OXFORD

ORIGINAL ARTICLE

Multicontact Co-operativity in Spike-Timing–Dependent Structural Plasticity Stabilizes Networks

Moritz Deger[1,2], Alexander Seeholzer[1] and Wulfram Gerstner[1]

[1]School of Computer and Communication Sciences and School of Life Sciences, Brain Mind Institute, École Polytechnique Fédérale de Lausanne, 1015 Lausanne EPFL, Switzerland and [2]Institute for Zoology, Faculty of Mathematics and Natural Sciences, University of Cologne, 50674 Cologne, Germany

Address correspondence to Postal address: Laboratory of Computational Neuroscience, LCN EPFL – Batiment AAB (Station 15), CH-1015 Lausanne EPFL, Switzerland. Email: wulfram.gerstner@epfl.ch

Abstract

Excitatory synaptic connections in the adult neocortex consist of multiple synaptic contacts, almost exclusively formed on dendritic spines. Changes of spine volume, a correlate of synaptic strength, can be tracked *in vivo* for weeks. Here, we present a combined model of structural and spike-timing–dependent plasticity that explains the multicontact configuration of synapses in adult neocortical networks under steady-state and lesion-induced conditions. Our plasticity rule with Hebbian and anti-Hebbian terms stabilizes both the postsynaptic firing rate and correlations between the pre- and postsynaptic activity at an active synaptic contact. Contacts appear spontaneously at a low rate and disappear if their strength approaches zero. Many presynaptic neurons compete to make strong synaptic connections onto a postsynaptic neuron, whereas the synaptic contacts of a given presynaptic neuron co-operate via postsynaptic firing. We find that co-operation of multiple synaptic contacts is crucial for stable, long-term synaptic memories. In simulations of a simplified network model of barrel cortex, our plasticity rule reproduces whisker-trimming–induced rewiring of thalamocortical and recurrent synaptic connectivity on realistic time scales.

Key words: barrel cortex model, cortical reorganization, dendritic spine motility, synaptic plasticity

Introduction

Most excitatory synapses to pyramidal cells of the mammalian neocortex project onto dendritic spines (Yuste 2011), visible as small protrusions that vary in size and shape and are subject to ongoing plasticity (Trachtenberg et al. 2002; Yasumatsu et al. 2008; Holtmaat and Svoboda 2009; Loewenstein et al. 2011). The volume of a spine is highly correlated with synaptic current amplitude, a measure of synaptic weight (Matsuzaki et al. 2001), and changes of spines have been linked to learning (Stepanyants et al. 2002; Hayashi-Takagi et al. 2015). Although all dendritic spines show considerable volatility and may be eliminated (Loewenstein et al. 2015), some spines are maintained

over longer periods of time than others (Holtmaat et al. 2005; Yang et al. 2009) and may contribute to long-term memories (Grutzendler et al. 2002).

There have been 3 puzzling observations that we would like to address in the present study. First, time-lapse imaging of spines on time scales of seconds, hours, and days to weeks has shown that small spines are removed and new spines formed (Matsuzaki et al. 2004; Holtmaat et al. 2005, 2006; Yasumatsu et al. 2008; Holtmaat and Svoboda 2009; Zito et al. 2009; Kwon and Sabatini 2011), consistent with functional connectivity measurements in slices (Le Be and Markram 2006; Wiegert and Oertner 2013). However, despite ongoing spine turnover, the

distribution of synaptic weights as measured with indicators for postsynaptic density molecules (Minerbi et al. 2009) or with electrophysiological methods (Le Be and Markram 2006) is long-term stable. In view of this result, we would like to pose a first question: "Can a synaptic and structural plasticity rule induce long-term stability of important features of wiring patterns and synaptic weight distributions in the presence of spine turn-over?" In contrast to models of synaptic memory consolidation, which invoke synaptic processes of "single" contacts on different time scales to explain long-term stability (Fusi et al. 2005; Clopath et al. 2008; Roxin and Fusi 2013; Zenke et al. 2015; Benna and Fusi 2016), we focus on a structural plasticity model with "multiple" potential contacts that can become active or inactive.

Second, cortical spines show remarkable dynamics after partial lesioning of input pathways (Trachtenberg et al. 2002; Holtmaat et al. 2006; Keck et al. 2008; Rose et al. 2016). In particular, after trimming of whiskers, the fraction of newly formed spines in somatosensory cortex that become persistent increases compared with control (Holtmaat et al. 2006). This leads us to pose a second question: "Can a combined model of synaptic and structural plasticity account for rewiring after input lesion?" The reconfiguration of synapses and receptive fields has been a topic of study in synaptic plasticity models with continuous weight dynamics (e.g., Bienenstock et al. 1982) and, separately, of structural synapse formation models (Butz and Ooyen 2013; Vlachos et al. 2013). However, in a "combined synaptic and structural" multicontact plasticity model, the time scale of rewiring of synaptic contacts and its interplay with homeostatic increases of synaptic weights (Toyoizumi et al. 2014) has so far not been considered.

Third, synapses between pyramidal neurons in the somatosensory cortex of the rat are sparse and the distribution of the number of putative anatomical synaptic contacts for pairs of pre- and postsynaptic neurons has a characteristic "bimodal" shape (Markram, Lübke, Frotscher, Roth et al. 1997a, 2015; Fares and Stepanyants 2009): Even if two neurons are close neighbors, the most likely outcome of paired recordings is that there is no synaptic contact; however, those neuron pairs that form synapses are likely to establish 4 or more contacts (Markram, Lübke, Frotscher, Roth et al. 1997a). The combination of both observations results in a distribution of putative synaptic contacts with a dominant peak for zero contacts, a trough for 1–2 contacts and a second peak around 4–8 contacts (Fares and Stepanyants 2009). This bimodal distribution (called distribution of "actual" contact multiplicity in the following) has to be compared with the number of "potential" synaptic contacts between a pair of neurons that has been estimated from the number of axon-dendritic appositions (near-touch points) in reconstructed cortical microcircuits (Fares and Stepanyants 2009). If near-touch points are defined whenever the distance between reconstructed dendritic and axonal branches is less than $2\,\mu m$ (Fares and Stepanyants 2009) or $2.5\,\mu m$ (Reimann et al. 2015), then the distribution of potential contact multiplicity between neighboring neurons in the same layer is always broadly distributed extending to 10 or more contacts (Fares and Stepanyants 2009; Markram et al. 2015; Reimann et al. 2015) and the probability of finding not a single near-touch point between close neighbors is nearly zero (Reimann et al. 2015). The difference between the distribution of actual and potential synaptic contact multiplicity implies that contacts driven by the same presynaptic neuron must "co-operate" (Fares and Stepanyants 2009), potentially via correlation detection in spike-timing–dependent plasticity (STDP) (Helias et al. 2008;

Deger et al. 2012). At the same time, synaptic plasticity rules that are useful for receptive field development and assembly formation must show "competition" between synapses from different presynaptic neurons, so that some connections grow at the expense of others (Bienenstock et al. 1982; Oja 1982; Miller and MacKay 1994; Kempter et al. 1999; Song et al. 2000; van Rossum et al. 2000; Shouval et al. 2002; Helias et al. 2008; Morrison et al. 2008; Clopath et al. 2010; Zenke et al. 2015). This observation leads us to pose two further questions: "Can multicontact synaptic and structural plasticity combining both co-operation and competition arise from a single mathematical rule?" In particular, "why does successful co-operation not lead to connections with a number of contacts equal to the potential maximum?" Competitive synaptic plasticity models developed for single-contact connections (Bienenstock et al. 1982; Kempter et al. 1999; Song et al. 2000; van Rossum et al. 2000; Helias et al. 2008; Bourjaily and Miller 2011; Miner and Triesch 2016) will drive, in a multicontact scenario, the number of active synaptic contacts to the maximum value, at least for those presynaptic neurons with the largest correlations (Fauth et al. 2015a). In contrast to earlier work with abstract Markovian state-transition models (Deger et al. 2012; Fauth et al. 2015a) or explicit algorithmic learning rules (Fares and Stepanyants 2009; Reimann et al. 2015), we propose an STDP-based learning rule for the multicontact scenario that leads to the emergence of "co-operativity" between contacts of one, but competition between contacts of multiple, presynaptic neurons with a bimodal distribution of actual synaptic contact multiplicity that peaks at values smaller than the number of potential contacts.

Our model introduces a novel combination of Hebbian and anti-Hebbian terms with several attractive features: synaptic and structural plasticity 1) stabilizes firing rates, 2) keeps network activity stable, 3) limits the correlation between pre- and postsynaptic neuron at single synaptic contact points, 4) limits weight growth without explicit upper bounds, 5) leads to co-operation of contacts of the same presynaptic neuron, 6) leads to a competition between different presynaptic neurons, and 7) enables us to study the network effects of structural plasticity. We thus demonstrate that a local spike-timing–dependent learning rule is sufficient to explain the bimodal distribution of contact numbers with a peak around 4–8 contacts. Furthermore, we demonstrate that multiple contacts are not only interesting from the perspective of information transmission (Laughlin et al. 1998; Fares and Stepanyants 2009) but also increase long-term stability of synaptic connections.

Materials and Methods

Synaptic Contacts

In the experiments of Figures 1–5, we model the plasticity of the synaptic connections from N presynaptic neurons onto a single postsynaptic neuron (Fig. 1a). A presynaptic neuron j may be connected to the postsynaptic neuron by several synaptic contacts k, up to a maximum of n_j potential synaptic contacts. Note that the number n_j of potential contacts is different for different neurons j and drawn in our model from a unimodal probability distribution $P(n)$ (Fig. 2a, blue line) taken from Fares and Stepanyants (2009). These authors estimated the distribution of potential contacts from reconstructed cortical microcircuits of rat barrel cortex. Potential contacts between 2 nearby (somatic distance less than $50\,\mu m$) layer 5 pyramidal neurons were defined as "near-touch" points where axons and dendrites pass at a distance of less than
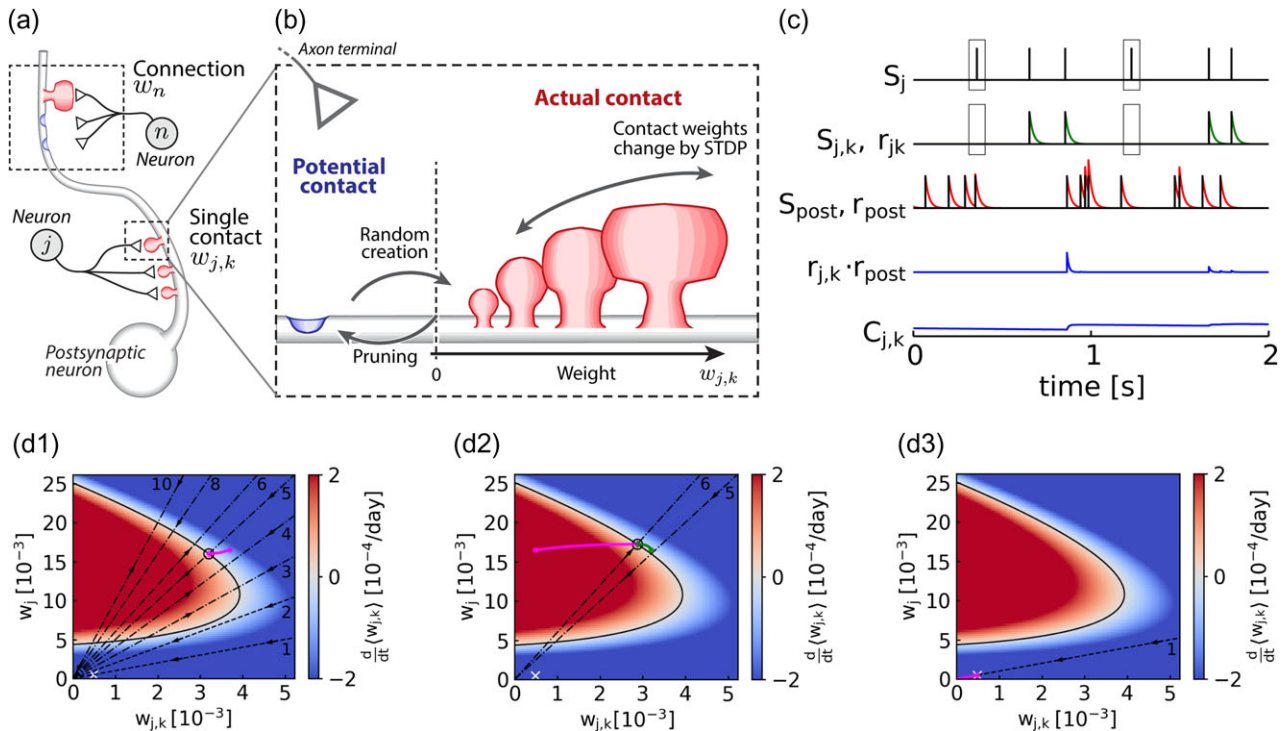
Figure 1. Model overview and analysis. (*a*) A synaptic connection between a presynaptic neuron j and the postsynaptic neuron consists of multiple contacts with weights $w_{j,k}$, summing to the total weight $w_j = \sum_k w_{j,k}$. Connections of similar total weight can be composed of a single large contact (connection from neuron n) or several smaller contacts (connection from neuron j). (*b*) Individual contact weights take continuous, positive values and change in time according to STDP. Small weights correspond to thin dendritic spines (with a small volume), and large weights correspond to large (mushroom-shaped) spines. Contacts with positive weight are actual contacts ($w_{j,k} > 0$). If they transition to $w_{j,k} \leq 0$, they are pruned (weights fixed at $w_{j,k} = 0$) and become potential (but inactive) contacts. These are transformed into actual contacts by setting $w_{j,k}$ to a positive value w_c at random times, with a rate $\lambda_c = 0.019$/day (creation). (*c*) Components of the plasticity model (top to bottom): presynaptic spike train S_j; transmitted spike train $S_{j,k}$ at the contact (black, random synaptic failures occur, indicated by black boxes), and its filtered trace $r_{j,k}$ (green); postsynaptic spike train S_{post} (black) and its filtered trace r_{post} (red); product (correlation) term $r_{j,k} \cdot r_{post}$ composed of pre- and postsynaptic trace (blue); low-pass–filtered trace $C_{j,k}$ of the product term $r_{j,k} \cdot r_{post}$. (*d*) Expected weight change $\frac{d}{dt}\langle w_{j,k}\rangle$ (positive in red; negative in blue) of a synaptic contact as a function of its weight $w_{j,k}$ and the total weight w_j, assuming constant postsynaptic rate. Black curves mark combinations of $w_{j,k}$ and w_j that have zero expected change ($w_{j,k}$ null-cline). Straight lines mark synapses where all m (small numbers) contacts have the same weight. Intersections of curves and lines mark stable fixed points (w_*/m, w_*) of the dynamics for m actual contacts. (*d1*) Expected trajectory after a single contact in a synaptic connection with $m = 5$ contacts is perturbed by w_Δ. The trajectory (pink) starts at ($w_*/5 + w_\Delta$, $w_* + w_\Delta$) and evolves back toward the fixed point. (*d2*) Five contacts each take a value $w_*/5$ consistent with the previous fixed point, when a new contact is created. The trajectory of the new contact (pink) starts at (w_c, $w_* + w_c$) and evolves toward the new fixed point for 6 contacts (circle), as do the other contacts (green). (*d3*) A new contact is created in a connection with no actual contacts. The new contact starts at (w_c, w_c) (white cross), approaches $w_{j,k} = 0$, and is removed (pink).

$2\,\mu$m (Fares and Stepanyants 2009). In the reconstructed microcircuits, a small fraction of neuron pairs would have more than 10 and up to 20 potential contacts (Fares and Stepanyants 2009). For computational reasons, we limited n_j to a maximum of 10. To do so, we truncated the distribution in Fares and Stepanyants (2009) at $n = 10$ and renormalized the distribution $P(n)$ thereafter to $\sum_{n=1}^{10} P(n) = 1$. In the model, we randomly assigned $P(n) \cdot N$ presynaptic neurons to each potential contact number n in order to exactly reproduce the distribution $P(n)$, except for Figure 4 (single-contact case) and Figure 6 (recurrent network). The number n_j of potential contacts of neuron j is then kept fixed thereafter. Our results in Figure 3a indicate that the model is insensitive to the truncation of the distribution at $n_j = 10$ because nearly none of the presynaptic neurons developed $n = 10$ actual contacts after 100 days of simulated time. We therefore conclude that a maximum value of *potential* contacts of $n_j = 20$ (Fares and Stepanyants 2009) (instead of our maximum of 10) would give equivalent results, albeit at increased computational cost.

The mean of the probability distribution $P(n)$ of the number of potential contacts (Fig. 2a, blue line) is 4.633. In the simulations of Figures 1–3 and 5, we use $N = 1000$ presynaptic neurons. Application of the initialization procedure for potential

contacts described above therefore leads in our multicontact model with $N = 1000$ presynaptic neurons to a total of 4633 potential synaptic contacts. In the single-contact model of Figure 4, we set $n_j = 1$ for all connections $1 \leq j \leq N$. To maintain the total number of potential synaptic contacts in the system with single-contact connections identical to that of the multi-contact system, we increase the number of presynaptic neurons to $N = 4633$, keeping all other parameters the same.

The synaptic contact k with $1 \leq k \leq n_j$ of a presynaptic neuron j onto the postsynaptic one is described by a unit-less weight $w_{j,k}$, which can change as a function of time. The total weight $w_j = \sum_{k=1}^{n_j} w_{j,k}$ of the connection from a presynaptic neuron j is given by the sum of the weights $w_{j,k}$ over all its contacts $1 \leq k \leq n_j$. The contact weight $w_{j,k}$ in our model describes how much the contact k contributes to firing the postsynaptic cell and can be viewed as representing the dendritic spine volume, which is in turn strongly correlated with the AMPA receptor content in the postsynaptic density (Matsuzaki et al. 2001, Holtmaat and Svoboda 2009). Note that in Figures 1–5, we have only a single postsynaptic neuron (except for the simulations of the recurrent network in Fig. 6), so that we do not need an index for the postsynaptic neuron.
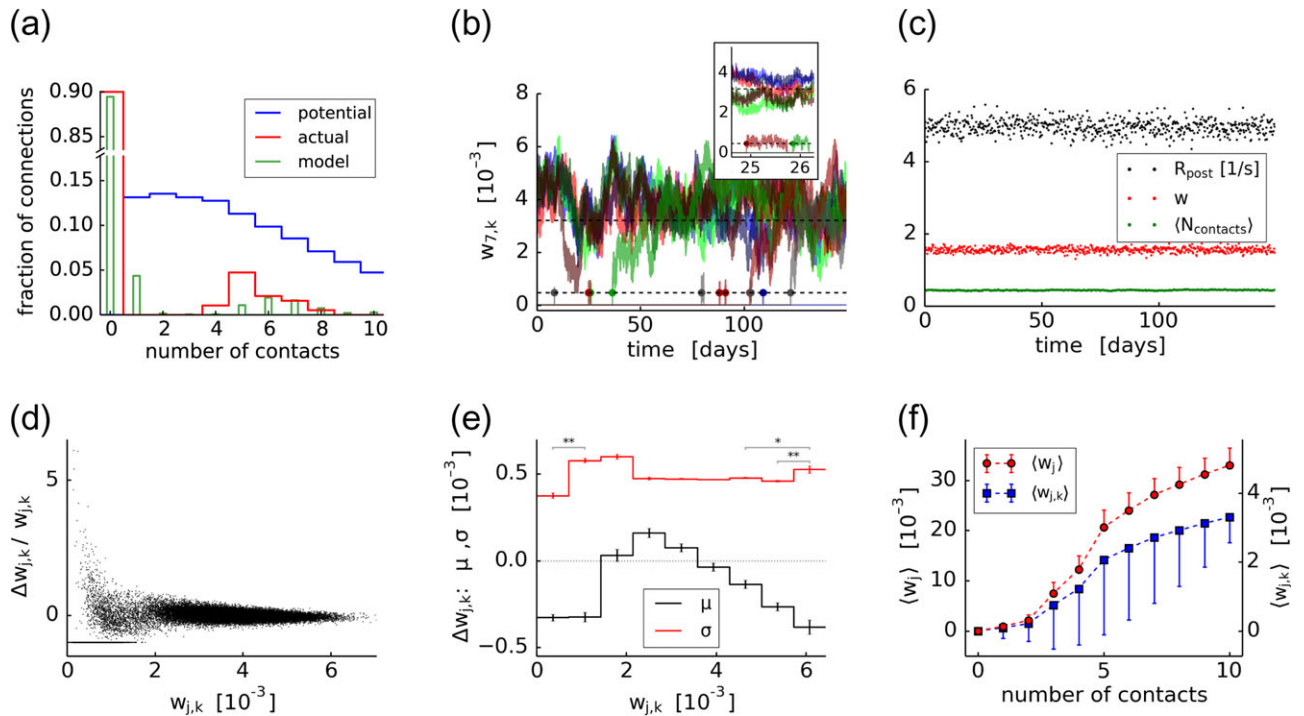
Figure 2. Dynamics of synaptic contacts in the steady state. (*a*) Reference distributions (data extracted from previous publications) of the number of potential (blue) (Fares and Stepanyants 2009) and putative anatomically functional (red, actual) (Markram, Lübke, Frotscher, Roth et al. 1997a) synaptic contacts for pairs of neurons in the adult somatosensory cortex (recurrent connections of nearby layer 5 pyramidal neurons, truncated to $n_j \leq 10$ and renormalized). The steady-state distribution generated by our model is shown in green (data pooled over 150 days of simulation); the distribution of potential contacts in the model is matched to the blue line. (*b*) The synaptic connection from presynaptic neuron $j = 7$ is formed by several contact weights $w_{j,k}$ (colored lines). New contacts (filled circles) are created with weight w_c given by the lower dashed line. Long-term stable contacts fluctuate around the upper dashed line, which is the fixed point of $w_c/5$ predicted by theory (see Materials and Methods). Inset shows zoom at the time of creation and pruning of 2 transient synaptic contacts. (*c*) Firing rate R_{post} of postsynaptic neuron (black), total synaptic input $w = \sum_j \sum_k w_{j,k}$ summed over all presynaptic neurons and contacts (red, unit-less), and average number of actual contacts per synaptic connection (green) over time, each sampled in intervals of 6 h. (*d*) Relative changes of synaptic contact efficacy $\Delta w_{j,k}$ within 1 day, dots for all contacts and 150 days of simulated time (sampled in intervals of 2 days). The horizontal line of dots at ordinate value -1 is due to contacts that were removed within 1 day. (*e*) Weight dependence of changes: Mean μ and standard deviation (STD) σ of the change of contact weight $\Delta w_{j,k}$ within 1 day, as a function of the contact weight $w_{j,k}$. Both μ and σ were estimated from the data shown in (*d*), grouped into $w_{j,k}$-intervals as indicated (solid lines). Error bars denote the standard error of the mean (SEM). For very small and very large contact weights, σ increases significantly (* indicates $p < 0.1$ in Welch's 2-sided t-test). (*f*) Total synaptic connection weight $w_j = \sum_k w_{j,k}$ (left axis, red) and contact weight $w_{j,k}$ (right axis, blue) as a function of the number of actual contacts ($w_{j,k} > 0$), averaged across all synaptic connections. Error bars denote the STD.

For the set-up of the recurrent network (Fig. 6), see the section "Recurrent Network Model."

Presynaptic Spike Trains and Transmission Failures

The activity of the $N = 1000$ presynaptic neurons $1 \leq j \leq N$ in our model (or $N = 4633$ for the single-contact network in Fig. 4) is described by the spike trains $S_j(t) = \sum_m \delta(t - t_j^m)$, where $\{t_j^1, t_j^2, ...\}$ are the spike times of neuron j. Except for the recurrent network of Figure 6, the presynaptic spike trains in all simulations are generated by independent Poisson processes with a constant firing rate of $\nu_{pre} = 5$Hz. This value is relatively high compared with the spontaneous firing rate *in vivo* (Kerr et al. 2005) and was chosen such that the synaptic activity would also represent some episodes of stimulation, such as active whisking (Crochet and Petersen 2006). In Supplementary Material S2.5, we show that a version of our computational model with slightly modified parameters yields qualitatively identical results at a presynaptic rate of 0.5 Hz when compared with the model with standard parameters at a presynaptic rate of 5 Hz.

Due to random transmission failures (Zador 1998), not all presynaptic spikes, however, are transmitted at each of the synaptic contacts that connect neuron j with the postsynaptic neuron. More precisely, the spike train transmitted at the contact k is given by

$$S_{j,k}(t) = \sum_m \delta(t - t_j^m)\, z_{j,k}(t_j^m)\,. \tag{1}$$

The spike train $S_{j,k}(t)$ differs from the presynaptic spike train $S_j(t)$ through multiplication with independent Bernoulli random variables $z_{j,k}(t_j^m) \in \{0,1\}$ that describe the stochastic failures of synaptic transmission of the spike fired at time t_j^m. In our model, synaptic failures occur randomly and independently with a probability of $p_f = 0.5$, so $z_{j,k}(t_j^m)$ is 1 (successful transmission) with probability $1 - p_f = 0.5$, except for Figure 4f where p_f is systematically varied.

Postsynaptic Spike Train

The Postsynaptic Neuron Emits a Spike Train

$$S_{post}(t) = \sum_m \delta(t - t_{post}^m), \tag{2}$$

where $\{t_{post}^1, t_{post}^2, ...\}$ are the spike times. For the simulation of the recurrent network in Figure 6, the postsynaptic spike train was generated by a leaky integrate-and-fire neuron model.
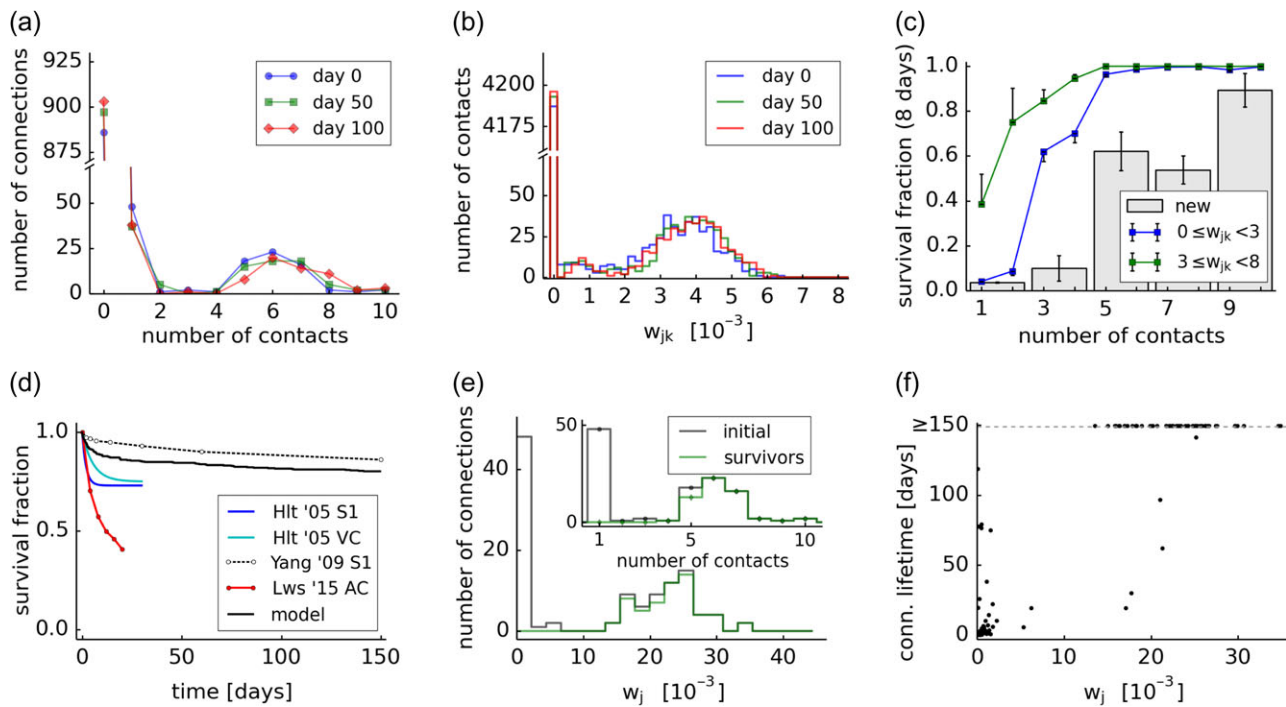
Figure 3. Statistics of synaptic contacts in the steady state. (*a*) Histogram of the number of actual contacts ($w_{j,k}>0$) of a connection, across all connections j. (*b*) Histogram of contact weights $w_{j,k}$, across all connections j and contacts k. The distribution at day 0 (measurements start after reaching a stable configuration to mimic the age of rats in experiments; see Simulation of the Plasticity Model in Materials and Methods) is not significantly different from day 100 (2-sample Kolmogorov–Smirnov test (2sKS), P-value 0.910). (*c*) Fraction of synaptic contacts that survive for 8 days, as a function of the number of actual contacts in the connection, for newly created contacts (gray bars) and existing contacts of weak or strong weight $w_{j,k}$ (colors, see legend) in units of 10^{-3}. Survival fractions are estimated separately per day and then averaged over days. Estimates from less than 5 samples are discarded; error bars denote SEM. (*d*) Fraction of surviving actual synaptic contacts that were present at time $t = 0$ (black: model) in comparison with published experimental dendritic spine survival data in mouse (Hlt '05 S1: somatosensory cortex L5B, age 6 months (exponential decay fit) (Holtmaat et al. 2005); Hlt '05 VC: visual cortex L5B, age 3–6 months (exponential decay fit) (Holtmaat et al. 2005); Yang '09 S1: somatosensory cortex L5, age 4–5 months (data estimated from Yang et al. (2009)); Lws '15 AC: auditory cortex (Loewenstein et al. 2015)). (*e*) Histogram of the total weights w_j of the connections present at $t = 0$ (black) and of the surviving connections at day 150 (green). Inset: Histogram of the number of contacts of the connections. The removed connections have small total weight and number of contacts; strong connections with many contacts are rarely removed. Note that new contacts created in the meantime are not considered in this analysis. (*f*) Lifetime of entire synaptic connections (dots) during the course of the simulation, as a function of the total connection weight.

However, in the remainder of the study, we use a linear Poisson neuron model (Kempter et al. 1999) so as to simplify the mathematical analysis of our synaptic and structural plasticity model. The probability of the linear Poisson neuron to fire a spike in a very short time step dt is $\lambda(t)dt$ where $\lambda(t)$ is the instantaneous firing rate. In the absence of synaptic input from the presynaptic neurons, the postsynaptic neuron fires with a baseline firing rate $\lambda(t) = \lambda_0 = 1/s$ (i.e., a homogeneous Poisson process). We further assume that synaptic inputs cause transient increases of the firing rate, which decay on the time scale τ of the membrane potential. We thus model the dynamics of the postsynaptic neuron's instantaneous firing rate $\lambda(t)$ (also called "stochastic intensity") as

$$\tau\frac{d}{dt}\lambda(t) = -(\lambda(t)-\lambda_0) + \sum_{j=1}^{N}\sum_{k=1}^{n_j} w_{j,k}(t)S_{j,k}(t-d), \quad (3)$$

where the second term sums all inputs across all n_j synaptic contacts over all N presynaptic neurons and $d = 1ms$ denotes the synaptic transmission delay.

Model of Synaptic Contact Plasticity

Synaptic contacts in our model follow a variant of STDP (see Fig. 1*b*), which we imagine to be realized by the biophysics of dendritic spines in combination with that of the presynaptic

terminal. Each contact is described by its efficacy (weight) $w_{j,k}$, which is the (unit-less) amplitude of the excitatory postsynaptic potential (EPSP) that the contact k evokes upon arrival of an action potential at the presynaptic terminal and in the absence of synaptic failure.

The synaptic weight evolves according to a local Hebbian learning rule defined by a differential equation $\frac{dw_{j,k}}{dt} = F[w_{j,k}; S_{j,k}, S_{post}]$, which depends only on the momentary value of the synaptic weight $w_{j,k}$ as well as on the recent history of pre- and postsynaptic spike trains. This history is summarized in our model by local "synaptic traces" of pre- and postsynaptic activity as well as pre–post correlations (Morrison et al. 2008). Specifically, we introduce variables $r_{j,k}(t)$ and $r_{post}(t)$ to describe low-pass filters of the pre- and postsynaptic spike trains $S_{j,k}(t)$ and $S_{post}(t)$, defined by the differential equations:

$$\tau\frac{d}{dt}r_{j,k}(t) = -r_{j,k}(t) + S_{j,k}(t) \quad (4)$$

$$\tau\frac{d}{dt}r_{post}(t) = -r_{post}(t) + S_{post}(t), \quad (5)$$

with a time constant $\tau = 20$ ms, typical for the duration an EPSP. The trace left by presynaptic spike arrival at contact k can, for example, be interpreted as the amount of glutamate bound to the postsynaptic receptor. Similarly, the trace left by a postsynaptic spike can be interpreted as the voltage time
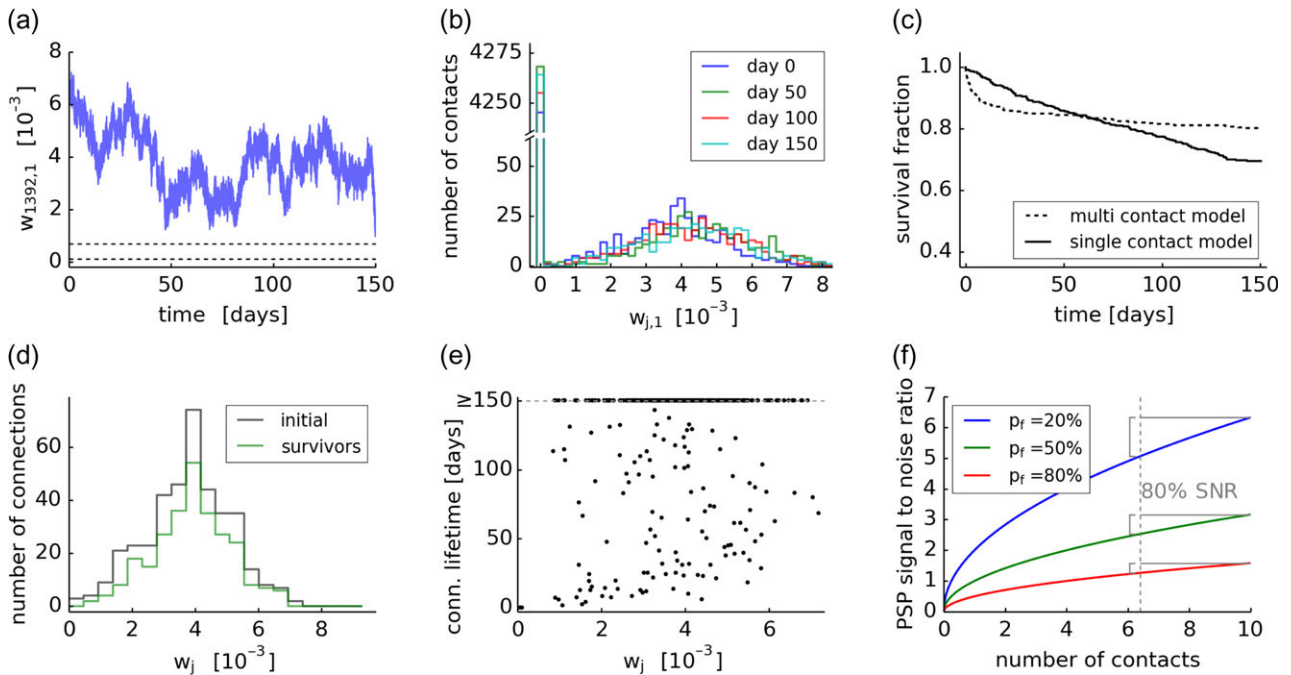
Figure 4. Single-contact synaptic connections are less stable. (a–e) Simulation of the plasticity model with only one potential contact per connection (n_j=1 for all j). (a) Example synaptic connection number j = 1392, blue line corresponds to contact weight $w_{j,1}$ over time, as in Figure 2b. (b) Histogram of contact weights $w_{j,1}$, across all connections j. A steady state is maintained for 100 days, just as in the case of multiple potential contacts (cf. Fig. 3b). (c) Fraction of surviving actual synaptic contacts (solid line) that were present at time $t = 0$ in comparison to the multicontact model (dashed line) of Figure 3d. (d) Histogram of the weights $w_j = w_{j,1}$ of the connections present at $t = 0$ (black) and of the surviving connections at day 100 (green). Connections of small and large weights are removed unspecifically. (e) Lifetime of synaptic connections during the course of the simulation, as a function of the connection weight; compare Figure 3f. (f) Theoretical signal-to-noise ratio (SNR) of the postsynaptic potential (PSP) (see Eq. (13)) in response to a presynaptic spike, in connections with multiple actual contacts and stochastic synaptic failures (probability p_f). Dashed line indicates the number of actual contacts for which a SNR of 80% of the maximum is achieved, if 10 contacts are considered to be the maximum.

course of the back-propagating action potential or the induced calcium transient.

As an estimate of the correlations of pre- and postsynaptic firing on a slower time scale, each synaptic contact computes also a correlation trace $C_{j,k}$ and a slow trace of the postsynaptic activity R_{post}. These slow traces are defined by the differential equations

$$\tau_{slow}\frac{d}{dt}C_{j,k}(t) = -C_{j,k}(t)+r_{j,k}(t)\cdot r_{post}(t), \qquad (6)$$

$$\tau_{slow}\frac{d}{dt}R_{post}(t) = -R_{post}(t) + S_{post}(t), \qquad (7)$$

with a time constant of τ_{slow} = 1 min.

Many synaptic plasticity rules in the literature can be formulated as a differential equation $\frac{dw_{j,k}}{dt} = F[w_{j,k};r_{j,k}(t), r_{post}(t), C_{j,k}(t), R_{post}(t)]$ with the momentary values of one or several "traces" on the right-hand side (Morrison et al. 2008). For example, if we suppress the index k and focus on classical single-contact plasticity rules, then a learning rule $\frac{dw_j}{dt} = a_2^{corr} C_j(t)$ with a positive constant a_2^{corr} leads to a Hebbian spike-based learning rule with a symmetric STDP window for long-term potentiation (Gerstner and Kistler 2002). Weight changes in such a rule will increase to infinity, unless a (soft or hard) bound on the weight is introduced. This is classically achieved by turning the positive constant a_2^{corr} into a weight-dependent parameter $a_2^{corr}(w_j)$. Alternatively, a slightly modified rule $\frac{dw_j}{dt} = a_2^{corr} C_j(t) - \alpha w_j$ limits growth of synapses by a weight-dependent decay. Inclusion of additional local traces with different time constants leads to various self-normalizing and asymmetric STDP models (Gerstner et al. 1996; Kempter et al. 1999, Song et al. 2000; van Rossum et al. 2000; Gerstner and Kistler

2002; Morrison et al. 2008). For example, a rate normalization can be introduced by a weight decay term $-\alpha w_j^n(R_{post}(t) - R_{target})$, where R_{target} is the target rate of the postsynaptic neuron and the weight enters with power of $n = 1$ (van Rossum et al. 2000) or $n = 2$ (Tetzlaff et al. 2012). A version of rate normalization that is particularly stable in recurrent networks is a decay term proportional to $-a_4^{post}R_{post}^4(t)$ with a (potentially weight-dependent) parameter a_4^{post} (Zenke et al. 2015), whereas a more traditional slow "homeostatic" control of firing rates is not sufficient to control synaptic plasticity in recurrent networks (Zenke et al. 2013). Calcium-dependent plasticity models (Shouval et al. 2002) lead to specific instantiations of correlation traces (Helias et al. 2008), and so do voltage-based plasticity models (Clopath et al. 2010).

Standard synaptic plasticity models with soft-bounds are intrinsically stable but induce only a weak competition between synapses (Morrison et al. 2007; Billings and van Rossum 2009) and have reduced memory retention capabilities, compared with synaptic memory models with hard bounds (Billings and van Rossum 2009), in which strong competition causes a subset of weights to reach the upper bound (Miller and MacKay 1994; Kempter et al. 1999; Song et al. 2000). The observation of a bimodal distribution of synaptic contact numbers suggests strong competition. On the other hand, in the experimental data, even the strongest connections with 4–8 synaptic contacts (Markram, Lübke, Frotscher, Roth et al. 1997a) do most likely not reach the maximum number of potential synaptic contacts as estimated by Fares and Stepanyants (2009).

In order to guarantee competition between synapses from different presynaptic neurons, we use the strong rate normalization of Zenke et al. (2015) together with Hebbian plasticity without soft or hard bounds. At the same time, in order to
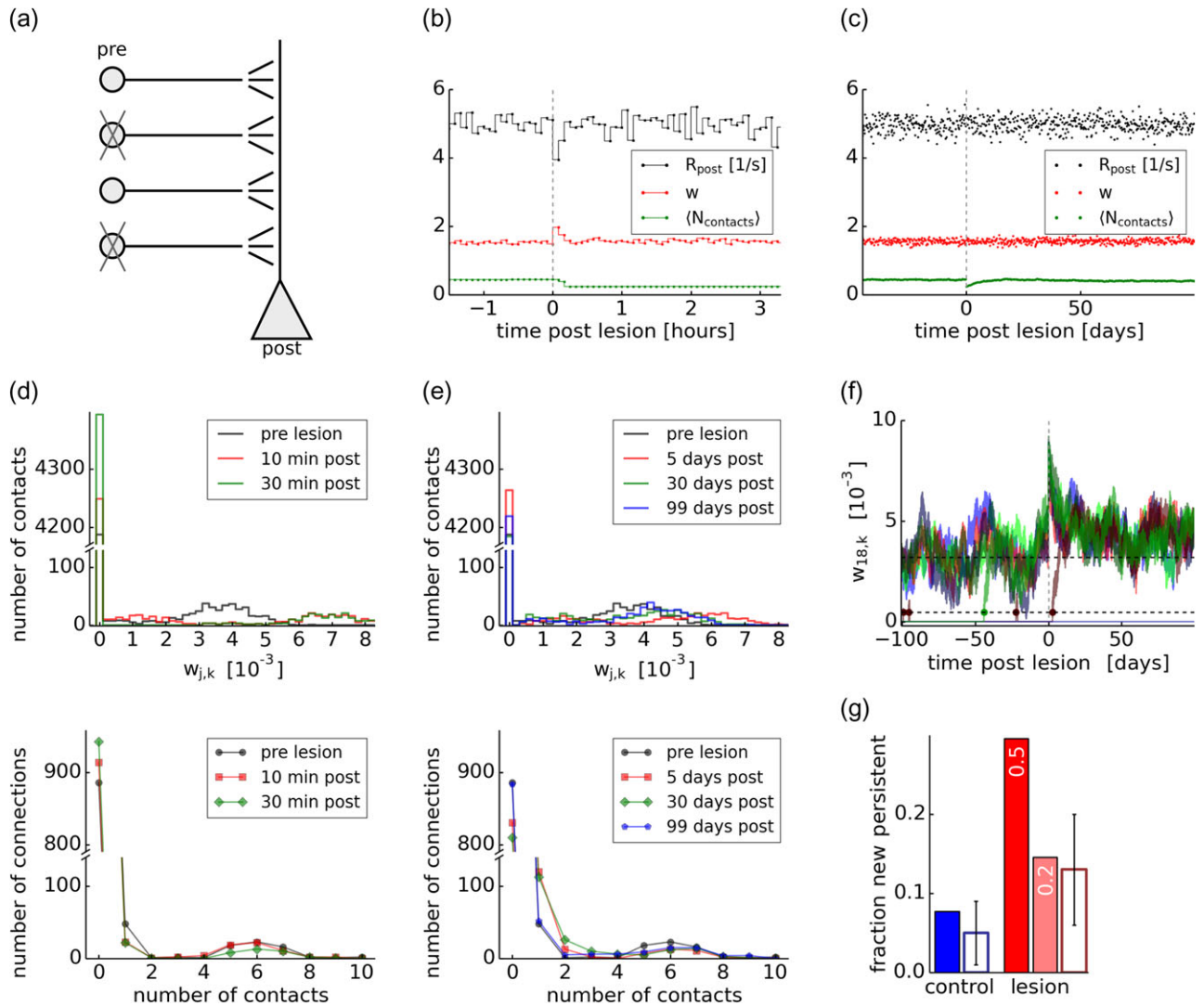
Figure 5. Rewiring in response to input lesion. (*a*) Schematic of the simulated lesion experiment. A substantial fraction of the presynaptic neurons that have actual synaptic contacts ($w_{j,k} > 0$) onto the postsynaptic cell are ablated (set to very low firing rate of 0.1/s; each connected neuron is ablated with probability $p_{lesion} = 0.5$; unconnected neurons are unaffected) at $t = 0$ (vertical dashes in *b–c*). (*b-c*) Firing rate R_{post} of postsynaptic neuron (black), total synaptic input $w = \sum_{j,k} w_{j,k}$ summed over all presynaptic neurons and contacts (red, unit-less) and average number of contacts per synaptic connection (green) over time. (cf. Fig. 2c). After the lesion, the average number of contacts (green) quickly drops by about 50% (*b*) and gradually recovers toward the steady state afterward (*c*). (*d-e*) Histograms of contact weights $w_{j,k}$ (top) and of the number of actual contacts (bottom), across all connections and contacts, before and after the lesion. Half of the actual synaptic contacts are removed within 30 min after the lesion (*d*, bottom, green), while the remaining half of the contacts potentiates (*d*, top, green). Within 99 days, the distributions gradually recover (*e*). (*f*) Example synaptic connection from presynaptic neuron with index $j = 18$, each colored line corresponds to a contact weight $w_{j,k}$ over time, as in Figure 2*b*. (*g*) Fraction of newly created persistent contacts (which survive for at least 8 days as defined in Holtmaat et al. (2006), control case (blue, cf. Fig. 3*c*) versus lesion model (red). Two lesion models are shown (red bars), marked with their respective value of p_{lesion}. For comparison, data from mouse whisker-trimming experiments (Holtmaat et al. 2006) are shown (open bars, error bars denote STD over observed cells). A smaller lesion with $p_{lesion}=0.2$ is in better agreement with the experimental data.

limit the maximal amount of correlation between pre- and postsynaptic neuron, we introduce a novel anti-Hebbian term that is proportional to the square of the correlation traces. This term, which limits weight growth if large correlations between pre- and postsynaptic neurons have been observed in the recent past, is loosely inspired by priming experiments of plasticity induction (Huang et al. 1992) and similar in spirit to an earlier model (El Boustani et al. 2012), albeit with a different mathematical formulation. Specifically, we study the plasticity model

$$\frac{d}{dt}w_{j,k}(t) = a_2^{corr} C_{j,k}(t) - a_4^{corr} C_{j,k}^2(t) - a_4^{post} R_{post}^4(t) - \alpha w_{j,k}(t), \quad (8)$$

with the parameters $a_2^{corr} = 1.94569 \cdot 10^{-6}$ s, $a_4^{corr} = 7.50642 \cdot 10^{-8}$ s^3, $a_4^{post} = 2.01605 \cdot 10^{-8}$ s^3, and $\alpha = 2 \cdot 10^{-6}$ s^{-1} (see Calibration of the System Dynamics below for details on parameter values). Weights evolve without an explicit upper bound, except for the experiment in Figure 7*b*, where we introduced an upper bound at twice the value of the equilibrium weight. However, we implement a lower absorbing boundary at zero, using the following procedure: If the evolution of the contact weight according to Equation (8) crosses (or hits) zero, the contact weight is removed. The contact now has the status of a potential, but inactive, contact (see Fig. 1*b*). Preferential removal of small weights is consistent with functional connectivity measurements in somatosensory (Le Be and Markram 2006) and hippocampal (Wiegner and Oertner 2013) slices.
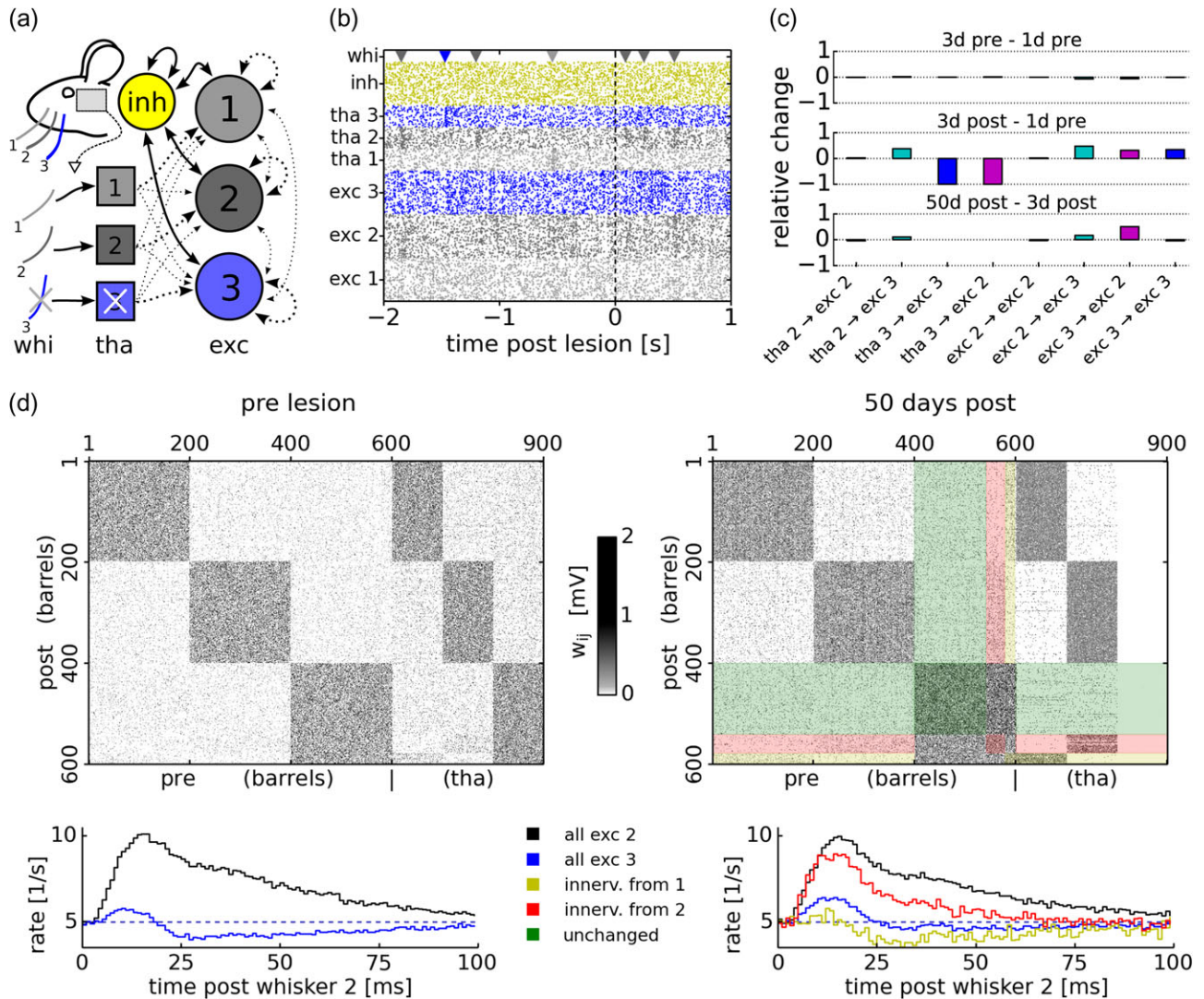
Figure 6. Structural plasticity in a thalamocortical network. (*a*) Schematic of the model. Thalamic neurons (tha) convey sensory input from whiskers (whi) to the recurrent cortical network. Each tha population (squares) projects to 1 of 3 cortical "barrels" (circles) of excitatory (exc) neurons. Cortical inhibitory (inh) neurons connect randomly to all barrels. Synapses with dotted arrows are modeled by the structural plasticity model, all other synapses (solid arrows) are static. Inh synapses have a constant weight and no synaptic failures. Tha neurons increase their firing rate transiently if the corresponding whisker is flicked, which happens randomly with rate 1/s. After 10 days of simulation, whisker 3 is trimmed ("lesion"), modeled as a progressive loss of firing of "tha 3" neurons (white cross). (*b*) Spike raster plot around time of lesion (dashed vertical line), colors as in a. If a whisker is flicked (whi, triangles), the corresponding tha and exc populations respond. (*c*) Relative changes of average connection weights $\langle \Delta w_{ij} \rangle / \langle w_{ij} \rangle$ between populations (here w_{ij} stands for the total weight of the connection from neuron j to neuron i). Changes before (top), around time of lesion (center), and long after (bottom). Strong changes during the first 3 days post-lesion (center) are followed by a slow restructuring process of "exc 3" over the following 47 days (bottom). (*d*) top: Exc synaptic connection weights (gray scale, w_{ij}) just before lesion (left) and 50 days after lesion (right). The comparison of the connection weight matrix before and after lesion shows that loss of the input to "exc 3" caused selective rewiring. "Exc 3" neurons (401 − 600) have been clustered in both graphs according to the inputs they receive at simulation end (assignments are indicated by shading). (*d*) bottom: Average spike response of "exc 2" (black) and "exc 3" (blue) neurons in response to whisker 2 flicks, just before trimming of whisker 3 (left) and at simulation end (right) (averaged over 60min of recording). Red and yellow colors denote subgroups of "exc 3" identified by clustering. Dashed lines: mean firing rate in recording episode. Before rewiring, "exc 3" neurons (blue) respond only weakly to flicks of whisker 2 (left); after rewiring (right) a subgroup (red) responds strongly.

Each potential, but inactive, contact may be re-activated again at random times t_c according to a Poisson process with rate $\lambda_c = .019/$day, which defines the rate of spontaneous spine formation (see Supplementary Material S4). In such an event, referred to as spine creation, the weight is set to $w_{j,k}(t_c)w_c$. Further details on the choice of w_c are described below. As suggested by previous models (Helias et al. 2008) and experiments (Le Be and Markram 2006), in our model, newly created spines first pass through a "period of grace" of duration $\tau_{gp} = 15$min, during which the weight is fixed to w_c. After the period of grace has passed (for $t \geq t_c + \tau_{gp}$), the weight dynamics again follow Equation (8). In the

event of synaptic spine creation (at t_c), the internal state variables $r_{j,k}(t_c)$, $r_{post}(t_c)$, $C_{j,k}(t_c)$, and $R_{post}(t_c)$ of the contact are each initialized to zero. The period of grace serves as a protected time interval for these variables to equilibrate to the current system state, that is, to obtain a good estimate of the present pre- and postsynaptic spike rates and correlations.

Expected Evolution of Synaptic Weights

We derive the expected dynamics of the weight of a single synaptic contact under the plasticity model defined by Equations
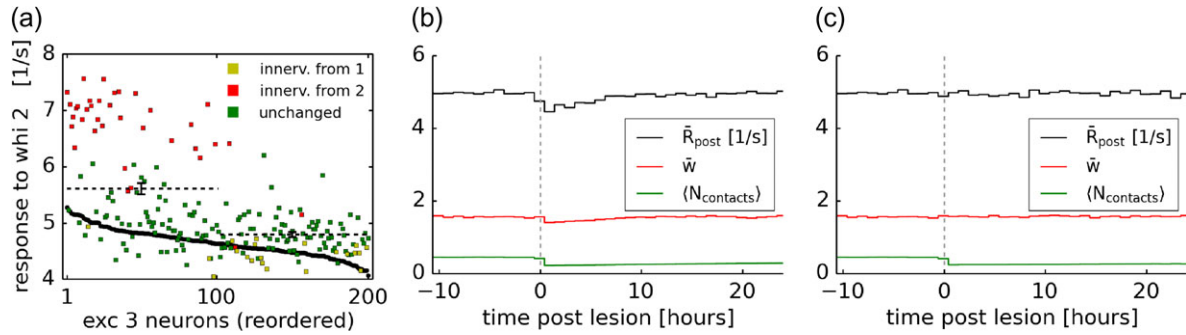
Figure 7. Rewiring can be predicted from initial responses/effect of bounded weights. (*a*) Spike rate in response to whisker 2 flicks for each neuron of "exc 3," before lesion (black dots) and 50 days post (colored squares), for the thalamocortical network simulation shown in Figure 6. Neurons are ordered according to their prelesion response (black). Rates are estimated by counting spikes in a time window of 25ms after each whisker flick. The 100 "exc 3" neurons that initially respond strongest to "whi 2" are more likely to increase their response to this whisker through structural plasticity (dashed horizontal lines with error bars show mean response 50 days post ± SEM) and more likely to participate in the cluster that is most strongly innervated by this whisker (red; colors correspond to the cluster assignments of the neurons in Fig. 6*d*, right). (*b*) Rewiring in response to input lesion (as in Fig. 5) with an upper bound of the contact weight so that contact weights cannot grow stronger than twice the fixed point value ($w_{j,k} \leq 6.4 \cdot 10^{-3}$). In this case, postsynaptic rate and total weight initially decrease in response to the lesion and recover on a time scale of about 10 h as new contacts are formed. Simulation data are averaged over consecutive time windows of 1 h. (c) For comparison, the data of the lesion simulation without upper bound of Figure 5*b*,c are replotted as in (*b*). Here, the remaining contacts quickly (within less than 1 h) compensate for the loss of input by strongly increasing their weights, so that the postsynaptic rate shows no visible change.

(1)–(8). To do so, we take the average, denoted as ⟨·⟩, of Equation (8) over realizations of the spike trains (S) and synaptic transmission failures (z). We obtain

$$\left\langle \frac{d}{dt}w_{j,k}(t) \right\rangle \approx a_2^{\text{corr}}\langle C_{j,k}(t)\rangle - a_4^{\text{corr}}\langle C_{j,k}(t)\rangle^2 - a_4^{\text{post}}\langle R_{\text{post}}\rangle^4 - \alpha w_{j,k}, \tag{9}$$

which is approximate because squaring and averaging of the terms $C_{j,k}$ and R_{post} have been interchanged. The approximation is expected to be good if the fluctuations of $C_{j,k}$ are small compared with its running average $\langle C_{j,k}(t)\rangle$. The terms in Equation (9) can be evaluated as (see Supplementary Material S1 for the derivation)

$$\langle R_{\text{post}}\rangle = \langle S_{\text{post}}\rangle = \langle \lambda \rangle = \lambda_0 + \nu_{\text{pre}}(1-p_f)\sum_j^N \sum_k^{n_j} w_{j,k}, \tag{10}$$

$$\langle C_{j,k}\rangle = \langle r_{j,k}r_{\text{post}}\rangle = \langle r_{j,k}\lambda \rangle = \nu_{\text{pre}}(1-p_f) \cdot \left[\frac{1}{2\tau}e^{-d/\tau}(p_f w_{j,k} + (1-p_f)\sum_l^{n_j} w_{j,l}) + \langle R_{\text{post}}\rangle \right]. \tag{11}$$

Equation (10) establishes that the postsynaptic rate is determined by the sum of all synaptic weights.

We note that, according to Equation (11), for small transmission failure probability $p_f \to 0$, the dynamics of the contact $w_{j,k}$ (Eq. (9)) is dominated by the total weight $w_j = \sum_l w_{j,l}$. This means that different contacts arising from the same presynaptic neuron interact with each other. For large $p_f \to 1$, the evolution of $w_{j,k}$ is independent of w_j; so increasing the failure probability p_f gradually decouples the dynamics of the contacts.

Co-operation and Competition

Equations (9)–(11) enable us to illustrate the process of co-operation and competition, closely linked to the stabilization of the postsynaptic rate and correlations. The postsynaptic rate does not depend on the weight of any specific synaptic contact, but only on the total input, summed over all weights and contacts; see Equation (10). By contrast, Equation (11) depends not only on the total input via the rate R_{post} but in addition also on the individual weight $w_{j,k}$ and the total weight $w_j = \sum_k w_{j,k}$

arising from the same presynaptic neuron. To study competition, let us consider a uniform state and suppose that all correlations $C_{j,k} = c$, and all momentary weights $w_{j,k} = w$ for all j, k are small but positive. With $\alpha \ll 1$ in Equation (9), the dominant evolution is therefore an increase of all weights, driven by the term $a_2^{\text{corr}}c$. However, as the weights increase, the firing rate does so as well and therefore the term $\langle R_{\text{post}}\rangle^4$ eventually stops further growth. This is the essential step of firing rate stabilization (for a mathematical demonstration, see Supplementary Material S2.1).

For the same firing rate $\langle R_{\text{post}}\rangle$, some weights will grow further at the expense of others, inducing competition via the instability of the uniform state, just as in other models (Oja 1982). The instability is caused by a positive feedback loop between $\langle dw_{j,k}(t)/dt\rangle$ on the left-hand side of Equation (9) and $w_{j,k}$ on the right-hand side of Equation (11). Going beyond standard plasticity models, Equation (11) shows that correlations $C_{j,k}$ driving the contact weight $w_{j,k}$ increase not only proportionally to this specific contact but also increase with the weight of other contacts $\sum_l w_{j,l}$ from the "same" presynaptic neuron. The positive dependence gives rise to co-operation between contacts arising from the same neuron. The optimal amount of correlation, and hence co-operation, however, is limited by the term $-a_4^{\text{corr}}\langle C_{j,k}\rangle^2$ in Equation (9). The interplay of Equations (9)–(11) therefore stabilizes the firing rate, or total input $\sum_j \sum_k w_{j,k}$, as well as the amount of correlations in actual contacts (see Supplementary Material S2.2 for a mathematical argument).

Calibration of the System Dynamics

Since the firing rate stabilizes (Supplementary Material S2.1) and because we assume pre- and postsynaptic neurons to be of the same type, we have chosen the parameters of the plasticity rule such that the stable firing rate of the postsynaptic neuron is on average $\langle R_{\text{post}}\rangle \approx \nu_{\text{post}} \approx \nu_{\text{pre}}$. Our simulations show that this value is tightly maintained; for example, after a lesion, synaptic plasticity re-adjusts the synaptic weights so that the postsynaptic rate ν_{pre} converges back to the value it had before the lesion (Fig. 5*c*). This implies that the sum of weights $\sum_j \sum_k w_{j,k}$ is normalized (Kempter et al. 1999; Song et al. 2000),

see Equation (10). Indeed, previous theoretical work (Zenke et al. 2015) has shown that terms like $-R_{post}^4$ in our learning rule lead to a normalization of the total weight. Regarding the calibration of other parameters of the plasticity rule, see the next section as well as Supplementary Material S4.

Model Analysis

Since the firing rate of the postsynaptic neuron stabilizes (Supplementary Material S2.1), the actual degrees of freedom of the multiconnection multicontact system reside in the configuration of contact weights $w_{j,k}$ under the constraint that the total input weight $\sum_j \sum_k w_{j,k}$ (which determines the firing rate via Eq. (10)) is fixed. For example, in a first configuration, all contact weights $w_{j,k}$ from all presynaptic neurons may have the same value while in a second configuration some weights could be relatively strong and others zero. In the following, we use the expected dynamics and show that there exist several potential configurations of contact weights that yield stationary solutions to the model equations.

Inserting Equations (10) and (11) into Equation (9) yields a closed, nonlinear dynamical system of ordinary differential equations in $w_{j,k}(t)$, with $1 \le k \le n_j$ and $1 \le j \le N$. To understand these dynamics, we proceed in two steps.

i. We focus on a solution where some presynaptic neurons j have no active contacts ($w_{j,k} = 0$ for $1 \le j \le n_j$) while other presynaptic neurons n have exactly $m > 0$ active synaptic contacts, each with the same weight $w_{n,k} = \frac{w_{n.act}}{m}$ for $1 \le k \le m$ (and $0 = w_{n,k}$ for $m < k \le n_n$). To be specific, we assume that there are $N_0 < N$ neurons with m active contacts each where N_0 is, so far, unknown. By definition, each of the N_0 neurons with active contacts has a total connection strength of $w_n = m\, w_{n,k} = w_{n,act}$. To construct this solution, we choose m, rewrite Equations (9)–(11) in terms of $w_{n,act}$ and search for a fixed point $w_{n,act} = w_*$ (see Eq. (S18), Supplementary Material). Note that the value of w_* will depend on m. Possible combinations w_*, m correspond to the crossing points between the curve and the straight dot-dashed lines in Figure 1d. Given our choice of m and our calculated value w_*, the postsynaptic firing rate is $\langle R_{post} \rangle = \lambda_0 + \nu_{pre}(1 - p_f)N_0 w_*$. Because of the firing rate stabilization at $\langle R_{post} \rangle = \nu_{pre}$, we can finally extract the parameter N_0. Therefore, we have shown that combinations of parameters N_0, w_*, m exist that give rise to stationary solutions of the model equations. Each of these parameter combinations characterizes one configuration of synapses – but for each of these configurations, there are many combinations of choosing N_0 active neurons out of a total of N. We cannot exclude that there are also other stationary states outside the class of configurations considered here.

ii. To analyze stability against weight perturbations within a group of m contacts, we also study the expected change of the weight of a "single" contact $\frac{d\langle w_{j,k}\rangle}{dt}$ in one specific synaptic connection of total weight w_j (which may change as well) assuming the total firing rate R_{post} to remain constant. The expected dynamics (Eq. (9)) allow us to predict 2D trajectories of $(w_{j,k}, w_j)$ (Fig. 1d and Supplementary Material S2.4). Parameter calibration was done such that a synaptic connection consisting of 3 or more contacts is stable in expectation but a connection with only 1 or 2 contacts is not (see also Supplementary Material S2.3 and Supplementary Fig. S4). A phase plane analysis of the dynamics of two contacts (Supplementary Fig. S4B) further confirms that, in

expectation, all connections of only 2 contacts are eventually removed because of the absence of stable fixed points. However, because of a region on the phase plane where the dynamics are slow, a connection with 2 contacts has a lifetime that is sufficiently long such that occasionally a third contact may be added. Still, due to the low creation rate, we expect this event to be rare.

Overall, our analysis shows that for the chosen combination of parameters, there are several stable fixed points of weight configurations. If we read off the value of $w_{j,k}$ from Figure 1d, we find that for $m = 3$ contacts, we would have $N_0 = 140$ active presynaptic connections, whereas with $m = 5$, we have about 100, and with $m = 10$ about 80 presynaptic connections. Overall, we expect a mix of several of these solutions. In Supplementary Material S2.6, we give arguments why a solution with $m = 10$ is less likely to occur than a solution with $m = 5$. Qualitatively, the analysis shows that, whatever the mix of different m-values, about 10% of the 1000 synaptic connections are active (consistent with Markram, Lübke, Frotscher, Roth et al. (1997a) and Fares and Stepanyants (2009)). These theoretical considerations suggest that the system is indeed calibrated to have a steady state that is qualitatively consistent with the experimental contact number distributions, in the sense that stable connections have at least 3 synaptic contacts.

Although all simulations have been performed with the same set of parameters, there is in fact a family of parameters that all lead to solutions with the following properties: connections with 5 or more contact points are stable fixed points, whereas those with 1 or 2-contact points are not (Supplementary Material S2.5 and Supplementary Fig. S6). Thus, the main qualitative properties of the model are independent of the specific choice of parameters. Moreover, while all the simulations have been performed with presynaptic and postsynaptic firing rates of 5 Hz, model parameters can be adjusted to firing rates of 0.5 Hz, such that we have stationary solutions with the following properties: 3 or more contact points are stable fixed points, whereas 1 or 2-contact points are not (Supplementary Material S2.5 and Supplementary Fig. S5). Thus, at low firing rates, our synaptic and structural plasticity model with appropriate parameters is expected to show the same qualitative behavior as the version at 5 Hz shown in the Results section.

SNR of Synaptic Responses

To better understand the effects of multiple synaptic contacts in the presence of stochastic transmission, let us analyze the postsynaptic response. To this aim, we force the presynaptic neuron j to emit an additional spike at time t_{pre}. For convenience, we neglect the synaptic transmission delay here ($d \to 0$), which has no effect on the following reasoning. Assuming constant weights on the short time scale of synaptic signaling τ, and by averaging Equation (S5) (see Supplementary Material) over all other presynaptic spikes and their synaptic transmission, we obtain the transient postsynaptic response:

$$L_j(t \mid t_{pre}) = \nu_{post} + \frac{1}{\tau}\theta(t - t_{pre})e^{-(t-t_{pre})/\tau} \sum_{k=1}^{n_j} w_{j,k} z_{j,k}(t_{pre}), \quad (12)$$

where we also inserted $\langle R_{post} \rangle \approx \nu_{post}$. Here $\theta(s)$ denotes the Heaviside step function, which is 1 for $s > 0$ and vanishes elsewhere. Note that $L_j(t \mid t_{pre})$ is still a stochastic quantity due to stochastic synaptic transmission $z_{j,k}(t_{pre})$ at the contacts k. We may obtain the mean spike-triggered response by averaging

over the remaining stochasticity of the transmission variable $z_{j,k}$:

$$\langle L_j(t \mid t_{\text{pre}})\rangle = \nu_{\text{post}} + \frac{1}{\tau}\theta(t - t_{\text{pre}})e^{-(t-t_{\text{pre}})/\tau}(1 - p_{\text{f}})w_j \ .$$

Thus, the average response only depends on the total synaptic weight w_j and not on the configuration of contact weights $w_{j,k}$. Similarly, the variance of the response can be derived as

$$\text{var}[L_j(t \mid t_{\text{pre}})] = p_{\text{f}}(1 - p_{\text{f}})\frac{1}{\tau^2}\theta(t - t_{\text{pre}})e^{-2(t-t_{\text{pre}})/\tau}\sum_{k=1}^{n_j} w_{j,k}^2 \ .$$

Here, we see that the contact configuration $w_{j,k}$ determines the variance of the postsynaptic response. To further understand these properties, consider a synaptic weight w_j that is made of m contacts of weight $w_{j,k} = w_j/m$ while all the remaining $n_j - m$ weights are zero. Then, the sum of squared weights term in $\text{var}[L_j(t)]$ becomes $\sum_{k=1}^{n_j} w_{j,k}^2 = w_j^2/m$. For this case, we evaluate the SNR of the synaptic response as

$$\text{SNR}_j = \frac{\langle L_j(t \mid t_{\text{pre}})\rangle - \nu_{\text{post}}}{\sqrt{\text{var}[L_j(t \mid t_{\text{pre}})]}} = \sqrt{\frac{1 - p_{\text{f}}}{p_{\text{f}}} \cdot m} \ . \tag{13}$$

Therefore, in the presence of synaptic transmission failures, multiple synaptic contacts increase the SNR of synaptic transmission, proportional to the square root of the number of contacts. Previously, a related result has been found numerically for the mutual information of synaptic inputs and neural outputs via multiple contacts (Zador 1998).

Simulation of the Plasticity Model

All simulations were performed using the Neural Simulation Tool (NEST) (Eppler et al. 2015); for further details, see Supplementary Material S3.

Recurrent Network Model

The recurrent thalamocortical network model presented in Figure 6 consists of $N_B = 3$ "barrel columns" of $N_E = 200$ excitatory (labeled "exc") neurons each, and $N_I = 200$ inhibitory (labeled "inh") neurons that connect without preference to all the barrels (random connections, see details below). All cortical neurons are modeled as leaky integrate-and-fire (LIF) neurons with alpha-function–shaped postsynaptic currents ('iaf_psc_alpha' neuron model in NEST simulator). The parameters of the LIF neurons are membrane time constant $\tau_{\text{LIF}} = 20.6$ms, reset and resting potential −70mV, action potential threshold −55mV, synaptic time constant 2ms, and refractory period 2ms. There are $N_B \cdot N_E + N_I = 800$ cortical neurons in total. All synaptic delays are 1ms.

Each barrel of excitatory neurons is further innervated by thalamic inputs that convey information from the whiskers. There are $N_T = 100$ thalamically driven input neurons (labeled "tha") per barrel. All $N_B \cdot N_T = 300$ input neurons are modeled as excitatory linear Poisson neurons according to Equation (3), with a baseline firing rate of $\lambda_0 = 4.5$/s. Each group of thalamically driven neurons modulates its firing rate in response to flicks of the corresponding whiskers (whi). The sequence of whisker flicks is stochastic and described by Poisson processes with rate $\nu_{\text{whi}} = 1$/s each. Each neuron in group "tha n" responds to each flick of the corresponding whisker number n. In response to a whisker flick, thalamically driven neurons of the receiving population increase their firing rate $\lambda(t)$

transiently by $w_{\text{whi}}/\tau = 25$/s, and subsequently their rate decays back to λ_0 with time constant τ (cf. Eq. (3)).

The structural plasticity model of Equation (8) describes all excitatory-to-excitatory connections, both thalamocortical and intracortical, and is continuously active (except for the first hour of simulation); autapses are excluded. All connections (both from "tha" neurons to cortical excitatory neurons and between cortical excitatory neurons) have a distribution of *potential* synaptic contacts, but initially the network is connected with "active" contacts in a whisker-specific manner as depicted in Figure 6a, see also below. Because the recurrent network contains many postsynaptic neurons, we need two indices to name a synaptic connection. A contact weight here is denoted by $w_{ij,k}$ instead of $w_{j,k}$ above, where i denotes the postsynaptic and j denotes the presynaptic neuron, and k the contact. Accordingly, the plasticity rule Equation (8) here reads

$$\frac{d}{dt}w_{ij,k}(t) = a_2^{\text{corr}}C_{ij,k}(t) - a_4^{\text{corr}}C_{ij,k}^2(t) - a_4^{\text{post}}R_i^4(t) - \alpha w_{ij,k}(t), \tag{14}$$

and the total weight of a synaptic connection is $w_{ij}(t) = \sum_{k=1}^{n_{ij}} w_{ij,k}(t)$, where n_{ij} is the number of potential contacts for the connection from j to i. For simplicity, all n_{ij} are drawn from the probability distribution $P(n_{ij})$ (Fig. 2a, blue line), irrespective of which group the neurons i and j belong to.

Because the postsynaptic neurons here are LIF neurons, synaptic efficacies $w_{ij,k}$ have to be expressed in units of the PSP (in contrast, above $w_{j,k}$ is a unit-less quantity). To match the impulse response function of the LIF neurons receiving an input spike with the fixed-point weight w_* (see Model Analysis) to the response of the linear Poisson neurons used above, we scale the synaptic weights as $w_{ij,k} = \gamma \cdot w_{ij,k}$, with $\gamma = 62.82$ mV, leading to a typical EPSP amplitude of $\gamma w_* = 1.01$ mV. Substituting $w_{ij,k}$ into Equation (14) implies that, to maintain the same plasticity dynamics as above, the parameters of the learning rule have to be rescaled according to $a_2^{\text{corr}} \mapsto \gamma a_2^{\text{corr}}$, $a_4^{\text{corr}} \mapsto \gamma a_4^{\text{corr}}$, $a_4^{\text{post}} \mapsto \gamma a_4^{\text{post}}$, and $w_0 \mapsto \gamma w_0$. We further inject additional Poisson excitatory and inhibitory input spikes to all LIF neurons, with synaptic weights γw_* (excitation) and $-4\gamma w_*$ (inhibition). Excitatory neurons receive input at rates 1519.2/s (excitation) and 328.3/s (inhibition), and inhibitory neurons receive 1391.0/s (excitation) and 351.2/s (inhibition). The scaling factor γ and the Poisson process input rates were numerically optimized to match the dynamics of the LIF neuron model to that of the linear Poisson neuron model used above. All other parameters take the same values as before. Note that, in order to enable a comparison with the single-neuron case, the network parameters here are chosen such that, under the assumption of uncorrelated Poisson spike train input to each neuron, synapses in the full network would operate approximately at the fixed point w_* of the plasticity dynamics (see Model Analysis above), derived for the single-neuron case.

Apart from its $(N_B[N_T + N_E])(N_B[N_E - 1]) \approx 5.4 \cdot 10^5$ plastic excitatory connections, our network also has static synapses. These we set as follows. We choose a connection probability of $p_{\text{conn}} = 1/3$. Each excitatory neuron receives $p_{\text{conn}}N_I$ synapses from randomly chosen inh neurons with a fixed weight of $-(1 - p_{\text{f}})\gamma g w_*$, with $g = 2.5$. Each inhibitory neuron also receives this number of inhibitory synapses, and $p_{\text{conn}}N_E$ excitatory synapses from randomly chosen excitatory neurons from each of the N_B cortical barrels, with a fixed weight of $(1 - p_{\text{f}})\gamma w_*$ (these synapses have no transmission failures; therefore, the weight is scaled down by the expected transmission rate $(1 - p_{\text{f}})$ of the plastic, stochastic ones).

We initialize the plastic synapses at the theoretically derived fixed point w_*, with an expected total of 100 active input connections with 5 active contacts each. So, for each excitatory-to-excitatory and each thalamocortical connection, we set $w_{ij}(0) = \gamma w_* q_{ij}$ if neuron i and neuron j belong to the same barrel, where q_{ij} is a Bernoulli random number that is 1 with probability p_{conn} and 0 else (in this way, we get $(N_T + N_E)p_{conn} = 100$ incoming connections per excitatory neuron in expectation). If i and j are part of different barrels, we set $w_{ij}(0) = 0$. If there are 5 or more potential contacts ($n_{ij} \geq 5$) in connection i, j, we set $w_{ij,k}(0) = w_{ij}(0)/5$ for $1 \leq k \leq 5$ and $w_{ij,k}(0) = 0$ for $k \geq 5$. If there are less contacts ($n_{ij} < 5$) but $w_{ij}(0) > 0$, we set $w_{ij}(0) = 0$ and look for a connection i', j' that connects the same 2 groups, has $w_{i'j'}(0) = 0$ and ($n_{i'j'} \geq 5$), and we set $w_{i'j'}(0) = \gamma w_*$ for this connection instead.

In Figure 6d, neurons of barrel column named "exc 3" are ordered according to labels obtained by clustering. We used feature agglomeration based on Ward's hierarchical clustering (Pedregosa et al. 2011) to assign 1 of 3 cluster labels to the vector $\{w_{i1}, w_{i2},...\}$ of connection weights (at simulation end) from any excitatory neuron onto each neuron i in the group "exc 3".

Results

Stable Bimodal Distribution of Synaptic Contacts

We devised a plasticity model linking STDP of dendritic spines with structural plasticity of synaptic contact formation and removal (Fig. 1a,b, see Materials and Methods for details). In this model, each pair of neurons is connected by several "potential synaptic contacts." The number of potential contacts in the model was broadly distributed as reported previously (Fares and Stepanyants 2009) so that each pair of neurons had a least 1 potential contact and at most 10 (Fig. 2a, blue line). In our model, individual contact weights are subject to STDP mediated by local traces of the pre- and postsynaptic spiking activity (Fig. 1c). If a contact weight decreases to zero, it is removed and the contact is labeled as an "inactive," but potential, contact. An inactive contact in the model can become spontaneously active at a constant rate of $\lambda_c = 0.019$/day.

Correlations between the traces left by successfully transmitted presynaptic spikes and postsynaptic spikes lead, in our model, to a symmetric STDP window for potentiation, the spike-based equivalent of standard Hebbian learning (Hebb 1949; Gerstner and Kistler 2002), but more elaborate models of plasticity (Bienenstock et al. 1982; Song et al. 2000; van Rossum et al. 2000; Shouval et al. 2002; Pfister and Gerstner 2006; Morrison et al. 2008; Clopath et al. 2010) could be implemented in the same modeling framework (see Materials and Methods and Discussion). We combined Hebbian potentiation with heterosynaptic plasticity that downregulates all synapses during episodes of large postsynaptic firing rates (Chen and Nedivi 2013; Chistiakova et al. 2015; Zenke et al. 2015). Mathematical analysis of our model shows that the combination of Hebbian potentiation with heterosynaptic plasticity leads to a stabilization of postsynaptic firing rates (Supplementary Material, S2.1). Furthermore, we postulate in our model a decrease of Hebbian potentiation for large correlations, which could manifest itself as a novel form of "anti-Hebbian" synaptic depression when correlations between pre- and postsynaptic firing become very strong (see Materials and Methods for model equations).

When a simulated postsynaptic neuron is stimulated with stochastic spike arrivals at many synapses, our model of synaptic plasticity causes a few presynaptic neurons to develop connections with 5–8 active contacts (Fig. 2a), whereas many other presynaptic neurons have zero active contacts. Our mathematical analysis (see Model Analysis in Materials and Methods) shows that synaptic configurations with zero contacts and those with 3 or more contacts are stable, whereas those with 1 or 2 contacts are not. For example, if one of the contact weights of a presynaptic neuron with 5 contacts is perturbed, the weight returns back to its stable value $w_*/5$ (Fig. 1d1), which we can predict analytically (Supplementary Material S2.3). If a new contact weight is spontaneously created in a connection with 5 contacts, the new connection with 6 contacts will become stable (Fig. 1d2). However, if a new contact weight is spontaneously created in a potential connection without existing contacts, the new connection with 1 contact is not stable and decays (Fig. 1d3). The stability of multicontact connections does not require an upper bound of the contact weight; indeed, our mathematical analysis (and simulations) were done without bounding the weight, in contrast to many standard STDP model (Kempter et al. 1999; Song et al. 2000; Morrison et al. 2007); see the "Discussion" for the effect of bounded weights.

We found that the model with the novel correlation-driven synaptic depression term leads, for a broad regime of parameter values, to a bimodal distribution of synapses (see Supplementary Material S2.3 and S2.5 for mathematical arguments). We emphasize that stability of multiple contacts is possible, even if the number of actual contacts (e.g., 5) is well below the maximal number of potential contacts (e.g., 10) for a given presynaptic neuron (Supplementary Material S2.3 and S2.6). The stability of weights in multicontact synapses disappears, if the novel anti-Hebbian correlation-driven synaptic depression is removed (see Eq. S25 in Supplementary Material with $a_4^{corr} \rightarrow 0$), indicating the importance of this term. Overall, our mathematical analysis (see Model Analysis in Materials and Methods and Supplementary Material S2) predicts a bimodal distribution of actual synaptic contact numbers in the following sense: most pairs of pre- and postsynaptic neurons have no active connection at all while the remaining pairs of neurons exhibit 3 or more active contacts. Thus, the distribution of actual contact numbers has a trough at 1 or 2 contacts, consistent with experiments in layer 5 pyramidal neurons (Markram, Lübke, Frotscher, Roth et al. 1997a), whereas the distribution of potential contact numbers is unimodal with a broad peak at $n = 2$ contacts.

The bimodal distribution of contact numbers predicted by our mathematical analysis has been confirmed in simulations of 1000 presynaptic neurons projecting onto a single postsynaptic neuron (Fig. 2a). Each presynaptic neuron had a fixed, but neuron-specific, number of potential synaptic contacts (Fig. 2a, blue line) with a distribution estimated from neuron reconstructions (Fares and Stepanyants 2009). The synaptic plasticity rule was driven by stochastic spike arrivals at 5 Hz and led to a bimodal distribution of the number of actual contacts (Fig. 2a, red line) qualitatively consistent with experiments (Markram, Lübke, Frotscher, Roth et al. 1997a). In the following, we keep these parameter values for the plasticity rule fixed and explore its properties in additional simulations.

STDP-Mediated Co-operation Stabilizes Contact Weights

During the simulation of 1000 presynaptic neurons firing stochastically at 5/s (independent Poisson processes) and projecting onto the same postsynaptic neuron, the synaptic contact weights fluctuate, after an initial transient, around a steady state (Fig. 2b), which can be predicted by our mathematical

analysis (Fig. 2b, dashed line). Occasionally, new contacts are formed, which mature or are quickly removed (Fig. 2b, inset). The firing rate of the postsynaptic neuron as well as the total synaptic weight and the average number of contacts are stabilized by the plasticity rule (Fig. 2c and Model Analysis in Materials and Methods and Supplementary Material S2.1). The average number of actual contacts $\langle N_{\text{contacts}} \rangle$, corresponding to the spine density, is stationary in the model (Fig. 2c) consistent with experiments (Trachtenberg et al. 2002; Holtmaat et al. 2005; Holtmaat and Svoboda 2009; Loewenstein et al. 2015), while individual contacts are created and pruned (Fig. 2b).

Within 1 day (1d), the relative change of typical contact weights $w_{j,k}$ (i.e., contact k from a presynaptic neuron j to the postsynaptic neuron) ranges from −100% to 500%, but fluctuations of $\Delta w_{j,k}/w_{j,k}$ decrease with increasing contact weight (Fig. 2d), consistent with long-term time-lapse imaging data of dendritic spine volume in vitro (Yasumatsu et al. 2008). We quantify the statistics of contact weight changes $\Delta w_{j,k}(t) = w_{j,k}(t) - w_{j,k}(t - 1d)$ by computing the mean and STD of these changes for different groups of contact weights $w_{j,k}(t - 1d)$, averaged over days of simulated time t. Consistent with experimental measurements of dendritic spine volume (Yasumatsu et al. 2008), the mean contact weight change (Fig. 2e, black) is positive for an intermediate range of contact weights (spine volumes) and becomes negative for large weights (large volumes). Contact removal over 1 day occurs only for weights smaller than some threshold, consistent with functional (Le Be and Markram 2006) and optical (Wiegert and Oertner 2013) connectivity measurements. The STD of the changes (Fig. 2e, red) is rather homogeneous for all weights (volumes) but increases slightly for very large weights. The mean change of very small contact weight is negative in our model, in contrast to previous experiments (Yasumatsu et al. 2008), but this difference might be due to experimental difficulties of observing very small spines (see Discussion).

As synaptic connections in our model consist of several synaptic contacts, we asked whether there is a systematic relation between synaptic weight and the number of contacts in the steady state (Fig. 2f). Indeed, the synaptic weight is strongly correlated with the number of actual contacts: Strong synaptic connections in our model consist of at least 5 synaptic contacts and connections with less than 3 contacts are physiologically weak (Fig. 2f, red, left axis). Moreover, we find that in our model individual contacts in connections made of more than 3 actual contacts tend to be stronger than those made of less than 3 (Fig. 2f, blue, right axis).

We further assess the statistical properties of the steady state by multiple measures. First, the distribution of actual synaptic contact numbers per connection (Fig. 3a) is bimodal, in line with experimental findings (Markram, Lübke, Frotscher, Roth et al. 1997a; Fares and Stepanyants 2009), see Figure 2a. In particular, the distribution of actual contacts goes through a clearly visible minimum at around 3 contacts (Fig. 3a), and this observation is stable over 100 days of simulated time. Second, the turnover ratio of synaptic contacts in the model is 0.176 ± 0.018/day (mean ± STD) consistent with the values found experimentally in somatosensory cortex (Holtmaat et al. 2005; Holtmaat and Svoboda 2009). Third, a stable distribution of contact weights (Fig. 3b) is formed, which also hardly changes over 100 days of simulated time.

The probability of a synaptic contact in our model simulations to survive for 8 consecutive days, irrespective of whether the contact is newly created, weak, or strong, depends strongly on the number of actual contacts in the connection that the contact is part of (Fig. 3c). In other words, contacts within a connection co-operate and stabilize each other. Tracking of individual synaptic contacts that existed at $t = 0$ for 150 days of simulated time (Fig. 3d) reveals a time course of the number of surviving synapses in our model that is qualitatively consistent with long-term in vivo imaging data of dendritic spines from mouse neocortex (Yang et al. 2009), even though the survival fraction in our model is slightly lower than their data and higher than in some other experimental measurements (Holtmaat et al. 2005; Keck et al. 2008; Loewenstein et al. 2015). Connections that consist of several contacts are stable in our model for long periods of time (here 150 days) (Fig. 3e, inset). The pruned connections in our model are almost exclusively connections that consist of a single contact and have relatively small total weight (Fig. 3e). Finally, the lifetime of synaptic connections increases strongly with the total weight of the connection (Fig. 3f), indicating that strong connections are protected against spine turnover by mutual co-operation of contacts. Overall, these findings from our simulation study of synaptic contacts are consistent with functional connectivity measures in layer 5 pyramidal neurons (Le Be and Markram 2006). The computational model enabled us to directly follow the weights, and statistics, of individual contacts over 100 h of simulated time, whereas Le Be and Markram had to infer the strength of individual contacts and their multiplicity indirectly from measurements over a much shorter amount of time.

Contact Multiplicity is Crucial for Synaptic Stability

To investigate the role of multiple synaptic contacts in our model, we compare its dynamics to the same model restricted to single-contact connections. In order to preserve the total amount of potential synaptic contacts of the full model, which is necessary for a fair comparison, we increase the number of presynaptic neurons in the single-contact model to compensate for the reduced number of potential contacts per presynaptic neuron (see Synaptic Contacts in Materials and Methods). Similar to the full model, the single-contact model (Fig. 4a) exhibits steady-state dynamics in which postsynaptic firing rate, total weight, and synaptic contact number are tightly regulated (data not shown), and the distribution of contact weights is stable over time (Fig. 4b). However, in contrast to the multi-contact model (Fig. 3), in the single-contact model, the temporal dynamics of synaptic contact survival do not indicate the presence of a subgroup of stable connections (Fig. 4c), and all connections are prone to random pruning, irrespective of whether they have a small or large synaptic weight (Fig. 4d,e). Thus, co-operation in multiple synaptic contacts from the same presynaptic neuron is crucial for the long-term stability of strong synaptic connections (Fig. 3e,f): in multicontact synaptic connections, contacts support each other via Hebbian synaptic plasticity and give rise to stable connectivity patterns despite turn over of spines (Le Be and Markram 2006). This co-operativity between synaptic contacts from the same presynaptic neurons is analogous to the well-known co-operativity of Hebbian learning in the presence of correlated input (Bienenstock et al. 1982; Oja 1982; Kempter et al. 1999; Song et al. 2000; Helias et al. 2008) but occurs here even in the absence of correlated inputs.

A theoretical analysis further explains the crucial role of contact multiplicity in synaptic transmission. We characterize the fidelity of transmission of a presynaptic spike by the ratio of the (trial-averaged) mean and STD of the evoked PSP, which we call the SNR (see Materials and Methods). In our model, failures of synaptic transmission occur randomly at each synaptic

contact, which causes variability of the PSP. If a synaptic connection consists of several contacts, a presynaptic spike is more reliably transmitted than in the case of a single contact (Fig. 4*f*). The mathematical analysis shows that the SNR is proportional to $\sqrt{\frac{1-p_f}{p_f}} \cdot m$, where m is the multiplicity of the contact and p_f the synaptic failure probability. If 10 contacts are considered to be the maximum number of actual contacts connecting a pair of neurons due to geometrical constraints of the tissue, about 6 contacts give a SNR of $\sqrt{6/10} \approx 80\%$ of the maximum value achievable – and this result is independent of the synaptic failure rate p_f (Fig. 4*f*). In electrophysiological experiments, the relatively high reliability of synaptic transmission in layer 5 neurons of the somatosensory cortex of the rat has previously been interpreted as a signature of multiple synaptic contacts per connection (Le Be and Markram 2006; Nawrot et al. 2009). Intuitively speaking, even if 3 out of 6 contacts in a multicontact connection fail to transmit, there will still be a large response of the postsynaptic neuron. Our model shows that this postsynaptic response in turn contributes to stabilizing all contacts in the multicontact connection via Hebbian synaptic plasticity – even those contacts that have failed to transmit (see Model Analysis in Materials and Methods).

Presynaptic Lesions Lead to Increased Formation of Persistent Contacts

Experimentally, structural plasticity can be induced by trimming whiskers (Trachtenberg et al. 2002; Holtmaat et al. 2006), lesions of the retina (Keck et al. 2008), or occlusion of one of the eyes (Rose et al. 2016). In our full model with multiple potential contacts for each of the 1000 presynaptic neurons, after ablating 50% of those presynaptic model neurons that have active synaptic contacts (Fig. 5*a*), we observe a loss of 50% of the synaptic contacts on the postsynaptic neuron after 30 min (Fig. 5*b*). However, this loss is rapidly compensated such that the firing rate and total weight (summed over all presynaptic neurons and all synaptic contacts) are hardly changed throughout the process (Fig. 5*b*,*c*). The compensation occurs on 2 different time scales. First, on the time scale of 10–30 min, existing synaptic contacts are upregulated from a prelesion value of $(3.3 \pm 1.3) \cdot 10^{-3}$ to a value of $(6.6 \pm 1.3) \cdot 10^{-3}$ measured 30 min after the lesion (Fig. 5*d*,*f*). Second, on the slow time scale of 10–30 days after the lesion, the number of presynaptic neurons without postsynaptic contact decreases from 886 prelesion to 810, 30 days after lesion while the number of presynaptic neurons with 1, 2, or 3 contact points transiently increases (Fig. 5*e*), suggesting that the plasticity rule "tests" new connection patterns. Competition between synaptic contacts from different presynaptic neurons and simultaneous co-operation of synaptic contacts arising from the same presynaptic neuron leads to pruning or strengthening, so that after 99 days, the distribution of contact numbers and synaptic weights is again very similar (but not yet completely identical) to the prelesion distributions (Fig. 5*e*). The gradual recovery of the contact numbers is due to an elevated probability of newly formed contacts to survive and become long-term stable compared with the control condition (Fig. 5*g*). In a simulated lesion experiment where 20% of the actual contacts are removed (instead of 50% in the simulations so far), 14.6% of newly created contacts survive for 8 days or more, consistent with experimental results on dendritic spines in the somatosensory cortex after whisker trimming (Holtmaat et al. 2006) and significantly above the 7.7% of surviving contacts in the control condition (Fig. 5*g*).

Rewiring of Input-Deprived Cortical Barrels Under Structural Plasticity

We wondered whether our model would generalize from a single postsynaptic neuron to predict structural changes in large recurrent networks of excitatory and inhibitory neurons. Therefore, in this section, we do no longer focus on the distribution of synaptic contact numbers but take a slightly more macroscopic view and ask how connectivity patterns develop in networks and how these patterns re-adapt after sensory deprivation.

Our network architecture (Fig. 6*a*) is inspired by rodent barrel cortex with 3 strongly connected cortical populations (representing the columns corresponding to different barrels) of excitatory neurons (labeled "exc" 1–3), each preferentially innervated by a thalamically driven population (labeled "tha" 1–3) that conveys sensory input from 1 of 3 whiskers (labeled "whi" 1–3) by strong connections. In the model, connections between different cortical populations and from thalamically driven populations to nonpreferred cortical barrel columns are random and weaker on average, but all excitatory connections, whether strong or weak, are subject to the same plasticity model. To simplify the simulations, we stay at the level of an abstract model and do not distinguish between different cortical layers, even though there are known differences in response properties; similarly, the thalamically driven populations in our model are not necessarily located in the thalamus but could instead represent layer 4 excitatory neurons. Note that although biological connections from thalamus to cortex are plastic during the critical period, they might not be thereafter (Crair and Malenka 1995), whereas connections from thalamically driven layer 4 neurons to excitatory neurons in other cortical layers might remain plastic similar to other cortical connections.

A short movement (flick) of the whisker is represented in our model by a small increase in the firing rate of the neurons in the thalamically driven population and results in a modest increase of the firing rate of the corresponding cortical population (Fig. 6*b*) above a spontaneous network activity of about 5 Hz (Fig. 6*d* bottom left). After an initial transient of 7 days of simulated time, we followed the mean connection weights of synapses from several groups (from tha 2 to exc 2, from tha 3 to exc 3, from tha 2 to exc 3, from exc 3 to exc 3, and from exc 2 to exc 2) during 3 days of simulated time and found no changes (Fig. 6*c*, top), indicating that the average connectivity pattern is globally robust during ongoing activity and random whisker stimulation, despite structural plasticity always being active. Note that, by model design, the neuronal activity inside each population is correlated. Since correlated inputs to a given neuron change the distribution of synaptic weights and contact numbers, we know that the distribution of contact numbers in the network neurons with correlated input is different from that of Figures 2 and 3 (Supplementary Material S2.7). Instead of retuning parameters to optimize contact numbers again, we decided that, for consistency with the earlier simulations, we would keep the parameters of the plasticity model unchanged and focus in the network simulations on the resulting connectivity patterns.

After 10 days of simulated time, the whisker corresponding to barrel 3 is trimmed. Whisker lesion is modeled by 1) absence of whisker flicks in the thalamically driven group "tha 3" while stimulation of the groups "tha 1" and "tha 2" continues and 2) an exponential decrease of the firing rate of the corresponding thalamic population with a time constant of 5 min to a new

baseline level of 0.1 Hz. We found that the spiking activity of the network after the lesion remains asynchronous and irregular (Fig. 6b). Within 3 days after the lesion, the recurrent connectivity of the barrel column 3 has increased (Fig. 6c, center) consistent with a recent experiment (Albieri et al. 2015). The average weight of connections within barrel column 2 is hardly affected by the lesion. The lateral connections between excitatory neurons of barrel columns 2 and 3 increase on average. Similarly, the average connection weight of the nonpreferred pathway from the group "tha 2" to the group "exc 3" increases, whereas the connections in the preferred pathway disappear after the lesion (Fig. 6c, center). We followed the synaptic changes for a total of 50 days after the lesion, during which the recurrent network slowly continues to reorganize itself into a new connectivity pattern. In particular, the average connectivity from excitatory neurons in barrel column 3 to those in 2 continues to slowly increase.

The detailed connectivity pattern in Figure 6d indicates that the increase in the average connection weight of the nonpreferred pathway (from the group "tha 2" to "exc 3") during the first 3 days after the lesion is due to a strong increase of a subset of the feed-forward connections. After clustering (see Materials and Methods) and color-coding the excitatory neurons in barrel column 3 according to their responsiveness (yellow shade: newly responsive to whisker 1; red shade: newly responsive to whisker 2; green shade: responsiveness unchanged), the relevance of this reorganization into subsets manifests itself in four consistent ways: 1) the subset of red-shaded neurons in population "exc 3" responds more strongly to stimulation of whisker 2 than the average of neurons in population "exc 3", nearly as strongly as neurons in barrel column 2 (Fig. 6d, bottom right), 2) the same subset has a larger fraction of strong lateral connections to excitatory neurons in barrel column 2 than other neurons in population "exc 3," 3) the same subset receives a larger fraction of strong lateral connections from excitatory neurons in barrel column 2 than from neurons in barrel column 1, 4) the same subset receives a larger fraction of strong feed-forward connections from thalamic neurons in group 2 than other neurons in "exc 3." Taken together, these four observations suggest that the subset of red-shaded neurons in barrel column 3 has been integrated in the information-processing stream of barrel column 2. The same observations can be repeated for the yellow-shaded subgroup of neurons in barrel column 3, except that these neurons have been integrated into barrel column 1. In both cases, the integration has been made possible by structural plasticity triggered by the lesion. Interestingly, the subset of those neurons of the deprived cortical column that are integrated into a new whisker-processing stream do not have to be physical neighbors but can be identified as those who, before the trimming, had already a stronger response to the adjacent whisker (Fig. 7a).

The simulation results mentioned under point 1) above are consistent with experience-dependent receptive field plasticity found experimentally in pyramidal neurons in mouse somatosensory cortex 3–4 days (Trachtenberg et al. 2002) or 20 days (Wilbrecht et al. 2010) after whisker trimming, where neurons that were part of a deprived barrel became responsive to the first-order surrounding whisker, in particular the subset of neurons located in the border region to the neighboring barrel (Wilbrecht et al. 2010). Observations (2)–(4) listed above are predictions of our model. Note that in our simulations, both barrel columns "exc 1" and "exc 2" are first-order surrounding columns of the deprived column (exc 3) since we have not introduced any distance-dependent connectivity.

We emphasize that the parameters of the structural plasticity model are kept fixed throughout the simulation, be it before, during, or after the lesion: First the network connection density was stationary before the lesion, then it changed significantly during 3 days after the lesion, and finally settled into a new state (Fig. 6c) while structural plasticity has always been "turned on." Indeed, individual synaptic contact points continue to grow or disappear even during the phases where the coarse connectivity pattern remains unchanged; see Figure 3.

Discussion

We linked structural dynamics of synaptic contacts in neuronal networks to a novel STDP rule with synaptic depression for strong correlations. In this plasticity model, weight changes depend on the relative timing of pre- and postsynaptic spiking, where firing of the postsynaptic cell could be transmitted by back-propagating action potentials to dendritic synapses. The implicit coupling between synaptic contacts onto the same postsynaptic neuron through back-propagating action potentials is sufficient to make synaptic contacts from one presynaptic neuron compete with contacts of other presynaptic neurons and co-operate with contacts of the same presynaptic neuron. Our phenomenological (as opposed to "molecular") model provides a mathematically tractable description of structural plasticity. We hypothesize that any synaptic and structural plasticity model must contain at least 3 mathematical terms (possibly corresponding to 3 biological mechanisms) so as to enable the maintenance of bimodal distributions of synaptic contacts and competition between presynaptic neurons. First, a simple Hebbian correlation term that drives learning. In our learning rule, this term is a somewhat naïve correlation detector, but this term could be replaced by more detailed plasticity models (Kempter et al. 1999; Song et al. 2000; van Rossum et al. 2000; Shouval et al. 2002; Pfister and Gerstner 2006; Morrison et al. 2007; Helias et al. 2008; Clopath et al. 2010). Second, a heterosynaptic plasticity term that stabilizes the postsynaptic activity on a rapid time scale. In our learning rule, we used a fourth-order term in the postsynaptic rate (Zenke et al. 2015), but other rate normalization rules (Oja 1982; Miller and MacKay 1994; van Rossum et al. 2000; Tetzlaff et al. 2011) have a similar spirit and could potentially work just as well if the feedback is fast enough to allow effective control of firing rates (Zenke et al. 2013) and if competition between presynaptic neurons is not removed. Third, and most importantly, a term that limits strong correlations between pre- and postsynaptic neurons. In our learning rule, we introduced an anti-Hebbian term triggering synaptic depression for strong pre- and postsynaptic correlation, which is sensitive to the amount of correlations squared, inspired by priming experiments of plasticity induction (Huang et al. 1992). The effects of this term are similar to the more traditional hard bounds on the synaptic weights (e.g., Kempter et al. 1999; Song et al. 2000; Gerstner and Kistler 2002; Morrison et al. 2008), but give, in a multi-neuron multicontact scenario, rise to a much broader variety of stable synaptic configurations. In particular, the combination of these 3 terms enables parameter settings where, for example, 2-contact connections are unstable, whereas zero-contact connections and 5-contact connections are stable. Moreover, in a scenario with N input neurons out of which N_0 have exactly 5 contacts each, there are $N!/(N − N_0)!N_0!$ different combinations of input neurons with active synapses. If any of the 3 terms in our model is omitted, our mathematical analysis suggests that the outcome will be qualitatively different. In particular, if firing rate normalization and weight bounds are implemented by the same mechanism, the number of available combinations of

input neurons with active synapses is likely to be much smaller (Tetzlaff et al. 2011, 2012; Fauth et al. 2015a).

The bimodal contact number distribution of synapses has been discussed in several theoretical papers. To explain the bimodal distribution, Fares and Stepanyants (2009) used an algorithmic 2-step approach: first they transformed a potential contact with probability P into a realized contact, then they probabilistically removed those connections that had less than N_s^c contacts (see also Reimann et al. 2015 for a related approach) but did not provide dynamical equations of how co-operativity could be implemented. Deger et al. (2012) showed in a multicontact model of a single pair of pre- and postsynaptic neurons that Markov transitions between 3 different states (inactive, small, and large contacts) can give rise to cooperative contact formation if plasticity is mediated by STDP. Fauth et al. (2015a) presented an elegant mathematical analysis for equilibrium distributions in a multicontact synapse model with predefined weight-dependent contact deletion rates. However, their analysis is restricted to a single pair of pre- and postsynaptic neurons and assumes that firing rate stabilization is achieved by adjusting the weights (and contact numbers) of this single neuron. If we re-interpret their model for multiple presynaptic neurons, their mathematical solution corresponds to a homogeneous configuration where connections of all presynaptic neurons behave identically; that is, they exclude solutions where different presynaptic neurons compete with each other. In contrast to the approach of Fauth et al. (2015a) and Deger et al. (2012), we study a model with multiple presynaptic neurons and show that stable solutions exist where a subset of presynaptic neurons develops strong multicontact connections, while other presynaptic neurons do not. Moreover, our approach highlights that stabilization of individual contact weights can be mathematically separated from firing rate stabilization. Finally, our approach does not assume a predefined weight dependence of the contact removal rate, but we simply remove contacts when they hit zero, similar to earlier single-contact STDP models that sometimes have a structural component of contact removal (Gerstner et al. 1996; Helias et al. 2008; Vlachos et al. 2013; Miner and Triesch 2016).

Co-operation and competition are both well-established concepts in theories of synaptic plasticity. For example, in classical rate-based plasticity models, synapses of different presynaptic neurons co-operate if they share correlated input (Bienenstock et al. 1982; Oja 1982). At the same time, synapses compete with each other if the firing rate (Bienenstock et al. 1982) or the vector of input weights (Oja 1982) is normalized. Analogous statements are valid for models of STDP (Gerstner et al. 1996; Kempter et al. 1999; Song et al. 2000; van Rossum et al. 2000; Gerstner and Kistler 2002; Helias et al. 2008; Morrison et al. 2008; Vlachos et al. 2013). Here, we have presented a multicontact plasticity model that shows competition of synapses from different neurons and co-operation of synaptic contacts arising from the same presynaptic neuron even if inputs are uncorrelated Poisson spike trains. The co-operation of synaptic contacts from the same neuron disappears in our model if the synaptic failure rate approaches 1. Note that several existing structural plasticity models rely on noncompetitive, purely homeostatic structural dynamics (Butz and Ooyen 2013; Diaz-Pier et al. 2016). However, in our opinion, competition is needed to make some synapses grow at the expense of others, in the tradition of Bienenstock-Cooper-Munro (1982).

Relations to, and Predictions for, Experiments

1. A synaptic contact with a given (potentially small) weight is more stable if it is part of a group of 4 or 5 synaptic contacts arising from the same presynaptic neuron than if it is isolated or part of a group of only 2 synaptic contacts (Fig. 3c). This result is generic in the sense that any Hebbian plasticity model with some normalization mechanism would predict this (Fauth et al. 2015b). The conditional stability of contact weights could be measured by correlating the survival time of a given synaptic contact (Holtmaat et al. 2005, Yasumatsu et al. 2008) with the total EPSP amplitude of the connection (Le Be and Markram 2006).

2. Our model predicts a substantial fraction of synaptic connections that have only one actual contact (see Fig. 3a). These synaptic contacts, however, are small and quickly removed (see Fig. 2f). Since weight, PSP amplitude, volume, and size are tightly correlated (Holtmaat and Svoboda 2009), these weak synapses might escape electrophysiological or visual detection with standard methods, which could explain the differences to experimental reports (Markram, Lübke, Frotscher, Roth et al. 1997a; Trachtenberg et al. 2002) but might be detectable using subdiffraction resolution imaging (Attardo et al. 2015).

3. After a lesion, the number of connections consisting of exactly 2 synaptic contacts increases transiently. While this might be expected, since the process of building a novel synaptic connection with 5 contacts has to pass through a transient state with only 2 contacts, our model predicts that only about a quarter of the 2-contact connections actually stabilize to a multicontact synapse, while the majority disappears again. Again, this result is generic and has been reported similarly in other models (Vlachos et al. 2013). As a consequence, the number of presynaptic neurons without a connection transiently decreases after a lesion before it increases again to a stable value (Fig. 5d and e).

4. After whisker trimming, the subset of neurons of the deprived cortical column, which will be integrated in the signal processing of an adjacent whisker, will establish stronger incoming and lateral projections to the cortical column in which they become integrated than other neurons in the deprived column (Fig. 6d). Again, we expect this result to be generic for a large class of Hebbian plasticity rules. As an aside, we note that the subset of those neurons of the deprived cortical column that are integrated into a new whisker-processing stream do not have to be physical neighbors but can be identified as those who, before the trimming, had already a stronger response to the adjacent whisker (Fig. 7a).

5. Our learning rule assumes that synaptic contact plasticity strongly depends on local traces of correlations that must be processed in the dendritic spines. Therefore, we expect experimental manipulations of biophysical activity traces, such as the calcium concentration, to crucially influence the development of dendritic spines (Lohmann et al. 2005).

Two Theoretical Insights

First, in our plasticity model, firing rate stabilization is separate, and significantly faster, than contact weight stabilization. The rapid time scale of firing rate stabilization is necessary in recurrent networks (Zenke et al. 2013) and has similar effects as a renormalization of the weight vector of afferent synapses in each time step (Miller and MacKay 1994). The separation of firing rate stabilization by a fast compensatory mechanism from other control terms that stabilize individual weights enables competition between synapses from different presynaptic neurons (Zenke et al. 2017), in contrast to models where firing rate stabilization is strong but mixed with weight stabilization (Tetzlaff et al. 2011). A separation similar to the one proposed here also occurs in models with hard bounds on the synaptic

weights (Kempter et al. 1999; Song et al. 2000). However, with hard bounds, weight values tend to get "stuck" at the upper and lower bounds so that fluctuation of contact weights as those highlighted in Figure 2b would not be possible. Soft-bound STDP models can show fluctuations but are less suitable for long-term storage of memories (Morrison et al. 2007; Billings and van Rossum 2009).

Second, our novel anti-Hebbian term in the synaptic plasticity rule, which causes the stabilization of weights without upper bounds, leads to an intrinsic stabilization of correlations between pre- and postsynaptic neuron at synaptic contact points (Supplementary Material S2.2). We hypothesize that there could be a novel class of "normative" learning rules in unsupervised learning, which do not maximize second-order correlations (as done by Oja's rule (Oja 1982)), but instead normalize these. While experimentally known forms of homeostasis have focused on the normalization of the mean firing rate (Ibata et al. 2008), we suggest that it is also worthwhile to study a normalization of correlations at each synaptic contact point. We speculate that in this class of correlation-normalizing learning rules, there could be rules that are sensitive to higher order correlations but suppress second-order correlations without the need of "pre-whitening" of data (Olshausen and Field 1996; Brito 2016).

The stabilization of the weights of individual contacts in the absence of explicit hard or soft bounds leads to a class of plasticity models with potentially interesting properties for information processing. On one side, for a given number m of contacts, the weights can be considered as bistable, and hence near-binary: our theory predicts that a contact weight is either zero (Brunel 2016) or approaches a finite value $w^*(m)/m$. On the other side, this finite value depends on the number of contacts m and on the correlations in the input – hence, there is information beyond purely binary weights. Moreover, the distribution of contact numbers may contain information about past plasticity events (Fusi and Abbott 2007).

Simplifications and Shortcomings

First, we limited the simulations and analysis to a point neuron model. We speculate that the synaptic plasticity rule used in this paper could also produce branch-specific synaptic plasticity and steady-state configurations (Druckmann et al. 2014) if we combine it with a more complex neuron model with nonlinear dendritic branches (see e.g., Schiess et al. (2016) and Kastellakis et al. (2016)). We expect local co-operation in such a model to stabilize preferentially several synaptic contacts from the same presynaptic neurons onto the same dendritic compartments, as observed in experiments (Toni et al. 1999; Fu et al. 2012) and earlier models (Kastellakis et al. 2016).

Second, the specific choice of STDP rule used in this paper shows, at low frequencies, a symmetric pre–post and post–pre learning window for LTP, which is not compatible with experiments on cortical STDP (Markram, Lübke, Frotscher et al. 1997b; Sjöström et al. 2001). However, important for our model is only that there are some synaptic traces sensitive to correlations. We have combined these traces in a specific manner so as to simplify the mathematical analysis, but it would be straightforward to extend our formalism to more realistic STDP models with phases of both potentiation and depression (Pfister and Gerstner 2006; Morrison et al. 2008). Importantly, our model assumes that, for high correlations between the activities of pre- and postsynaptic neurons, there is a transition from strong LTP to much weaker LTP or even LTD. The effects of such a

term should be comparatively weak, and escape analysis in standard protocols, but appear compatible with results that found LTP to be weaker if the synapses were previously subjected to strong correlations (Huang et al. 1992). Similar principles might explain why, depending on experimental preparations, variable STDP rules have been reported in experiments (Sjöström et al. 2001).

Third, whereas for our recurrent network simulation of Figure 6, we have used leaky integrate-and-fire neurons, for the mathematical analysis of the system dynamics, we chose to describe the activity of the postsynaptic neuron by a Poisson process with a linearly modulated rate. However, we expect very similar mathematical results in the case of nonlinear neurons linearized around the operating point, which changes little on the time scale of hours or days (see Fig. 2c, black).

Fourth, the parameters of our synaptic plasticity rule have been adapted so that a bimodal distribution of contact numbers is seen for independent Poisson spike trains at 5 Hz, whereas cortical neurons *in vivo* show typically a small amount of correlations and are active at lower frequencies. Indeed, a bimodal distribution of contact numbers is also possible in the presence of correlations and at lower firing rates, but with a different set of parameters (Supplementary Material S2). Nevertheless, for the sake of consistency of parameters across the paper, we performed our network simulations with the same set of synaptic plasticity parameters as the single-neuron simulations even though we knew that with this set of parameters synaptic contact distribution numbers are not bimodal (Supplementary Material S2.7). The question then arises whether, in the brain, bimodality is specific to a few sensory cortex areas where experiments have been done or truly generic across brain areas and neuron types with potentially different firing rates and correlations. In the latter case, our model cannot account for bimodality unless additional metaparameters are introduced that would automatically retune the parameters.

Fifth, the time scale of effective firing rate homeostasis in our model is too fast compared with experimental results. Homeostatic synaptic scaling in response to experimental blocking of postsynaptic firing occurs on the time scale of hours (Ibata et al. 2008) or, in response to sensory deprivation (likely to correspond to a reduction in presynaptic firing), on the time scale of days (Toyoizumi et al. 2014). In an earlier structural model, weights were given an explicit time dependence after a lesion (Vlachos et al. 2013), while the dynamic model of Toyoizumi et al. (2014) captures these time scales intrinsically, but is nonstructural. In our model, lesion-induced changes of synaptic contact weights occur rather quickly. While the average synaptic contact number recovers slowly within about 20 days, within less than 10 min after the lesion, the contact weights from spared presynaptic neurons are upregulated to compensate for the loss of input (see Figs. 5b-d and 6c). We propose to extend the model presented in this paper by a combination of hard bounds (Morrison et al. 2008) and multiplicative interaction of Hebbian and explicit homeostatic processes (Toyoizumi et al. 2014). For example, with hard bounds set to twice the fixed-point weight, the stationary distributions and results of Figures 1–3 remain unchanged while the recovery of the firing rate after a lesion is slower since individual synaptic contacts cannot grow beyond the hard bound (see Fig 7b). A more complete model could combine the hard bounds with the essential features of the model of Toyoizumi et al. (2014). More generally, the interactions of our rule with manifold plasticity mechanisms such as short-term plasticity (Markram and Tsodyks 1996), homeostasis (Ibata et al. 2008; Butz and Ooyen

2013; Toyoizumi et al. 2014; Zenke et al. 2017), or intrinsic excitability (Daoudal and Debanne 2003; Miner and Triesch 2016) pose interesting challenges for future research.

Our synaptic and structural plasticity model results in multiple contacts that form stable clusters. The contact weights are binary in the sense that the 2 stable fixed points of a contact weight are either zero or some finite value, but this finite value depends on the input correlations and firing rates (Supplementary Material S2). Interestingly, a model with multiple binary contacts is one of the potential implementations of a synaptic consolidation model of Benna and Fusi (2016). In contrast to other binary synapse models, our approach does not rely on predefined weight values for strong weights, which may give additional flexibility to neural systems while preserving their long-term stability.

Authors' Contributions

M.D. and W.G. conceived and designed the study. M.D., A.S., and W.G. wrote the paper. M.D. conducted the simulation experiments and analyses. A.S. contributed to the mathematical analysis as well as model implementation and testing.

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

Funding

Notes

Conflict of Interest: None declared.

References

Albieri G, Barnes SJ, de Celis Alonso B, Cheetham CEJ, Edwards CE, Lowe AS, Karunaratne H, Dear JP, Lee KC, Finnerty GT. 2015. Rapid bidirectional reorganization of cortical microcircuits. Cereb Cortex. 25:3025–3035.

Attardo A, Fitzgerald JE, Schnitzer MJ. 2015. Impermanence of dendritic spines in live adult ca1 hippocampus. Nature. 523: 592–596.

Benna MK, Fusi S. 2016. Computational principles of synaptic memory consolidation. Nat Neurosci. 19:1697–1706.

Bienenstock EL, Cooper LN, Munro PW. 1982. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J Neurosci. 2:32–48.

Billings G, van Rossum MCW. 2009. Memory retention and spike-timing-dependent plasticity. J Neurophysiol. 101: 2775–2778.

Bourjaily MA, Miller P. 2011. Excitatory, inhibitory, and structural plasticity produce correlated connectivity in random networks trained to solve paired-stimulus tasks. Frontiers in Computational Neuroscience. 5:1–24.

Brito C (2016) Theory of Representation Learning in Cortical Networks. PhD thesis École polytechnique fédérale de Lausanne EPFL, n° 6948 doi:10.5075/epfl-thesis-6948

Brunel N. 2016. Is cortical connectivity optimized for storing information? Nat Neurosci. 19:749–755.

Butz M, Van Ooyen A. 2013. A simple rule for dendritic spine and axonal bouton formation can account for cortical reorganization after focal retinal lesions. PLoS Comput Biol. 9: e1003259.

Chen JL, Nedivi E. 2013. Highly specific structural plasticity of inhibitory circuits in the adult neocortex. Neuroscientist. 19: 384–393.

Chistiakova M, Bannon NM, Chen J-Y, Bazhenov M, Volgushev M. 2015. Homeostatic role of heterosynaptic plasticity: models and experiments. Front Comput Neurosci. 9:89.

Clopath C, Büsing L, Vasilaki E, Gerstner W. 2010. Connectivity reflects coding: a model of voltage-based STDP with homeostasis. Nat Neurosci. 13:344–352.

Clopath C, Ziegler L, Vasilaki E, Büsing L, Gerstner W. 2008. Tag-trigger-consolidation: a model of early and late long-term-potentiation and depression. PLoS Comput Biol. 4: e1000248.

Crair MC, Malenka RC. 1995. A critical period for long-term potentiation at thalamocortical synapses. Nature. 375: 325–328.

Crochet S, Petersen CCH. 2006. Correlating whisker behavior with membrane potential in barrel cortex of awake mice. Nat Neurosci. 9:608–610.

Daoudal G, Debanne D. 2003. Long-term plasticity of intrinsic excitability: learning rules and mechanisms. Learn Mem. 10: 456–465.

Deger M, Helias M, Rotter S, Diesmann M. 2012. Spike-timing dependence of structural plasticity explains cooperative synapse formation in the neocortex. PLoS Comput Biol. 8: e1002689.

Diaz-Pier S, Naveau M, Butz-Ostendorf M, Morrison A. 2016. Automatic generation of connectivity for large-scale neuronal network models through structural plasticity. Front Neuroanat. 10:57.

Druckmann S, Feng L, Lee B, Yook C, Zhao T, Magee JC, Kim J. 2014. Structured synaptic connectivity between hippocampal regions. Neuron. 81:629–640.

El Boustani S, Yger P, Fregnac Y, Destexhe A. 2012. Stable learning in stochastic network states. J Neurosci. 32:194–214.

Eppler JM, Pauli R, Peyser A, Ippen T, Morrison A, Senk J, Schenck W, Bos H, Helias M, Schmidt M, et al. 2015. NEST 2.8.0. doi: 10.5281/zenodo.32969

Fares T, Stepanyants A. 2009. Cooperative synapse formation in the neocortex. Proc Natl Acad Sci USA. 106:16463–16468.

Fauth M, Wörgötter F, Tetzlaff C. 2015a. The formation of multisynaptic connections by the interaction of synaptic and structural plasticity and their functional consequences. PLoS Comput Biol. 11:e1004031.

Fauth M, Wörgötter F, Tetzlaff C. 2015b. Formation and maintenance of robust long-term information storage in the presence of synaptic turnover. PLoS Comput Biol. 11: e1004684.

Fu M, Yu X, Lu J, Zuo Y. 2012. Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. Nature. 483:92–95.

Fusi S, Drew PJ, Abbott LF. 2005. Cascade models of synaptically stored memories. Neuron. 45:599–611.

Fusi S, Abbott LF. 2007. Limits on the memory storage capacity of bounded synapses. Nat Neurosci. 10:485–493.

Gerstner W, Kempter R, van Hemmen JL, Wagner H. 1996. A neuronal learning rule for sub-millisecond temporal coding. Nature. 383:76–78.

Gerstner W, Kistler WM. 2002. Spiking neuron models: single neurons, populations, plasticity. Cambridge, UK: Cambridge University Press.

Grutzendler J, Kasthuri N, Gan W-B. 2002. Long-term dendritic spine stability in the adult cortex. Nature. 420:812–816.

Hayashi-Takagi A, Yagishita S, Nakamura M, Shirai F, Wu YI, Loshbaugh AL, Kuhlman B, Hahn KM, Kasai H. 2015. Labelling and optical erasure of synaptic memory traces in the motor cortex. Nature. 525:333–338.

Hebb DO. 1949. The organization of behavior; a neuropsychological theory. New York: Wiley.

Helias M, Rotter S, Gewaltig M-O, Diesmann M. 2008. Structural plasticity controlled by calcium based correlation detection. Front Comput Neurosci. 2:7.

Holtmaat A, Svoboda K. 2009. Experience-dependent structural synaptic plasticity in the mammalian brain. Nat Rev Neurosci. 10:647–658.

Holtmaat A, Wilbrecht L, Knott GW, Welker E, Svoboda K. 2006. Experience-dependent and cell-type-specific spine growth in the neocortex. Nature. 441:979–983.

Holtmaat AJGD, Trachtenberg JT, Wilbrecht L, Shepherd GM, Zhang X, Knott GW, Svoboda K. 2005. Transient and persistent dendritic spines in the neocortex in vivo. Neuron. 45: 279–291.

Huang YY, Colino A, Selig DK, Malenka RC. 1992. The influence of prior synaptic activity on the induction of long-term potentiation. Science. 255:730–733.

Ibata K, Sun Q, Turrigiano GG. 2008. Rapid synaptic scaling induced by changes in postsynaptic firing. Neuron. 57: 819–826.

Kastellakis G, Silva AJ, Poirazi P. 2016. Linking memories across time via neuronal and dendritic overlaps in model neurons with active dendrites. Cell Reports. 17:1491–1504.

Keck T, Mrsic-Flogel TD, Vaz Afonso M, Eysel UT, Bonhoeffer T, Hübener M. 2008. Massive restructuring of neuronal circuits during functional reorganization of adult visual cortex. Nat Neurosci. 11:1162–1167.

Kempter R, Gerstner W, Hemmen JL van. 1999. Hebbian learning and spiking neurons. Phys Rev E. 59:4498–4514.

Kerr JND, Greenberg D, Helmchen F. 2005. Imaging input and output of neocortical networks in viso. Proc Natl Acad Sci USA. 102:14063–14068.

Kwon H-B, Sabatini BL. 2011. Glutamate induces de novo growth of functional spines in developing cortex. Nature. 474:100–104.

Laughlin SB, de Ruyter van Steveninck RR, Anderson JC. 1998. The metabolic cost of neural information. Nat Neurosci. 1: 36–41.

Le Be J-V, Markram H. 2006. Spontaneous and evoked synaptic rewiring in the neonatal neocortex. Proc Nat Acad Sci USA. 103:13214–13219.

Loewenstein Y, Kuras A, Rumpel S. 2011. Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. J Neurosci. 31: 9481–9488.

Loewenstein Y, Yanover U, Rumpel S. 2015. Predicting the dynamics of network connectivity in the neocortex. J Neurosci. 35:12535–12544.

Lohmann C, Finski A, Bonhoeffer T. 2005. Local calcium transients regulate the spontaneous motility of dendritic filopodia. Nat Neurosci. 8:305–312.

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

Markram H, Lübke J, Frotscher M, Roth A, Sakmann B. 1997a. Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. J Physiol. 500:409–440.

Markram H, Lübke J, Frotscher M, Sakmann B. 1997b. Regulation of synaptic efficacy by coincidence of postysnaptic AP and EPSP. Science. 275:213–215.

Markram H, Muller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA, Ailamaki A, Alonso-Nanclares L, Antille N, Arsever S, et al. 2015. Reconstruction and simulation of neocortical microcircuitry. Cell. 163:456–492.

Matsuzaki M, Ellis-Davies GC, Nemoto T, Miyashita Y, Iino M, Kasai H. 2001. Dendritic spine geometry is critical for AMPA receptor expression in hippocampal CA1 pyramidal neurons. Nat Neurosci. 4:1086–1092.

Matsuzaki M, Honkura N, Ellis-Davies GCR, Kasai H. 2004. Structural basis of long-term potentiation in single dendritic spines. Nature. 429:761–766.

Miller KD, MacKay DCJ. 1994. The role of constraints in Hebbian learning. Neural Comput. 6:100–126.

Miner D, Triesch J. 2016. Plasticity-driven self-organization under topological constraints accounts for non-random features of cortical synaptic wiring. PLoS Comput Biol. 12: e1004759.

Minerbi A, Kahana R, Goldfeld L, Kaufman M, Marom S, Ziv NE. 2009. Long-term relationships between synaptic tenacity, synaptic remodeling, and network activity. PLoS Biol. 7: e1000136.

Morrison A, Aertsen A, Diesmann M. 2007. Spike-timing-dependent plasticity in balanced random networks. Neural Comput. 19:1437–1467.

Morrison A, Diesmann M, Gerstner W. 2008. Phenomenological models of synaptic plasticity based on spike timing. Biol Cybern. 98:459–478.

Nawrot MP, Schnepel P, Aertsen A, Boucsein C. 2009. Precisely timed signal transmission in neocortical networks with reliable intermediate-range projections. Front Neural Circuits. 3:1.

Oja E. 1982. A simplified neuron model as a principal component analyzer. J Math Biol. 15:267–273.

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

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. 2011. Scikit-learn: machine learning in Python. J Mach Learn Res. 12:2825–2830.

Pfister J-P, Gerstner W. 2006. Triplets of spikes in a model of spike-timing dependent plasticity. J Neurosci Nurs. 26: 9673–9682.

Reimann MW, King JG, Muller EB, Ramaswamy S, Markram H. 2015. An algorithm to predict the connectome of neural microcircuits. Front Comput Neurosci. 9:120.

Rose T, Jaepel J, Hübener M, Bonhoeffer T. 2016. Cell-specific restoration of stimulus preference after monocular deprivation in the visual cortex. Science. 352:1319–1322.

Roxin A, Fusi S. 2013. Efficient partitioning of memory systems and its importance for memory consolidation. PLoS Comput Biol. 9:e1003146.

Schiess M, Urbanczik R, Senn W. 2016. Somato-dendritic synaptic plasticity and error-backpropagation in active dendrites. PLoS Comput Biol. 12:e1004638.

Shouval HZ, Bear MF, Cooper LN. 2002. A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. Proc Natl Acad Sci USA. 99:10831–10836.

Sjöström PJ, Turrigiano GG, Nelson SB. 2001. Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron. 32:1149–1164.

Song S, Miller KD, Abbott LF. 2000. Competitive hebbian learning through spike-timing-dependent synaptic plasticity. Nat Neurosci. 3:919–926.

Stepanyants A, Hof PR, Chklovskii DB. 2002. Geometry and structural plasticity of synaptic connectivity. Neuron. 34:275–288.

Tetzlaff C, Kolodziejski C, Timme M, Worgotter F. 2011. Synaptic scaling in combination with many generic plasticity mechanisms stabilizes circuit connectivity. Front Comput Neurosci. 5:47.

Tetzlaff C, Kolodziejski C, Timme M, Worgotter F. 2012. Analysis of synaptic scaling in combination with Hebbian plasticity in several simple networks. Front Comput Neurosci. 6:36.

Toni N, Buchs P-A, Nikonenko I, Bron CR, Muller D. 1999. LTP promotes formation of multiple spine synapses between a single axon terminal and a dendrite. Nature. 402:421–425.

Toyoizumi T, Kaneko M, Stryker MP, Miller KD. 2014. Modeling the dynamic interaction of hebbian and homeostatic plasticity. Neuron. 84:497–510.

Trachtenberg JT, Chen BE, Knott GW, Feng G, Sanes JR, Welker E, Svoboda K. 2002. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. Nature. 420:788–794.

van Rossum MC, Bi GQ, Turrigiano GG. 2000. Stable hebbian learning from spike timing-dependent plasticity. J Neurosci. 20:8812–8821.

Vlachos A, Helias M, Becker D, Diesmann M, Deller T. 2013. NMDA-receptor inhibition increases spine stability of denervated mouse dentate granule cells and accelerates spine density recovery following entorhinal denervation in vitro. Neurobiol Dis. 59:267–276.

Wiegert JS, Oertner TG. 2013. Long-term depression triggers the selective elimination of weakly integrated synapses. PNAS. 110:E4510–E4519. doi:10.1073/pnas.1315926110.

Wilbrecht L, Holtmaat A, Wright N, Fox K, Svoboda K. 2010. Structural plasticity underlies experience-dependent functional plasticity of cortical circuits. J Neurosci. 30:4927–4932.

Yang G, Pan F, Gan WB. 2009. Stably maintained dendritic spines are associated with lifelong memories. Nature. 462:920–924.

Yasumatsu N, Matsuzaki M, Miyazaki T, Noguchi J, Kasai H. 2008. Principles of long-term dynamics of dendritic spines. J Neurosci. 28:13592–13608.

Yuste R. 2011. Dendritic spines and distributed circuits. Neuron. 71:772–781.

Zador A. 1998. Impact of synaptic unreliability on the information transmitted by spiking neurons. J Neurophysiol. 79:1219–1229.

Zenke F, Henniquin G, Gerstner W. 2013. Synaptic plasticity in neural networks needs homeostasis with a fast rate detector. PLOS Comput Biol. 9:e1003330.

Zenke F, Agnes EJ, Gerstner W. 2015. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. Nat Commun. 6:6922.

Zenke F, Gerstner W, Ganguli S. 2017. The temporal paradox of Hebbian learning and homeostatic plasticity. Curr Op Neurobiol. 43:166–176.

Zito K, Scheuss V, Knott G, Hill T, Svoboda K. 2009. Rapid functional maturation of nascent dendritic spines. Neuron. 61:247–258.