($PwC_{STN} = 0.51 \pm 0.08$; **Figure 4B**). In the non-oscillatory state, the cross-correlograms were largely non-oscillatory in both STN and GPe populations (**Figures 4C,D** blue traces). In the oscillatory state, the cross-correlograms were also oscillatory (**Figures 4C,D** orange traces). Interestingly, in the GPe population ∼15% cross-correlograms showed $OI_{cc} \leq 0.5$.

In animal models of PD, a large fraction of STN neuron pairs showed oscillatory cross-correlograms (Levy et al., 2002), whereas only a relatively small number of GPe neuron pairs (16.7%) exhibited oscillatory cross-correlograms (Heimer et al., 2002). At first sight, our model is inconsistent with these experimental data, as we predict a larger number of oscillatory cross-correlograms in the GPe network. However, a direct comparison of our model with the experimental data is not possible because the oscillation index of the GPe activity was not made available. Evidently, the number of oscillatory cross-correlograms depends on the strength of oscillations in network activity. Furthermore, the pertinent publications do not provide quantitative numbers on the strength of the cross-correlation and its oscillation index, making a direct comparison difficult.

QUENCHING OF OSCILLATIONS IN THE BASAL GANGLIA NETWORK

According to the network model proposed here, there are different ways to effectively interfere with the oscillatory modes. In particular, it would be sufficient to reduce the excitatory input to STN

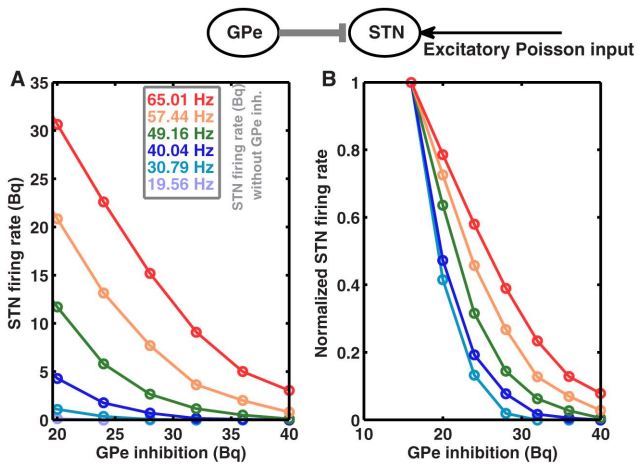

FIGURE 3 | GPe to STN coupling is not increased in high striatal output regime. Top Schematic of a single GPe and STN neuron coupled by a single inhibitory synapse. **(A)** Firing rate of an STN neuron as a function GPe inhibition, for different baseline firing rates in the STN. STN baseline firing rates were tuned to different by values by increasing the rate of input Poisson spike trains. Similarly, GPe spike trains were modeled as homogeneous Poisson process. **(B)** Same as in **(A)**, but normalized to the maximum firing rate for each STN baseline firing rates, to show the relative decrease in the STN activity as a function of increasing GPe inhibition.

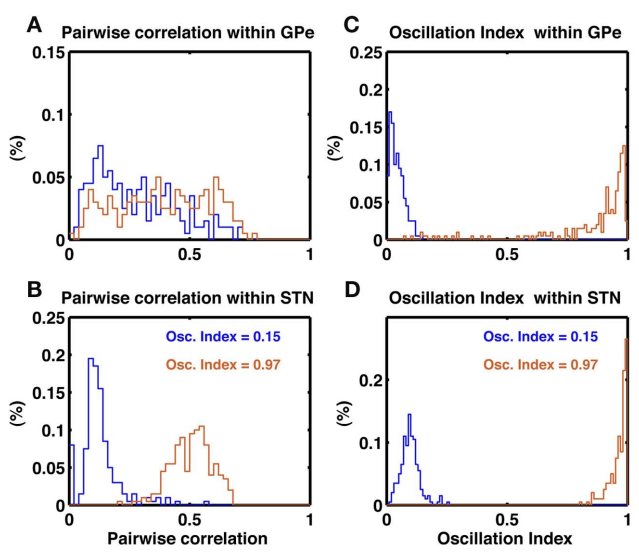

FIGURE 4 | Pairwise correlations in the STN and GPe populations. (A) Histogram of pairwise correlations estimated from 100 unique pairs of GPe neuron in a non-oscillating (blue trace) and a oscillating state (brown trace). **(B)** Same as in **(A)**, but for 1000 unique pairs of STN neurons. Oscillation induce stronger synchrony (correlations) in the STN than GPe. **(C)** Oscillation index of the cross-correlogram of GPe neuron pairs. In a non-oscillatory state the OI_{corr} is small (blue trace). In an oscillatory state despite of global oscillations some neuron pairs do not show oscillations in their cross-correlogram (brown trace). **(D)** Same as in **(C)**. Unlike GPe network, STN more neuron pairs show oscillations in their cross-correlogram (brown trace).

neurons to efficiently quench the oscillations, because increased activity of the STN neurons is the main cause of oscillations. To test this general mechanism, we provided additional inhibitory synaptic inputs to all STN neurons in the model. Indeed, even weak uncorrelated inhibitory input to STN neurons was sufficient to quench the oscillations (**Figures 1B–E**; region covered by green horizontal bar). Increasing the rate of additional inhibitory input to STN neurons progressively attenuated the oscillations in both STN and GPe neurons (**Figures 2E–H**). This mechanism quenched in the oscillation in all the different networks studied here, irrespective of the strength of oscillations. Note, however, that, although this type of interference was effective in reducing the network oscillations, it failed to restore the healthy state of firing in GPe neurons (**Figures 1B–C** and **E–H2**). We also tested if it is necessary to provide external inhibitory inputs to all the STN neurons. For a different STN-GPe network we systematically increased the faction of STN neurons that received extra inhibition. For some networks it was sufficient to inhibit ∼50% of STN neurons (**Figure 6A**) to obtain OI ≤ 0.4. On average, at a stimulation frequency of 50 Bq, ∼75% neurons were needed to be inhibited to quench the oscillations (OI ≤ 0.3; **Figure 6A**, black trace). The fraction of STN neurons to be stimulated can be decreased for higher rate of Poisson type stimulation (**Figures 2D,E**). Previously, gamma distributed stimulation pulses with a mean rate of 130 Hz have been used to stimulate the STN. Such stimulation, however, failed to alleviate the symptoms of bradykinesia (Dorval et al., 2010). It is not clear from the published literature whether aperiodic DBS reduced or quenched the oscillations or not (Dorval et al., 2010). Our model is not detailed enough to reproduce the specific firing patterns observed during bradykinesia. Here, we focus on the β oscillations which according to our model can be quenched by Poisson type stimulation at a stimulation frequency as low as 50 Bq.

EFFECT OF STN LESIONS ON OSCILLATIONS

One of the earliest surgical treatments of PD motor symptoms was to lesion the STN (Dorval et al., 2010). Surprisingly, it was recently reported that selective silencing of a fraction of STN neurons by optogenetic methods was not sufficient in suppressing oscillatory activity and motor dysfunction in a mouse model of PD (Gradinaru et al., 2009). Interestingly, Gradinaru et al. (2009) found that eNpHR virus was effective in about 95% STN neurons and optical stimulation reduced the firing rate of the STN neurons by ∼80%. If we assumes the same decrease for all of the 95% eNpHR infected neurons, the overall firing rate in the STN will be ∼25% of the initial one. This would suggest that silencing of ∼75% of STN neural activity is not sufficient to quench the oscillations. To find out how many of STN neurons need to be stimulated to quench oscillations (OI ≤ 0.3) we investigated how silencing a fraction of randomly chosen neurons in the STN affects oscillations in our model.

By systematically increasing the fraction of silenced neurons we found that $> 40\%$ STN neurons must be completely silenced before the amplitude of oscillations is reduced to a sufficiently small value (OI ≤ 0.5; **Figures 5A–D**). Silencing a large enough fraction of STN neurons reduced the excitatory input to the GPe to the extent that the GPe firing pattern remains largely dominated by the striatal (non-oscillatory) input and, hence, remained non-oscillatory. By contrast, when only a small fraction of STN neurons was silenced, GPe neurons still received enough excitatory input to maintain the oscillations. It is not clear what level of oscillation index is compatible with the absence of Parkinsonian symptoms. If a very low oscillation index is needed, then silencing of 75% of the neurons would be not enough to suppress Parkinsonian symptoms, thus explaining the experimental findings (Gradinaru et al., 2009).

MINIMAL NEURONAL MECHANISMS OF DBS FUNCTIONING

In the last two decades, DBS has developed into an effective clinical approach to alleviate PD symptoms in certain patients. Yet, in spite of the phenomenal success of DBS in PD and several other neurological disorders (Benabid, 2003; Kringelbach et al., 2007), the biophysical and neuronal mechanisms underlying DBS functioning are still only poorly understood (Benabid, 2003; McIntyre et al., 2004; Kringelbach et al., 2007; Nambu, 2008). Also, it is well established that, while periodic high-frequency stimulation (HFS) of the STN is efficient for the treatment of PD symptoms, periodic low-frequency stimulation (LFS) may even aggravate motor impairment (Eusebio et al., 2008).

PERIODIC DBS

To understand the neural mechanism underlying the efficacy or inefficacy of periodic stimulation it is worth to consider it as a abstract dynamical system (Bender and Orszag, 1999) with negative feedback, exhibiting periodic activity. In such systems, a periodic input with frequency approximately matched to the resonant frequency of the system can entrain it in a new oscillatory mode, possibly increasing oscillations amplitude. If, in contrast, the input frequency is much higher than the resonant frequency, the system will not be entrained and will behave as subject to the temporal average of the periodic input.

To test this idea, we implemented two different methods of periodic stimulation in our large-scale spiking network simulations of the STN-GPe network.

Periodic blanking of the excitatory inputs to STN

Here we periodically switched the excitatory afferents to STN neurons on and off. This is equivalent to a scenario of repeated electrical stimulation of axons into the STN, resulting in a cessation of spiking in these axons due to adaptation effects. The frequency of such periodic blanking was systematically varied to study its efficacy in quenching the oscillations (**Figures 5F–J**). We found that this stimulation protocol was only effective in measurably quenching the oscillations at stimulus frequencies larger than 100 Hz (**Figure 5J**), as is evident both in the raster diagram (**Figure 5I**) and in the broadband spectrum of the STN-GPe network activity (**Figure 5I**). Interestingly, low-frequency blanking of the STN input generated harmonics of the stimulation frequency in the STN-GPe network activity, analogous to recent observations by Deger et al. (2010). This emergence of harmonics at low-frequency stimulation could explain why low-frequency DBS in many cases impairs movements in PD patients (Eusebio et al., 2008). Finally, it was possible to quench oscillations even when input to up to 60% of the STN neurons was silenced periodically at high-frequency (**Figure 6B**).

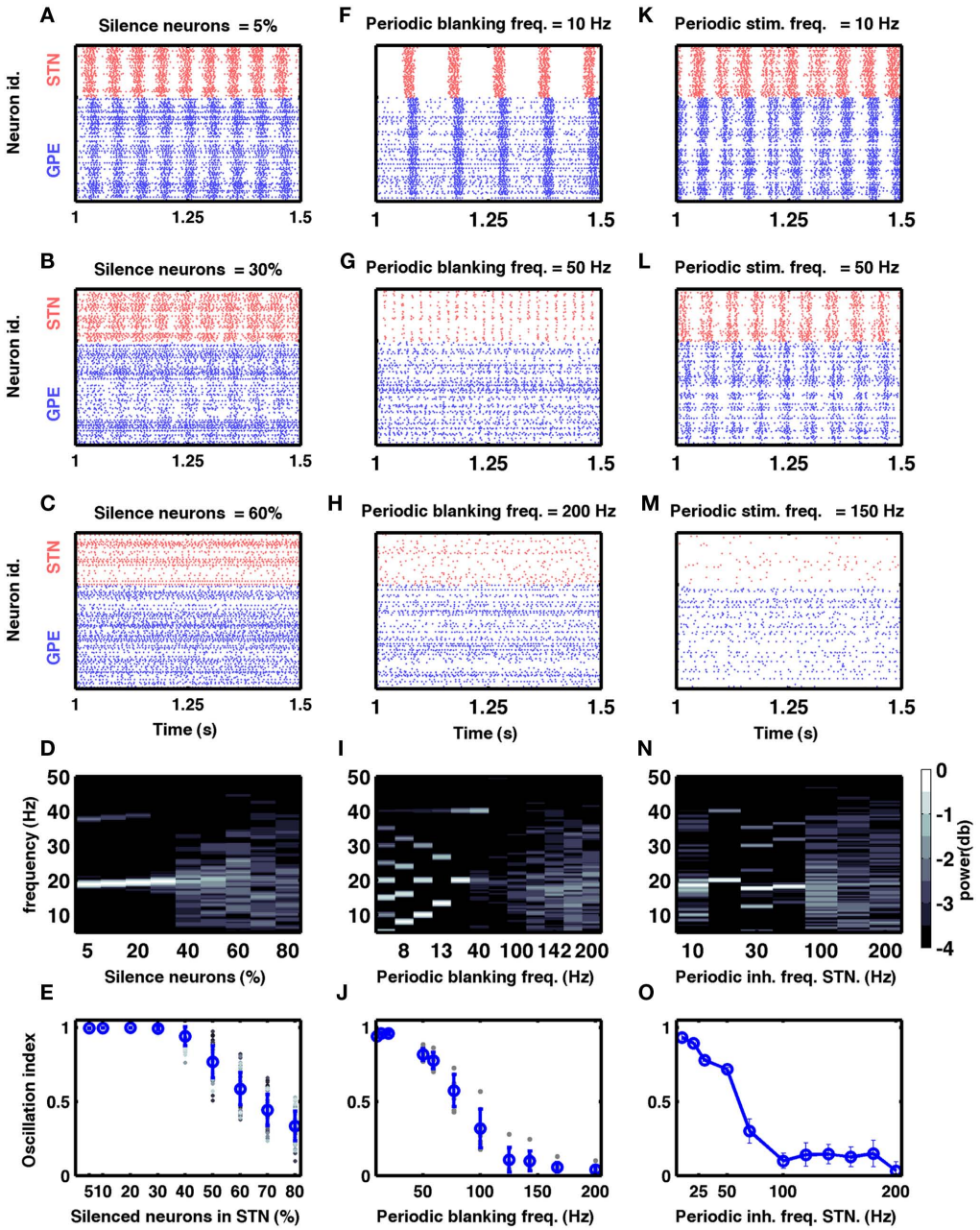

FIGURE 5 | Models of DBS. (A–E) Effect of silencing a fraction of STN neurons on oscillations in the STN-GPe network. **(A–C)** Three examples of spiking activity when a fraction of STN neurons was silenced, as indicated above the respective plots. **(D)** Spectrum of the STN population activity as a function of the fraction of silenced STN neurons. **(E)** Oscillation index measured in the STN as a function of the silenced fraction of STN neurons. Here and in **(J)**, blue dots and vertical bars represent mean and SD, respectively, whereas different gray dots refer to individual simulations with different network parameters (see **Table 2**). Note that only with more than 60% of all neurons silenced, the oscillation index was reduced. However, even with 70% silenced neurons, some networks still exhibited oscillations. **(F–J)** Effect of periodic blanking of excitatory afferents to the STN neurons on the oscillations in the STN-GPe network. **(F–H)** Three examples of spiking activity when afferents of the STN neurons were blanked at frequencies indicated above the respective plots. **(I)** Spectrum of the STN population activity as a function of the frequency at which the excitatory inputs to STN neurons were blanked. **(J)** Oscillation index in the STN, as a function of the frequency of periodic blanking of the excitatory inputs to STN neurons. Observe that only blanking frequencies above 100 Hz were effective in reducing the network oscillations. **(K–O)** Effect of periodic inhibition of excitatory afferents to the STN neurons on the oscillations in the STN-GPe network. **(K–M)** Three examples of spiking activity of STN neurons when periodic inhibitory spiking activity was injected into the STN neurons at frequencies indicated above the respective plots. **(N)** Spectrum of the STN population activity as a function of the frequency of periodic inhibitory input to the STN neurons. **(O)** Oscillation index in the STN, as a function of the frequency of periodic inhibitory stimulation of STN neurons.

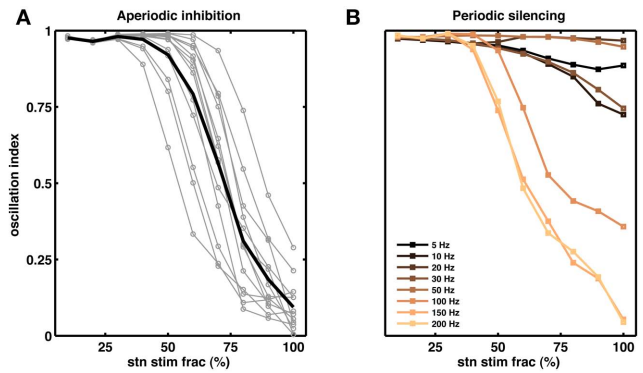

FIGURE 6 | Partial stimulation of STN neurons can quench oscillations. **(A)** Effect of extra inhibitory input to a fraction of STN neurons. Each gray line corresponds to a network with unique set of parameters. The black trace is the average of the gray traces. For all networks we injected extra Poisson type inhibitory inputs (50 Hz). Out simulations suggests that for some networks inhibition of even 50% of the STN neurons is sufficient to reduce the oscillations (OI $\leq$ 0.3), on average however, $\geq$70% STN neurons should be inhibited to quench the oscillations (OI $\leq$ 0.3). **(B)** Effect of periodic silencing (**Figures 5F–J**) of excitatory input to a fraction of STN neurons. At high frequencies ($\geq$100 Hz) even partial stimulation of STN neurons ($\geq$50%) is enough to reduce the oscillations in the STN-GPe loop.

Periodic inhibitory synaptic input to STN neurons

In this hypothetical realization of DBS we provided additional inhibitory synaptic input to STN neurons in a periodic fashion. We systematically varied the frequency of these inputs to study its efficacy in quenching the oscillations (**Figures 5K–O**). We found that also with this protocol the efficacy of stimulation in quenching the oscillations increased with stimulus frequency (**Figure 5O**). Unlike the previous protocol, however, it did not generate any harmonics during stimulation at low frequencies (**Figure 5N**).

Our large-scale network simulations further revealed that the two protocols differed in their fine-scale spectral composition: while periodic blanking induced harmonics for low-frequency stimulation, periodic inhibitory inputs increased the power in a broader frequency range (compare **Figures 5I,N**). Thus, measurement of LFPs during stimulation can be used to reveal how the current periodic stimulation protocols work.

QUENCHING OSCILLATIONS WITH APERIODIC STIMULATION

Prolonged high-frequency periodic stimulation can induce synaptic plasticity (Malenka and Bear, 2004), which could possibly impair the efficacy of DBS in PD patients. However, as described above, under sustained, periodic low-frequency stimulation the system becomes entrained to the stimulation frequency. Therefore, we tested whether aperiodic stimulation can also quench oscillations.

As an example, we delivered aperiodic stimulation in the form of 10 ms wide blanking pulses at random inter-pulse intervals, with an upper bound on the maximum interval. We chose this particular implementation because it amounts to only a minor modification of existing periodic protocols and could be easily incorporated into currently available DBS systems (cf. Materials and Methods).

Interestingly, this type of aperiodic stimulation was very effective in quenching the oscillations (**Figure 7A**). In fact, such aperiodic stimulation was more effective than periodic stimulation at the same mean pulse rate. For instance, periodic blanking at 100 Hz resulted in an average oscillation index of $\approx$0.3, whereas aperiodic blanking at the same mean rate reduced the oscillation index to $\approx$0.1 (compare **Figures 5J** and **7C**).

Such aperiodic protocol would, in fact, have another advantage beyond avoiding both entrainment (LFS) and undesired plasticity (HFS). Periodic stimulation at 100 Hz or higher is quite energy consuming. As a result, the battery life of state-of-the-art DBS systems is limited to a few years only (Kringelbach et al., 2007). Thus, in addition to being more effective, our proposed aperiodic stimulation could prolong the battery life time and, hence, increase recharging and maintenance intervals – clearly an advantage for the patient carrying the battery.

EFFECT OF EXCITABILITY OF STN NEURONS ON OSCILLATIONS

Our model suggests that oscillations are unleashed because increased inhibitory input to GPe neurons releases STN neurons from inhibition and, in turn, increased activity in the STN initiates oscillations. Consequently, oscillations could, in principle, also be quenched by decreasing the excitability of STN neurons. For the simple neuron model used here, neural excitability could conveniently be manipulated by altering the spike threshold, but biologically more realistic neuron models provide other options to modify the excitability. Indeed, we found that it is possible to effectively quench the oscillations by reducing neuronal excitability in STN (**Figure 8**).

ROLE OF TRANSIENT INHIBITION OF GPE NETWORK

Our model suggests that increased striatal output can induce oscillation in the STN-GPe network. In normal state, striatum neurons that project to the GPe are thought to be active during inhibition tasks in the classical Go-NoGo task (Frank and O'Reilly, 2006). This means that GPe may be receiving increased inhibitory input from the striatum even in a healthy state and such transient increase in the GPe inhibition could initiate transient oscillations in the STN-GPe network. To test for this possibility, we recorded the output of the GPe network when it was stimulated with a short (20 ms) Poisson type inhibitory input. We varied the number of GPe neuron stimulated and the stimulation firing rate. When baseline firing rates were high ($\sim$40 Bq), as is observed in healthy animals, stimulation of 10% GPe neurons was sufficient to induce transient oscillations (**Figures 9A–C**). In general, oscillation strength of the GPe response increased with the strength of input and fraction of neuron stimulated. Furthermore, when baseline firing rate were low ($\sim$17 Bq), as is observed in dopamine-depleted animals, inhibition-induced transient oscillations in GPe required stimulation of a large number of neurons (**Figure 9D**). In summary, we suggest:

1. In healthy patients, NoGo behavior is mediated by transient oscillations in basal ganglia output.
2. As PD advances, spontaneous oscillations emerge in the basal ganglia network. This persistent inhibition signal from GPe impairs movement initiation.

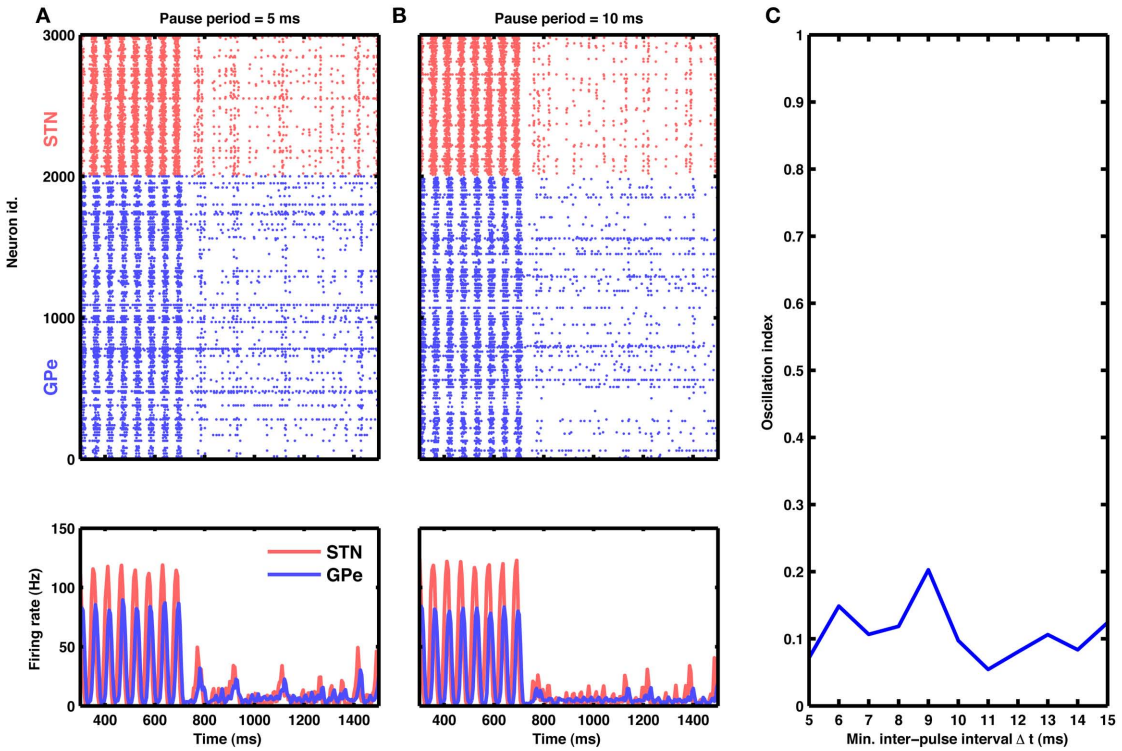

FIGURE 7 | Aperiodic stimulation protocol. (A) Raster of spiking activity (top) and population activity (bottom) in the STN-GPe network for STN (red) and GPe (blue) neurons. At time 750 ms, input to STN neurons was blanked aperiodically for 10 ms. Successive inter-pulse intervals were chosen randomly from a finite set of possible values: 5, 10, and 15 ms (cf. Materials and Methods). **(B)** Same as in **(A)**, but with inter-pulse intervals double the size (possible values: 10, 20, and 30 ms). **(C)** Oscillation index as a function of minimal inter-pulse interval. Observe that this aperiodic stimulation is at least as effective in quenching oscillations as a comparable periodic HFS.

3. DBS treatment eliminates oscillations. The persistent oscillatory signal from GPe is quenched and movement initiation could be restored.
4. However, DBS treatment is not able to restore proper firing rates in GPe. Therefore, inhibition-induced transient oscillations are more difficult to initiate and NoGo behavior is impaired.

Taken together, the above arguments can provide a putative explanation for the motor initiation impairment in Parkinson and reduced response inhibition in DBS implanted patients, reflected as enhanced stop-signal reaction time in NoGo tasks (van den Wildenberg et al., 2006).

DISCUSSION

NEURAL MECHANISMS UNDERLYING EMERGENCE AND QUENCHING OF OSCILLATIONS IN PD

We demonstrated theoretically that inhibitory inputs from upstream brain areas (e.g. striatum) to the STN-GPe network are sufficient to both unleash and quench oscillations in the basal ganglia.

Our model predicts that in the STN-GPe network both, an increase in inhibitory input to the GPe and an increase in excitatory input to the STN can generate oscillations in the basal ganglia network. An increase in inhibitory input to GPe can occur due to increased firing rates in the D2-expressing medium-spiny neurons in the striatum. Indeed, several experimental studies have reported either direct (Liang et al., 2008; Mallet et al., 2008) or indirect (Raz et al., 1996; Tseng et al., 2001) evidence for increase in striatal firing rates in animal models of PD. This increase in striatal firing rates can be caused by a potentiation of glutamatergic synapses in the striatum (Smith et al., 2009). Tseng et al. (2001) hypothesized that an altered functional state of striatal projection neurons could facilitate the transmission of cortical rhythms to the basal ganglia, resulting in the motor symptoms of PD. Here we have shown that increased activity in the striatum can induce oscillations in the basal ganglia. However, it is not necessary that striatal activity is increased because of local changes in the striatal network. An increased input to striatum from cortex could also increase firing rates in the striatum and thereby initiate oscillations in the STN-GPe network according to the "volleyball effect." Thus, our model is consistent with this hypothesis and, in fact, provides a computational explanation of it. Alternatively, increased activity in the cortex can also increase striatal firing rates. Moreover, increased cortical activity would imply that STN neurons also receive strong input, which, according to our model, could be another reason for the emergence of oscillations in the basal ganglia.

Taken together, our theory provides a unified explanation for the absence of oscillations in the normal state, for the origin of

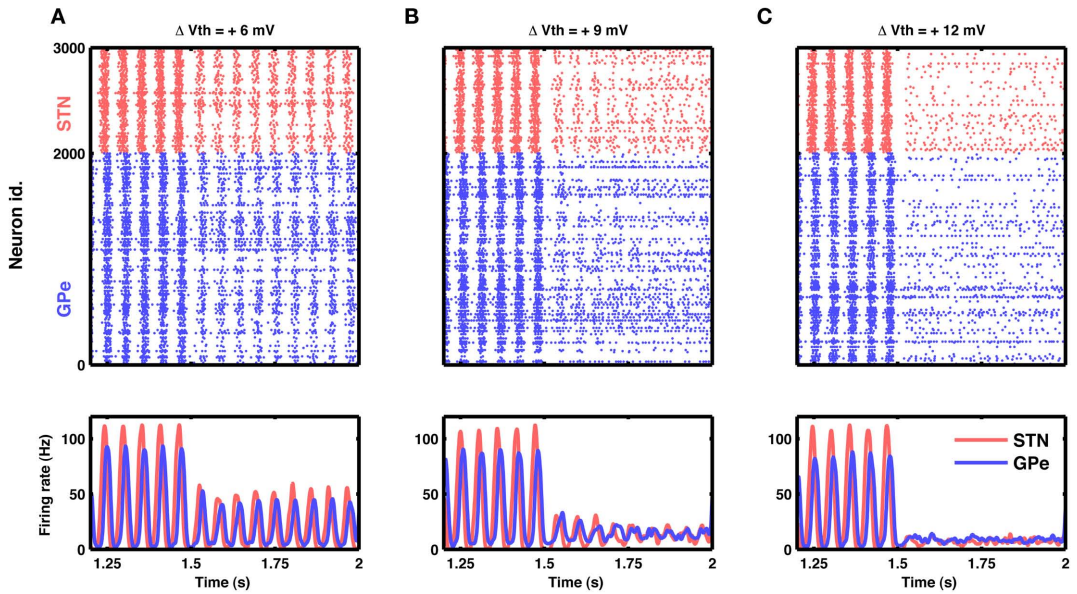

FIGURE 8 | Reduction in excitability of STN neurons can quench oscillations. (A) Raster of spiking activity (top) and population activity (bottom) in STN (red) and GPe (blue). At time 1.5 s, the excitability of STN neurons was reduced by increasing their spike threshold by 6 mV. Such small reduction in excitability attenuated STN-GPe oscillations only by some 50%. **(B)** Same as in **(A)**, but with a spike threshold increase of 9 mV. Here the attenuation of oscillations was considerable, down to some 20%. **(C)** Same as in **(A)**, but with a spike threshold increase of 12 mV. Now the oscillations were nearly abolished, while both STN and GPe remained active at a low firing rate.

oscillations in the dopamine-depleted state, and for the efficacy of DBS in PD.

Most previous models of oscillations in the basal ganglia have explicitly focused on the mutual interactions between the STN and GPe networks (Terman et al., 2002; Humphries et al., 2006; Holgado et al., 2010). In these models, either a change in strength (Terman et al., 2002; Holgado et al., 2010) or time constant and delay (Holgado et al., 2010) of synaptic coupling between the STN and GPe caused the oscillations. Striatum input to the GPe was assumed to have increased several times compared to the non-oscillatory states, however, the authors did not elaborate on this issue (Terman et al., 2002; Humphries et al., 2006; Holgado et al., 2010). Another model (Leblois et al., 2006) ignored the role of the GPe and suggested that the change in the overall strengths of the direct pathway (involving striatum and globus pallidus internal, GPi) and the hyper-direct pathway (involving the STN and GPi) is the key factor in inducing oscillations in the basal ganglia. In that model of oscillations, the indirect pathway (involving striatum, GPe, and GPi) played no role in the emergence of oscillations.

Although all these models were able to produce oscillations in the basal ganglia, the evidence in favor of potentiated synapses with unaltered effective connectivity is rather weak. In fact, in a coupled system with plastic synapses, an increase in effective connectivity between the STN and GPe networks could be both cause and effect of oscillatory dynamics. Moreover, these models imply that the synaptic couplings between the STN and the GPe are carefully tuned to avoid oscillations. Very importantly, none of these models predicted or explained the significance of increased firing rates of striatal neurons projecting to the GPe, as have been experimentally observed in dopamine-depleted striatum. Finally, according to previous models (Terman et al., 2002; Humphries et al., 2006), the putative function of DBS is to reduce the synaptic coupling between STN and GPe, for which, again, there is currently no experimental evidence.

Alternatively, the inhibition-induced increase in oscillations in the basal ganglia has been studied using abstract firing rate based models by Gillies et al. (2002) and, more recently, van Albada et al. (van Albada and Robinson, 2009; van Albada et al., 2009). These authors have made specific suggestions about the role inhibition may play in PD related oscillation in the STN-GPe network. Here, using a more realistic model of the STN-GPe network we not only confirm previous suggestions from abstract models (Gillies et al., 2002; van Albada and Robinson, 2009; van Albada et al., 2009) but also extend these insights to understand mechanisms underlying the functioning of DBS.

Here, given the mechanism of oscillations, STN inhibition emerges as a most natural mechanism to quench the oscillations. Indeed, experimental data suggests that DBS may inhibit STN neurons (Benabid, 2003; Kringelbach et al., 2007) and, thus, also reduce firing rates in the GPe. However, at the same time some experiments indicate that during DBS, firing rate in GPe neurons remains unaffected (Moran et al., 2011) or increases (Hashimoto et al., 2003). According to our model, oscillations in the STN-GPe network are unleashed when firing rates in the GPe neurons are reduced (by increased activity of D2 neurons in the striatum). Therefore, quenching of oscillations and alleviation of PD symptoms by an increase in the firing rates of GPe neurons by DBS is consistent without model.

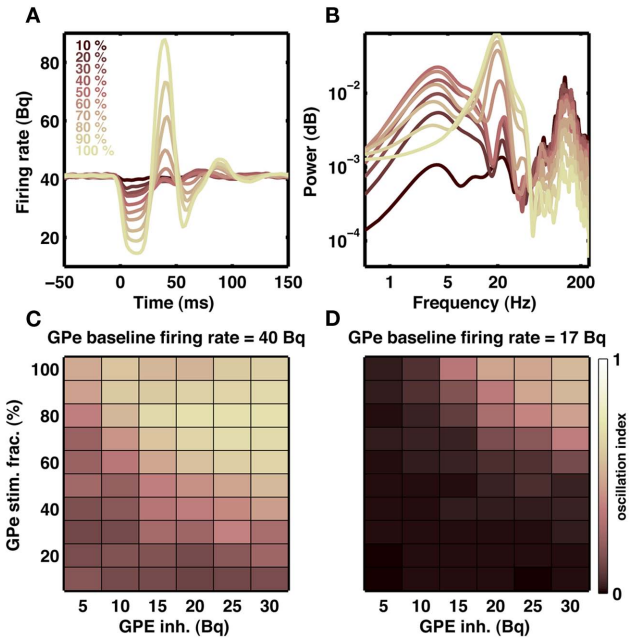

FIGURE 9 | Transient oscillation in STN-GPe loop. (A) Example of a transient (20 ms) inhibitory input to a fraction of GPe neurons indicated by different colors. Each trace is an average of 20 stimulations. Transient inhibitory input induces a damped oscillation in the population activity of the GPe network. **(B)** Spectra of GPe activity shown in **(A)**. Transient inhibition of even 10% neurons generates a weak but visible oscillation at 20 Hz. **(C)** Effect of strength of inhibitory input to the GPe (*x*-axis) and fraction of GPe neurons inhibited (*y*-axis) on the oscillation strength. **(D)** Same as in **(C)**, but with high baseline firing rate in the GPe network. Only when the baseline activity in the GPe is high transient inhibitory inputs to the GPe can initiate short lasting oscillations.

PUTATIVE FUNCTIONAL SIGNIFICANCE OF TRANSIENT OSCILLATIONS

Previous experiments and computational models have suggested that the STN-GPe network can impose a global modulatory signal in the basal ganglia to influence action facilitation and action suppression (Frank and O'Reilly, 2006; Frank et al., 2007). Our model suggests that strong activation of the indirect pathway of the basal ganglia is the primary cause for the oscillations reported in PD. It is possible, though, that transient oscillations are, in fact, instrumental for functional computations in the basal ganglia. In the STN-GPe network, transient oscillations can be induced both by phasic inhibition of GPe neurons and/or phasic excitation of STN neurons (**Figure 9**). Such phasic inputs may recruit the indirect and/or hyper-direct pathway, causing transient synchronous activity and phase-locking in the STN-GPe network. Further theoretical work is needed to determine whether and how such transient oscillatory modes can play a role as a global modulatory signal in action selection or action suppression, as suggested by Frank et al. (Frank and O'Reilly, 2006; Frank et al., 2007).

SUGGESTIONS FOR NOVEL, MORE EFFECTIVE DBS PROTOCOLS

Beyond its general relevance for understanding basal ganglia function and dysfunction, our model also offers a computational framework for developing and testing innovative DBS protocols (Hauptmann et al., 2009). In this framework, we made specific suggestions for new strategies of therapeutic intervention in PD and other diseases exhibiting abnormal oscillatory activity in the STN-GPe network: (1) STN-GPe oscillations can be quenched more effectively by reducing the activity of D2-expressing striatal neurons projecting to the GPe (**Figures 2A–D**). (2) Aperiodic stimulation of the STN is more effective in quenching the oscillations than periodic stimulation at the same mean rate (**Figures 2E–H** and **7**). (3) Because the oscillations are likely caused by disinhibition of STN neurons, reducing the excitability of STN neurons may also be effective in quenching the oscillations (**Figure 8**). Among the above mentioned novel DBS protocols, the aperiodic stimulation protocol, in addition, promises to reduce the energy consumption of the DBS device. With little effort, it can be implemented using the existing, clinically approved DBS systems.

MODEL PREDICTIONS AND LIMITATIONS

Our model suggest that excitatory inputs to the STN from the cortex would lead to an increase in the β band oscillations. This seems at odds with the experiments which show that during behavior when strong excitatory input arrives in the STN, β band oscillations are reduced, for example in a high-conflict decision making task (Frank et al., 2007). However, task related activity will arrive in the basal ganglia through both striatum and the STN. We have not investigate how such input will interact and whether they will enhance or weaken the β band oscillations. This will also require the knowledge of spatio-temporal structure of the task related input to the basal ganglia.

The spectrum of correlations in our model during oscillatory and non-oscillatory states is not in complete match with the experimentally measured correlation in monkey with Parkinson's disease (Brunel, 2000; Heimer et al., 2002; Levy et al., 2002). The available experimental data does not provide quantitative estimates of cross-correlation and strength of the oscillations in the cross-correlogram, which makes a direct comparison between our model and the data difficult.

The observation of transient oscillations in the healthy state upon short stimulation of the GPe motivated us to suggest that task inhibition may involve transient oscillations. With DBS such transients oscillations are hard to generate, therefore we suggest that patients with DBS will have difficulty in controlling their actions. This however, does not explain why and how patients with DBS may speed up their decisions under high-conflict condition (Frank et al., 2007). The model is too simple and therefore not suitable to address the behavioral performance of the basal ganglia in complex task such as high-conflict decision making.

Finally, our simple model needs to be extended to include more realistic neuron models to explain the observed increased bursting in the GPe neurons during dopamine-depleted state (Raz et al., 2000).

Nevertheless, despite its simplicity the model is powerful enough to suggest

- Striatum inhibition is sufficient to generate oscillations associated with PD.
- DBS should be effectively inhibitory in nature.
- Transient oscillations in the STN-GPe network which can play a crucial role in decision making depend on the baseline firing rate in the GPe.

Together these three main results provide novel insights about the origin of the oscillations in the basal ganglia, their quenching due to deep-brain-stimulation (DBS), and provides a putative explanation of certain cognitive deficits in the PD patients with or without DBS.

ACKNOWLEDGMENTS

We thank Hagai Bergman and Izhar Bar-Gad for helpful discussions. All simulations were carried out using the NEST simulation software (www.nest-initiative.org). Partial funding by the German Federal Ministry of Education and Research (BMBF grant 01GQ0420 to the BCCN Freiburg, BMBF GW0542 Cognition, and BMBF 01GW0730 Impulse Control) is gratefully acknowledged.

REFERENCES

Baufreton, J., Kirkham, E., Atherton, J. F., Menard, A., Magill, P. J., Bolam, J. P., and Bevan, M. D. (2009). Sparse but selective and potent synaptic transmission from the globus pallidus to the subthalamic nucleus. *J. Neurophysiol.* 102, 532–545.

Benabid, A. L. (2003). Deep brain stimulation for Parkinson's disease. *Curr. Opin. Neurobiol.* 13, 696–706.

Bender, C. M., and Orszag, S. M. (1999). *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory*. New York: Springer.

Bergman, H., Wichmann, T., and DeLong, M. R. (1990). Reversal of experimental parkinsonism by lesions of the subthalamic nucleus. *Science* 249, 1436–1438.

Bergman, H., Wichmann, T., Karmon, B., and DeLong, M. R. (1994). The primate subthalamic nucleus. II. Neuronal activity in the MPTP model of parkinsonism. *J. Neurophysiol.* 72, 507–520.

Bevan, M. D., Clarke, N. P., and Bolam, J. P. (1997). Synaptic integration of functionally diverse pallidal information in the entopeduncular nucleus and subthalamic nucleus in the rat. *J. Neurosci.* 17, 308–324.

Brown, P., and Williams, D. (2005). Basal ganglia local field potential activity: character and functional significance in the human. *Clin. Neurophysiol.* 116, 2510–2519.

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci.* 8, 183–208.

Brunel, N., and Wang, X. J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* 90, 415–430.

Coffey, R. J. (2009). Deep brain stimulation devices: a brief technical history and review. *Artif. Organs* 33, 208–220.

Deger, M., Moritz, H., Stefano, C., Fatichan, A., and Stefan, R. (2010). Nonequilibrium Non-equilibrium dynamics of stochastic point processes with refractoriness. *Phys. Rev. E* 82, 021129.

Dorval, A. D., Kuncel, A. M., Birdno, M. J., Turner, D. A., and Grill, W. M. (2010). Deep brain stimulation alleviates parkinsonian bradykinesia by regularizing pallidal activity. *J. Neurophysiol.* 104, 911–921.

Eusebio, A., Chen, C. C., Lu, C. S., Lee, S. T., Tsai, C. H., Limousin, P., Hariz, M., and Brown, P. (2008). Effects of low-frequency stimulation of the subthalamic nucleus on movement in Parkinson's disease. *Exp. Neurol.* 209, 125–130.

Frank, M. J., and O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behav. Neurosci.* 120, 497–517.

Frank, M. J., Samanta, J., Moustafa, A., and Moustafa, S. (2007). Hold your horses: impulsivity, deep brain stimulation and medication in Parkinsonism. *Science* 318, 1309–1312.

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.

Gillies, A., Willshaw, D., and Li, Z. (2002). Subthalamic-pallidal interactions are critical in determining normal and abnormal functioning of the basal ganglia. *Proc. Biol. Sci.* 269, 545–551.

Gradinaru, V., Mogri, M., Thompson, K. R., Henderson, J. M., and Deisseroth, K. (2009). Optical deconstruction of parkinsonian neural circuitry. *Science* 324, 354–359.

Hammond, C., Bergman, H., and Brown, P. (2007). Pathological synchronization in Parkinson's disease: networks, models and treatments. *Trends Neurosci.* 30, 357–364.

Hammond, C., and Yelnik, J. (1983). Intracellular labelling of rat subthalamic neurones with horseradish peroxidase: computer analysis of dendrites and characterization of axon arborization. *Neuroscience* 8, 781–790.

Hashimoto, T., Elder, C. M., Okun, M. S., Patrick, S. K., and Vitek, J. L. (2003). Stimulation of the subthalamic nucleus changes the firing pattern of pallidal neurons. *J. Neurosci.* 23, 1916–1923.

Hauptmann, C., Roulet, J. C., Niederhauser, J. J., Döll, W., Kirlangic, M. E., Lysyansky, B., Krachkovskyi, V., Bhatti, M. A., Barnikol, U. B., Sasse, L., Bührle, C. P., Speckmann, E. J., Götz, M., Sturm, V., Freund, H. J., Schnell, U., and Tass, P. A. (2009). External trial deep brain stimulation device for the application of desynchronizing stimulation techniques. *J. Neural Eng.* 6, 066003.

Heimer, G., Bar-Gad, I., Goldberg, J. A., and Bergman, H. (2002). Dopamine replacement therapy reverses abnormal synchronization of pallidal neurons in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine primate model of parkinsonism. *J. Neurosci.* 22, 7850–7855.

Holgado, A. J. N., Terry, J. R., and Bogacz, R. (2010). Conditions for the generation of beta oscillations in the subthalamic nucleus-globus pallidus network. *J. Neurosci.* 30, 12340–12352.

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.

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.

Kita, H., and Kitai, S. T. (1994). The morphology of globus pallidus projection neurons in the rat: an intracellular staining study. *Brain Res.* 636, 308–319.

Kita, H., Tachibana, Y., Nambu, A., and Chiken, S. (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.

Krack, P., Hariz, M. I., Baunez, C., Guridi, J., and Obeso, J. A. (2010). Deep brain stimulation: from neurology to psychiatry? *Trends Neurosci.* 33, 474–484.

Kravitz, A. V., Freeze, B. S., Parker, P. R. L., Kay, K., Thwin, M. T., Deisseroth, K., and Kreitzer, A. C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature* 466, 622–626.

Kringelbach, M. L., Jenkinson, N., Owen, S. L. F., and Aziz, T. Z. (2007). Translational principles of deep brain stimulation. *Nat. Rev. Neurosci.* 8, 623–635.

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

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.

Levy, R., Ashby, P., Hutchison, W. D., Lang, A. E., Lozano, A. M., and Dostrovsky, J. O. (2002). Dependence of subthalamic nucleus oscillations on movement and dopamine in Parkinson's disease. *Brain* 125, 1196–1209.

Liang, L., DeLong, M. R., and Papa, S. M. (2008). Inversion of dopamine responses in striatal medium spiny neurons and involuntary movements. *J. Neurosci.* 28, 7537–7547.

Magill, P. J., Bolam, J. P., and Bevan, M. D. (2001). Dopamine regulates the impact of the cerebral cortex on the subthalamic nucleus-globus pallidus network. *Neuroscience* 106, 313–330.

Malenka, R. C., and Bear, M. F. (2004). LTP and LTD: an embarrassment of riches. *Neuron* 44, 5–21.

Mallet, N., Ballion, B., Le Moine, C., and Gonon, F. (2006). Cortical inputs and GABA interneurons imbalance projection neurons in the striatum of Parkinsonian rats. *J. Neurosci.* 26, 3875–3884.

Mallet, N., Pogosyan, A., Márton, L. F., Bolam, J. P., Brown, P., and Magill, P. J. (2008). Parkinsonian beta oscillations in the external globus pallidus and their relationship with subthalamic nucleus activity. *J. Neurosci.* 28, 14245–14258.

McIntyre, C. C., Savasta, M., Walter, B. L., and Vitek, J. L. (2004). How does deep brain stimulation work? Present understanding and future questions. *J. Clin. Neurophysiol.* 21, 40–50.

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

Nambu, A. (2008). Seven problems on the basal ganglia. *Curr. Opin. Neurobiol.* 18, 595–604.

Parent, A., and Hazrati, L. N. (1995). Functional anatomy of the basal ganglia. II. The place of subthalamic nucleus and external pallidum in basal ganglia circuitry. *Brain Res. Rev.* 20, 128–154.

Raz, A., Feingold, A., Zelanskaya, V., Vaadia, E., and Bergman, H. (1996). Neuronal synchronization of tonically active neurons in the striatum of normal and Parkinsonian primates. *J. Neurophysiol.* 76, 2083–2088.

Raz, A., Vaadia, E., and Bergman, H. (2000). Firing patterns and correlations of spontaneous discharge of pallidal neurons in the normal and the tremulous 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine vervet model of parkinsonism. *J. Neurosci.* 20, 8559–8571.

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.

Sato, F., Parent, M., Levesque, M., and Parent, A. (2000). Axonal branching pattern of neurons of the subthalamic nucleus in primates. *J. Comp. Neurol.* 424, 142–152.

Shen, K. Z., and Johnson, S. W. (2005). Dopamine depletion alters responses to glutamate and GABA in the rat subthalamic nucleus. *Neuroreport* 16, 171–174.

Smith, Y., Villalba, R., and Raju, D. (2009). Striatal spine plasticity in Parkinson's disease: pathological or not? *Parkinsonism Relat. Disord.* 15, S156–S161.

Surmeier, D. J., Mercer, J. N., and Chan, C. S. (2005). Autonomous pacemakers in the basal ganglia: who needs excitatory synapses anyway? *Curr. Opin. Neurobiol.* 15, 312–318.

Tass, P., Smirnov, D., Karavaev, A., Barnikol, U., Barnikol, T., Adamchic, I., Hauptmann, C., Pawelcyzk, N., Maarouf, M., Sturm, V., Freund, H. J., and Bezruchko, B. (2010). The causal relationship between subcortical local field potential oscillations and Parkinsonian resting tremor. *J. Neural Eng.* 7, 16009.

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.

Tseng, K. Y., Kasanetz, F., Kargieman, L., Riquelme, L. A., and Murer, M. G. (2001). Cortical slow oscillatory activity is reflected in the membrane potential and spike trains of striatal neurons in rats with chronic nigrostriatal lesions. *J. Neurosci.* 21, 6430–6439.

van Albada, S. J., Gray, R. T., Drysdale, P. M., and Robinson, P. A. (2009). Mean-field modeling of the basal ganglia-thalamocortical system. II Dynamics of parkinsonian oscillations. *J. Theor. Biol.* 257, 664–688.

van Albada, S. J., and Robinson, P. A. (2009). Mean-field modeling of the basal ganglia-thalamocortical system. I Firing rates in healthy and Parkinsonian states. *J. Theor. Biol.* 257, 642–663.

van den Wildenberg, W. P. M., van Boxtel, G. J. M., van der Molen, M. W., Bosch, D. A., Speelman, J. D., and Brunia, C. H. (2006). Stimulation of the subthalamic region facilitates the selection and inhibition of motor responses in Parkinson's disease. *J. Cogn. Neurosci.* 18, 626–636.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 May 2011; accepted: 03 October 2011; published online: 24 October 2011.

Citation: 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*

OPEN ACCESS Freely available online

PLoS COMPUTATIONAL BIOLOGY

Context-Dependent Encoding of Fear and Extinction Memories in a Large-Scale Network Model of the Basal Amygdala

Ioannis Vlachos[1]*, Cyril Herry[2,3], Andreas Lüthi[4], Ad Aertsen[1,5], Arvind Kumar[1,5]*

1 Bernstein Center for Computational Neuroscience Frieburg, Freiburg, Germany, **2** Neurocentre Magendie, Bordeaux Cedex, France, **3** INSERM U862, Bordeaux Cedex, France, **4** Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland, **5** Department of Neurobiology and Biophysics, Faculty of Biology, University of Freiburg, Freiburg, Germany

Abstract

The basal nucleus of the amygdala (BA) is involved in the formation of context-dependent conditioned fear and extinction memories. To understand the underlying neural mechanisms we developed a large-scale neuron network model of the BA, composed of excitatory and inhibitory leaky-integrate-and-fire neurons. Excitatory BA neurons received conditioned stimulus (CS)-related input from the adjacent lateral nucleus (LA) and contextual input from the hippocampus or medial prefrontal cortex (mPFC). We implemented a plasticity mechanism according to which CS and contextual synapses were potentiated if CS and contextual inputs temporally coincided on the afferents of the excitatory neurons. Our simulations revealed a differential recruitment of two distinct subpopulations of BA neurons during conditioning and extinction, mimicking the activation of experimentally observed cell populations. We propose that these two subgroups encode contextual specificity of fear and extinction memories, respectively. Mutual competition between them, mediated by feedback inhibition and driven by contextual inputs, regulates the activity in the central amygdala (CEA) thereby controlling amygdala output and fear behavior. The model makes multiple testable predictions that may advance our understanding of fear and extinction memories.

Citation: Vlachos I, Herry C, Lüthi A, Aertsen A, Kumar A (2011) Context-Dependent Encoding of Fear and Extinction Memories in a Large-Scale Network Model of the Basal Amygdala. PLoS Comput Biol 7(3): e1001104. doi:10.1371/journal.pcbi.1001104

Editor: Tim Behrens, John Radcliffe Hospital, United Kingdom

Received September 22, 2010; **Accepted** February 7, 2011; **Published** March 17, 2011

Funding: This work was supported in parts by the German Federal Ministry of Education and Research (BMBF grant 01GQ0420 to BCCN Freiburg), the Swiss National Science Foundation, the Novartis Research Foundation, and a Neurex+ grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: vlachos@bcf.uni-freiburg.de (IV); arvind.kumar@biologie.uni-freiburg.de (AK)

Introduction

In classical fear conditioning an animal learns to associate an initially neutral stimulus (the conditioned stimulus, CS) with an aversive stimulus (the unconditioned stimulus, US) after paired exposure to the CS and the US. Subsequent repeated non-reinforced presentations of the CS alone result in a decline of the conditioned response, a process called fear extinction [1]. Fear extinction is a highly context-dependent process: the conditioned fear response returns when the animal is exposed to an extinguished CS outside the extinction context [2,3].

Studies over the last decades have identified the amygdaloid complex as a key brain structure involved in both fear conditioning and extinction [4–6]. In the lateral nucleus of the amygdala (LA), signals carrying information about the CS and the US converge onto the same neurons where they become associated through activity-dependent plasticity mechanisms [7–9]. The LA can directly or indirectly influence activity in the central nucleus (CEA) [10], the major output nucleus of the amygdala that can trigger fear responses via its projections to the hypothalamus and to the brainstem [11]. The basal nucleus of the amygdala (BA) has been suggested to play an important role in contextual fear conditioning [12,13], cued fear conditioning [14], fear extinction [15–17] and context-dependent fear renewal [17].

Recently, two distinct fear and extinction specific neuronal sub-populations in the BA have been identified [17]. The balance of activity between fear and extinction neurons was correlated with states of high and low fear, respectively. Moreover, pharmacological inactivation of the BA blocked the acquisition of fear extinction and context-dependent fear renewal, suggesting that BA fear and extinction neurons may underlie the induction of behavioral changes and contribute to the formation of fear and extinction memories.

These findings raise the question of what the potential mechanisms underlying the differential activation of these two neuronal sub-populations are. Here, we used a modeling approach based on *in vivo* physiological data to address this specific question and to draw more general conclusions on potential neural mechanisms involved in fear and extinction memories in the BA. In vivo stimulation of identified fear and extinction neurons revealed that the two neuronal populations receive differential functional input from the hippocampus and from the medial prefrontal cortex (mPFC) [17]. This finding could reflect anatomical specificity of inputs and/or selective functional plasticity of non-specific inputs. Independently of these two possibilities, in our model, we assume that anatomically and/or functionally distinct inputs from the hippocampus or the mPFC modulate the activity of BA fear and extinction neurons in a context-specific manner. That is, sub-populations of BA neurons are innervated by hippocampus/mPFC efferents that represent the current context. In addition, all BA neurons receive inputs from

Author Summary

The amygdaloid complex is one of the key brain structures involved in fear-related processes. A typical way to study neural correlates of fear expression (e.g. freezing response) in the amygdala is to perform a fear conditioning paradigm, which yields a conditioned fear response. This response can be reversed by another procedure called fear extinction. Thanks to the experimental approaches to date we have some understanding about the putative roles of specific subnuclei within the amygdala in the formation of these fear and extinction memories. Here, we complement the experimental studies by providing a computational model that addresses the question of how fear and extinction memories are encoded in the amygdala, and specifically, in the basal nucleus (BA). We propose a specific neural mechanism to explain how the BA may integrate information about a salient, conditioned stimulus and the environment, thereby enabling it to switch the state of the animal from low to high fear and vice versa. We also provide possible explanations for various other behavioral findings, such as the recovery of fear after it had been extinguished (renewal). Finally, we make specific, experimentally testable predictions that need to be addressed in future work.

US/CS responsive LA neurons during conditioning and extinction. Those sub-populations of BA neurons that receive simultaneous LA and context-specific inputs become responsive during conditioning or extinction and, thus, emulate the "fear" and "extinction" neurons reported by Herry et al. [17]. Activation of BA neurons per se, however, is not sufficient to cause or prevent a behavioral response, but the selective activation of BA neurons conveys important information about the context-CS relation to the CEA. Although we do not model here the CEA, we stipulate that context-dependent BA activity provides an instructive signal to CEA neurons. In the CEA, it is likely that conditioning [18] and possibly extinction learning-induced changes act upon this signal in order to activate or suppress a fear response. If more experimental data, sufficient to constrain the possible parameter space, become available, then our present model of the BA could be extended to study the impact of context-dependent BA activity on learning-induced changes in the CEA as well.

We test the plausibility of context-dependent activation of BA neurons in two different approaches: first, in an abstract firing rate model; second, in a more realistic spiking neuron network (SNN) model of the BA. Based on the results of our model we provide plausible explanations for several experimental observations in fear and extinction learning and make specific, experimentally testable predictions.

Models

Experimental observations

The description of the evolution of the firing rates of BA neurons during fear conditioning and extinction reported by [17] provide certain simple, yet important, indications on the underlying dynamics in the BA network:

1. In naive animals, the ongoing activity of BA neurons does not predict the existence of different sub-populations of neurons.
2. As the animal learns to associate the CS with the US in $\mathrm{context}_A$ (the conditioning context), the activity of a sub-population of neurons in BA (fear neurons) increases, in correlation with the animal's freezing behavior (Fig. 1). The most parsimonious explanation suggests that the strength of a fraction of afferents carrying information on the CS (CS inputs from hereon) has been increased. Alternatively, changes in single neuron properties, e.g. excitability, or alterations in network activity states, e.g. reduced global inhibition, could also account for this observation.
3. During extinction training in a different context ($\mathrm{context}_B$), CS-induced activity of fear neurons progressively diminishes while the activity of a new sub-population of neurons (extinction neurons) increases. This suggests that during extinction, the strength of a new subgroup of CS inputs is strengthened, leading to the increase in response of the extinction neurons. Furthermore, the second and third observations highlight the importance of the context in the selective increase of the CS inputs. They also suggest that the strength of the contextual inputs to BA neurons may increase as well.
4. The sudden and selective increase of activity of fear neurons when the animal is put back in $\mathrm{context}_A$ after extinction learning (i.e. renewal), reveals that extinction cannot be merely unlearning. Thus, the most simple explanation for the response reduction of fear neurons in $\mathrm{context}_B$ is local inhibition generated by the increased response of extinction neurons. It cannot be excluded, though, that partial unlearning - i.e. depotentiation of a certain fraction of previously strengthened synapses - may occur in parallel.

To test the feasibility of the above observations and their inferences in explaining the emergence of fear and extinction neurons in BA, we first studied the dynamics of a mean-field (or firing rate) model of the BA. Subsequently, we constructed a spiking neuron network (SNN) model to examine our hypotheses and their implications under more realistic conditions.

Mean-field model of the BA

The mean-field model of BA consisted of two neuron populations, A and B, described by Wilson-Cowan type rate dynamics [19] (Fig. 2A). Both populations were *identical* in their properties (Eqs. 1–2) and received both CS input and non-specific background input. There is ample experimental evidence that in different contexts, different sets of hippocampal neurons (e.g. in CA1) are active [20–22]. Thus, to mimic context-specific inputs - either directly from hippocampus or indirectly via the mPFC or other brain structures such as entorhinal cortex - we provided population A with additional input CTX_A reflecting $\mathrm{context}_A$, and likewise, population B with additional input CTX_B reflecting $\mathrm{context}_B$. Populations A and B were mutually interconnected with inhibitory synapses. The system of differential equations describing the activity of the populations A and B is as follows:

$$\tau_A \frac{dA}{dt} = -A + \eta(t) + (k_A - r_A A)\cdot S(j_{A,CS}\cdot CS + j_{A,B}\cdot B + j_{A,CTX_A}\cdot CTX_A), \quad (1)$$

$$\tau_B \frac{dB}{dt} = -B + \eta(t) + (k_B - r_B B)\cdot S(j_{B,CS}\cdot CS + j_{B,A}\cdot A + j_{B,CTX_B}\cdot CTX_B), \quad (2)$$

where $S(x) = \frac{1}{1+e^{-p(x-\theta)}}$.

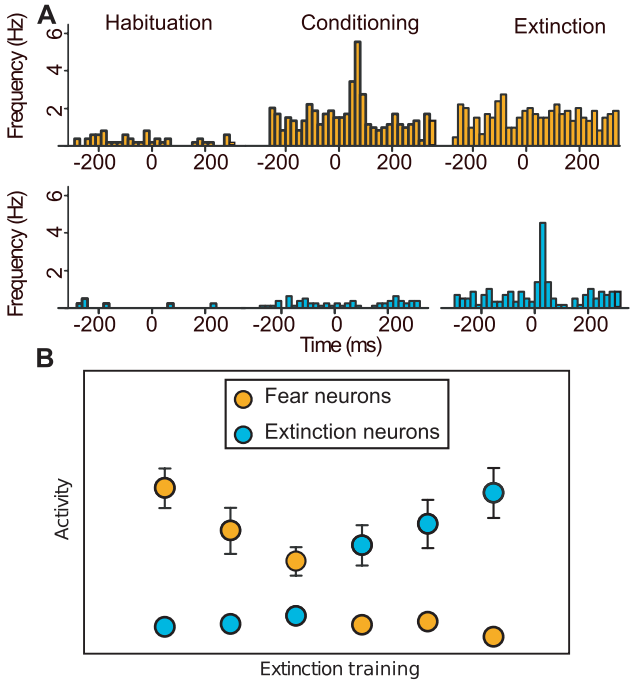

Figure 1. Fear and extinction neurons in rodents. (A) CS-evoked activity in the BA in pre-conditioning (left), post-conditioning (center) and post-extinction (right). After conditioning one subpopulation of neurons within the BA (fear-neurons, amber) increased their firing rates in response to the CS. This subpopulation did not show any CS evoked response after extinction. A different subpopulation (extinction neurons, cyan) did not respond to the CS during or after fear conditioning, but showed a CS evoked response after extinction training. (B) Population activities of fear and extinction neurons during extinction training for different blocks of CS presentations. In a different context, extinction training resulted in a progressive decrease in the response of fear neurons and increase in the response of extinction neurons. The switch of activity was correlated with a shift in behavior from high to low freezing. Figure adapted from [17].
doi:10.1371/journal.pcbi.1001104.g001

The evolution of the connection strengths is given by

$$\frac{dj_{A,CS}}{dt} = \alpha_A \cdot CS\ CTX_A, \tag{3}$$

$$\frac{dj_{B,CS}}{dt} = \alpha_B \cdot CS\ CTX_B. \tag{4}$$

Here, $j_{X,Y}$ $(X, Y \varepsilon \{A,B\})$ represents the connection strength from population (or external input) Y to population X, τ_X is the time constant governing the dynamics of population X, k_X is the maximum firing rate of population X, and r_X captures the refractoriness of neurons in X. The transfer function S is a sigmoid function, integrating all inputs to population X in a non-linear fashion and producing a bounded output rate. The parameters p and θ of the sigmoid function determine the steepness and the position of its maximum slope, respectively. The term $\eta(t)$, with zero mean, reflects the stochastic input to the two populations, mimicking the background activity in the BA.

Equations 3 and 4 describe the dynamics of the connection strengths of the CS afferents onto populations A and B respectively. These weights were increased in an additive way whenever the respective CS and CTX inputs were present simultaneously and remained constant otherwise. The parameters a_A and a_B specify the learning rates (see also Eqs. 6–8).

We simulated fear conditioning and extinction by applying CS input to both populations in the form of short pulses of 50 ms duration each, based on the experimental design used in [17]. Contextual input was provided continuously. Note that we did not make any explicit distinction between the unconditioned stimulus (US) and conditioned stimulus (CS). Instead, we assumed that during conditioning, neurons in the LA initially responded to the US and eventually to the CS, while continuing to respond to the CS during extinction [23]. The output of these LA neurons was then fed downstream to the BA. In addition, US or CS inputs from the thalamus or the primary sensory cortex may directly target BA neurons [24]. In our model, we represented those inputs,

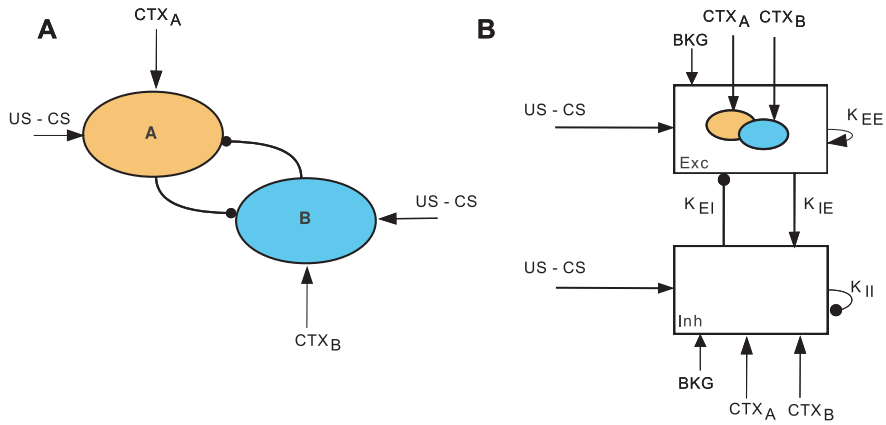

Figure 2. Schematic network model diagrams of the BA. (A) Firing rate model. Two neuron populations A and B are mutually coupled with negative weights. Both populations receive US-CS and context-specific (CTX) inputs. These external inputs can exhibit LTP. (B) Spiking neural network model. The network consists of 3400 excitatory and 600 inhibitory LIF neurons. The neurons are interconnected in a recurrent fashion. US-CS input is provided to all neurons. CTX input is fed only to two subpopulations of excitatory neurons. The external inputs (CS, US and CTX) are modeled as rate-modulated Poisson spike trains.
doi:10.1371/journal.pcbi.1001104.g002

independently of their origin, as CS-US in the conditioning context and CS in the extinction context.

Spiking neuron network model of BA

For the description of the SNN we adopted the *good model description practice* proposed by [25], which provides guidelines for a standardized way of describing complex neural networks. We share the authors' belief that such model description facilitates reproducibility and direct comparisons between models. Within this framework, we organized the description in different subsections, complemented by additional information on the model parameters. This collected information is presented in an easily accessible, tabular form in the Supplementary Materials (Table S1).

Our choice to use leaky-integrate-and-fire (LIF) neurons was motivated by four major arguments: (i) multiple combinations of sub-cellular parameters can result in the same network state [26]; (ii) even simple neuron models such as LIF with minor modifications are sufficient to reproduce complex *in vivo* spike patterns [27]; (iii) realistically-sized large scale networks of LIF neurons can now be simulated with the currently available simulation technology [28]; this is hardly possible for similarly large networks built of detailed compartmental models and, finally, (iv) the extent to which sub-cellular properties of individual neurons influence the global network dynamics is presently not clear. Most importantly, however, here we are interested in understanding the key *network* level properties of the BA which play a critical role in the formation of fear and extinction memories. For this purpose, the LIF neurons, although they are reduced models of a biological cell, provide an adequate level of biophysical realism, sufficient to identify these key network properties.

Network composition and connectivity

We modeled the BA as a random recurrent network, consisting of $N_E = 3400$ excitatory (EXC) and $N_I = 600$ inhibitory (INH) neurons [24,29]. A total number of 4000 neurons corresponds roughly to 10% of all neurons in the rat BA [30]. The schematic diagram of the network is shown in Fig. 2B. Each connection from a pre- to a post-synaptic neuron had an assigned probability, the value of which depended on the types of pre- and postsynaptic neurons involved (EXC and INH, respectively): $p_{EE} = 0.01$, $p_{EI} = 0.15$, $p_{IE} = 0.15$, and $p_{II} = 0.1$. Thus, each EXC neuron received on average $p_{EE} * N_E = 0.01 * 3{,}400 = 34$ excitatory and $p_{EI} * N_E = 0.15 * 600 = 90$ inhibitory connections. Likewise, each INH neuron received $p_{IE} * N_E = 0.15 * 3{,}400 = 510$ excitatory and $p_{II} * N_I = 0.10 * 600 = 60$ inhibitory connections. Neurons were allowed to form recurrent connections to themselves. For the simulations shown in the last figure, we systematically varied the connection probability of the recurrent inhibition from 0.1 to 1.0.

Input, output, and free parameters

EXC and INH neurons received inputs encoding information on the CS. Similarly to the rate model, these inputs represented initial responses of LA neurons to combined CS and US presentations, later only to the CS. They might also reflect more peripheral, thalamic or cortical responses to CS-US. A fraction of BA EXC neurons (20%, randomly chosen) received inputs representing CS and context_A. Similarly, another 20% of BA EXC neurons received inputs representing CS and context_B. Thus, similar to the rate model, we assumed that BA EXC neurons receive contextual information directly from the HPC (or entorhinal cortex) and/or via the mPFC. Crucially, CS-US and contextual inputs converged onto the same neurons [8]. Furthermore, EXC and INH neurons received unspecific background inputs (BKG), representing activity originating in other areas, either within or outside the amygdaloid complex. The BKG inputs accounted for the baseline spiking activity of EXC and INH neurons at <1 Hz and 10–15 Hz, respectively [24].

The exact temporal and spatial patterns of the spiking inputs to the BA are not known. Here, we used independent Poisson spike generators with different firing rates to produce the specific inputs. Contextual and BKG inputs provided a tonic drive to BA neurons. By contrast, the CS input had a short duration of 50 ms, based on the experimental design used in [17]. All external inputs formed excitatory synapses onto their target neurons.

Neurons, synapses, and plasticity

Neurons were modeled as leaky-integrate-and-fire (LIF) neurons. The subthreshold dynamics of each LIF neuron were governed by the following equation

$$\tau_m \frac{dV}{dt} = (E_0 - V) + g_{exc}(E_{exc} - V) + g_{inh}(E_{inh} - V). \quad (5)$$

A spike was generated whenever the membrane potential crossed a predefined static threshold θ in upgoing direction. The potential was then reset to a value E_k and clamped for t_{ref} ms before the synaptic integration started again (Table S1F). Neurons made either excitatory or inhibitory connections onto their postsynaptic targets via conductance-based synapses [31–33].

The synapses of all connections were non-modifiable, except those providing CS and contextual input to EXC neurons. These latter, plastic synapses were modified according to the following phenomenological rule:

$$w_+ = \begin{cases} w_- + \alpha_1 \cdot h \cdot m \cdot |w_{max} - w_-| \cdot c & \text{if CS and CTX temporally overlapped} \\ w_- - \alpha_2 \cdot m \cdot |w_{min} - w_-| \cdot c & \text{otherwise} \end{cases}, \quad (6)$$

$$\dot{c} = -\frac{c}{\tau_c} + A \cdot \delta(t_{pre}), \quad (7)$$

$$\dot{h} = -\frac{h}{\tau_h} + B \cdot \delta(t_{pre}). \quad (8)$$

Note that three variables were used: the synaptic weight w and the auxiliary variables c and h. Each time a presynaptic neuron fired, the value of c increased by a fixed amount. Afterward, this value relaxed towards zero. Thus, variable c acted as a synaptic tag, encoding the recent activity in the synapse receiving CS input. Likewise, variable h encoded information about recent activity in neighboring synapses receiving contextual input.

At the offset of each CS presentation, the variables c and h were probed in the synapses of all EXC neurons and the strength of each synapse was modified accordingly. The synaptic strengths before and after the update are denoted by w_- and w_+, respectively. If CS and contextual inputs at the same neuron coincided within a temporal window of ~ 100 *ms*, then both

synapses were strengthened [34]. By contrast, if only one of the inputs was present, both synapses were weakened (Eq. 6). This decrease of synaptic strength was based on studies reporting that synapses in LA, which had been strengthened during fear conditioning, depotentiated after extinction training [35,36]. We assumed a similar mechanism to hold for the BA. This type of bidirectional plasticity rule implemented in our model is similar to the BCM rule [37], the "calcium-control hypothesis" [38–40] and the ABS rule [41,42]. Common in all these rules is the specification that the level of postsynaptic Ca^{2+} determines the direction of plasticity (for review see [43]). A large increase in Ca^{2+} causes LTP, whereas a moderate increase results in LTD. Low levels of Ca^{2+} do not cause any modification at all. We essentially incorporated this bidirectional induction of plasticity in our rule using fixed thresholds (Fig. 3C), rather than sliding ones, as is the case e.g. in the BCM rule. The parameters a_1 and a_2 denote the learning rates for potentiation and depotentiation of the synapses, respectively.

Ca^{2+} influx depends on NMDA receptor activation and sufficient postsynaptic depolarization. The latter can be caused by coincident presynaptic input or by a backpropagating action potential (BAP). However, in our model, a BAP was not required. That is, we assumed that if the total presynaptic firing rates were high enough, they could cause sufficient depolarization to unblock NMDA receptors. This assumption is supported by experimental evidence showing that a BAP is neither necessary nor sufficient for synaptic plasticity [44,45].

Note that this plasticity rule is also compatible with changes induced purely in the presynaptic terminal. In fact, experimental evidence suggests that presynaptic induction, completely independent of postsynaptic activity, occurs in the LA [46]. Thus, the plasticity rule implemented in our model incorporates both changes that are dependent on post-synaptic depolarization, but not postsynaptic spiking, and changes that are presynaptic and entirely independent of post-synaptic depolarization or spiking.

Because in our model the presynaptic spiking was caused by CS and contextual inputs, their total activity encoded in the variables c and h, respectively, determined the direction of plasticity. Thus, both c and h functioned as eligibility traces for synaptic modification [34,47]. They could be interpreted as describing any relatively slow process associated with the effects of Ca^{2+}, e.g. autophosphorylation of CaMK-II [39,48].

The terms $|w_{max} - w_-|$ and $|w_{min} - w_-|$ in the update rule were introduced to provide upper and lower bounds to the synaptic weights, such that they did not increase or decrease indefinitely. They also controlled the step-size with which synapses were modified: the closer a weight was to w_{max} (w_{min}) the smaller were its increments (decrements).

The parameter m represented the action of neuromodulators released during fear conditioning and extinction. It is known that many neuromodulators target the BA [5], possibly affecting synaptic plasticity in a complex way. Among the possible candidates are norepinephrine (NE) [49–52], dopamine (DA) [53,54] and opioids [5]. Here, however, lacking more detailed experimental data, we cannot be more specific about which exact neuromodulators are involved and how they interact. Fortunately, this lack of knowledge does not pose a problem for the plasticity rule we propose, because it is general enough to accommodate any combination of neuromodulators that may turn out to be involved in BA fear processing.

The dynamics of the mean-field model were simulated in MATLAB. The SNN simulations were written in python (http://www.python.org), using the PyNN interface [55, http://neuralensemble.org/trac/PyNN] to the NEST simulation environment [56, http://www.nest-initiative.org].

Results

Firing-rate model of BA

Fig. 4 shows the response of the mean-field model, i.e. the firing rate model, of BA during fear conditioning and extinction. To simulate fear conditioning in $\mathrm{context}_A$, we stimulated the population A five times with CS, US and CTX_A inputs (Eqs. 1,2). This resulted in a progressive strengthening of CS synapses onto population A ($j_{cs,A}$) (Fig. 4C), accompanied by a corresponding increase in the response of population A (Fig. 4A). To simulate fear extinction training in $\mathrm{context}_B$, we stimulated population A with CS input and population B with CS and CTX_B input six times to mimic a different context. Now, in $\mathrm{context}_B$, the synaptic strength of the CS input synapses ($j_{cs,B}$) onto population B progressively got stronger, whereas $j_{cs,A}$ remained unchanged (Fig. 4D). The slow increase in the response of population B resulted in a small decrease in the response of population A, due to the recurrent inhibition. When

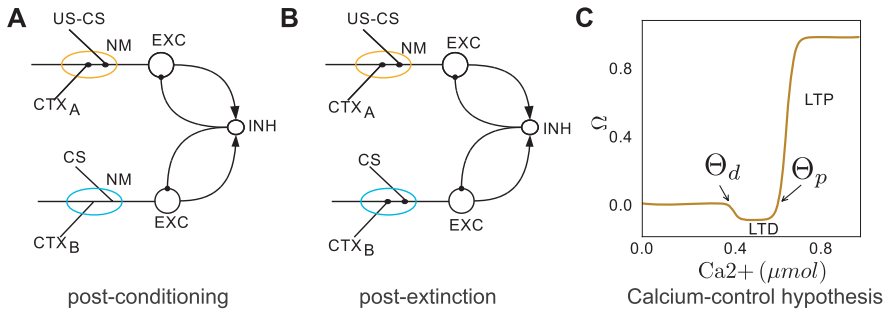

Figure 3. Synaptic plasticity. (A) and (B) Schematic connectivity diagrams of typical fear and extinction neurons in the BA network model (cf. Fig. 2B). We assume that US-CS and CTX synapses are plastic and spatially clustered on the same BA excitatory neuron (amber and cyan ellipses; NM: neuromodulator). (A) During fear conditioning, co-activation of CS, US and CTX afferents, strengthened the synapses (black dots) on fear neurons in the presence of a neuromodulator. (B) During fear extinction training, co-activation of CS and CTX afferents strengthened the synapses (black dots) on extinction neurons. Lack of CTX inputs resulted in a small depotentiation of CS synapses onto fear neurons (cf. Models). (C) The plasticity mechanism that drives the change in synaptic strength is essentially an implementation of the calcium-control hypothesis [Eqs. 6–8]. If CS and CTX inputs temporally coincide, the calcium influx in the neuron crosses threshold Θ_p and LTP is induced in both CS and CTX synapses. By contrast, if only one of the CS and CTX inputs is active, the total calcium influx lies between Θ_d and Θ_p and the CS or CTX synapses exhibit LTD. If none of the inputs is active, the total calcium level stays below threshold Θ_d and the synapses remain unaltered.
doi:10.1371/journal.pcbi.1001104.g003

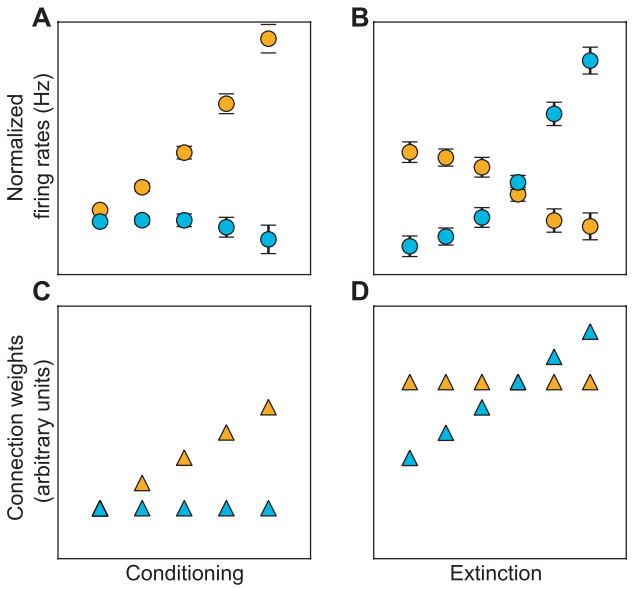

Figure 4. Dynamics of the firing rate model of BA. Five US-CS stimulations were used for conditioning and six CS stimulations for extinction. (A) During conditioning, LTP at US-CS and CTX afferents yielded increased activity of one of the sub-populations (amber dots). The increase in the activity of this sub-population resulted in a weak inhibition of the second sub-population (cyan dots). (B) During extinction, the same plasticity mechanism resulted in a gradual increase in the activity of the second subpopulation (cyan dots). This increase inhibited the first subpopulation, which also received less excitation in the extinction training. The normalized firing rates represent the average over 30 simulations. Note that the firing rates, although they have an upper bound, determined by the refractory term in Eqs. 1–2, remain far from saturation. (C) The evolution of synaptic weights $j_{CS,A}$ governing the increase in the firing rate of the fear neurons during fear conditioning (amber Δ). (D) Evolution of synaptic weights $j_{CS,B}$ during extinction training (cyan Δ). A steady increase in the strength of synapses $j_{CS,B}$ resulted in a steady increase in the firing rate of extinction neurons. In the firing rate model no depotentiation of synapses was implemented.
doi:10.1371/journal.pcbi.1001104.g004

the strength of $j_{cs,B}$ became larger than $j_{cs,A}$ (Fig. 4D), the activity of population B dominated and, hence, the response of population A was suppressed (Fig. 4B).

The differential activation of two neuronal sub-populations in two different contexts can be interpreted as fear (population A) and extinction (population B) neurons as observed in [17]. This is purely a functional characterization of the two sub-populations, which are identical otherwise. That is, we used exactly the same parameters for both sub-populations and the differential activation results solely from differences in contextual inputs they receive. Thus, the two populations were not different in terms of their intrinsic properties. Of course, cases where the two subpopulations do have different properties can be easily accommodated in the model resulting in an enhancement of the differential activation. To be consistent with [17], we used the terms fear and extinction neurons to refer to those subpopulations that are active in context_A and context_B respectively.

Note that we did not include any component that imitates behavioral output, i.e. freezing. Instead, we assume, in agreement with experimental findings [17], that high activity of fear neurons directly corresponds to a high level of freezing whereas high activity of extinction neurons and low activity of fear neurons corresponds to low levels of freezing.

Spiking neuron network model of BA

Although a simple firing rate model was able to account for the dynamic emergence of fear and extinction neurons, such mean-field models have only limited explanatory and predictive power. For instance, they assume uncorrelated activity in the underlying neuronal populations and, thus, cannot be used to predict any correlations in firing rate or spike timing that may emerge in the network. In addition, these models cannot be used to predict the spike patterns of individual neurons. Thus, to understand the dynamics of the BA network beyond average firing rates only, we simulated a biologically realistic large-scale network composed of spiking neurons. Again, fear conditioning and extinction were simulated by applying five CS-US presentations in context_A and six CS presentations in context_B respectively. In the two different environments tonic contextual input was provided to EXC neurons (cf. Models).

The results of the simulation are presented in Fig. 5. Initially, all EXC neurons spiked at very low firing rates. Presentations of the CS-US led to a steady increase in the firing rates of one sub-population (fear neurons) within the EXC population, which peaked at the end of conditioning (Figs. 5A, E amber dots). The increase in activity of fear neurons was a direct consequence of the potentiation of CS and contextual inputs onto fear neurons (Figs. 5 G, I; amber triangles). In context_B, the fear neurons still responded with high firing rates upon the first CS presentation, even though they did not receive contextual inputs (Figs. 5A, F). With further CS presentations, however, $j_{CS,B}$ synapses became potentiated (Eq. 6, Figs. 5H, J; cyan dots), causing a steady increase in the firing rate of the second sub-population of neurons (extinction neurons) (Figs. 5A, F; cyan dots). The increased recurrent inhibition in the network then caused a decrease in the activity of the fear neurons (Figs. 5A–C, F). At the end of extinction, the population rate of the extinction neurons peaked, whereas the firing rate of the fear neurons had returned to the initial, pre-conditioning values. The reduction of fear neurons activity was further facilitated by small depotentiation of CS and contextual input synapses onto the fear neurons (Eq. 6, Figs. 4H, J; amber triangles). Note that depotentiation of CS synapses onto extinction neurons also occurred during conditioning (Fig. 5G) as described by the learning rule. By contrast, CTX synapses were not decreased during conditioning, because their initial values were close to the lower bound (w_-) (Fig. 5I). During conditioning and extinction the baseline firing rates increased as well (Fig. 5A). This increase was induced by the strengthening of the contextual inputs (Figs. 5H, I), providing an explanation for contextual conditioning. However, because only a small percentage of neurons exhibited this increase in firing rates, this could make it difficult to measure it experimentally. This fact reveals a key advantage of network models which allow for simultaneously sampling a large number of neurons. Based on this, predictions can be inferred which otherwise would not have been possible.

Note that, again, the assignment of BA EXC neurons in fear and extinction sub-populations is purely a functional one. That is, neurons were characterized post-experiment as fear or extinction cells depending on whether they responded to the CS after conditioning or after extinction training respectively. In particular, they were not predetermined in terms of their intrinsic properties and the two sub-populations resulted solely from the differences in the contextual inputs they received.

Also, it is important to emphasize that whereas the population rates of fear and extinction neurons increased gradually during conditioning and extinction training respectively, this was not the case for individual neurons. Instead, they changed their state quite abruptly from non-responding to responding (Fig. 6A). The further

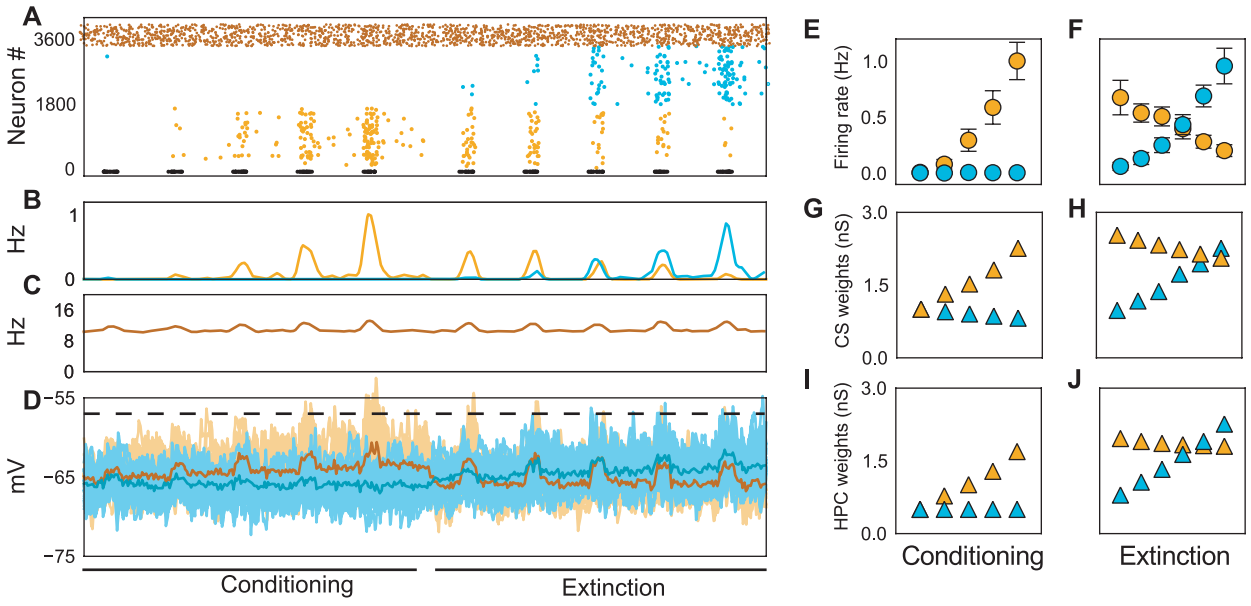

Figure 5. Dynamics of the spiking network model of BA. Five US-CS stimulations were used for conditioning and six CS stimulations for extinction. (A) Raster plot of the spiking activity in the SNN. Two different subpopulations of excitatory neurons (EXC) within the BA emerged during conditioning and extinction simulation, respectively. Neurons 1–1,700 correspond to the fear neurons (amber) and neurons 1,701–3,400 to the extinction neurons (cyan). Dark amber dots show the activity of inhibitory neurons. Short horizontal lines (black) at the bottom of the raster plot represent US-CS presentations (50ms) in fear conditioning and only CS input during extinction training. Observe that activity increased in the fear neurons during conditioning and was then suppressed during extinction training by a steady increase in the activity of the extinction neurons. Note that in contrast to the firing rate model, inhibitory neurons were explicitly simulated and played a critical role in the mutual competition between fear and extinction neuron subpopulations. The decline of fear neurons activity reflects the combined effects of active inhibition and unlearning (LTD at CS and CTX afferents). The increase of the baseline rates provides an explanation for contextual conditioning. Only 5 out of 6 CS stimulations shown for extinction training. (B) Firing rates of fear (amber) and extinction (cyan) neurons. (C) Firing rate of inhibitory neurons. Similar to the rate model, the rates of excitatory and inhibitory neurons remained low and far below the saturation level, determined by the refractoriness of the neurons. (D) Average free membrane potentials (spiking prohibited) of 100 randomly recorded fear (amber) and extinction (cyan) neurons. Twenty traces for individual neurons from fear (light amber) and extinction neurons (light cyan). The dashed black line shows the average spiking threshold (−57 mV) of the excitatory neurons. (E,F) Evolution of the firing rates of fear and extinction neurons during fear conditioning (E) and extinction (F). Firing rates were averaged over 30 simulations. (G,H) Evolution of synaptic weights of CS afferents onto EXC neurons during fear (G) and extinction (H) training, for the simulation shown in (A). (I,J) Same as in (G) and (H) but for the weights of CTX afferents to EXC neurons.
doi:10.1371/journal.pcbi.1001104.g005

the training advanced, the more neurons started to respond. Hence, the gradual increase in population rates (Figs. 5E, F) reflects the growing recruitment of responding neurons, rather than a gradual increase of single neuron activity itself (Fig. 6). The responsive neurons fired maximally two spikes per CS presentation. The baseline firing rates for the inhibitory population were normally distributed with a mean of 10 Hz, whereas the CS-evoked rates shifted their distribution towards a mean of 20 Hz. This is consistent with the neuronal firing patterns *in vivo* reported by [17].

Persistent neurons

Although we performed our main simulations using separated contextual inputs to distinct neuronal subpopulations within the BA (cf. Models), this is not a necessary requirement of the model. In fact, performing simulations with varying amounts of contextual input overlap showed that fear and extinction neurons still existed as distinct populations, even when contextual inputs had an overlap of around 50% (Fig. S1). In addition, the simulations revealed the existence of a third sub-populations of neurons. These were the neurons receiving inputs in both contexts and, thus, were active during both fear conditioning and fear extinction (so called persistent neurons). Note that, similar to the case of fear and extinction neurons, the characterization of cells as "persistent" is functional and denotes the fact that these neurons were responding

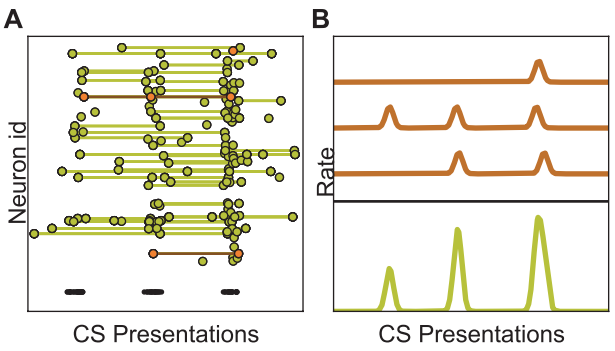

Figure 6. Single-neuron activity during fear and extinction. (A) Spikes (dots) of excitatory neurons for three consecutive CS presentations (short black lines). Spikes from the same neurons are connected by thin lines. Different neurons started to respond to the CS with 1–2 spikes at different points in time. Thus, the state of individual neurons did not change gradually, but quite abruptly from non-responding to responding. (B) The firing rates of three neurons from (A) (amber) that started responding to the CS at three different times. The average response of these three neurons yielded a gradual increase of activity (green), which does not reflect the abrupt response changes at the single-neuron level.
doi:10.1371/journal.pcbi.1001104.g006

to the CS during both conditioning and extinction. Moreover, these neurons had much stronger CS and CTX synapses, which resulted in higher firing rates. This observation of the model is consistent with the experimental data [17], suggesting that conditioning and extinction are not affected by overlapping inputs, unless the overlap is high (>50%).

Renewal and extinction over-training

Following extinction training in context_B, presentations of the CS in the original fear conditioning context (context_A) resulted in context-dependent renewal (ABA renewal) of conditioned fear responses [2]. This renewal phenomenon points at two important aspects of possible neural mechanisms underlying fear extinction: (i) extinction is mainly new learning and only partly unlearning of previously acquired fear memories ([57]; see also Discussion), (ii) extinction learning is context-dependent.

We simulated ABA renewal by changing context at the end of extinction (Fig. 7). This resulted in a sudden switch of activity between fear and extinction neuronal subpopulations. That is, although the activity of extinction neurons was high after extinction learning, the contextual switch caused the activation of fear neurons and a significant drop in the extinction neurons activity. These results are in complete accordance with the experimental findings reported by [17].

It is important to note that this rapid activity switch is purely a network phenomenon and not an effect of synaptic plasticity, as the change is much too fast for the plasticity mechanisms to act. We illustrate this point by depicting the average membrane potentials of 100 randomly selected fear and extinction neurons (Figs. 5D, 7D; amber and cyan traces respectively). It is evident that in either context there was a clear difference between the membrane potentials of the two cell populations, stemming from the fact that one of the populations continuously received a higher excitatory drive due to the additional contextual input. Switching contexts led to a corresponding instantaneous switch in the assignment of the contextual input and, hence, in opposite shifts in the average membrane potentials of the two sub-populations, which was immediately reflected in corresponding shifts in the firing rates.

We also modeled the case where the renewal context was different from both the conditioning and the extinction context (ABC renewal). The results of the simulations revealed that if after extinction training the CS was presented in a third, different context_C, fear neurons became rapidly active again and suppressed extinction neurons (Fig. S2 middle). However, our model also indicated that the absolute response of fear neurons - and thus the magnitude of the fear response- would be weaker than in the ABA case. The reason is that in context_C CTX synapses had not been strengthened during the conditioning procedure. This provides an account for the experimentally observed ABC renewal [58,59] explaining why ABC renewal may occur in the first place and also why the effect may be weaker compared to ABA renewal.

Moreover, our simulations also suggested that massive extinction (extinction over-training) in context_B can abolish ABC renewal, because depotentiation of CS and CTX_A afferents onto BA neurons yield less excitatory input to these neurons. Extinction over-training can also impair ABA renewal, although to a lesser extent (Fig. S2 right). The reason that ABA renewal is more robust and ABC renewal more vulnerable to massive extinction stems from the fact that in the latter case not only CS synapses onto fear neurons are weakened, but also potentiated CTX synapses are entirely missing. These findings are in agreement with and provide a possible explanation for the experimentally observed effects of massive extinction [60].

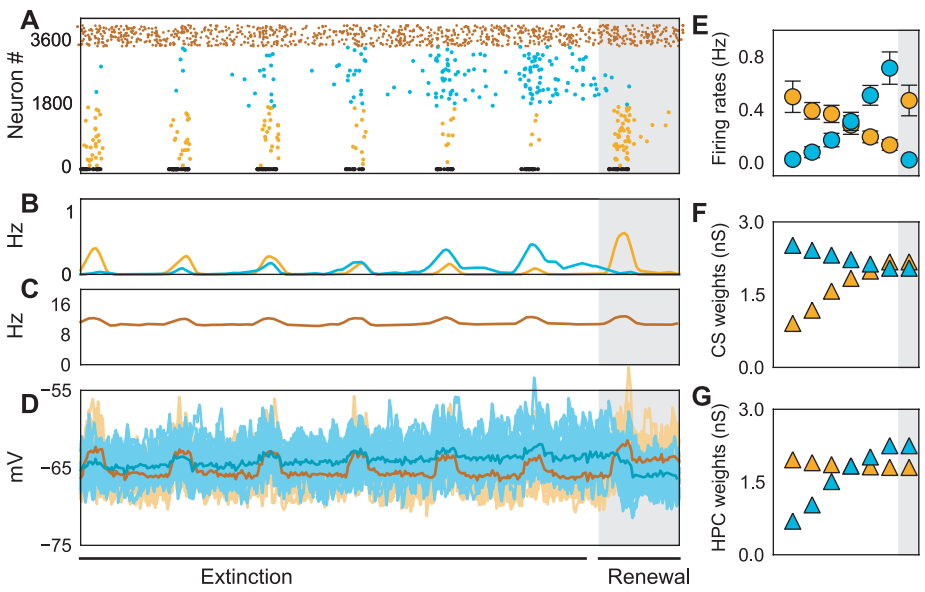

Figure 7. ABA fear renewal. (A) Spiking activity of fear (amber), extinction (cyan) and inhibitory (dark amber) neurons during extinction and renewal (gray shaded region). (B) Average firing rate of fear (amber) and extinction (cyan) neurons. (C) Average firing rate of inhibitory neurons. (D) Free membrane potential of fear and extinction neurons (cf. Caption Fig. 5 for more details). (E) Average firing rates of fear and extinction neurons during extinction and renewal (gray shaded region). (F) Evolution of synaptic strengths of CS afferents onto fear (amber) and the extinction (cyan) neurons. (G) Evolution of synaptic strengths of CTX_A afferents onto fear (amber) and CTX_B afferents onto the extinction (cyan) neurons. Switching context after extinction led to an instantaneous switch of activities between fear and extinction neurons (shaded gray regions in panels (A,B,E). In the initial conditioning context_A fear neurons dominated (due to context_A specific additional excitatory drive) and suppressed extinction neurons. Note that there was no change of weights during renewal (gray shaded area in panels F,G) revealing that the switch of activity was purely a network effect.
doi:10.1371/journal.pcbi.1001104.g007

Extinction of contextual conditioning

Although we did not focus on extinction of contextual fear, it is important to note that our model also accounts for this specific conditioning phenomenon. Indeed, the plasticity rule dictates that in the absence of the CS synaptic weights will decay. That is, CTX synapses, which had been strengthened during conditioning in context_A and encode contextual fear, will depotentiate in the same context if the CS is not present. This will yield decreased fear neuron activity and, thus, extinction of contextual fear. Note that within the framework of our model, this form of extinction is truly unlearning and not masking of contextual fear memories.

High connectivity introduces gamma oscillations

The experimentally reported connection probabilities from excitatory to inhibitory neurons as well as among inhibitory neurons in the BA are around 0.5 [61]. This is a much higher value than the ones we used in our initial simulations (Figs. 5–7, Table S1E). To test the effects of such higher connectivity, we performed additional simulations adopting the experimentally reported values for the connection probabilities. The qualitative behavior of the model did not change (data not shown). However, a new aspect in the network dynamics emerged. High frequency oscillations - typically in the gamma range (30–80Hz) - occurred throughout the simulation in both excitatory and inhibitory populations. These oscillations were present already in the ongoing activity patterns and CS-US presentation enhanced them even further (Fig. 8A). They resulted from the high shared connectivity and, hence, large amount of shared inputs that caused correlated spiking in the neurons. The oscillation frequency was determined by synaptic time constants and delays in the network.

Gamma oscillations in networks of excitatory and inhibitory neurons have been reported in many experiments [62–67] and discussed in multiple theoretical studies [68–75]. Moreover, several studies have reported gamma oscillations in the amygdala under certain conditions, e.g. in anesthetized animals [76], in slow wave-sleep [77], in the presence of reward predicting stimuli [78] and in paradigms involving consolidation of emotional memories [79]. Therefore, there is at least partial experimental and theoretical support for the gamma range oscillations observed here in high connectivity BA network simulations.

Yet, in networks with high mutual connectivity between excitatory and inhibitory neurons and within inhibitory neuron populations such as in the BA, oscillations should be a prevailing feature and should, therefore, be readily identifiable *in vivo* recordings under all conditions and not only in the special cases mentioned above. It is, thus, possible that certain mechanisms operate in the BA that could dampen gamma oscillations (but see Discussion). We, therefore, used our network model to investigate this issue in further simulations by exploring the parameter space of the network properties that could quench oscillations. Two mechanisms proved to be effective in reducing the power of gamma oscillations. The first one was the introduction of heterogeneity in the inhibitory population [80,81]. This approach was motivated by experimental data showing that interneurons in the BA exhibit a large diversity in terms of their morphological and electrophysiological properties [24], similar to interneurons in the cortex [82] and hippocampus [83]. In the latter case, the diversity was expressed in a wide range of values for synaptic rise times, reversal potentials, response latencies etc. In a preliminary study [84], we introduced heterogeneity in one of the neuronal properties in our model, the spiking threshold, by drawing values from a bimodal distribution with peaks at -35 mV and -28 mV. This corresponds to the experimentally measured threshold values of two subclasses of parvalbumin-expressing interneurons in the BA: the fast-spiking (FS) and the delayed-firing (DF) interneurons [24,61]. In such heterogeneous networks, oscillations were indeed reduced, but not totally eliminated [84].

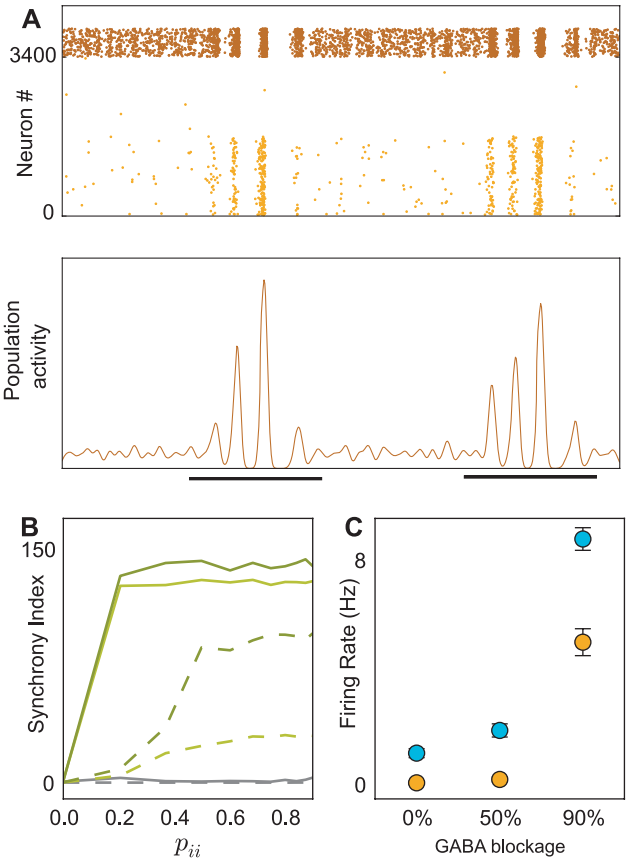

Figure 8. Synchrony and oscillations in the BA network model. (A) Spiking activity of a densely connected spiking neural network. A high connection probability from EXC to INH neurons as well as among INH neurons resulted in a population wide synchronization and gamma range oscillations, which were more salient during CS stimulation (black lines). (B) The effects of synaptic delays on the synchrony within the inhibitory population shown for synaptic strengths of 1 nS, 2 nS and 3 nS in gray, light and dark green, respectively. When synaptic delays were drawn from a uniform distribution [1–2 ms], increasing connectivity beyond 0.2 enhanced the oscillations and synchrony to their maximum (solid lines). Only for weaker synapses (1 nS) synchrony did not increase with connection probabilities, because the network was mainly input driven. Synchrony was significantly reduced for all synaptic strengths when synaptic delays were drawn from a uniform distribution in a lower delay range [0.2–1 ms] (dashed lines). However, note that for high connectivity (>0.4) and strong synaptic couplings, synchrony was always present. (C) Inactivation of INH neurons during extinction led to increased firing rates in both fear and extinction neurons (x-axis, percentage of inactive INH neurons). In context_B, extinction neurons received additional contextual input and, thus, their activity was much higher (cyan) than that of fear neurons (amber). Independently of the behavior of the animal, the results in panel C suggest a simple experiment (i.e. blocking GABAergic synapses) to test the validity of the mechanism we propose here (see text, blockage of inhibition).
doi:10.1371/journal.pcbi.1001104.g008

A second, more effective way to reduce the network oscillations was to decrease the synaptic delays between inhibitory neurons (Fig. 8B). First, we studied the oscillations dynamics for different connection probabilities in a network of homogeneous neurons, with synaptic delays drawn from a uniform distribution (1–2 ms). In such networks, increasing connectivity (>0.2) enhanced the oscillations and synchrony to their maximum (Fig. 8B, green solid

lines). Only for very weak synapses (1 nS), that is, when the network was mainly driven by external inputs, increasing the connectivity did not add to the oscillations (Fig. 8B, gray solid line). Increasing the width of the synaptic delay distribution did not reduce the synchrony and oscillations in high-connectivity networks (data not shown, Vlachos et al. in prep.). However, choosing short delays from a narrow uniform distribution (0.2–1 ms) considerably reduced the oscillations, up to connection probabilities of 0.4 (Fig. 8B, green and gray dashed lines). Thus, in a recurrent network, smaller delays have a powerful effect in reducing synchrony and oscillations. This finding is in agreement with a previous numerical study [69] and also with more recent analytical approaches [74,75,85]. At first sight, synaptic delays less than 1 ms might appear unrealistically small. However, delays as short as 0.5 ms have been reported among inhibitory neurons in the hippocampus [66]. Moreover, the delays between inhibitory neurons in the BA have been reported to be around 0.7 ms [61], or even smaller (Lüthi, unpublished data). Therefore these short delays, might indeed account for the lack of gamma band oscillations observed under baseline conditions in experimental recordings.

Blockage of inhibition

Because inhibition plays a critical role in our model, we tested the effects of partial inactivation of inhibitory neurons. For this, we performed two additional sets of simulations, in which, during acquisition of extinction, we deactivated 50% and 90% of INH neurons, respectively. The results are shown in Fig. 8C. As expected, with reduced inhibition the activity of both fear and extinction neurons increased. The increase of activity of the latter population was more pronounced, due to the fact that it received additional excitatory drive from contextual inputs in context_B. This suggests that blockage of inhibitory activity should lead to enhanced, context-specific extinction. This is consistent with the finding that GABA blockage enhances extinction of contextual freezing [86]. However, there is a potential caveat here. Activity of both fear and extinction neurons is increased upon blockage of inhibition and it is not clear how downstream structures, specifically CEA neurons, would respond to this. If the relative difference between fear and extinction neuron activity matters, then extinction should be facilitated by impaired inhibition. If, by contrast, the ratio between fear and extinction neuron activity is more relevant, then extinction might be impaired. Note that these two possibilities apply to both blocking of inhibition during acquisition of extinction training and blocking of inhibition during expression of fear extinction.

Removal of contextual input

Because contextual input is one of the key aspects in our model we tested how removal of these inputs would affect the behavior of the network. The simulations yielded two different results depending on the exact time point of removal of contextual inputs. When contextual inputs were removed after fear conditioning, fear neurons remained active during extinction training and no extinction neurons emerged (Fig. S3; left). This result is a direct consequence of our synaptic learning rule, because strengthening of synapses requires temporal overlap of CS and CTX inputs. Note that although fear neurons remained active, their firing rates were reduced, because they now lacked contextual input. Thus, our model suggests that lesions of hippocampal or prefrontal areas after fear conditioning may result in impaired extinction. This conclusion is supported by experimental evidence [87].

By contrast, when contextual inputs were removed after fear extinction, activity of neither fear nor extinction neurons was sufficiently strong to suppress the other neuron group (Fig. S3; right). That is, because the decisive contextual input was lacking and, thus, both groups were simultaneously active, although to a lesser degree than in case either group was active alone. The behavioral consequences of these results are beyond the scope of our model, because here we did not model any downstream structures such as the central amygdala that presumably further process output from fear and extinction neurons. Thus, at present, we can only speculate that lesions of hippocampal or prefrontal areas after extinction training may result in impaired renewal, because fear neuron activity will be both decreased and also counteracted by simultaneous extinction neuron activity. In fact, experimental evidence supports this conclusion [88]. However, it is important to point out a subtle difference between our model and certain lesion experiments. In our model, removing CTX input means that the BA network does not receive any contextual input *at all*. By contrast, in some experiments in which the hippocampus had been lesioned or inactivated, contextual information may still have been accessible, because the context in which the CS was presented was still decisive for the behavioral outcome [89,90].

Predictions

Our model enables us to make a number of specific predictions that can be tested experimentally:

1. We predict that there is convergence of CS and CTX inputs onto the same BA EXC neurons. This prediction can be tested in multiple ways. One way is to look for anatomical connectivity. This can be achieved by employing pathway tracing studies. Note, however, that this approach can reveal stimulus convergence only on non-specific BA neurons, because fear and extinction neurons are only behaviorally determined in vivo. Alternatively, one could use optogenetic tools to activate/inactivate the LA, HPC or PFC while simultaneously performing intracellular recordings of BA neurons. If there is stimulus convergence, then activation (inactivation) of the connected structures would result in an increase (decrease) of the mean and/or variance of the membrane potential. A second way would be to test for functional connectivity, e.g. by using an imaging technique akin to the one used in the LA [91]. Again, these approaches can only reveal convergence on non-specific BA neurons. However, using electrophysiological techniques, the covariance of the spike rates between LA, HPC/PFC and *identified* BA fear and extinctions neurons could be measured while, again, selectively activating/inactivating LA or HPC/PFC. Functional convergence of CS and CTX inputs to BA fear and extinction neurons should be reflected in associated changes of the covariances.
2. Extinction over-training has a dual effect: (i) CS and context_B afferents to extinction neurons will become very strong, thereby enhancing the suppression of fear neurons. (ii) At the same time, depotentiation of CS and context_A afferents onto fear neurons will substantially decrease the excitatory input to these neurons. This should be visible during renewal, where presentation of the CS in context_A or context_C should lead to only a weak fear response (Fig. S2; right). Existing experimental results seem to support this potential effect of massive extinction [60]. Here we predict, in addition to these behavioral findings, the state of 'hidden' or not directly observable variables (large synaptic weights) and explain how they may affect direct observables (enhanced extinction).

3. Conditioning and extinction training increase the excitatory and inhibitory inputs to principal neurons in the BA. This results in stronger fluctuations of their membrane potentials. Thus, the variance of the recorded membrane potential fluctuations of BA neurons in trained animals should be higher than in naive animals.
4. Blockage of inhibition in $\mathrm{context}_B$ results in elevated activity in both fear and extinction neurons, with extinction neurons spiking at a higher rate than fear neurons. It is not clear how this may impact on downstream structures such as the CEA. Irrespective of the behavioral outcome, a specific experiment addressing the effects of GABA blockage on fear and extinction neurons activity would provide more insights into this issue.
5. The strength of fear (extinction) memories is directly proportional to the strength of contextual inputs to the BA during conditioning (extinction) training, respectively. If salience of environmental features translates to higher firing rates or increased number of contextual inputs to the BA, then renewal (extinction) will be enhanced. Along the same lines, the bigger the differences between conditioning and extinction context, the more extinction should be facilitated, because the overlap of contextual inputs to the BA between the two contexts will be minimized.

Discussion

Here we presented for the first time a large-scale network model of the BA addressing the question how contextual inputs may shape the activity of distinct sub-populations of BA neurons. Although we started from a very specific experimental data-set [17], we implemented a network model that has more far-reaching implications. That is, the results of the simulations together with the model architecture provide, non-trivial and experimentally testable new insights into potential neural mechanisms underlying cued and contextual fear conditioning and extinction, ABA and ABC renewal, and extinction over-training. In addition, a specific and important function of inhibition is sketched as a mechanism that could enable mutual competition between fear and extinction memories. These results allow us to provide a synthesis of several experimental findings and to propose a role for the BA as a nucleus that integrates information about the CS and the context. This brings it into the position to provide a context-dependent instruction to downstream structures, enabling the switching of states from low to high fear and vice versa.

Specifically, we propose that context modulates neuronal activity within the BA, resulting in the formation of associations between CS, US and context in this nucleus. During fear conditioning the $\mathrm{CS}-\mathrm{US}-\mathrm{context}_A$ representation signals danger and causes a high fear state. During extinction, the newly formed $\mathrm{CS}-\mathrm{context}_B$ representation signals safety and suppresses the fear state. Back in the conditioning context, the initial representation dominates again (renewal). Thus, as far as neural mechanisms within the BA are concerned, conditioning and extinction could be understood as mutual competition between different representations of fear and safety. Partial unlearning or erasure may also occur, although to a limited degree.

Memories are assumed to be stored in a distributed manner in the brain [92]. Consistent with this view, fear-related memories may also be distributed among different nuclei within the amygdala and brain regions connected to it [10]. Our model suggests that context-related features of these distributed fear memories are represented in the BA. Thus, inactivation of the BA would impair context-related aspects of fear and extinction memories, whereas non-contextual features, represented in LA or CEA, would remain unaffected [17].

Sources of contextual inputs

One core feature of our model is that contextual inputs are gated to the BA. In this framework, the precise origin of these inputs does not matter; as long as the BA neurons receive differential inputs in two different contexts, the model behavior remains unaltered. However, there are strong indications from anatomical [29,93,94] and physiological [93] studies that the HPC is a major source of contextual information to the BA. In addition, a previous report showed context-dependent modulation of neuronal activity in the LA [95]. By designing our model to have contextual input directly influencing the activity of excitatory neurons in the BA, we have essentially postulated a similar mechanism for this subnucleus. This assumption is further supported by the finding that fear neurons show orthodromic responses to HPC stimulation [17].

A second source of contextual input may be the mPFC. There is anatomical evidence that the mPFC projects to the BA [29]. Moreover, [17] reported that mPFC stimulation induces orthodromic responses in identified extinction neurons. Here, we suggest that part of the information conveyed by these projections might be contextual. This assumption is based on evidence reporting extinction-related induction of LTP on hippocampus-mPFC afferents [96]. In our model both fear and extinction neurons receive context-specific information either directly from hippocampus or indirectly via the mPFC. This may also explain the ambiguous results that the hippocampus may or may not interact with the mPFC during extinction learning [97].

The context-specific modulation of activity in the BA presented here provides a general framework that can explain experimental findings on the involvement of the hippocampus in the acquisition, encoding, and context-dependent retrieval of both conditioning [13,98,99] and extinction memories [3,87]. Future refinements of the model, in combination with new experimental data are necessary for a better understanding of the detailed interactions between hippocampus, mPFC and amygdala.

Gamma oscillations

We showed that high connectivity between excitatory and inhibitory and within inhibitory neuron populations results in robust oscillations in the gamma range, characterized by high activity correlation among neurons. The main cause of these oscillations was the high degree of shared inputs among neurons as a result of the dense connectivity. We suggested two different, biologically plausible ways to reduce these oscillations: by either introducing heterogeneity in neuron properties and/or by reducing synaptic delays to sub-millisecond time scales. Yet another way would be to have synapses exhibit a certain transmission failure rate [100,101], resulting in activity dependent reduction of the effective connectivity. However, we do not wish to imply that gamma oscillations do not exist in the BA. In fact, as noted earlier, gamma oscillations have been reported in the amygdala under various conditions [76–79]. Here, we want to emphasize the point that in networks with high connectivity, gamma range oscillations are a salient feature of the network dynamics. Therefore, they should be visible even in the ongoing activity, unless suppressing mechanisms, such as those elaborated here, are in effect.

Several suggestions for a specific role of gamma oscillations have been made in the past. For instance, it has been proposed that in the cortex or the hippocampus oscillations might contribute to temporal encoding [102], sensory binding [103], attentional

selection [104] and memory formation or retrieval [105,106]. It is currently unclear whether these hypotheses also apply to the amygdala. Oscillations in lower frequency ranges (delta and theta) have also been reported. For example, increased theta oscillations - that synchronized with hippocampal theta activity - were shown to be related to conditioned freezing [107,108], whereas delta oscillations have been implicated in gating aversive stimuli [109]. Gamma oscillations, on the other hand, have been suggested to facilitate interactions between the amygdala and connected structures [78,110].

Here, because we modeled only the BA, we cannot give any informed predictions about how gamma oscillations may affect those various interactions. Moreover, in our current model, we have used plasticity only in the input connections and those are not affected by oscillatory activity in the recurrent network. However, before addressing the effects of gamma oscillation on the dynamics of the BA network, it is of key importance to resolve experimentally whether gamma oscillations are indeed present in BA activity and, if so, under which conditions.

Conditioned inhibition

A well-known behavioral phenomenon is conditioned inhibition, referring to the ability of a second CS (CS−) to suppress the conditioned response, after it has been paired several times with the first CS (CS+) in the absence of a US [57,111]. It is possible that the CS−, referred to as conditioned inhibitor, employs similar mechanisms to those described in our model to suppress the conditioned response. That is, neural subpopulations in the BA encoding the CS− might, similar to extinction neurons, use local inhibitory circuits to suppress fear neuron activity. Future work is needed to explore further this interesting line of reasoning.

Conditioning and extinction in the same context

Our model accounts for experimental paradigms that use a different extinction context from the conditioning one, but not for those in which fear conditioning and extinction occur in the same context. For instance, if conditioning and extinction both occur in context_A, then only those neurons that receive inputs in this context will be active. Thus, downstream structures will not be able to differentiate between fear conditioning and extinction training solely from spiking activity in the BA. It is evident that performing conditioning and extinction in the same context *per se* increases ambiguity about the meaning of the context. Thus, it is likely that circuits within the BA alone are not sufficient to solve this computational problem. Both, a detailed description of neural activity during this type of extinction and a more detailed analysis of interactions between the BA and downstream structures are required to address this behavioral phenomenon.

Relation to other models

Although a wealth of experimental studies exist on the amygdala and its role in fear conditioning and extinction, computational or theoretical approaches to study amygdala function are largely lacking. Most of the previous theoretical studies involve symbolic models [112,113], mainly based on the Rescorla-Wagner rule [114]. These models have their merit in describing behavioral findings such as generalization, blocking etc. However, since these models treat the amygdala as a "black-box", it is not within their scope to account for neuroanatomical or electrophysiological data, therefore providing little insight into the underlying neuronal mechanisms involved. Despite these apparent differences, it is still possible to draw some parallels to symbolic models. For instance, in our model, potentiation of synapses occurs only if CS and CTX inputs temporally overlap. This is similar to the SOP model, where US and CS have to coincide for strengthening of associations to take place [115,116].

Connectionist or parallel-distributed (PDP) models of fear related processes go one step further than symbolic models by introducing networks composed of multiple, mutually connected computational units. One such model was successful in capturing certain features observed in fear conditioning studies [117]. Its main limitation, however, is the fact that it does not take into account the different substructures within the amygdala, nor do the computational units used in the model map to any biophysically realistic counterparts.

Fortunately, the computational power presently available allows us to improve these models and to overcome many of their limitations. The model presented here is to our knowledge the first large-scale spiking neuron network model that investigates the mechanisms of fear conditioning and extinction within the amygdala using biologically realistic neurons in adequate detail. The model closest to this is a compartmental model introduced by [118] to investigate the function of the LA in fear conditioning and extinction. However, [118] used a small network composed of only eight two-compartment neurons and focused on role of the kinetics of multiple ionic currents in fear conditioning and extinction. By contrast, we modeled the BA using a large network of 4000 LIF neurons, which enabled us to identify the network level interactions involved in the formation of fear and extinction memories.

Conclusions

The present model provides a plausible explanation for the neural mechanisms underlying fear conditioning and extinction within the BA. We did not address the question of how the neural activity within the BA impacts on downstream structures, such as CEA or mPFC. We neither attempted to model the interactions between hippocampus and mPFC in conditioning and extinction, which would require additional experimental data to constrain the possible models. Given these restrictions, we provided a plausible mechanism of how contextual inputs may affect the activity of distinct neuronal subpopulations in the BA, enabling them to control downstream structures such as the CEA. We proposed that context-related aspects of fear and extinction memories are partially stored in the BA and that they provide a context-dependent instruction for the triggering or blocking of the fear-response. In addition, we showed how extinction training may mask previously acquired fear memories and, thus, provided an account for renewal. Finally, our model, next to yielding several interesting predictions discussed above, raises the important question of how downstream structures such as the CEA or mPFC discriminate the activity of the distinct neuronal subpopulations within the BA. Is this problem solved purely on an anatomical level, e.g. by differential projections of the BA subpopulations to specific target neurons? Or do specific features in the activity of the BA subpopulations, e.g. the statistical structure of pairwise or higher-order correlations, also play a role, providing downstream networks with a mechanism to distinguish between them? These questions need to be addressed in future work combining experimental and theoretical approaches.

Supporting Information

Figure S1 Overlapping contextual inputs. (A) Venn diagram illustrating overlap of CTX inputs. (B) Activity of BA neurons at the end of extinction training. Varying the overlap of CTX inputs to BA neurons from 0–100% resulted in a third subgroup, which was active both during conditioning and extinction (persistent

neurons). The higher the overlap, the more neurons in this group were active, as is reflected in the population rate. Note, that fear and extinction neurons still existed with the latter ones suppressing the former ones, even for an overlap of >50%.
Found at: doi:10.1371/journal.pcbi.1001104.s001 (0.13 MB EPS)

Figure S2 ABC renewal and extinction over-training. After fear extinction in context B the activity of extinction neurons (cyan) was high, thereby suppressing the activity of fear neurons (amber) (left). Presenting the CS in the conditioning context-A resulted in a rapid switch of activity, with fear neurons suppressing extinction neurons activity (ABA renewal; see also Figure 7 and main text). When the CS was presented in a different context-C, then again fear neurons were active (ABC renewal). However, their activity was decreased compared to ABC renewal, because potentiated contextual input was now missing (middle). Our model suggests that extinction over-training can abolish both ABA and ABC renewal (right; see text).
Found at: doi:10.1371/journal.pcbi.1001104.s002 (0.06 MB EPS)

Figure S3 Removal of CTX input. (left) Effects of removal of contextual input on fear and extinction neurons activity. When contextual inputs were removed after conditioning, then during extinction simulation no new group of neurons became active and, therefore, fear neurons remained active, as they were the only group of neurons for which CS inputs had been potentiated during conditioning. Thus, the model predicts that lesions of hippocampal or prefrontal areas after conditioning may result in impaired extinction. (right) When contextual inputs were removed after extinction, then both fear and extinction neurons were active and no group was able to suppress the other. However, the firing rates of both groups were decreased, because contextual input was lacking. The behavioral implication of these results may be impaired ABA as well as ABC renewal (see main text).
Found at: doi:10.1371/journal.pcbi.1001104.s003 (0.12 MB EPS)

Table S1 Model parameters. Parameters used in the rate and SNN model.
Found at: doi:10.1371/journal.pcbi.1001104.s004 (0.13 MB PDF)

Acknowledgments

We thank Philippe Gastrein and Francois Grenier for providing data regarding the electrophysiological properties of BA neurons. We also thank Man-Yi Yim and Ajith Padmanabhan for proof-reading the manuscript. Finally, we would like to thank the two anonymous referees for their valuable comments throughout the reviewing process.

Author Contributions

Conceived and designed the experiments: IV CH AL AA AK. Performed the experiments: IV AK. Analyzed the data: IV CH AK. Wrote the paper: IV CH AL AA AK.

References

1. Pavlov I (1927) Conditioned reflexes. Oxford, UK: Oxford University Press.
2. Bouton ME, Bolles RC (1979) Contextual control of the extinction of conditioned fear. Learn Motiv 10: 445–466.
3. Bouton ME, Westbrook RF, Corcoran KA, Maren S (2006) Contextual and Temporal Modulation of Extinction: Behavioral and Biological Mechanisms. Biol Psychiatry 60: 352–360.
4. LeDoux JE (2000) The amygdala and emotion: a view through fear. Oxford University Press, 2nd edition. 690 p.
5. Myers KM, Davis M (2007) Mechanisms of fear extinction. Mol Psychiatry 12: 120–150.
6. Quirk GJ, Mueller D (2007) Neural Mechanisms of Extinction Learning and Retrieval. Neuropsychopharmacology 33: 56–72.
7. Romanski LM, Clugnet MC, Bordi F, LeDoux JE (1993) Somatosensory and auditory convergence in the lateral nucleus of the amygdala. Behav Neurosci 107: 444–450.
8. Barot SK, Kyono Y, Clark EW, Bernstein IL (2008) Visualizing stimulus convergence in amygdala neurons during associative learning. Proc Natl Acad Sci U S A 105: 20959–20963.
9. Sigurdsson T, Doyère V, Cain CK, LeDoux JE (2007) Long-term potentiation in the amygdala: A cellular mechanism of fear learning and memory. Neuropharmacology 52: 215–227.
10. Paré D, Quirk GJ, Ledoux JE (2004) New Vistas on Amygdala Networks in Conditioned Fear. J Neurophysiol 92: 1–9.
11. LeDoux JE, Iwata J, Cicchetti P, Reis DJ (1988) Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. J Neurosci 8: 2517–2529.
12. Muller J, Corodimas KP, Fridel Z, Ledoux JE (1997) Functional inactivation of the lateral and basal nuclei of the amygdala by muscimol infusion prevents fear conditioning to an explicit conditioned stimulus and to contextual stimuli. Behav Neurosci 111: 683–691.
13. Goosens KA, Maren S (2001) Contextual and Auditory Fear Conditioning are Mediated by the Lateral, Basal, and Central Amygdaloid Nuclei in Rats. Learn Mem 8: 148–155.
14. Anglada-Figueroa D, Quirk GJ (2005) Lesions of the Basal Amygdala Block Expression of Conditioned Fear But Not Extinction. J Neurosci 25: 9680–9685.
15. Maren S, Poremba A, Gabriel M (1991) Basolateral amygdaloid multi-unit neuronal correlates of discriminative avoidance learning in rabbits. Brain Res 549: 311–316.
16. Muramoto K, Ono T, Nishijo H, Fukuda M (1993) Rat amygdaloid neuron responses during auditory discrimination. Neuroscience 52: 621–636.
17. Herry C, Ciocchi S, Senn V, Demmou L, Muller C, et al. (2008) Switching on and off fear by distinct neuronal circuits. Nature 454: 600–606.
18. Ciocchi S, Herry C, Grenier F, Wolff SBE, Letzkus JJ, et al. (2010) Encoding of conditioned fear in central amygdala inhibitory circuits. Nature 468: 277–282.
19. Wilson HR, Cowan JD (1972) Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. Biophys J 12: 1–24.
20. Knierim JJ (2002) Dynamic interactions between local surface cues, distal landmarks, and intrinsic circuitry in hippocampal place cells. J Neurosci 22: 6254–6264.
21. Kentros CG, Agnihotri NT, Streater S, Hawkins RD, Kandel ER (2004) Increased attention to spatial context increases both place field stability and spatial memory. Neuron 42: 283–295.
22. Leutgeb JK, Leutgeb S, Moser MB, Moser EI (2007) Pattern separation in the dentate gyrus and CA3 of the hippocampus. Science 315: 961–966.
23. Repa JC, Muller J, Apergis J, Desrochers TM, Zhou Y, et al. (2001) Two different lateral amygdala cell populations contribute to the initiation and storage of memory. Nat Neurosci 4: 724–731.
24. Sah P, Faber ESL, Armentia MLD, Power J (2003) The amygdaloid complex: anatomy and physiology. Physiol Rev 83: 803–834.
25. Nordlie E, Gewaltig MO, Plesser HE (2009) Towards Reproducible Descriptions of Neuronal Network Models. PLoS Comput Biol 5: e1000456.
26. Prinz AA, Bucher D, Marder E (2004) Similar network activity from disparate circuit parameters. Nat Neurosci 7: 1345–1352.
27. Gerstner W, Naud R (2009) How Good Are Neuron Models? Science 326: 379–380.
28. Morrison A, Mehring C, Geisel T, Aertsen A, Diesmann M (2005) Advancing the Boundaries of High-Connectivity Network Simulation with Distributed Computing. Neural Comput 17: 1776–1801.
29. McDonald AJ (1998) Cortical pathways to the mammalian amygdala. Prog Neurobiol 55: 257–332.
30. Tuunanen J, Pitkänen A (2000) Do seizures cause neuronal damage in rat amygdala kindling? Epilepsy Res 39: 171–176.
31. Kuhn A, Aertsen A, Rotter S (2004) Neuronal Integration of Synaptic Input in the Fluctuation-Driven Regime. J Neurosci 24: 2345–2356.
32. Kumar A, Schrader S, Aertsen A, Rotter S (2008) The High-Conductance State of Cortical Networks. Neural Comput 20: 1–43.
33. Kumar A, Rotter S, Aertsen A (2008) Conditions for Propagating Synchronous Spiking and Asynchronous Firing Rates in a Cortical Network Model. J Neurosci 28: 5268–5280.
34. Izhikevich EM (2007) Solving the distal reward problem through linkage of STDP and dopamine signaling. Cereb Cortex 17: 2443–2452.
35. Kim J, Lee S, Park K, Hong I, Song B, et al. (2007) Amygdala depotentiation and fear extinction. Proc Natl Acad Sci U S A 104: 20955–20960.
36. Hong I, Song B, Lee S, Kim J, Kim J, et al. (2009) Extinction of cued fear memory involves a distinct form of depotentiation at cortical input synapses onto the lateral amygdala. Eur J Neurosci 30: 2089–2099.
37. Bienenstock E, Cooper L, Munro P (1982) Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J Neurosci 2: 32–48.
38. Bear M, Cooper L, Ebner F (1987) A physiological basis for a theory of synapse modification. Science 237: 42–48.
39. Lisman J (1989) A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. Proc Natl Acad Sci U S A 86: 9574–9578.

40. Shouval HZ, Bear MF, Cooper LN (2002) A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. Proc Natl Acad Sci U S A 99: 10831–10836.
41. Artola A, Brocher S, Singer W (1990) Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. Nature 347: 69–72.
42. Artola A, Singer W (1993) Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. Trends Neurosci 16: 480–487.
43. Malenka RC, Bear MF (2004) LTP and LTD: An Embarrassment of Riches. Neuron 44: 5–21.
44. Lisman J, Spruston N (2005) Postsynaptic depolarization requirements for LTP and LTD: a critique of spike timing-dependent plasticity. Nat Neurosci 8: 839–841.
45. Hardie J, Spruston N (2009) Synaptic Depolarization Is More Effective than Back-Propagating Action Potentials during Induction of Associative Long-Term Potentiation in Hippocampal Pyramidal Neurons. J Neurosci 29: 3233–3241.
46. Humeau Y, Shaban H, Bissiere S, Luthi A (2003) Presynaptic induction of heterosynaptic associative plasticity in the mammalian brain. Nature 426: 841–845.
47. Frey U, Morris RGM (1997) Synaptic tagging and long-term potentiation. Nature 385: 533–536.
48. Barria A, Muller D, Derkach V, Griffith LC, Soderling TR (1997) Regulatory phosphorylation of AMPA-type glutamate receptors by CaM-KII during long-term potentiation. Science 276: 2042–2045.
49. Harley CW (2004) Norepinephrine and dopamine as learning signals. Neural Plast 11: 191–204.
50. McGaugh JL (2004) The amygdala modulates the consolidation of memories of emotionally arousing experiences. Annu Rev Neurosci 27: 1–28.
51. Tully K, Li Y, Tsvetkov E, Bolshakov VY (2007) Norepinephrine enables the induction of associative long-term potentiation at thalamo-amygdala synapses. Proc Natl Acad Sci U S A 99104: 14146–50.
52. Faber ESL, Delaney AJ, Power JM, Sedlak PL, Crane JW, et al. (2008) Modulation of SK Channel Trafficking by Beta Adrenoceptors Enhances Excitatory Synaptic Transmission and Plasticity in the Amygdala. J Neurosci 28: 10803–10813.
53. Pinard C, Muller J, Mascagni F, McDonald A (2008) Dopaminergic innervation of interneurons in the rat basolateral amygdala. Neuroscience 157: 850–863.
54. Muller J, Mascagni F, McDonald A (2009) Dopaminergic innervation of pyramidal cells in the rat basolateral amygdala. Brain Struct Funct 213: 275–288.
55. Davison AP, Brüderle D, Eppler J, Kremkow J, Muller E, et al. (2008) PyNN: A Common Interface for Neuronal Network Simulators. Front Neuroinformatics 2: 11.
56. Gewaltig MO, Diesmann M (2007) Nest (neural simulation tool). Scholarpedia 2: 1430.
57. Bouton ME (2004) Context and Behavioral Processes in Extinction. Learn Mem 11: 485–494.
58. Harris JA, Jones ML, Bailey GK, Westbrook RF (2000) Contextual control over conditioned responding in an extinction paradigm. J Exp Psychol Anim Behav Process 26: 174–185.
59. Bouton ME (2002) Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. Biol Psychiatry 52: 976–986.
60. Denniston JC, Chang RC, Miller RR (2003) Massive extinction treatment attenuates the renewal effect. Learn Motiv 34: 68–86.
61. Woodruff AR, Sah P (2007) Networks of parvalbumin-positive interneurons in the basolateral amygdala. J Neurosci 27: 553–563.
62. Gray CM, Singer W (1989) Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. Proc Natl Acad Sci U S A 86: 1698–1702.
63. Buzsaki G, Horvath Z, Urioste R, Hetke J, Wise K (1992) High-frequency network oscillation in the hippocampus. Science 256: 1025–1027.
64. Bragin A, Jando G, Nadasdy Z, Hetke J, Wise K, et al. (1995) Gamma (40–100 Hz) oscillation in the hippocampus of the behaving rat. J Neurosci 15: 47–60.
65. Whittington MA, Traub RD, Jefferys JGR (1995) Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. Nature 373: 612–615.
66. Bartos M, Vida I, Frotscher M, Meyer A, Monyer H, et al. (2002) Fast synaptic inhibition promotes synchronized gamma oscillations in hippocampal interneuron networks. Proc Natl Acad Sci U S A 99: 13222–13227.
67. Bartos M, Vida I, Jonas P (2007) Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. Nat Rev Neurosci 8: 45–56.
68. Lytton WW, Sejnowski TJ (1991) Simulations of cortical pyramidal neurons synchronized by inhibitory interneurons. J Neurophysiol 66: 1059–1079.
69. Erb M, Aertsen A (1992) Dynamics of activity in biology-oriented neural network models: stability at low firing rates. In: Aertsen A, Braitenberg V, eds. Information Processing in the Cortex: Experiments and Theory. Berlin: Springer. pp 477.
70. Aertsen A, Arndt M (1993) Response synchronization in the visual cortex. Curr Opin Neurobiol 3: 586–594.
71. Vreeswijk C, Abbott LF, Ermentrout GB (1994) When inhibition not excitation synchronizes neural firing. J Comput Neurosci 1: 313–321.
72. Traub RD, Whittington MA, Stanford IM, Jefferys JGR (1996) A mechanism for generation of long-range synchronous fast oscillations in the cortex. Nature 383: 621–624.
73. Brunel N, Hakim V (1999) Fast Global Oscillations in Networks of Integrate-and-Fire Neurons with Low Firing Rates. Neural Comp 11: 1621–1671.
74. Brunel N, Wang XJ (2003) What Determines the Frequency of Fast Network Oscillations With Irregular Neural Discharges? I. Synaptic Dynamics and Excitation-Inhibition Balance. J Neurophysiol 90: 415–430.
75. Maex R, Schutter ED (2003) Resonant Synchronization in Heterogeneous Networks of Inhibitory Neurons. J Neurosci 23: 10503–10514.
76. Collins DR, Pelletier JG, Paré D (2001) Slow and fast (gamma) neuronal oscillations in the perirhinal cortex and lateral amygdala. J Neurophysiol 85: 1661–1672.
77. Ponomarenko AA, Korotkova TM, Haas HL (2003) High frequency (200 Hz) oscillations and firing patterns in the basolateral amygdala and dorsal endopiriform nucleus of the behaving rat. Behav Brain Res 141: 123–129.
78. Popescu AT, Popa D, Pare D (2009) Coherent gamma oscillations couple the amygdala and striatum during learning. Nat Neurosci 12: 801–807.
79. Paré D, Collins DR, Pelletier JG (2002) Amygdala oscillations and the consolidation of emotional memories. Trends Cogn Sci 6: 306–314.
80. Neltner L, Hansel D, Mato G, Meunier C (2000) Synchrony in Heterogeneous Networks of Spiking Neurons. Neural Comp 12: 1607–1641.
81. Denker M, Timme M, Diesmann M, Wolf F, Geisel T (2004) Breaking Synchrony by Heterogeneity in Complex Networks. Phys Rev Lett 92: 74103.
82. Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, et al. (2004) Interneurons of the neocortical inhibitory system. Nat Rev Neurosci 5: 793–807.
83. Jonas P, Bischofberger J, Fricker D, Miles R (2004) Interneuron Diversity series: Fast in, fast out–temporal and spatial signal processing in hippocampal interneurons. Trends Neurosci 27: 30–40.
84. Vlachos I, Kumar A, Luthi A, Aertsen A (2009) Dynamical emergence of fear and extinction cells in the amygdala - a computational model. BMC Neurosci 10: P142.
85. Roxin A, Brunel N, Hansel D (2005) Role of Delays in Shaping Spatiotemporal Dynamics of Neuronal Activity in Large Networks. Phys Rev Lett 94: 238103–4.
86. Berlau DJ, McGaugh JL (2006) Enhancement of extinction memory consolidation: The role of the noradrenergic and GABAergic systems within the basolateral amygdala. Neurobiol Learn Mem 86: 123–132.
87. Corcoran KA, Desmond TJ, Frey KA, Maren S (2005) Hippocampal Inactivation Disrupts the Acquisition and Contextual Encoding of Fear Extinction. J Neurosci 25: 8978–8987.
88. Ji J, Maren S (2005) Electrolytic lesions of the dorsal hippocampus disrupt renewal of conditional fear after extinction. Learn Mem 12: 270–276.
89. Corcoran KA, Maren S (2004) Factors Regulating the Effects of Hippocampal Inactivation on Renewal of Conditional Fear After Extinction. Learn Mem 11: 598–603.
90. Ji J, Maren S (2007) Hippocampal involvement in contextual modulation of fear extinction. Hippocampus 17: 749–758.
91. Barot SK, Chung A, Kim JJ, Bernstein IL (2009) Functional Imaging of Stimulus Convergence in Amygdalar Neurons during Pavlovian Fear Conditioning. PLoS One 4: e6156.
92. Fuster JM (1998) Distributed Memory for Both Short and Long Term. Neurobiol Learn Mem 70: 268–274.
93. Maren S, Fanselow MS (1995) Synaptic plasticity in the basolateral amygdala induced by hippocampal formation stimulation in vivo. J Neurosci 15: 7548–7564.
94. Pitkänen A (2000) Connectivity of the rat amygdaloid complex. New York: Amygdala A functional analysis Oxford University Press. pp 31–115.
95. Hobin JA, Goosens KA, Maren S (2003) Context-Dependent Neuronal Activity in the Lateral Amygdala Represents Fear Memories after Extinction. J Neurosci 23: 8410–8416.
96. Hugues S, Garcia R (2007) Reorganization of learning-associated prefrontal synaptic plasticity between the recall of recent and remote fear extinction memory. Learn Mem 14: 520–524.
97. Farinelli M, Deschaux O, Hugues S, Thevenet A, Garcia R (2006) Hippocampal train stimulation modulates recallof fear extinction independently of prefrontalcortex synaptic plasticity and lesions. Learn Mem 13: 329–334.
98. Anagnostaras SG, Gale GD, Fanselow MS (2001) Hippocampus and contextual fear conditioning: Recent controversies and advances. Hippocampus 11: 8–17.
99. Rudy JW, Huff NC, Matus-Amat P (2004) Understanding contextual fear conditioning: insights from a two-process model. Neurosci Biobehav Rev 28: 675–685.
100. Hessler NA, Shirke AM, Malinow R (1993) The probability of transmitter release at a mammalian central synapse. Nature 366: 569–572.
101. Gulyás AI, Miles R, Sk A, Tóth K, Tamamaki N, et al. (1993) Hippocampal pyramidal cells excite inhibitory neurons through a single release site. Nature 366: 683–687.
102. Singer W, Gray CM (1995) Visual Feature Integration and the Temporal Correlation Hypothesis. Annu Rev Neurosci 18: 555–586.
103. von der Malsburg C (1995) Binding in models of perception and brain function. Curr Opin Neurobiol 5: 520–526.

104. Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of Oscillatory Neuronal Synchronization by Selective Visual Attention. Science 291: 1560–1563.
105. Lisman JE, Idiart MA (1995) Storage of 7 +/− 2 short-term memories in oscillatory subcycles. Science 267: 1512–1515.
106. Montgomery SM, Buzsáki G (2007) Gamma oscillations dynamically couple hippocampal CA3 and CA1 regions during memory task performance. Proc Natl Acad Sci U S A 104: 14495–14500.
107. Seidenbecher T, Laxmi TR, Stork O, Pape HC (2003) Amygdalar and hippocampal theta rhythm synchronization during fear memory retrieval. Science 301: 846–850.
108. Pape HC, Narayanan RT, Smid J, Stork O, Seidenbecher T (2005) Theta activity in neurons and networks of the amygdala related to long-term fear memory. Hippocampus 15: 874–880.
109. Crane JW, Windels F, Sah P (2009) Oscillations in the Basolateral Amygdala: Aversive Stimulation Is State Dependent and Resets the Oscillatory Phase. J Neurophysiol 102: 1379–1387.
110. Bauer EP, Paz R, Pare D (2007) Gamma Oscillations Coordinate Amygdalo-Rhinal Interactions during Learning. J Neurosci 27: 9369–9379.
111. Rescorla R (1969) Pavlovian conditioned inhibition. Psychol Bull 72: 77–94.
112. Newell A, Simon HA (1976) Computer science as empirical inquiry: symbols and search. Commun ACM 19: 113–126.
113. Schmajuk NA (2008) Computational models of classical conditioning. Scholarpedia, 3: 1664.
114. Rescorla R (2008) Rescorla-Wagner model. http://www.scholarpedia.org/article/Rescorla-Wagner.
115. Wagner A (1981) SOP: A model of automatic memory processing in animal behavior. In: Spear NE, Miller RR, eds. Information processing in animals: Memory mechanisms. Hillsdale: Erlbaum. pp 5–47.
116. Brandon SE, Vogel EH, Wagner AR (2003) Stimulus representation in SOP: I: Theoretical rationalization and some implications. Behav Processes 62: 5–25.
117. Armony JL, Servan-Schreiber D, Cohen JD, LeDoux JE (1997) Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning. Trends Cogn Sci 1: 28–34.
118. Li G, Nair SS, Quirk GJ (2009) A Biologically Realistic Network Model of Acquisition and Extinction of Conditioned Fear Associations in Lateral Amygdala Neurons. J Neurophysiol 101: 1629–1646.