ARTICLE =

# **Emergence of Optimal Decoding of Population Codes Through STDP**

## Stefan Habenschuss

habenschuss@igi.tugraz.at **Helmut Puhr** hwp@aatc.at **Wolfgang Maass** maass@igi.tugraz.at Institute for Theoretical Computer Science, Graz University of Technology, A-8010 Graz, Austria

The brain faces the problem of inferring reliable hidden causes from large populations of noisy neurons, for example, the direction of a moving object from spikes in area MT. It is known that a theoretically optimal likelihood decoding could be carried out by simple linear readout neurons if weights of synaptic connections were set to certain values that depend on the tuning functions of sensory neurons. We show here that such theoretically optimal readout weights emerge autonomously through STDP in conjunction with lateral inhibition between readout neurons. In particular, we identify a class of optimal STDP learning rules with homeostatic plasticity, for which the autonomous emergence of optimal readouts can be explained on the basis of a rigorous learning theory. This theory shows that the network motif we consider approximates expectation-maximization for creating internal generative models for hidden causes of high-dimensional spike inputs. Notably, we find that this optimal functionality can be well approximated by a variety of STDP rules beyond those predicted by theory. Furthermore, we show that this learning process is very stable and automatically adjusts weights to changes in the number of readout neurons, the tuning functions of sensory neurons, and the statistics of external stimuli.

## 1 Introduction

Uncertainty accompanies us in almost all situations in life. Whether we try to recognize a distant object, decide which path to take on a mountain hike, or read a person's face in an important negotiation, the environment often

Stefan Habenschuss and Helmut Puhr contributed equally to this work. A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/NECO\_a\_00446.

provides us with many cues, but each single cue is too unreliable to inform a decision on its own. Thus, we are forced to combine different cues in a meaningful manner in order to gain sufficient certainty. The theoretical framework for solving such tasks in an optimal way is Bayesian inference. Notably, behavioral and psychophysical studies strongly support the picture that the brain implements this strategy: in numerous experiments, human subjects have been shown to take into account uncertainty in a near-optimal way (Griffiths & Tenenbaum, 2006).

At the level of neural coding in early sensory areas, uncertainty is a particularly well-studied phenomenon: individual neurons that encode certain stimulus properties show significant trial-to-trial variability, making it difficult to infer the original stimulus from single-neuron responses. This observation has led to the notion of population coding: the value of a single variable is encoded by a whole population of neurons, each noisy and typically broadly tuned to the external variable (Pouget, Dayan, & Zemel, 2000). Experimental data suggest this coding strategy as a candidate for understanding how important variables are represented in different areas across cortex, for example, sound location in auditory cortex (Miller & Recanzone, 2009) or stimulus location in somatosensory cortex (Petersen, Panzeri, & Diamond, 2002). Just how the brain reliably decodes information from populations of neurons in a near-optimal way remains one of the key open questions in computational neuroscience.

In experimental neuroscience, the computation of robust readouts from population codes has become indispensable in the analysis and interpretation of neural data; population vector analysis and maximum likelihood (ML) estimation (Pouget et al., 2000) are two frequently used methods. More recently, attempts have been made to model neural networks that exhibit near-optimal decoding capabilities (Deneve, Latham, & Pouget, 1999; Jazayeri & Movshon, 2006; Chaisanguanthum & Lisberger, 2011). The hope is that such models will advance our understanding of the neural substrates of perceptual judgments: how the noisy and broadly tuned representations found in sensory areas can be efficiently used and transformed by downstream populations to allow near-optimal performance in a variety of perceptual tasks. The theoretical framework that has been guiding the search for suitable models is Bayesian inference. Here, in contrast to previous, static models, we will use this perspective for the analysis of an adaptive cortical microcircuit, featuring spike-timing-dependent plasticity.

In the Bayesian framework, an observed response of a population of M neurons, which we will write here as  $\mathbf{x} = (x_1, \dots, x_M)^T$ , is interpreted as the result of an underlying cause  $\theta$  that cannot be observed directly. In visual processing, this external cause could correspond to the true stimulus orientation or motion direction at some retinal location. The population code (i.e., the relation between the external cause  $\theta$  and the observed population response  $\mathbf{x}$ ) can then be represented by the conditional distribution  $p(\mathbf{x}|\theta)$ . This captures deterministic dependencies on  $\theta$ , usually specified in

terms of tuning functions, and neuronal noise. The optimal way of inferring the external variable  $\theta$  from noisy observations **x** is then given by Bayes' theorem:

$$\underbrace{p(\theta|\mathbf{x})}_{\text{posterior}} \propto \underbrace{p(\theta)}_{\text{prior}} \cdot \underbrace{p(\mathbf{x}|\theta)}_{\text{likelihood}}$$
(1.1)

In the context of population codes, the prior distribution is often assumed to be uniform, such that inference reduces to computing the likelihood function  $L(\theta) = p(\mathbf{x}|\theta)$ , which indicates how likely different  $\theta$ are to have caused a given observation x. This results in a quite straightforward decoding strategy: compute the likelihood for each possible cause  $\theta$ that could have given rise to the observed population response and choose the one with maximal likelihood. Somewhat surprisingly, the computational requirements for implementing such a readout are minimal: under a few simplifying assumptions, including Poisson firing statistics and zero noise correlations in the population pool, it was shown recently that the likelihood  $L(\theta)$  of a stimulus can be written as a weighted sum of sensory responses x (Jazaveri & Movshon, 2006). Based on this observation, the authors argued that a readout neuron that specializes in detecting a particular stimulus value from a set of possible values can compute the corresponding likelihood by integrating its synaptic inputs in a feedforward manner (see the feedforward path in Figure 1). The synaptic weights required for this operation depend on the tuning functions  $f_i(\theta)$  of the sensory neurons  $x_i$ and the preferred orientation  $\theta_k$  of the readout neuron:

$$w_{ki} = \log f_i(\theta_k) + \text{const.} \tag{1.2}$$

This establishes an important link between the response properties of the sensory population (the tuning functions) and the optimal weights to decode information from it and gives clear instructions on how to construct an optimal readout network. However, equation 1.2 also highlights that each population of neurons in the brain must be read out differently, depending on the particular tuning functions of the neurons.

The preceding research leaves the question open of how readout neurons could acquire the theoretically optimal synaptic weights (see equation 1.2). The importance of this question is underlined by recent findings suggesting that the brain constantly retunes its circuits in order to improve probabilistic inference (Bejjanki, Beck, Lu, & Pouget, 2011). How experimentally observed plasticity mechanisms at the synaptic level could account for such an improvement remained an open question.

Here we present a learning theory for spike-timing-dependent plasticity (STDP) rules in the context of spiking neurons and lateral inhibition that addresses this question. Building on the analysis of previous work (Nessler,

Pfeiffer, & Maass, 2010) we show that optimal likelihood decoding of noisy population codes emerges automatically through STDP in a winner-take-all (WTA) circuit, a ubiquitous network motif in cortical microcircuits (Douglas & Martin, 2004). In particular, we theoretically analyze the weight dynamics of a particular form of STDP with homeostatic plasticity and show that it can be described as an attractor dynamics, where the centers of the attractors are the weight values (see equation 1.2) that are optimal from the perspective of probabilistic inference and learning. In this way, we create a direct link between simple local rules for synaptic plasticity and theoretically optimal inference and learning on the network level. Whereas the analysis of Nessler et al. (2010) was restricted to the case of multinomial input variables (each encoded by a population of neurons of which at any time *t*, exactly one has fired within the time window  $[t - \tau, t]$  and rectangular excitatory postsynaptic potentials (EPSPs) (modeled by a step function, rather than by a function with smooth decay), we show here that the underlying learning theory can be extended to cover the biologically more realistic case where each input neuron fires according to some Poisson-like statistics and EPSPs have a smooth decay. This new learning theory makes it possible to derive learning curves for STDP and dependencies between current weight values and the amount of weight potentiations or depressions under STDP that are optimal for a given input statistic and EPSP shape. These analytically derived predictions for details of STDP match currently available data quite well in a number of aspects. Furthermore, we show that in practice, variants of these optimal rules, and in particular plasticity rules based on typical STDP curves, approximate this optimal behavior well.

We test predictions of the new learning theory in computer simulations and show that previously derived optimal weights for likelihood decoding (see equation 1.2) emerge autonomously through STDP in conjunction with homeostatic plasticity. Based on this learning theory, we also show that an adaptive architecture for reading population codes provides attractive benefits compared to previously considered static models. In particular, we demonstrate stable and predictable behavior under challenging but biologically realistic dynamic scenarios, like changes in sensory tuning functions or neuron growth and loss. We furthermore demonstrate that selective modulation of learning leads to effects reminiscent of perceptual learning (Gilbert, Li, & Piech, 2009), where the accuracy of neural codes is selectively enhanced for behaviorally relevant stimuli.

The article is structured as follows. First, we introduce the canonical microcircuit model we consider. In this circuit model, we demonstrate STDPbased autonomous learning on exposure to population-coded stimuli and show that in such a setup, the synaptic weights converge to a setting that is optimal from the perspective of likelihood decoding. We then introduce the theory that underlies this optimality and derive a rigorous link between local synaptic learning and theoretical optimality in the Bayesian framework. We complement these theoretical results by an extensive performance



Figure 1: Network motif in which STDP generates optimal decoding of population codes. A time-varying external variable  $\theta(t)$  is encoded by a population of sensory neurons. Each sensory neuron  $x_j$  has a characteristic average response for each value of  $\theta$ : its tuning function. Based on these tuning functions, they emit spikes according to a Poisson process. Readout neurons receive these spikes via feedforward connections with synaptic weights  $w_{kj}$ . A stochastic WTA circuit induces competition among readout neurons k. As a consequence, those readout neurons, which receive the greatest stimulation from the sensory neurons, fire preferentially.

comparison of different optimal and near-optimal STDP-based learning rules. We present results of simulations that test further predictions of the theory and demonstrate the versatility and robustness of the considered model. We conclude with a simulation in which effects reminiscent of perceptual learning are reproduced.

#### 2 Results

In order to make accurate perceptual judgments, the brain must use the information provided by sensory areas as efficiently as possible. Since sensory neurons tend to be noisy and broadly tuned, the computation of a sparse representation of the most likely external stimulus ("hidden cause") is a nontrivial task for a network of neurons. However, Jazayeri and Movshon (2006) showed that for the case of a one-dimensional hidden cause  $\theta$ , which could, for example, represent the current orientation, speed, or direction of a visual stimulus, this task could in principle be solved by an array of linear readout neurons k = 1, ..., N with spiking outputs  $z_k$  that receive synaptic inputs  $x_j$  from sensory neurons j = 1...M (see Figure 1). It was shown that the readout neurons k can compute in their membrane potential the log likelihood that a particular hidden cause  $\theta_k$  (which was assigned externally to readout neuron k in Jazayeri & Movshon, 2006) had caused the current

spike output **x** of the sensory neurons. This occurs if the weights  $w_{kj}$  from sensory neuron *j* to readout neuron *k* are set according to equation 1.2. We show here that these weights emerge as fixed points (equilibria) of a class of theoretically optimal STDP rules with homeostatic plasticity, provided that the readout neurons are subject to lateral inhibition. Through the same learning process, each readout neuron *k* implicitly develops a preferred stimulus  $\psi_{k'}$  which then allows the reconstruction of an external input variable  $\theta(t)$  at any moment in time.

**2.1** Adaptive, Stochastic WTA Architecture. For the readout neurons k, a stochastic neuron model is used similar to the model proposed by Jolivet, Rauch, Lüscher, and Gerstner (2006), which has been shown to explain neural data well. The neuron model is characterized by an exponential dependence of the firing probability on the current membrane potential  $u_k$ ,

$$p(k \text{ fires in } [t, t + \Delta t]) = \Delta t \exp(u_k(t)), \qquad (2.1)$$

for small  $\Delta t$ . The membrane potential of a readout neuron *k* consists of an excitatory and an inhibitory contribution:

$$u_k(t) = \sum_{j=1}^{M} w_{kj}(t) x_j(t) - I(t).$$
(2.2)

Excitation comes from the feedforward connections originating in the sensory population  $x_1, \ldots, x_M$ . In particular, the term  $w_{kj}(t)x_j(t)$  represents the contribution of previous spikes from the *j*th sensory neuron to the membrane potential  $u_k$  of the readout neuron *k* at time *t*. The term  $x_j(t)$  hence models the unweighted output spike train of the *j*th sensory neuron after filtering according to the low-pass filtering properties of the postsynaptic membrane.  $\mathbf{x}(t)$  represents the collection of all input signals available to the readout neurons at time *t*. We will refer to this as the current (sensory) population response.

I(t) denotes the contribution to the membrane potential due to lateral inhibition. I(t) is common to all readout neurons in the circuit and can thus be viewed as a global inhibitory signal that controls the total gain of the circuit. In computer simulations, we modeled I(t) such that the total firing activity of the readout circuit remained approximately constant. The resulting effect of inhibition is a normalization of circuit responses, reminiscent of normalization models in cortex (Simoncelli & Heeger, 1998; Zoccolan, Cox, & DiCarlo, 2005; Ohshiro, Angelaki, & DeAngelis, 2011; Louie, Grattan, & Glimcher, 2011). Note that the lateral inhibition introduces competition among the readout neurons, since a readout neuron with strong feedforward input will claim a large fraction of the total firing rate, thereby suppressing other readout neurons. We refer to the resulting network as a (stochastic) winner-take-all (WTA) circuit.

We focus in this article on the following theoretically motivated spiketiming-dependent plasticity rule (but see Figure 3 for other forms of STDP we consider),

$$\Delta w_{kj} = \eta \cdot z_k \cdot (x_j \cdot \alpha e^{-w_{kj}} - 1), \tag{2.3}$$

where  $\alpha$  is a positive constant that controls the balance between potentiation and depression and  $\eta$  is a learning rate that is constant unless otherwise stated. We define  $z_k = 1$  at spike times of the WTA neuron k and  $z_k = 0$ otherwise. The learning rate  $\eta$  is chosen small throughout this article such that learning takes place on a (much) longer timescale than stimulus and network dynamics.

The first term of equation 2.3 corresponds to a prototypical spike-timingdependent long-term potentiation (LTP) window (with the same shape and time constant as PSPs) with weight dependence. The second term depends on postsynaptic spikes only and can be interpreted as a form of homeostatic plasticity providing negative feedback to the postsynaptic rate. Both terms fit well into the relatively broad phenomenological framework of STDP rules (Gerstner & Kistler, 2002; Gilson, Burkitt, & van Hemmen, 2010). But to distinguish it from more common STDP rules (which typically feature spike-timing-dependent depression), we will refer to this rule as theoretically optimal STDP.

The motivation for first considering this form of STDP comes from its notable and provable theoretical properties, which will be developed in this article: equation 2.3 leads to stable equilibrium weight settings that can be analyzed and have a clear interpretation from the perspective of probability theory. In particular, we will show that equation 2.3 is an instance of a family of learning rules that can be directly derived from the principle of adapting and optimizing an implicit generative model of the input statistics.

With regard to biological plausibility, equation 2.3 is in many aspects consistent with experimental studies on STDP. First, the strength and direction of learning depend on the timing difference between pre- and postsynaptic spike. For pre-before-post pairings,  $x_j$  is large at the time of the postsynaptic spike and equation 2.3 will typically lead to potentiation. For post-beforepre pairings, the negative part dominates and leads to depression (Bi & Poo, 1998; Sjöström, Turrigiano, & Nelson, 2001). Also, the strength of potentiation correlates inversely with the synaptic weight before pairing. This feature is also consistent with a number of experimental studies (Bi & Poo, 1998; Sjöström et al., 2001; Liao, Jones, & Malinow, 1992; Montgomery, Pavlidis, & Madison, 2001). Furthermore, the amount of depression is independent of the current weight. This is consistent with the experimental results of Jacob, Brasier, Erchova, Feldman, and Shulz (2007), which is to the best of our knowledge the only study of this effect in vivo. Moreover, when measured under a typical pairing STDP protocol, equation 2.3 will automatically shift toward LTP for higher pairing frequencies, since for each postsynaptic spike, the  $x_j$  will accumulate more presynaptic spikes in the causal STDP window while the negative term remains constant. This effect is hence reminiscent of the tendency toward LTP (and abolishment of LTD) for higher frequencies found experimentally (Sjöström et al., 2001). Altogether, the agreement with these experimental STDP data is in a sense remarkable, given that equation 2.3 can be derived from purely statistical principles.

However, there are also a few deviations from STDP data. Most important, the negative contribution of equation 2.3 is activated for every postsynaptic spike regardless of presynaptic input. Hence, long-term depression (LTD) is not spike timing dependent and can be triggered by postsynaptic spikes alone. We address this potential issue later, when we consider a variation of equation 2.3 with timing-dependent depression in Figure 3; it turns out that the results are qualitatively very similar. Another deviation is with regard to the frequency dependence mentioned above: despite the general shift toward LTP, the exact form of frequency dependence in experimental data (see Figure 7 in Sjöström et al., 2001) cannot be reproduced by equation 2.3.

For the described network architecture, a computer simulation was carried out in which a fluctuating external variable  $\theta \in [0, 2\pi]$  with a periodic boundary condition (like any angle) was presented (see Figure 2). Concretely one could interpret this external variable as a motion direction. The changes in  $\theta$  were designed to be slightly slower than the network dynamics, corresponding to the assumption that the sensory stimulus remains stable for the duration of a few PSPs. A number of broadly tuned sensory neurons with Poisson firing statistics represented the value of this external variable over time. Initially synaptic readout weights were randomly initialized, leading to rather unspecific responses in the readout circuit when stimulated with different  $\theta$ . At t = 100 s of exposure to different population responses, readout neurons started to specialize on different stimuli. This was reflected by both selective firing of individual readout neurons to certain  $\theta$  and the gradual specialization of synaptic readout weights to different patterns. This specialization can be understood as follows. During exposure to population responses, the readout network continuously emits spikes. After each readout spike, the current population response leaves a small trace in the synaptic weights of the readout neuron k, which has fired. In particular, as expected from STDP, synapses that are not active at the time of the postsynaptic spike are depressed, and those with strong activity are enhanced. The synaptic changes induced by plasticity increase the probability that the same neuron fires again when a similar population

response is registered, while making it less likely to fire when significantly different patterns are active. Due to the inhibitory circuit that promotes competition, different neurons start to pick up different patterns. As a consequence, a specific pattern emerges in the synaptic weights of each readout neuron. This pattern reflects the history of population responses that made the readout neuron fire. At t = 2000 s the readout weights reached a setting that was stable with minor fluctuations around some target values. One can now compare this converged weight setting with the results by Jazayeri and Movshon (2006) on optimal population decoding. As shown in Figure 2C, the weights, which have autonomously emerged during exposure to unknown population responses, have recovered the previously identified theoretical optima for likelihood decoding, in particular those obtained through equation 1.2.

What is the overall effect of this learning process? Most important, individual readout neurons have become specialists for individual hidden causes (or stimuli). Hence, after, but not before, learning, each readout spike can be considered an indication for a particular hidden cause. As a consequence, one can achieve an increasingly faithful reconstruction of the original stimulus  $\theta$  from readout spikes, as shown in Figure 2E.

2.2 Learning Theory for STDP in a Network with Lateral Inhibition. We have shown in Figure 2 that the STD rule 2.3 in a network with lateral inhibition can regulate the synaptic readout weights such that near-optimal decoding of population responses can take place. We demonstrated that a main feature of this learning process was the emergence of a sparse spike code, optimized to reveal the hidden causes of high-dimensional input spike trains. Here we will show that this behavior can be understood in terms of a rigorous learning theory. Our theoretical analysis relies on the notion of an implicit generative model implemented in the network-a powerful statistical tool for the extraction of hidden causes from highdimensional data. Using basic assumptions about the input distribution, we arrive at a performance measure favoring sparse spiking in which different readout neurons fire for distinct hidden causes (with clearly distinct sensory representations). This theoretical tool allows us to track and analyze the performance of the network, in particular, small changes in performance due to a single application of the learning rule. The main result is that an application of the local learning rule, equation 2.3, is always expected to increase global performance. This increase can be understood in terms of an attractor dynamics in the weight space induced by the plasticity rule, in that the attractor centers are weight settings that are stable under the dynamics of the network. These equilibria are characterized by

$$w_{kj} = \log \langle x_j | k \text{ fires} \rangle + \text{const.},$$
 (2.4)



where we write  $\langle \cdot | k \text{ fires} \rangle$  for the empirical average over all spike times of neuron *k* during the operation of the circuit (for a more precise definition, see section A.4 in the appendix). Importantly, we show that these weight settings are not only the attractors of the network dynamics but also correspond to locally optimal weights settings from the perspective of the implicit generative model. Hence, we establish a direct link between local synaptic learning rules and theoretically optimal performance of the network.

A common problem in data analysis is to extract hidden causes from high-dimensional input **x** without supervision, that is, without the help of

a teacher signal, which provides the desired outcome **z** during training. Generative models are arguably the most powerful paradigm of unsupervised learning, with examples including (probabilistic) PCA, ICA, gaussian mixture models, and hidden Markov models (Bishop, 2006). The rationale behind generative models is "analysis by synthesis": if you aim to understand (decode) something, learn to build (generate) it yourself. The starting point is a model  $p(\mathbf{x}|\mathbf{W})$  describing the statistics of an input stream **x** as the

Figure 2: Spike-timing-dependent plasticity regulates synaptic weights towards theoretical optima for likelihood decoding. (A) An external variable  $\theta(t)$  was presented to the network of Figure 1A, here shown over a period of 6 s. (B) One hundred sensory neurons  $(x_1, \ldots x_M)$  with broad tuning functions and Poisson firing statistics provided a noisy population code of this stimulus. Through plastic synaptic readout weights, these sensory spikes were fed into a readout network (stochastic WTA) consisting of 20 spiking neurons with lateral inhibition. The readout network produces a sparse output code. An external observer can associate each readout neuron with a preferred stimulus  $\psi_k(t)$ , which develops over the course of learning. Based on the preferred stimuli, the original stimulus can be reconstructed from the sparse readout spikes. (C) The development of the weight matrix through theoretically optimal STDP is shown over 2000 s during exposure to a randomly changing stimulus  $\theta(t)$ . Colors range from dark blue (0) to dark red (2.4). Initially the weight matrix was randomly initialized. At t = 100 s, the effects of learning became visible: sensory neurons that typically fired together had developed a preference for activating the same readout neuron. At t = 2000 s, the weights had stabilized and settled on a configuration that closely matches the theoretical Bayes-optimal solution, equation 1.2. Note that the readout neurons in panel C (and D) are ordered by their preferred stimuli  $\psi_k$  at t = 2000 s, to facilitate visual comparison with the optimal solution. (D). The development of readout responses is shown over the course of learning. Initially, responses were unspecific to different stimuli. At t = 100 s, readout neurons had developed noticeable preferences for certain stimuli. At t = 2000 s, each readout neuron had become responsible for representing a particular preferred stimulus. The firing activity of a readout neuron reflects the posterior probability of the neuron's preferred stimulus, given the current population response. The characteristic sparse code that emerged is predicted by theory (see Bayes optimal, bottom). (E) The readout spikes were used to compute a reconstruction  $\theta_{rec}(t)$  (shown in red) of the external stimulus  $\theta(t)$  (shown in blue), in order to compare the quality of the readout at different stages of learning. The reconstruction signal at time t was obtained based on the preferred stimuli  $\psi(t)$  of those readout neurons, which spiked in a 100 ms time window before *t* (see the text). As shown, initially each readout spike conveyed rather unspecific information about the stimulus. The following specialization of readout neurons to different stimuli allowed for an increasingly accurate reconstruction.

result of a generative random process, typically involving hidden causes z,

$$p(\mathbf{x}|\mathbf{W}) = \sum_{z} p(\mathbf{z}|\mathbf{W}) \cdot p(\mathbf{x}|\mathbf{z}, \mathbf{W}).$$
(2.5)

This generative model can be understood as a two-step process. First, a hidden configuration is drawn according to the hidden probabilities  $p(\mathbf{z}|\mathbf{W})$ . Then the hidden states generate the actual data  $\mathbf{x}$  according to  $p(\mathbf{x}|\mathbf{z}, \mathbf{W})$ . The sum in equation 2.5 reflects the fact that the same data  $\mathbf{x}$  can be generated by different hidden configurations  $\mathbf{z}$ .

Then the goal of learning is to find parameters **W** that bring the distribution  $p(\mathbf{x}|\mathbf{W})$  generated by the model as close as possible to the actual data distribution, which we write here as  $p^*(\mathbf{x})$ . This is done by adjusting the parameters **W** of the model, usually in small steps, until the Kullback-Leibler divergence KL( $p^*(\mathbf{x})||p(\mathbf{x}|\mathbf{W})$ ), a quantity that measures the distance between the real distribution and the model, becomes minimal. During this process, an efficient representation for the hidden causes **z** emerges, which is optimized to "explain" the data in  $p^*(\mathbf{x})$ . This hidden representation can be retrieved by evaluating the posterior distribution  $p(\mathbf{z}|\mathbf{x}, \mathbf{W})$ .

Here, we adopt the perspective of a generative model to analyze learning for the network shown in Figure 1. But what is a good choice for a generative model? First, we restrict ourselves to a sparse representation of readout neurons, meaning that only a few hidden causes should be sufficient to explain any given input. A simple way to achieve this is a mixture model, in which at any time, exactly one hidden cause (or expert) is active and each expert is associated with a different set of input patterns. Second, since inputs to the neural network consist of spike trains with Poisson characteristics, a reasonable choice is to assign to each expert a characteristic pattern of input firing rates to which it is specialized. Formally, the generative model then takes the form

$$p(\mathbf{x}|\mathbf{W}) = \frac{1}{N} \sum_{k=1}^{N} \prod_{j=1}^{M} \text{Poisson}(x_j; \alpha^{-1} e^{w_{kj}}),$$
(2.6)

where Poisson(x;  $\lambda$ ) denotes the Poisson distribution over x with "rate"  $\lambda$  and  $\alpha$  is an arbitrary positive constant (note that our analysis extends beyond Poisson distributions; see below). The model can be understood as follows: the probability of cause k being active is  $\frac{1}{N}$ . If the cause k is active, the number of spikes for the sensory neuron  $x_j$  within a rectangular PSP window of length  $\tau$  is Poisson distributed with a rate that is encoded in the parameter  $w_{kj}$ . Hence, each cause k is an expert for a particular pattern of Poisson rates in the input, and this preferred pattern is encoded in the parameters  $w_{k1}, \ldots, w_{kM}$ . Note that prior parameters instead of a fixed prior  $\frac{1}{N}$  could be easily incorporated into the model. These prior parameters could

then be mapped onto neural excitabilities in the neural implementation (Nessler et al., 2010).

Several steps are required to establish a direct correspondence between the network implementation and the generative model: first, the network elements need to be related to the variables of the generative model. We already implicitly linked the synaptic weights to the model parameters in a straightforward fashion by using the weights  $w_{ki}$  in the definition of the generative model. Similarly, we can link the readout neurons to the hidden causes of the model: each hidden cause k is represented by one readout neuron, k. Second the network should support Bayesian inference, that is, it should respond to an input  $\mathbf{x}$  by inferring likely hidden states  $\mathbf{z}$ according to the posterior probabilities  $p(\mathbf{z}|\mathbf{x}, \mathbf{W})$ . Indeed we will show that each output spike from the readout network is generated according to the correct posterior distribution. Third, network plasticity should optimize the parameters of the generative model over time; the synaptic weights should come to reflect the input statistics  $p^*(\mathbf{x})$  through learning. This is the main focus of this article. We will prove this by relating STD learning to a standard algorithm for generative model learning, expectation-maximization (EM). One operation of the generative model that will not be required from the network is the actual generation of data x. Since, in contrast to other models (see Dayan, Hinton, Neal, & Zemel, 1995; Rao & Ballard, 1999), such an explicit generation of data is not needed here, we will refer to our model as an implicit generative model.

In the spirit of Jazayeri and Movshon (2006) and Nessler et al. (2010), Bayesian inference in the model defined by equation 2.6 can be related to a feedforward neural implementation with synaptic weights  $w_{kj}$  between input and readout layer. This is because the log likelihood of an input **x** under the cause *k* is given by,

$$\log p(\mathbf{x}|k, \mathbf{W}) = \sum_{j=1}^{M} w_{kj} x_j - \sum_{j=1}^{M} \alpha^{-1} e^{w_{kj}} - \sum_{j=1}^{M} \log(x_j!).$$
(2.7)

The second term can be ignored in practice if the input representation is homogeneous (Jazayeri & Movshon, 2006). The third term is independent of k and hence drops out later in the normalization of the posterior distribution  $p(k|\mathbf{x}, \mathbf{W})$ . This leaves only the first term, which is simply the feedforward sum of inputs, which a neuron k can compute in its soma. As a consequence, we can show that each spike from a readout neuron in Figure 1 can be interpreted as a sample from the correct posterior distribution  $p(k|\mathbf{x}, \mathbf{W})$  (see the appendix).

The key component of the learning theory is the connection between STD learning and the implicit generative model, in particular the optimization of its parameters **W**. In machine learning, the best-known algorithm for

performing this optimization is EM (Bishop, 2006). What we will show here is that the operation of STD learning in the cortical microcircuit of Figure 1 can be understood as a stochastic online version of EM. This allows us to view learning as an attractor dynamics in the weight space, where the attractors correspond to local optima in the generative model perspective.

The EM framework provides a general tool for deriving and analyzing the learning dynamics of autonomously learning systems that fit an implicit or explicit generative model to an external distribution  $p^*(\mathbf{x})$ . Of particular interest is the family of online EM algorithms (Sato, 1999; Cappé & Moulines, 2009), which operate at one input at a time, thereby gradually absorbing the input statistics into the parameters of the model. The fundamental steps of online EM are inference (expectation) and learning (maximization). During inference, the posterior probabilities  $p(\mathbf{z}|\mathbf{x}, \mathbf{W})$  of hidden configurations are evaluated for the current input x. The subsequent learning step uses both x and the inferred hidden configurations to perform a small update  $\Delta W$ , which increases the model's likelihood for the current input **x**. As indicated above, each firing of a neuron k in the WTA circuit of the network provides a sample from the current posterior distribution: hence, this is equivalent to a stochastic E-step. The corresponding postsynaptic spike then triggers the plasticity rule, equation 2.3, in the synapses  $w_{k1}, \ldots, w_{kM}$ . Through an analysis of this STDP-based update, one can show that this provides a step in the direction of a correct M-step in online EM. In particular, one can prove that the expected application of the rule with a sufficiently small learning rate will always improve the performance of the network until a local optimum is reached,

$$\operatorname{KL}(p^*(\mathbf{x})||p(\mathbf{x}|\mathbf{W} + \Delta \mathbf{W})) \le \operatorname{KL}(p^*(\mathbf{x})||p(\mathbf{x}|\mathbf{W})), \tag{2.8}$$

where  $\overline{\Delta W}$  is the expected update for a randomly chosen input pattern from  $p^*(\mathbf{x})$  (see the appendix).

This allows us to view learning in the considered cortical microcircuit as an attractor dynamics: the trajectories in weight space induced by STD learning are attracted to specially distinguished weight settings which have the property that for all synapses  $\overline{\Delta w_{kj}} = 0$ , that is, the system is in equilibrium with respect to equation 2.3. These attractors in weight space are optimal from the perspective of the implicit generative model in the sense that they correspond to locally optimal solutions to the problem of fitting the internal model  $p(\mathbf{x}|\mathbf{W})$  to the input distribution  $p^*(\mathbf{x})$ .

Before moving on to the generalization of this main result to non-Poisson input statistics, a remark is in order concerning the connection between the optimization of the implicit generative model  $p(\mathbf{x}|\mathbf{W})$  and the optimal readout weights for population coding, as they were derived in Jazayeri and Movshon (2006). As already suggested by the simulation results shown in Figure 2, these two optimality criteria are intimately connected. This is

because the "true" statistics of the population code  $p(\mathbf{x}|\theta)$  from which the optimal readout weights in Jazayeri and Movshon (2006) are derived can be mapped with arbitrary precision onto a mixture model representation in the form of equation 2.6. The mapping is most faithful if the Kullback-Leibler divergence between the real population code and the mixture model is lowest: this is precisely what the plasticity rule optimizes. By increasing the number of readout neurons, the theoretically achievable divergence can be made arbitrarily small, and the two optimality criteria become effectively equivalent.

The assumption of Poisson variability is a reasonable and popular approximation to the firing statistics of real cortical neurons. At the same time, there exists ample experimental evidence for deviations from this rule, such as bursty or regularly firing neurons (Shinomoto et al., 2009). This raises the question of whether our learning theory can also be applied to other firing statistics. It turns out that the theory can indeed be generalized to the exponential family of distributions, which contains many well-known parametric models like the Poisson, normal, gamma, and negative binomial distributions. The main result of this generalization is that each firing statistic is associated with a different, small variation of optimal STDP. Through the generalization to exponential families, the presented theory for STDP in the context of lateral inhibition becomes applicable to a wide range of biologically relevant firing statistics (see the online supplement). The power of exponential family distributions also allows incorporating more realistic EPSP shapes: given the firing statistics of sensory neurons, together with the EPSP shape, one can construct a tailor-made exponential family distribution that accounts for the variability encountered in the inputs  $\mathbf{x}$  (see the online supplement). This makes the theory applicable to virtually arbitrary EPSP shapes, in particular, those derived from electrophysiological experiments.

**2.3 A Family of STDP Rules Leads to Near-Optimal Readouts.** These theoretical results are quite encouraging, since they establish a link between a particular form of STDP, on the one hand, and the powerful statistical framework of EM, on the other hand. But how relevant are these results for the study of autonomous learning processes in the cortex? After all, the theoretically postulated mechanism for depression in equation 2.3, which leads to LTD even in the absence of presynaptic spikes, is quite speculative. It would thus be highly interesting to know whether this term could also be replaced by timing-dependent depression without loss of functionality. Furthermore, plasticity mechanisms in the brain appear to be heterogeneous and noisy (Sjöström et al., 2001; Caporale & Dan, 2008), in contrast to precise theoretical rules. If one assumes a given PSP shape and Poisson variability, is equation 2.3 really the only rule that can be guaranteed to lead to the emergence of optimal weights?

In this section, we provide further results to address these questions. First, the convergence results (and the link to EM) that hold for equation 2.3

can in fact be easily generalized to a whole family of optimal STDP rules with homeostatic plasticity,

$$\Delta w_{ki} = \eta \cdot z_k \cdot (f(w_{ki}) \cdot x_j \cdot \alpha e^{-w_{kj}} - f(w_{ki})), \tag{2.9}$$

where  $f(w_{kj})$  can be any strictly positive function. For  $f(w_{kj}) = 1$ , the original learning rule is recovered, which yields LTD proportional to the postsynaptic rate  $\langle z_k \rangle$ . This form of LTD is thus reminiscent of synaptic scaling (Turrigiano, 2010), except that in synaptic scaling, updates scale with the current weight (Abbott & Nelson, 2000). Indeed, a learning rule with synaptic scaling can be obtained by letting  $f(w_{kj}) = w_{kj}$ . We verified the correctness of this (optimal) variant of equation 2.3 in Figure 3B.

Furthermore, we tested two variations that are not provably optimal from the perspective of the presented learning theory but approximate the functionality of the optimal rules well. As shown in Figure 3C, in another variation of equation 2.3, the time constant of the causal STDP window was chosen twice as long as the PSP decay constant (while its magnitude was halved). The resulting performance is virtually optimal. Finally, we tested the performance of an STDP rule with common timing-dependent depression (see Figure 3D), with additional superimposed noise. The resulting weights are slightly less pronounced and noisier, but the readout performance is comparable to the other variants (compare RMSE values in the bottom row). Notably, this behavior is also quite robust against variations of parameters of this STDP rule (see Figure S1 in the online supplement).

Although these variants cover only a few cases of special interest, they make clear that the main finding of this article, the emergence of optimal readouts, is not an artifact of a specific definition of the learning rule, but in practice a property of a whole family of STDP learning rules. The most important features that appear to characterize this family of STDP rules are (1) the pronounced weight dependence of STD potentiation (in accordance with experimental data) and (2) some form of homeostatic regulation of weights, implemented either explicitly (e.g. through synaptic scaling) or implicitly through STD depression.

2.4 Maintenance of Optimality in Spite of Drastic Changes in Downstream and Upstream Neurons. To our knowledge, this article provides the first adaptive neural approach to optimal decoding from sensory populations in the literature. As a consequence, the canonical microcircuit considered here is substantially more flexible than previous static models for optimal decoding. In the following computer simulations, we have selected two biologically motivated dynamic scenarios that demonstrate its functional advantages. The first scenario is based on the observation that plasticity in sensory cortex appears to persist throughout adulthood, as suggested by numerous studies in different species and cortical areas



Figure 3: A variety of STDP rules leads to (near-) optimal readouts. Top row: Illustration of the four STDP rules that were tested. To illustrate weight dependence, the STDP curves are shown for different weights (blue: 0.5, green: 1, red: 1.5). (A, B) Theoretically optimal rules: equation 2.3 and a variant with synaptic scaling, respectively. (C) A modification of equation 2.3 with the time constant of the causal STDP branch doubled. (D) STDP with common timing-dependent depression and superimposed noise. The black dots in the top row represent noisy samples (weights for these samples are drawn randomly from a normal distribution with mean 1 and standard deviation 0.4). Middle row: Synaptic weights after 500 s of learning. As a performance measure, the reconstruction RMSE is shown (the root mean squared error of reconstruction signal in a 2000 s test run). Bottom row: weights after 3000 s. The RMSE of the optimal decoder is 0.58. All variants yield near-optimal performance.

(Trachtenberg et al., 2002; Goel & Lee, 2007). Hence, the tuning functions of sensory neurons, especially in superficial layers, are not fixed but subject to constant change. This constitutes a challenge to downstream populations that rely on the representation provided by sensory areas. We will show that the canonical microcircuit considered here gracefully handles a scenario of changing sensory tuning functions: any change in the input distribution is automatically detected by the network, resulting in an appropriate adjustment of the weight settings. As a consequence of this constant readaptation, the readout representation of the external stimulus can remain remarkably

invariant to changes in the sensory representation. The second scenario deals with neurogenesis and neuron death, two phenomena that have been reported repeatedly in adult cortex (Morrison & Hof, 2007). We studied the effects of neuron growth and death in the readout population, showing that through learning, the readout network can minimize the detrimental effects of cell death and maximize the gain in representation accuracy brought by neurogenesis.

In the first scenario, we studied the consequences of changing tuning functions in the sensory population on the readout network (see Figure 4). For this, we divided the sensory population into three groups (G1–G3). Initial sensory tuning functions were chosen such that the whole range of the input stimulus was represented uniformly in each group. The simulation was then split into three phases: a static phase (0 s–1000 s), a dynamic phase (1000 s–2000 s), and a consolidation phase (2000 s–3000 s). During the static phase, sensory tuning functions were kept constant, and the readout network learned an optimal readout representation based on the initial sensory code. The resulting weight matrices are qualitatively identical to those obtained in Figure 2C. In the dynamic phase, tuning functions in all three groups of sensory neurons started to change. In the first group, tuning functions broadened, in the second group they narrowed, and in the third group tuning functions narrowed and simultaneously started to develop a second mode (see Figure 4B). This change in the sensory population code was automatically "detected" by the readout network, causing the synaptic readout weights to take up pursuit of a quickly moving target: the attractor centers of the learning dynamics that depend on and thus drift in parallel with sensory tuning functions. A snapshot of the readout weight matrix at time 1600 s in the middle of the dynamic phase is shown in Figure 4C. In the consolidation phase, sensory tuning functions were fixed again. This allowed the readout network to converge to the new optimal setting, thereby reinterpreting all sensory neurons according to their new properties. This is particularly visible in group 3 (G3), where each sensory neuron had developed two modes in the tuning function. The corresponding optimal decoding strategy is to connect each sensory neuron in G3 strongly with the readout neurons that correspond to the two modes. As shown in Figure 4C, this strategy was found autonomously through learning. Figure 4D illustrates the preferred stimuli of five exemplary readout neurons during the dynamic phase. One should note that in spite of drastic changes in the sensory representation (e.g., the development of bimodal tuning functions; see Figure 4B), the preferred stimuli of readout neurons barely change during the adaptation period. Hence, learning facilitates the decoupling between sensory and readout representations with respect to the external stimulus.

In a second scenario (see Figure 5) we studied the effect of neuron growth and death in the readout population. In a computer simulation, a microcircuit was set up with five readout neurons, and the weights were allowed



Figure 4: Canonical microcircuit maintains optimality in spite of changing tuning functions in the sensory population. (A) The sensory population was divided into three groups, G1–G3, each responsive to the whole range of input stimuli. (B) The simulation had a static phase (0 s-1000 s), a dynamic phase (1000 s-2000 s) and a consolidation phase (2000 s–3000 s). In the static phase (first row), constant sensory tuning functions were used analogous to Figure 2. In the dynamic phase (second row), tuning functions started to change. Each group of sensory neurons developed different tuning functions. In G1, tuning functions broadened, in G2 they narrowed, and in G3 they became bimodal. In the consolidation phase (third row), tuning functions were fixed again. (C) The change in the sensory population code during the dynamic phase (second row) resulted in a shift in the attractor centers of the learning dynamics. This became visible in the synaptic readout weights, which slowly tracked the changes in the sensory population (see the weight matrix at t = 1600 s). During the consolidation phase, the synaptic weights then converged to a new optimal representation. Colors range from dark blue (0) to dark red (1.9). (D) The preferred directions of five representative z-neurons are shown over the course of learning. Although the sensory representation drastically changes over the considered period (B), the readout representation remains remarkably stable.



Figure 5: STDP enables robust decoding despite neuron growth and death. (A) The preferred stimuli  $\psi_k$  of readout neurons are shown over time. The circuit was initially set up with five readout neurons. Through learning, these readout neurons became experts with uniformly distributed preferred stimuli (t = 800 s). At t = 800 s, five readout neurons were added to the circuit with randomly initialized weights. They quickly specialized on the most weakly represented ranges of the stimulus space (t = 1600 s). A similar effect could be observed after adding another five neurons, leading to a refined representation (see 1600 s–2400 s). At t = 2400 s, the cell death of 10 readout neurons was simulated. This resulted in large representation gaps: poorly represented regions of the input stimulus. Through learning, the preferred stimuli of WTA neurons rearranged toward an optimized configuration again (2400 s-3200 s). (B) Mean-squared reconstruction error over time. Performance plunged whenever randomly initialized readout neurons were inserted (at t = 800 s and t = 1600 s), before recovering again through learning. Neuron loss at t = 2400 s also degraded performance. Although learning minimized the ensuing detrimental effects, the final MSE with 5 neurons (t = 3200 s) is markedly higher than with 15 neurons (t = 2400 s). Blue dots correspond to selected time points in panel C. (C) Weight matrices at different stages of the simulation (time points indicated by blue dots in panel B). Colors: dark blue (0)-dark red (2.1). An optimized weight setting for five readout neurons is shown at t = 800 s. At t = 900 s, the synaptic weights of five newly grown neurons, amid previously existing neurons, are depicted shortly after random initialization. Optimized weight settings for 10 and 15 neurons are shown at t = 1600 s and t = 2400 s, respectively.

to converge to a stable setting on stimulation with input stimuli analogous to Figure 2. At t = 800 s and t = 1600 s, the growth of five new readout neurons with randomly initialized weights was simulated, yielding a jump

in the mean squared error (MSE) of the reconstruction (see Figure 5B). In theory, a larger number of readout neurons allows a more finely grained representation, and hence a more accurate readout. However, this requires that the weights of all neurons are adjusted appropriately. In particular, the newly formed neurons should learn to respond to the most poorly represented regions of the external stimulus. Figure 5A shows that this optimal strategy automatically emerges through learning: the preferred stimuli  $\psi_k$  of the newly grown neurons quickly learned to fill in the "spaces" in between the existing neurons. This is also reflected by a lower mean squared error of the reconstruction signal (Figure 5B; compare at t = 800 s and t = 1600 s). A similar effect is observed after simulating the new growth of another five neurons with random weight initialization. At t = 2400 s, the sudden death of 10 neurons was simulated. Theoretically, this substantially reduces the achievable readout accuracy. However, the loss in accuracy is smallest if the surviving neurons rearrange such that their preferred stimuli  $\psi_{k}$ are uniformly distributed over the input range. As shown in Figure 5A, this is the strategy automatically implemented by learning in the network (2400 s-3200 s). Hence, the detrimental effects of neuron loss on the readout representation are minimized through learning (see the corresponding reduction of reconstruction MSE in Figure 5B).

2.5 Improved Representation of Behaviorally Relevant Inputs. The ability to dynamically allocate resources for representing important peripheral inputs in a use-dependent manner is a hallmark of cerebral cortex (Buonomano & Merzenich, 1998). Indeed, cortical representations are highly plastic, and the dynamic changes in cortical representation on manipulation in input or during task learning are thought to underlie the phenomenon of perceptual learning: the improvement of sensory abilities during training (Seitz & Watanabe, 2005). Although the exact neural mechanisms that give rise to task-related improvements are largely unknown, experimental evidence suggests an important implication of synaptic plasticity of intracortical connections (Buonomano & Merzenich, 1998) in conjunction with top-down "relevance" or reward signals (Seitz & Watanabe, 2005). The latter have been hypothesized to be communicated to local cortical circuits via diffuse neuromodulatory signals, such as "acetylcholine, norepinephrine or dopamine, which gate learning and thus restrict sensory plasticity" (Seitz, Kim, & Watanabe, 2009). Here we will show that by incorporating such a modulatory signal into the considered cortical microcircuit model, the internal representation of stimuli becomes automatically focused on behaviorally relevant inputs-those inputs that are consistently paired with high levels of the modulatory signal. The resulting allocation of cortical resources in proportion to relevance is strongly reminiscent of experimentally observed training-dependent cortical map plasticity (Buonomano & Merzenich, 1998).



Figure 6: Improved representation of behaviorally relevant inputs. (A) For 1000 s, weights were regulated by equation 2.3 with a constant learning rate (no modulation, blue line). For the following 3000 s, the learning rate was modulated in dependence of the external variable (red curve), thereby simulating the stimulus selective modulation of learning. The external variable  $\theta(t)$  and the tuning functions of sensory neurons were simulated as in Figure 2. (B) The simulated cortical microcircuit responded to the modulation by increasing the representation density for those regions paired with strong modulation, at the cost of reduced representation density for other regions. (C) The weight setting that emerged when gating was present (t = 4000 s) differs from the unmodulated case (t = 1000 s) by a stronger focus on population responses that coincided with periods of high "relevance." In particular, after modulated learning, more readout neurons receive their strongest synaptic connections from sensory neurons that respond to relevant stimuli (~sensory neurons 40-60). These readout neurons correspond to the dense region of preferred directions in panel B. As a consequence of the reallocation of resources, a greater readout resolution is achieved for regions of the high-dimensional input space that are behaviorally relevant.

In a computer simulation, we incorporated a global modulatory signal into the cortical microcircuit considered above, acting as a soft gate for learning. The gating was implemented as a multiplicative factor on the learning rate. During an initial reference phase, modulation remained constant during exposure to stimuli, leading, as expected, via learning to a uniform internal representation (see Figures 6B and 6C at t = 1000 s). Modulation was activated at t = 1000 s. The modulation was chosen in dependence of the current external stimulus, such that certain ranges of  $\theta$  were paired with higher levels of the modulatory signal than others (see Figure 6A, red

curve). As shown, population responses that coincided with higher levels were more strongly imprinted in the synaptic network weights, and consequently, the internal representation started to shift toward a denser concentration around more "relevant"  $\theta$ s (see Figures 6B and 6C at t = 4000 s). In contrast, stimuli paired with lower levels of modulation became more crudely represented by the network. Hence, the resulting density of representation in the distribution of preferred stimuli  $\psi_k$  is directly connected to the strength of modulation imposed for different  $\theta$ .

Such use-dependent allocation of cortical resources is a ubiquitous feature of cortical plasticity and has been shown to correlate with significant performance improvements on corresponding tasks (Buonomano & Merzenich, 1998). Our results indicate that the modulation of learning in a local cortical microcircuit model through a diffuse global signal is indeed sufficient to reproduce relevance-dependent allocation of representational resources. This provides a simple mechanism that could underlie the characteristic improvements in sensory abilities associated with perceptual learning.

#### 3 Discussion

Ensembles of pyramidal cells with lateral inhibition on layers 2/3 and layer 5/6 (Fino & Yuste, 2011) constitute a universal network motif of cortical microcircuits in many different cortical areas and different species. We propose that these network motifs acquire through STDP a rather universal computational function that appears to be essential for many, if not all, cortical areas: the compression of high-dimensional noisy spike inputs into lower-dimensional sparse representations of the most likely hidden causes of these high-dimensional spike inputs. This yields in particular an emergent optimal decoding of population codes in the sense of Jazaveri and Movshon (2006). This emergent computational function of stochastic WTA circuits is very stable because a rigorous learning theory shows that the theoretically optimal values of synaptic weights are attractors in the dynamics of synaptic weights under the considered class of STDP rules with homeostatic plasticity. A remarkable feature of the underlying learning theory is that it creates a link between local synaptic learning rules such as STDP and the most powerful known abstract method for autonomous (i.e., unsupervised) learning: EM. This link holds provably for the class of optimal STDP rules and, in practice, also approximately for STDP rules that are not directly covered by theory (see Figure 3).

Furthermore the learning theory for STDP that we have introduced provides a new benchmark for analyzing and understanding from a functional perspective the large variety of parameters and learning curves for STDP that have been found at different synapses (Dan & Poo, 2006). This learning theory proposes that the theoretically optimal version of STDP (from the perspective of autonomous generation of implicit generative models for high-dimensional noisy spike inputs through EM) depends on both the firing statistics of pre-synaptic neurons and the average shape of EPSPs at the soma in a systematic and predictable manner.

In particular, the learning theory for STDP that we have presented throws new light on the question of downstream decoding of information conveyed by populations of noisy spiking neurons. Jazayeri and Movshon (2006) had already shown that simple linear neurons could in principle carry out a theoretically optimal maximum likelihood decoding, provided that suitable values are chosen for their synaptic weights. The learning theory that we have presented shows that these optimal weight values are in fact attractors with regard to the dynamics of synaptic weights under STDP in stochastic WTA circuits. This learning theory for STDP therefore also provides a possible explanation for the problem of how downstream neurons can maintain optimal decoding of population codes of sensory neurons in spite of ubiquitous changes in the tuning functions of sensory neurons (see Figure 4) and changes in the number of neurons involved in this decoding (see Figure 5). Furthermore if one assumes that synaptic plasticity is gated by neuromodulators or network activity so that the learning rate of STDP is increased for behaviorally relevant stimuli, the downstream decoding network automatically adapts its resolution in order to achieve a finer representation of behaviorally relevant ranges of external stimuli (see Figure 6).

An interesting aspect, which we did not study in this article, is whether self-organizing maps, similar to the well-documented orientation maps in cat visual cortex (Hubel & Wiesel, 1963), could emerge in the presented model. Indeed, one could think of clever changes in the model to induce such effects, for example, adding lateral excitatory connections between neighboring neurons, which would likely facilitate the emergence of locally smooth maps. But it is not clear at all whether a downstream "user" of the network output actually requires such a smooth map representation; indeed, many rodents, including rats, mice, and squirrels, can live without smooth maps (Van Hooser, Heimel, Chung, Nelson, & Toth, 2005). Furthermore, from a purely functional standpoint, neurons in the presented model with similar tuning (e.g., specialized on similar motion directions) will generally have nonnegligible signal correlation (see Figure 2D). Hence, downstream neurons that receive connections from the WTA neurons could quite easily detect groups of neurons that code for similar input stimuli and establish selective synaptic connections to "functionally neighboring" neurons even if their spatial arrangement is scrambled.

The WTA mechanism in this article relies on an exponential input-output nonlinearity of the form  $e^u$  to fit the requirements of Bayesian inference. Indeed, this is precisely the input-output relation found empirically by Jolivet et al. (2006), who fit a stochastic spike response model to predict spike timings in pyramidal cells from L5 rat somatosensory cortex. Other authors have suggested a power relation of the form  $u^p$  between membrane potential

and output rate (Priebe, Mechler, Carandini, & Ferster, 2004). Although we expect that power relations would suffice to induce the desired competition, it would be interesting to study the effect of different nonlinearities in future work.

One of the model assumptions is conditional independence of inputs, and hence learning is guaranteed to be optimal only if there are no noise correlations in the input (although convergence is guaranteed regardless). Many studies have found moderate noise correlations in cortex (Smith & Kohn, 2008; Huang & Lisberger, 2009). On the other hand, a recent study by Ecker et al. (2010) found that in primary visual cortex, even nearby neurons with similar tuning show close to zero correlation (and the authors argue that earlier findings may need to be reevaluated under more precise experimental paradigms). Nonetheless, we tested how nonzero noise correlations affect network performance. Figure S2 in the online supplement shows that for a moderately correlated input population code, the performance remains quite high although not optimal. This has to be expected from a learner that relies on the independence assumption, and is in accordance with the findings of Graf, Kohn, Jazayeri, and Movshon (2011) who showed that linear readouts can be further improved if input correlations are taken into account. Whether input correlations could be efficiently exploited in the context of this article-whether and how the correlation structure in the inputs could be learned autonomously in an unsupervised manner (Graf et al., 2011, used supervised learning) by a spiking network with local plasticity rules-remains an open question for future work.

Another model assumption we made is that the stimulus distribution is uniform. Note that this does not invalidate the presented results for nonuniformly distributed inputs: even in a nonuniform stimulus setting, learning is still guaranteed to converge to an optimal solution with respect to the fixed prior distribution (since the assumption is made on the level of the implicit generative model, not on the input distribution  $p^*(\mathbf{x})$ ). The solution produced by the network for a nonuniform stimulus distribution resembles Figure 6: a greater number of neurons specialize on the more likely region of the stimulus (see Figure S3 in the online supplement). This is precisely what is expected from a maximum likelihood learner with a fixed prior model distribution. Another important question is whether the network also adapts its implicit prior distribution over the stimulus in an optimal manner. In other words, does the WTA network automatically become biased toward activating those neurons whose preferred directions occur more often in the input? Indeed, this turns out to be the case. Since highprobability stimuli will automatically attract more WTA neurons during the specialization process, the implicit prior distribution of the network favors high-probability stimuli after learning. Indeed this phenomenon is also quantitatively consistent with the prediction of a Bayesian framework (see Figure S3).

3.1 Related Work. As a model for decoding of high-dimensional population codes in the brain, the key novelty of this work is plasticity. In particular, this article demonstrates to the best of our knowledge for the first time that population codes not only can be read out efficiently (this had already been shown by Deneve et al., 1999; Jazayeri & Movshon, 2006; and Chaisanguanthum & Lisberger, 2011), but that this readout can also be learned optimally by a stereotypical cortical microcircuit motif in a process that is entirely autonomous and self-organizing. An interesting difference from previous models for population decoding lies in the output code of the readout circuit, which can represent a whole distribution (in contrast to Deneve et al., 1999), while being sparse for typical stimuli (as opposed to the predictions of Jazaveri & Movshon, 2006). Hence, the information extracted about the inputs is represented and conveyed by the whole network population rather than by the identity of a single neuron. Strong, unambiguous stimuli will elicit sharp responses, whereas weaker, low-contrast, or ambiguous stimuli will lead to a distributed code. Furthermore, our model predicts that neural activity represents samples from a probability distribution, a coding scheme that has recently attracted considerable attention, as it appears to be particularly suitable for probabilistic representations subject to learning and adaptation (Fiser, Berkes, Orban, & Lengyel, 2010). As a consequence, our model is not only consistent with the experimentally observed trial-to-trial variability of neuronal responses, but in fact requires that neurons respond in a stochastic manner. Hence, one can view this article also as a contribution to the growing literature on computational properties of networks of stochastically firing neurons.

The link to EM extends the model's generality beyond population coding to input representations, which cannot be easily described in terms of a fixed number of external variables. In this more general context, different authors have proposed neural network models that are able to carry out Bayesian inference (Doya, Ishii, Pouget, & Rao, 2007), and some also considered the question how probabilistic representations could emerge autonomously through learning, for example, Dayan et al. (1995), Rao and Ballard (1999), and Keck, Savin, and Lücke (2012), who used a related approach to the one developed here but required more approximations and did not model spiking neurons (and hence also not the relation between EM and STDP). While these models focused on rather artificial neural networks based on abstract (either binary or continuous-valued rate-based) neural units, recent studies have started to address whether more realistic spiking neural networks are also capable of acquiring optimized probabilistic representations through autonomous learning. A model by Deneve (2008) focused on a single spiking neuron and showed that such a neuron can in principle learn an efficient code for its inputs in a basic temporal generative model. Whether parameter learning in the model can be scaled up to multiple neurons, which would be required to cope with complex but biologically realistic input distributions, remains an interesting open issue. Finally, the recent model by Nessler et al. (2010) showed that spiking neurons can learn to represent a mixture of multinomial input variables, assuming that exactly one input neuron in a group is active at a time. The generalization and extension of this result to biologically realistic firing statistics and EPSP shapes, required for the decoding of realistic population codes, was described in this article.

Finally, Law and Gold (2009) showed in a related work that near-optimal population readouts can also be learned autonomously through a simple reinforcement learning (RL) rule. More precisely, they demonstrate that for a given decision problem, for example, whether the motion direction encoded by the input is less than or more than 180 degrees, near-optimal linear readouts from an input population can be found through RL. The learning scheme relies on a feedback signal that conveys the correctness of the decisions computed from the readout. In principle, for any particular decision problem, a different set of readout weights is required. Our work complements these findings insofar as we show that an (external) feedback signal is not needed for the emergence of optimal readouts. Instead, optimal readout weights in a downstream population of neurons can emerge through a purely self-supervising process. Furthermore, in principle, any decision problem can then be reduced to simply counting spikes emitted by the network population, for example, checking whether those output neurons with preferred directions below 180 degrees produced more spikes than those with preferred directions of above 180 degrees. Given the attractive properties of both approaches and the prevalence of both STDP and reward-modulated learning in the brain, it is not unreasonable to assume that the cortex employs a combination of these learning mechanisms.

**3.2 Experimentally Testable Predictions.** Our results predict that abolishment of STDP during a critical period prevents the emergence of sparse codes for frequently occurring sensory stimuli. They also predict that with intact STDP, the coding properties of pyramidal cells will change in a predictable manner in response to changes in the distribution of external stimuli or their behavioral relevance (see Figure 6), since this change will be tracked by the implicit generative model of ensembles of pyramidal cells (probably on layers 2/3 and layers 5/6). In addition, a lesion of some neurons within this ensemble will cause a redistribution of neural codes among the remaining neurons (see Figure 5). Our theoretical analysis of optimal versions of STDP predicts a specific dependence of features of STDP on the firing statistics of presynaptic neurons that can in principle be tested experimentally.

Furthermore, our model predicts that in WTA neurons, that is, pyramidal cells in layers 2/3 and layers 5/6, the tuning of firing rates should be sharper than the tuning of membrane potentials. In fact, this is the well-known experimentally observed iceberg effect (Carandini & Ferster, 2000). But in our model, the iceberg effect also has a novel functional interpretation from

the perspective of Bayesian inference: the instantaneous firing rate must depend on the current membrane potential via a sharpening exponential activation function in order to ensure that neurons encode probabilities. Hence, the iceberg effect is a prerequisite for correct learning of optimal decoding weights. Another direct prediction of our model is that the stimulus selectivity in ensembles of pyramidal cells with lateral inhibition is sharper when familiar stimuli are presented compared to novel stimuli. This effect has been recently reported in monkey inferior temporal cortex (ITC) by Freedman, Riesenhuber, Poggio, and Miller (2006).

#### 4 Conclusion \_

In summary, we have shown that in conjunction with STDP, a common network motif of cortical microcircuits acquires a generic computational function: it creates a sparse representation of complex high-dimensional spike inputs to a local microcircuit. Although we have discussed in this article only the application to decoding of information from populations of noisy sensory neurons, the generation of sparse representations for complex high-dimensional inputs, which converge onto a microcircuit from many different brain areas, is a candidate for a generic computational operation that is meaningful for microcircuits in any cortical area.

#### Appendix: Methods \_

**A.1 Spike-Timing-Dependent Plasticity Rules.** All STDP rules used in this article (see Figure 3 for a visual comparison) fit into the phenomenological framework of STDP rules by Gerstner and Kistler (2002). Using their notation,  $S_j(t) = \sum_f \delta(t - t_j(f))$ , and  $S_k(t) = \sum_f \delta(t - t_k(f))$  denoting preand postsynaptic spike trains, respectively, STDP can be expressed as

$$\frac{d}{dt}w(t) = \eta S_j(t) \left[ \int_0^\infty a_2^{\text{pre,post}}(s, w) S_i(t-s) \, ds \right]$$
(A.1)

$$+ \eta S_i(t) \left[ a_1^{\text{post}}(w) + \int_0^\infty a_2^{\text{post,pre}}(s, w) S_j(t-s) \, ds \right], \quad (A.2)$$

where  $a_2^{\text{pre,post}}(s, w)$  and  $a_2^{\text{post, pre}}(s)$  denote the pre-before-post and postbefore-pre kernels, respectively, and  $a_1^{\text{post}}(w)$  the post-only contributions. The kernels are given by (double-)exponential decays throughout this article:

$$a_2^{\text{pre,post}}(s, w) = A_+(w) \cdot [\exp(s/\tau_+) - \exp(s/\tau_{+\text{rise}})],$$
 (A.3)

$$a_{2}^{\text{post,pre}}(s,w) = A_{-}(w) \cdot [\exp(s/\tau_{-}) - \exp(s/\tau_{-\text{rise}})].$$
(A.4)

| Rules                            | $\tau_+(\tau_{+\rm rise})$ | $A_+(w)$               | $\tau_{-}(\tau_{-\rm rise})$ | $a_1^{\mathrm{post}}(w)$ | Noise |
|----------------------------------|----------------------------|------------------------|------------------------------|--------------------------|-------|
| A. Optimal rule, equation 2.3    | 20 ms (2)                  | $\alpha e^{-w}$        | 0 ms                         | -1                       | No    |
| B. Optimal with synaptic scaling | 20 ms (2)                  | $\alpha w e^{-w}$      | 0 ms                         | -w                       | No    |
| C. Longer causal window          | 40 ms (4)                  | $\alpha e^{-w}/2$      | 0 ms                         | -1                       | No    |
| D. Common STDP curve             | 20 ms (0)                  | $e^{\beta - \gamma w}$ | 60 ms (0)                    | 0                        | Yes   |
|                                  |                            |                        |                              |                          |       |

Table 1: Comparison of STDP Rules Used.

Notes: Labels A–D are as in Figure 3.  $\tau_{-} = 0$  ms in A–C corresponds to the absence of STD depression. For these rules, weight stabilization is achieved through the homeostatic plasticity term  $a_{1}^{\text{post}}(w)$ , which is triggered for every postsynaptic spike.

Post-before-pre effects were weight independent for all tested rules, that is,  $A_{-}(w) \equiv -1$ . The remaining parameters and functions are specified in Table 1 for each of the four rules used in this article.

We set  $\alpha = 4.3$ ,  $\beta = 1.58$ , and  $\gamma = 0.59$  for the simulations in the main text. Figure S1 shows the impact on performance of varying  $\beta$  and  $\gamma$ . For the common STDP curve (see Figure 3D), additional zero-mean gaussian noise was superimposed for each learning update triggered by a postsynaptic spike. The standard deviation  $\sigma_{\text{STDP}}$  of this noise was dependent on the magnitude of the deterministic update: for each deterministic update of magnitude *M*, gaussian noise with standard deviation  $\sigma_{\text{STDP}} = \kappa M + \zeta$  was added. The noise parameters were set to  $\kappa = 0.3$  and  $\zeta = 10^{-4}$ . The optimal rule, equation 2.3, was used for simulations in Figures 2 to 6. Rules B to D were used in Figure 3.

**A.2 Implicit Generative Model.** Neural activity and synaptic learning in the considered cortical microcircuit model can be understood from the perspective of an underlying implicit generative model. This generative model takes the form of a mixture model with  $k \in \{1, ..., N\}$  hidden causes, uniform priors, and conditionally independent Poisson variables  $x_i$ :

$$p(\mathbf{x}|\mathbf{W}) = \frac{1}{N} \sum_{k=1}^{N} p(\mathbf{x}|k, \mathbf{w}_k),$$
(A.5)

$$p(\mathbf{x}|k, \mathbf{w}_k) = \prod_{j=1}^{M} \text{Poisson}(x_j; \lambda_{kj} = \alpha^{-1} e^{w_{kj}})$$
(A.6)

$$=h(\mathbf{x})\cdot\exp(\mathbf{w}_{k}^{T}\mathbf{x}-A(\mathbf{w}_{k})), \qquad (A.7)$$

where  $A(\mathbf{w}_k) = \alpha^{-1} \sum_j e^{w_{kj}}$ , and  $h(\mathbf{x}) = \prod_j \frac{1}{x_j!}$ . We associate each hidden cause *k* with one readout neuron, *k*, and the parameters  $w_{kj}$  with the synaptic

weights of the network. This allows us to relate inference and learning in the generative model to the operation of the microcircuit.

A.3 Stochastic Winner-Take-All Circuit and Inference. Building on Nessler et al. (2010), we show here that the WTA circuit in the cortical microcircuit implements inference in the implicit generative model defined by equations A.5 to A.7. According to Bayes' rule, inferring the hidden cause k from input data x in the generative model can be written as

$$p(k|\mathbf{x}, \mathbf{W}) = \frac{\exp(\mathbf{w}_k^T \mathbf{x} - A(\mathbf{w}_k))}{\sum_l \exp(\mathbf{w}_l^T \mathbf{x} - A(\mathbf{w}_l))}.$$
(A.8)

In case of a homogeneous input representation, for example, a population code in which the sum of sensory activations is constant  $\sum_j x_k = A_0$ , this reduces to (see the online supplement),

$$p(k|\mathbf{x}, \mathbf{W}) = \frac{\exp(\mathbf{w}_k^T \mathbf{x})}{\sum_l \exp(\mathbf{w}_l^T \mathbf{x})}.$$
(A.9)

Note that the restriction  $\sum_j x_k = A_0$  is necessary to make the theory tractable. However, as demonstrated in the simulations through this article (see Figures 2–6) in which sensory input neurons fire randomly and only the total input population rate was kept constant as in Jazayeri and Movshon (2006), this is not required for the functionality of the model in practice.

Now consider a population of stochastically spiking readout neurons k that fire at an instantaneous rate  $\rho_k = e^{u_k}$ , depending on their current membrane potential  $u_k = \mathbf{w}_k^T \mathbf{x} - I$ . We make two basic assumptions about the inhibitory contribution: that I is common to all readout neurons and that the inhibitory circuit that provides I ensures that an approximately constant total target firing rate  $\rho_{total} = \sum_k \rho_k$  is maintained. This form of divisive inhibition introduces competition among the readout neurons, since a strongly activated readout neuron will claim a large fraction of the total target firing rate, thereby suppressing other readout neurons.

In an ideal stochastic WTA circuit, the inhibitory circuit keeps  $\rho_{total}$  constant. The ideal inhibitory contribution is given by

$$I = \log \sum_{j} \exp(\mathbf{w}_{j}^{T} \mathbf{x}) - \log \rho_{total}.$$
(A.10)

Then the readout neuron k will respond to an input **x** with an instantaneous rate,

$$\rho_k = \rho_{total} \cdot \frac{\exp(\mathbf{w}_k^T \mathbf{x})}{\sum_l \exp(\mathbf{w}_l^T \mathbf{x})}.$$
(A.11)

By comparison with equation A.9, one can verify that this is proportional to the posterior probability of the hidden cause *k* in the generative model. Note that since the relative spiking probabilities match the posterior probabilities, each spike produced by the readout circuit can be interpreted as a sample from the posterior distribution, independent of the total firing rate  $\rho_{total}$  of the circuit.

A.4 Equilibria of Theoretically Optimal STDP Rules and Maximum Likelihood. Here we discuss two important results. First, we derive conditions for the equilibrium points of STDP: stable weight settings for which the expected STDP update is zero. The significance of these points derives from the fact that they are attractors in the weight dynamics. Hence, STDP will always drive the weights into the neighborhood of such an equilibrium point. Second, we show that these equilibria also have a special interpretation from the generative model perspective: they correspond to locally optimal parameter settings for the generative model in the maximum like-lihood sense.

An equilibrium point of STDP is, by definition, invariant to the average STDP update. Our first goal here is to derive from this global definition more concrete statements for single synapses. To this end, we first define the empirical joint distribution over **x** and the activation of a readout neuron,  $z_k \in \{0, 1\}$ :

$$\tilde{p}(\mathbf{x}, z_k | \mathbf{W}) = p^*(\mathbf{x}) \cdot p(z_k | \mathbf{x}, \mathbf{W}), \tag{A.12}$$

where  $p^*(\mathbf{x})$  is the input distribution and  $p(z_k = 1 | \mathbf{x}, \mathbf{W}) = p(k | \mathbf{x}, \mathbf{W})$ , using the fact that the WTA circuit implements inference according to the generative model.

Then, at equilibrium, the following must hold for all *k* and *j*,

$$\langle \Delta w_{kj} \rangle_{\tilde{p}(\mathbf{x}, z_k | \mathbf{W})} = 0, \tag{A.13}$$

where  $\langle \cdot \rangle_{q(\cdot)}$  denotes the average (or expectation) taken over the distribution  $q(\cdot)$ .

After substituting the STDP learning rule specified in equation 2.3 and rearranging terms, one obtains

$$w_{kj} = \log \alpha + \log \langle x_j \rangle_{\tilde{p}(\mathbf{x}|z_k=1,\mathbf{W})}.$$
(A.14)

Indeed, this result is obtained for any learning rule from the optimal family, equation 2.9. Hence, at equilibrium, each synaptic weight  $w_{kj}$  is set to the logarithm of the average presynaptic activity  $x_{j}$ , the average taken over those input patterns that make the postsynaptic neuron k fire (plus some constant).

The relation of this result to the generative model becomes apparent when considering the maximization of the marginal likelihood  $(\log p(\mathbf{x}|\mathbf{W}))_{p^*(\mathbf{x})}$ , which is equivalent to the minimization of the KL-divergence  $\mathrm{KL}(p^*(\mathbf{x})||p(\mathbf{x}|\mathbf{W}))$ . For local maxima of the log-likelihood  $(\log p(\mathbf{x}|\mathbf{W}))_{v^*(\mathbf{x})}$  one has for all *k* and *j*,

$$\partial_{w_{ki}} \langle \log p(\mathbf{x}|\mathbf{W}) \rangle_{p^*(\mathbf{x})} =$$
 (A.15)

$$\left\langle \frac{1}{p(\mathbf{x}|\mathbf{W})} \partial_{w_{kj}} \sum_{k} p(\mathbf{x}, k|\mathbf{W}) \right\rangle_{p^*(\mathbf{x})} =$$
(A.16)

$$\left\langle \frac{1}{p(\mathbf{x}|\mathbf{W})} \sum_{k} p(\mathbf{x}, k|\mathbf{W}) \partial_{w_{kj}} \log p(\mathbf{x}, k|\mathbf{W}) \right\rangle_{p^{*}(\mathbf{x})} =$$
(A.17)

$$\left\langle \sum_{k} p(k|\mathbf{x}, \mathbf{W}) \,\partial_{w_{kj}} \log p(\mathbf{x}, k|\mathbf{W}) \right\rangle_{p^{*}(\mathbf{x})} =$$
(A.18)

$$\langle \partial_{w_{kj}} \log p(\mathbf{x}|k, \mathbf{W}) \rangle_{\tilde{p}(\mathbf{x}, z_k = 1|\mathbf{W})} = 0.$$
 (A.19)

By evaluating the derivative, one can verify that this is indeed equivalent to equation A.14. Hence, all STDP equilibrium points are automatically local optima with respect to the implicit generative model.

A.5 Link Between Optimal STDP and Expectation-Maximization. EM is a powerful and widely used algorithm for optimizing generative models and extracting hidden causes from high-dimensional input data. In the operating microcircuit model, an online version of this algorithm can be identified. As we showed above, each time a readout neuron fires, a sample from the posterior distribution is drawn. Together with the current input pattern, this sample forms a pair (**x**, *k*). Theoretically optimal STDP, equation 2.3, or more generally any rule from the family 2.9, then increases the log-likelihood  $p(\mathbf{x}, k | \mathbf{W})$  of this pair in the model. The main difference from standard EM is the fact that STDP can access only temporally local information (at least the simple model considered here). Hence, just as in online EM (Sato, 1999), STDP does not maximize the likelihood over the whole data set in each step, but rather makes a small update in the right direction after each postsynaptic spike. To show this, consider the expected STDP update  $\overline{\Delta w_{ki}}$  for a synapse if the input distribution is  $p^*(\mathbf{x})$ :

$$\overline{\Delta w_{kj}} = \eta f(w_{kj}) \left\langle p(k|\mathbf{x}, \mathbf{W}) (x_j e^{-w_{kj}} - \alpha^{-1}) \right\rangle_{p^*(\mathbf{x})}, \tag{A.20}$$

where  $\eta$  is some small learning rate. At the same time, the log-likelihood gradient with respect to that synaptic weight is

$$\partial_{w_{ki}} \langle \log p(\mathbf{x}|\mathbf{W}) \rangle_{p^*(\mathbf{x})} = e^{w_{kj}} \langle p(k|\mathbf{x},\mathbf{W}) (x_j e^{-w_{kj}} - \alpha^{-1}) \rangle_{p^*(\mathbf{x})}.$$
(A.21)

Comparing equations A.20 and A.21, one can see that the expected update of STDP drives the weights toward a skewed version of the true gradient. Importantly, the dot product between update direction and gradient is strictly nonnegative:

$$\sum_{k,j} \overline{\Delta w_{kj}} \cdot \partial_{w_{kj}} \langle \log p(\mathbf{x}|\mathbf{W}) \rangle_{p^*(\mathbf{x})} \ge 0.$$
(A.22)

This means that the average STDP update will never decrease the performance of the generative model. In fact, one can easily verify that it will always increase performance unless the weight setting already constitutes a local optimum of the log likelihood (see also section A.4). Altogether this implies that the average effect of STDP can be described as an attractor dynamics in the weight space in which the attractors are the equilibrium points of STDP and at the same time the local optima of the generative model. Stochastic deviations from these dynamics are zero mean and can thus be suppressed to arbitrary precision via the learning rate  $\eta$  (at the cost of convergence speed).

**A.6 Computer Simulations.** To abbreviate explanations, computer experiments will be referred to by numbers: E1 (Figure 2), E2 (Figure 2), E3 (Figure 4), E4 (Figure 5), and E5 (Figure 6). The simulation time step throughout the experiments was  $\Delta t = 2.5$  ms. The external stimulus  $\theta(t)$  was generated by a random noise process in which the value of  $\theta$  changed every 100 ms, according to

$$\theta(t) = (\theta(t - \Delta t) + \mathcal{N}(0, 1)) \operatorname{mod} 2\pi, \qquad (A.23)$$

where we used the modulo operation on real numbers to enforce periodic boundary conditions on  $\theta$ .

Each sensory neuron  $j \in \{1, ..., M\}$  was assigned a tuning function  $f_j(\theta)$ , determining the firing rate of neuron j in response to stimulus  $\theta$ . For experiments E1, E2, E4, and E5, the tuning functions were given by

$$f_i(\theta) = c \cdot \exp(k(\cos(\theta - \Theta_i) - 1)), \tag{A.24}$$

where each sensory neuron was associated with a different preferred stimulus  $\Theta_j$ , equally spaced over  $[0, 2\pi]$  across the population. The concentration factor *k* regulates the sharpness of the tuning function and was chosen k = 1for E1, E2, E4, and E5. In all experiments, the scaling constant was set to c = 40, yielding a maximal firing rate of 40 Hertz for the preferred stimulus.

The final tuning functions for groups G1 and G2 in E3 were set up with concentration parameters k = 0.5 and k = 3, respectively. G3 developed a

bimodal tuning function,

$$f_j(\theta) = c \cdot \exp(k(\cos(\theta - \Theta_j) - 1)) + c \cdot \exp(k(\cos(\theta - \Theta_j - 2) - 1)),$$
(A.25)

with k = 6. The transition between initial and final tuning functions in E3 was done by linear interpolation of the tuning functions.

Input spike trains were generated as inhomogeneous Poisson processes with instantaneous rates  $f_j(\theta(t))$ . The resulting spike trains were then filtered by the shape of a double-exponential EPSP,

$$\epsilon(t) = D \cdot (\exp(-t/\tau_1) - \exp(-t/\tau_2)), \tag{A.26}$$

to yield the filtered unweighted spike trains  $x_j(t)$ . The constant *D* was chosen such that  $x_j(t)$  has equal mean and variance for a fixed firing rate of the sensory neuron. The EPSP time constants were  $\tau_1 = 20$  ms and  $\tau_2 = 2$  ms.

Synaptic weights  $w_{kj}$  were initialized by drawing independently from a normal distribution with mean  $\mu_w$  and standard deviation  $\sigma_w = 0.1$ , where we used  $\mu_w = 2$  for E1, E2, E4, and E5, and  $\mu_w = 1.81$  for E3. The filtered spike trains  $x_j(t)$  were then used as inputs to the readout neurons. The WTA circuit was implemented according to a discrete time approximation of equation A.11; at each discrete time step, the firing probability for each neuron *k* is given by  $\Delta t \cdot \rho_k(t)$ . A spike occurs for neuron *k* if a draw from a Bernoulli distribution with this probability is successful. In all simulations, the total firing rate  $\rho_{\text{total}}$  of the WTA circuit was chosen beforehand to achieve an average firing rate of 3 Hertz per readout neuron (in E4,  $\rho_{\text{total}}$  was computed for N = 15 and kept constant throughout the simulation). Whenever a readout neuron spiked, STDP was applied according to equation 2.3. The learning rate was set to  $\eta = 0.001$  in E1, E3, and E4, and to  $\eta = 0.0005$  for E2. For E5, the learning rate  $\eta(t)$  depended on  $\theta(t)$ , as indicated in Figure 6.

At regular intervals, the preferred stimuli  $\psi_k(t)$  of readout neurons were computed. This was done in a separate offline simulation, in which we swept over all input stimuli and selected for each readout neuron k the stimulus  $\theta$  that elicited the greatest average response from that neuron. As a performance indicator, a reconstruction of the input signal was computed from the output spikes (shown in Figure 2E). First, we computed a raw reconstruction signal  $\hat{\theta}_{rec}(t)$ , which jumps to the preferred stimulus  $\psi_k$  of a readout neuron k whenever  $z_k$  emits a spike. From this, we obtained the reconstruction  $\theta_{rec}(t)$  as the population vector average of the raw reconstruction values  $\hat{\theta}_{rec}(t)$  within a [t - 20 ms, t] window. After training in E1, E2, E3, and E5, readout neurons were sorted for visualization according to their preferred stimuli at the end of the simulation. In E4, sorting was done individually for each visualization point, since different neurons were involved at each time.

## Acknowledgments \_\_\_\_

This article was written under partial support by the European Union: projects FP7-269921 (BrainScaleS), FP7-216593 (SECO), FP7-216886 (PAS-CAL2), and FP7-243914 (Brain-i-Nets).

### References \_

- Abbott, L. F., & Nelson, S. B. (2000). Synaptic plasticity: Taming the beast. Nature Neuroscience, 3, 1178–1183.
- Bejjanki, V., Beck, J., Lu, Z., & Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. *Nature Neuroscience*, 5, 642–648.
- Bi, G., & Poo, M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience*, *18*(24), 10464–10472.
- Bishop, C. M. (2006). Pattern recognition and machine learning. New York: Springer.
- Buonomano, D., & Merzenich, M. (1998). Cortical plasticity: From synapses to maps. Annual Reviews of Neuroscience, 21(1), 149–186.
- Caporale, N., & Dan, Y. (2008). Spike timing-dependent plasticity: A Hebbian learning rule. Annual Reviews of Neuroscience, 31, 25–46.
- Cappé, O., & Moulines, E. (2009). Online EM algorithm for latent data models. *Journal* of the Royal Statistical Society Series B (Statistical Methodology), 71(3), 593–613.
- Carandini, M., & Ferster, D. (2000). Membrane potential and firing rate in cat primary visual cortex. *Journal of Neuroscience*, 20(1), 470–484.
- Chaisanguanthum, K., & Lisberger, S. (2011). A neurally efficient implementation of sensory population decoding. *Journal of Neuroscience*, 31(13), 4868–4877.

Dan, Y., & Poo, M.-M. (2006). Spike timing-dependent plasticity: From synapse to perception. *Physiological Reviews*, 86, 1033–1048.

- Dayan, P., Hinton, G., Neal, R., & Zemel, R. (1995). The Helmholtz machine. Neural Computation, 7(5), 889–904.
- Deneve, S. (2008). Bayesian spiking neurons II: Learning. *Neural Computation*, 20(1), 118–145.
- Deneve, S., Latham, P., & Pouget, A. (1999). Reading population codes: A neural implementation of ideal observers. *Nature Neuroscience*, 2(8), 740–745.
- Douglas, R. J., & Martin, K. A. (2004). Neuronal circuits of the neocortex. Annual Reviews of Neuroscience, 27, 419–451.
- Doya, K., Ishii, S., Pouget, A., & Rao, R. P. N. (2007). *Bayesian brain: Probabilistic approaches to neural coding*. Cambridge, MA: MIT Press.
- Ecker, A. S., Berens, P., Keliris, G. A., Bethge, M., Logothetis, N. K., & Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science*, 327, 584–587.
- Fino, E., & Yuste, R. (2011). Dense inhibitory connectivity in neocortex. Neuron, 69(6), 1188–1203.

- Fiser, J., Berkes, P., Orban, G., & Lengyel, M. (2010). Statistically optimal perception and learning: From behavior to neural representation. *Trends in Cognitive Sciences*, 14(3), 119–130.
- Freedman, D., Riesenhuber, M., Poggio, T., & Miller, E. (2006). Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cerebral Cortex*, 16(11), 1631–1644.
- Gerstner, W., & Kistler, W. (2002). *Spiking neuron models*. Cambridge: Cambridge University Press.
- Gilbert, C. D., Li, W., & Piech, V. (2009). Perceptual learning and adult cortical plasticity. *Journal of Physiology*, 387, 2743–2751.
- Gilson, M., Burkitt, A., & van Hemmen, J. (2010). STDP in recurrent neuronal networks. Frontiers in Computational Neuroscience, 4(23).
- Goel, A., & Lee, H. (2007). Persistence of experience-induced homeostatic synaptic plasticity through adulthood in superficial layers of mouse visual cortex. *Journal* of Neuroscience, 27(25), 6692–6700.
- Graf, A., Kohn, A., Jazayeri, M., & Movshon, J. (2011). Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature Neuroscience*, 14(2), 239–245.
- Griffiths, T., & Tenenbaum, J. (2006). Optimal predictions in everyday cognition. *Psychological Science*, 17(9), 767–773.
- Huang, X., & Lisberger, S. (2009). Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smoothpursuit eye movements. *Journal of Neurophysiology*, 101(6), 3012–3030.
- Hubel, D., & Wiesel, T. (1963). Shape and arrangement of columns in cat's striate cortex. *Journal of Physiology*, 165, 559–568.
- Jacob, V., Brasier, D., Erchova, I., Feldman, D., & Shulz, D. E. (2007). Spike timingdependent synaptic depression in the in vivo barrel cortex of the rat. *Journal of Neuroscience*, 27(6), 1271–1284.
- Jazayeri, M., & Movshon, J. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, 9(5), 690–696.
- Jolivet, R., Rauch, A., Lüscher, H.-R., & Gerstner, W. (2006). Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *Journal of Computational Neuroscience*, 21, 35–49.
- Keck, C., Savin, C., & Lücke, J. (2012). Feedforward inhibition and synaptic scaling two sides of the same coin? *PLoS Computational Biology*, 8(3), e1002432.
- Law, C., & Gold, J. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature Neuroscience*, 12, 655–663.
- Liao, D., Jones, A., & Malinow, R. (1992). Direct measurement of quantal changes underlying long-term potentiation in CA1 hippocampus. *Neuron*, 9(6), 1089–1097.
- Louie, K., Grattan, L., & Glimcher, P. (2011). Reward value-based gain control: Divisive normalization in parietal cortex. *Journal of Neuroscience*, 31(29), 10627–10639.
- Miller, L., & Recanzone, G. (2009). Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. *Proceedings of the National Academy of Sciences*, 106(14), 5931–5935.
- Montgomery, J., Pavlidis, P., & Madison, D. (2001). Pair recordings reveal all-silent synaptic connections and the postsynaptic expression of long-term potentiation. *Neuron*, 29(3), 691–701.

- Morrison, J., & Hof, P. (2007). Life and death of neurons in the aging cerebral cortex. International Review of Neurobiology, 81, 41–57.
- Nessler, B., Pfeiffer, M., & Maass, W. (2010). STDP enables spiking neurons to detect hidden causes of their inputs. In Y. Bengio, D. Schuurmans, J. Lafferty, C.K.I. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems*, 22 (pp. 1357–1365). Cambridge, MA: MIT Press.
- Ohshiro, T., Angelaki, D., & DeAngelis, G. (2011). A normalization model of multisensory integration. *Nature Neuroscience*, 14, 775–782.
- Petersen, R., Panzeri, S., & Diamond, M. (2002). Population coding in somatosensory cortex. Current Opinion in Neurobiology, 12(4), 441–447.
- Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nature Reviews Neuroscience*, 1(2), 125–132.
- Priebe, N., Mechler, F., Carandini, M., & Ferster, D. (2004). The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nature Neuroscience*, 7, 1113–1122.
- Rao, R., & Ballard, D. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87.
- Sato, M. (1999). Fast learning of on-line EM algorithm (Tech. Rep.). Kyoto, Japan: ATR Human Information Processing Research Laboratories.
- Seitz, A., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron*, 61(5), 700–707.
- Seitz, A., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, 9(7), 329–334.
- Shinomoto, S., Kim, H., Shimokawa, T., Matsuno, N., Funahashi, S., Shima, K., et al. (2009). Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Computational Biology*, 5(7), e1000433.
- Simoncelli, E., & Heeger, D. (1998). A model of neuronal responses in visual area MT. Vision Research, 38(5), 743–761.
- Sjöström, P., Turrigiano, G., & Nelson, S. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32(6), 1149–1164.
- Smith, M., & Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. *Journal of Neuroscience*, 28(48), 12591–12603.
- Trachtenberg, J., Chen, B., Knott, G., Feng, G., Sanes, J., Welker, E., et al. (2002). Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. *Nature*, 420(6917), 788–794.
- Turrigiano, G. (2010). The self-tuning neuron: Synaptic scaling of excitatory synapses. *Cell*, 135(3), 422–435.
- Van Hooser, S., Heimel, J., Chung, S., Nelson, S., & Toth, L. (2005). Orientation selectivity without orientation maps in visual cortex of a highly visual mammal. *Journal of Neuroscience*, 25(1), 19–28.
- Zoccolan, D., Cox, D., & DiCarlo, J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *Journal of Neuroscience*, 25(36), 8150–8164.

Received June 5, 2012; accepted December 23, 2012.