



## The Cerebral Cortex: An Evolutionary Breakthrough

Wolf Singer

Ernst Strüngmann Institute for Neuroscience in Cooperation with Max Planck Society; Max Planck Institute for Brain Research; Frankfurt Institute for Advanced Studies, Frankfurt am Main, Germany

### The evaluation and encoding of perceptual relations

Living systems have to establish models of the world in which they evolve in order to be able to predict the outcome of actions and to thereby assure survival and reproduction. Establishing a good model of the world requires the detection of relevant and consistent relations between features of the environment and the efficient storage of these relations (rules). The simplest solution, found in virtually all neuronal systems, are relation encoding feed-forward circuits. Neurons tuned to respond to particular features of the environment converge on common target cells and these respond selectively to particular conjunctions of features provided that the gain of the input connections to these conjunction specific neurons are appropriately adjusted (Barlow, 1972). A particular relation among features gets encoded in the discharge rate of a neuron responding selectively to this relation. Because this neuron encodes always the same relation, one talks about a “labelled line code”. In order to evaluate and encode combinations of relations (relations of higher order) this process of input recombination and gain adjustment is iterated across successive layers. This basic principle for the evaluation, encoding and classification of relational constructs has been implemented in numerous versions of artificial neuronal networks (Rosenblatt, 1958; Hopfield, 1987; DiCarlo and Cox, 2007; LeCun et al., 2015). The highly successful recent developments in the field of “deep learning” (LeCun et al., 2015), capitalize on the iteration of this principle in large multilayer architectures. As far as feed-forward connections are concerned, these artificial multilayer systems resemble the organization of sensory systems in the brain. Marked differences exist, however, with respect to other essential features. Feedback or recurrent lateral connections are prominent in brains (Markov et al., 2014; Bastos et al., 2015) but implementation of these architectural features is still rare in artificial systems. Moreover, the training mechanism used in technical systems for the supervised adjustment of the synaptic gain of connections, the so called “back-propagation algorithm” is biologically implausible and differs from both unsupervised and supervised learning mechanisms implemented in brains (Feldman, 2012; Singer, 2016).

A complementary way to detect and encode relations between signals is to evaluate temporal contingencies: If event A consistently precedes event B, event A is likely to be the cause of B, if A and B often coincide the two events most likely have a common cause and if A and B are uncorrelated in time they are most likely unrelated.

The learning rules implemented in nervous systems are adapted to evaluate such temporal relations and to translate them into lasting changes of coupling. Both the traditional Hebbian rules (Hebb, 1949) and the more recently discovered mechanisms (Stiefel et al., 2005; Holthoff et al., 2006; Carvalho and Buonomano, 2011; Grienberger et al., 2015) evaluate temporal relations among converging inputs as well as between pre- and postsynaptic activity (spike timing dependent plasticity – STDP), (Markram et al., 1997; Bi and Poo, 1998). The molecular mechanisms underlying these use-dependent synaptic modifications operate with a temporal precision in the millisecond range. This has two important implications: First, it implies that the precise timing of spikes in converging pathways matters in determining the occurrence and polarity of synaptic gain changes. Second, the mechanism subserving synaptic modifications not only evaluates simple covariations between pre- and postsynaptic firing rates, but also evaluates causal relations. It increases the gain of excitatory connections whose activity can be causally related to the activation of the postsynaptic neuron and it weakens connections whose activity could not have contributed to the postsynaptic response. Thus, temporal relations reflecting semantic relations among events are evaluated by time sensitive mechanisms and converted into lasting changes of the coupling strength of interacting neurons. In this way statistical contingencies among features of the sensory environment are translated into synaptic weight distributions in neuronal networks.

This time sensitivity of synaptic plasticity mechanisms has deep implications for signal processing. If the known plasticity mechanisms are used for the storage of relations in general, all relations eventually have to be expressed as temporal relations among distributed neuronal responses. Thus, for the association of responses that lack temporal structure or are offset in time by intervals longer than those bridgeable by the time constants of the molecular processes, mechanisms are required that endow neuronal responses with

temporal structure and permit bridging temporal gaps. Otherwise rather different and still unknown mechanisms of synaptic plasticity have to be postulated.

### **Mechanisms for the generation of temporally structured activity**

Results from an initially completely different line of research suggest the existence of mechanisms capable of imposing temporal structure on neuronal activity and of making perceptually related responses coherent in time. It had been discovered with multisite recordings from the visual cortex that cortical circuits have a high propensity to engage in oscillatory activity and that these intrinsically generated oscillations can become synchronized, leading to correlated firing of the synchronously oscillating neurons. Of particular importance is the fact that this temporal coordination is dynamically regulated. It is context sensitive and reflects meaningful relations among encoded features (Gray and Singer, 1989; Gray et al., 1989). One reason for the synchronization of neurons encoding features that should be bound together is that the reciprocal connections between the neurons are adaptive and undergo Hebbian modifications. As a consequence they strengthen between neurons that encode features which have a high probability of co-occurring in natural environments. As increased coupling among oscillators enhances the probability that they synchronize (Kuramoto et al., 1992), synchronization probability reflects the probability of feature contingencies. Thus, neurons encoding features that often co-occur, e.g. because they are constitutive for a particular object, have an increased likelihood to synchronize and to form a coherently active cell assembly. The saliency of their responses is enhanced jointly because synchronous discharges have a stronger impact on target neurons (Bruno and Sakmann, 2006). Thus, synchronously oscillating cells convey the message that the features they encode should be bound together because they have a high probability to be related in a meaningful way, e.g. because they are constitutive for a particular object and therefore have often co-occurred in the past (for review see Singer, 1999; Engel et al., 2001; Fries, 2009; Uhlhaas et al., 2009). Initially, the synchronization was seen as a relation defining mechanism mainly in the context of low-level visual processes such as feature binding and figure ground segregation. The reason was that synchronization probability reflected well the common Gestalt criteria for perceptual grouping and also reflected the architecture of the recurrent connections in the visual cortex that couple preferentially neurons coding for features which tend to be bound perceptually (Löwel and Singer, 1992). However, it soon turned out that synchronization of oscillatory activity is not confined to the visual system but a ubiquitous phenomenon (for review see Buzsáki et al., 2013). What makes these dynamic phenomena particularly interesting is the fact that they result from highly dynamic self-organizing processes that enable rapid reorganization of the temporal coherence of the responses of widely distributed groups of neurons. For this reason synchronization of oscillatory activity is now considered by many to serve a large number of different functions that have in common the requirement for temporal coordination of distributed neuronal responses. Examples are the enhancement of the saliency of responses (Fries et al., 1997; Biederlack et al., 2006), the dynamic formation of functional networks (Siegel et al., 2015; Fries, 2005; Deco and Kringelbach, 2016), the selection of responses by attention mechanisms (Fries et al., 2001a), the matching of top down signals with sensory input (Bastos et al., 2015), the parsing of speech signals (Ding et al., 2016) and the definition of relations in the context of learning and memory (Siapas et al., 2005; Fell et al., 2011; Yamamoto et al., 2014; for review see Singer, 2016).

### **Complex dynamics**

As more laboratories engaged in multisite recordings, a prerequisite for the analysis of the correlation structure of neuronal dynamics, it became clear that oscillations with constant frequency sustained over long time intervals and synchronization of these oscillations with stable phase relations occur only under specific stimulation conditions. Especially the high frequency oscillations in the beta and gamma frequency range were found to exhibit a much more complex and variable dynamics than reported in the early days of their discovery. In the visual cortex, the frequency of stimulus-induced oscillations increases with the energy and the complexity of the stimuli and with their motion speed (Gray et al., 1990; Kayser et al., 2003; Ray and Maunsell, 2015; Lima et al., 2011). The amplitude of stimulus-induced oscillations decreases with the complexity of the inducing stimuli and increases with attention and expectancy (Lima et al., 2011; Fries et al., 2001a). Moreover, in awake behaving animals the oscillations are usually transient, occur as brief bursts (Pipa and Munk, 2011; Lundqvist et al., 2016) and are often coupled with the phase of other oscillations that have lower frequency (cross frequency coupling, Canolty et al., 2010). Accordingly the pairwise correlations between oscillating cell populations are also highly variable. They are transient and exhibit phase shifts that vary over time (for review see Fries et al., 2001b; Maris et al., 2016).

It has been argued that this high degree of variability of oscillations and synchrony is incompatible with a functional role of these dynamic phenomena (Ray and Maunsell, 2015). This critique concerns both the initial postulate that temporal coherence serves to encode relations and the formation of assemblies in distributed coding regimes (Singer, 1999) as well as the extension of this concept known as the Communication Through Coherence (CTC) hypothesis (Fries, 2005). However, others have argued that variability and non-stationarity of

the dynamics are necessary properties for flexible processing in order to comply with the speed and versatility of cognitive operations (Roberts et al., 2013) and with the requirement to configure on the fly functional networks on the fixed backbone of the cortical connectome (Deco et al., 2016).

### **A unifying concept**

These facts and arguments urge for a novel framework that attributes specific functions to the various manifestations of cortical dynamics and provides a cohesive interpretation of both low-dimensional states characterized by sustained frequency-stable oscillations and high-dimensional states with complex and rapidly changing correlation structure. The hypothesis proposed here is that the cortex exploits the high dimensional state space provided by the non-linear dynamics of recurrently coupled networks in order to perform flexible and efficient computation. In this framework, emphasis is placed on the characteristic parameters of self-organizing complex systems with non-linear dynamics. These parameters include changes in correlation structure, the entropy and dimensionality of distributed activity, network oscillations, synchronisation phenomena and phase shifts. The proposed computational strategy is likely to account for a number of hitherto poorly understood functions: The encoding of temporal sequences, the storage of vast amounts of information about the environment in the networks of sensory cortices, the ultrafast retrieval of information in processes requiring comparison between input signals and stored knowledge and the fast and effective classification of complex spatio-temporal input patterns.

Early theories of perception (von Helmholtz, 1867) have suggested that the brain interprets sparse input signals on the basis of an internal model of the world and these early ideas have received substantial support by studies on active sensing and predictive coding. The internal model is thought to build on inherited, genetically transmitted information and on knowledge acquired by experience. The information provided by this model is used to reduce redundancy in sensory signals and to facilitate perceptual grouping, feature binding, classification and identification. Because of the daunting complexity of the visual world, the store containing such an elaborate model must have an immense capacity. Moreover, read-out must be extremely fast to comply with behavioural evidence. When primates, including humans, scan their visual environment, they change the direction of their gaze on average four times a second. Thus, the prior knowledge required for the interpretation of a particular input pattern needs to be accessible within fractions of a second. The proposal is that these conditions can only be met if encoding, storage and processing of information take place in the *high-dimensional state space provided by a complex system with non-linear dynamics*.

### **The hypothesis**

Neocortex, especially its supragranular compartment, is ideally suited to provide such a high-dimensional coding space. It is a recurrently coupled network (Gilbert and Wiesel, 1989; Stettler et al., 2002), whose nodes are feature selective and have a high propensity to oscillate (Gray and Singer, 1989). This network, so the assumption, provides the high-dimensional state space required for the storage of statistical priors, the fast integration with input signals and the representation of the results in a classifiable format. Statistical priors are supposed to be stored in the functional architecture of long-range horizontal connections and their synaptic weight distributions.

In this framework a number of experimentally testable predictions can be formulated. Spontaneous activity should reflect the dynamics of the structured network harbouring the entirety of latent internal priors and therefore exhibit very high dimensionality. Input signals are supposed to drive in a graded way the feature sensitive nodes *and* thereby constrain the network dynamics. If the evidence provided by the input patterns matches priors stored in the network architecture, the network dynamics will collapse to a specific substate, corresponding to a particular perceptual experience. Such a substate is expected to have a lower dimensionality than the resting activity, exhibit specific correlation structures and be metastable due to reverberation among nodes supporting the respective substate. Because these processes occur within a very high-dimensional state space, substates induced by different input patterns should be able to coexist (superposition of information), outlast the duration of the stimuli because of reverberation and be well segregated and therefore easy to classify. They can then either serve as input to the next cortical processing stage, where the same matching process is iterated, albeit with different, more global and abstract priors, - or they can be classified by local readout units that directly feed into executive centres. According to this concept every cortical area has its own model of the world and these models differ in granularity and the degree of abstraction because of the mapping rules specifying the distribution and recombination of input signals across different processing stages.

### **Experimental evidence**

Preliminary evidence is already available for some of these predictions. Developmental studies support the notion that the statistics of feature conjunctions in the outer world gets translated into cortical connectivity. Both feed forward as well as the reciprocal tangential connections in the visual cortex have been shaped during

evolution and get further refined by experience dependent pruning to match the statistical properties of visual scenes (Hubel and Wiesel, 1962; Smith et al., 2015; Pecka et al., 2014; Eysel et al., 1998; Gilbert et al., 2009) according to a Hebbian mechanism (Singer and Tretter, 1976; Rauschecker and Singer, 1981; Löwel and Singer, 1992). In agreement with the hypothesis, the covariance structure of resting activity reflects the anisotropic layout of these connections (Kenet et al., 2003; Fries et al., 2001b; Bosking et al., 1997; Löwel and Singer, 1992; Gilbert and Wiesel, 1989), is modified by learning (Lewis et al., 2009; Kundu et al., 2013) and reveals hallmarks of an internal model of the environment (Berkes et al., 2011).

Ample evidence is also available for the ability of cortical circuits to engage in oscillatory activity in a wide range of frequencies and for stimulus dependent changes of correlations mediated by intracortical connections, both being hallmarks of recurrently coupled networks (for reviews see Singer, 1999; Buzsáki et al., 2013).

However, much less is known about how the ensuing oscillatory responses depend on the particular properties of natural stimuli, both in the spatial and temporal domain, how particular Gestalt principles of grouping translate into informative neuronal dynamics and how noise or ambiguity affect the efficiency of this encoding.

There are also indications that both sensory stimulation and top-down mechanisms related to attention induce changes in the dimensionality of states, because they can enhance synchronized oscillatory activity in distinct frequency bands (Gray et al., 1989; Lima et al., 2011; Churchland et al., 2010; Fries et al., 2001a). However, no direct analysis of dimensionality changes were performed in these studies. In Lima et al., 2011, the “attended” stimulus evoked gamma band oscillations of much higher amplitude than the “non-attended” stimulus. Thus, the expectancy of having to respond to a particular stimulus changed the correlation structure of the activity induced by this stimulus towards enhanced coherence. In other terms, anticipatory top down signals constrained the dynamics of an early visual area – most likely leading to a reduction of dimensionality.

Evidence is also available that cortical dynamics exhibit a fading memory for recent inputs. This is a hallmark of recurrent networks (Buonomano and Maass, 2009; Bertschinger and Natschläger, 2004; Lukoševičius and Jaeger, 2009) that greatly facilitates encoding and classification of sequences. As demonstrated in Nikolic et al. (2009) information about a briefly presented stimulus could persist for up to one second in the distributed responses of cortical neurons, could superimpose with information about subsequent stimuli and remain classifiable with high fidelity. We presented sequences of visual stimuli (letters and numbers) while recording from a large number of neurons in the visual cortex and trained a linear classifier on short segments (5-100 ms) of the high dimensional vector of responses obtained from a training set and then used the same classifier to identify the nature of the presented stimuli in a test set. We found that i) the information about a particular stimulus persists in the activity of the network for up to a second after the end of the stimulus, ii) the information about sequentially presented stimuli superimposes so that two subsequent stimuli can be correctly classified some time after the end of the second stimulus and iii) the information about stimulus identity is distributed across neurons and encoded both in the discharge rate of the neurons and in the precise timing of the spikes.

Finally, we have preliminary evidence that repeated exposure to a set of images changes the response properties of populations of neurons in the primary visual cortex, such that stimulus classification improves over time: we observe changes in the dynamics of the network through the state-space that favor the segregation of responses into stimulus specific substates. Hence the network “learns” in an unsupervised way about the statistics of feature constellations in frequently presented stimuli and this leads to enhanced segregation and classifiability of substates in the high-dimensional state space (Lazar and Singer, in preparation).

## **Concluding remarks**

Despite considerable effort there is still no unifying theory of cortical processing and therefore numerous experimentally identified phenomena lack a cohesive theoretical framework. This is particularly true for the dynamic phenomena reviewed here because they cannot easily be accommodated in the prevailing concepts that emphasize serial feedforward processing and labelled line codes. Here we have proposed a concept that assigns specific functions to recurrent coupling and to features of the emerging dynamics. This concept is fully compatible with the robust evidence for labelled line codes and extends this notion by the proposal that precise temporal relations among the discharges of coupled neurons also serve as code for the definition of relational constructs both in signal processing and learning. We proposed a computational strategy that capitalizes on the high-dimensional coding space offered by reciprocally coupled networks. In this conceptual framework, information is distributed and encoded both in the discharge rate of individual nodes (labelled lines) and to a substantial degree also in the precise temporal relations among the discharge sequences of distributed nodes. The core of the hypothesis is that the dynamic interactions within recurrently coupled oscillator networks i) endow responses with the temporal structure required for the recoding of semantic relations into temporal relations, ii) exhibit complex, high dimensional correlation structures that reflect the signatures of an internal model stored in the weight distributions of the coupling connections and iii) permit fast convergence towards stimulus specific substates that are easy to classify because they occupy well segregated loci in the high-

dimensional state space. The analysis of the correlation structure and consistency of these high-dimensional response vectors is still at the very beginning. However, methods are now available for massive parallel recording from large numbers of network nodes in behaving animals. It is to be expected, therefore, that many of the predictions formulated above will be amenable to experimental testing in the near future.

Irrespective of the outcome of these tests, available evidence suggests that nature – with the evolution of the cerebral cortex – succeeded to realize an extremely powerful, scalable and versatile computational strategy, that probably differs in some crucial aspects from algorithms implemented presently in artificial devices – and is probably not yet fully understood. As the intrinsic organization of cortical modules is strikingly similar across the whole cortical mantle, this strategy must be of a very general nature and capable of serving a wide spectrum of seemingly different cognitive and executive functions. This versatility is the likely reason for the tremendous evolutionary success of this structure. Its expansion is the hallmark of the evolutionary changes that distinguish the human species from its nearest neighbours, the great apes, and the cognitive functions resulting from the addition of cortical modules ultimately enabled humans to initiate cultural evolution.

#### *Acknowledgments*

This work was supported by the Max Planck Society and the Ernst Strüngmann Institute for Neuroscience in Cooperation with Max Planck Society, the DFG grant NI 708/51 and the Frankfurt Institute for Advanced Studies.

#### **References**

- Barlow, H.B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception* 1, 371-394.
- Bastos, A.M., Vezoli, J., Bosman, C.A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J.R., De Weerd, P., Kennedy, H., and Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* 85, 390-401.
- Bi, G.Q., and Poo, M.M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18, 10464-10472.
- Biederlack, J., Castelo-Branco, M., Neuenschwander, S., Wheeler, D.W., Singer, W., and Nikolic, D. (2006). Brightness induction: Rate enhancement and neuronal synchronization as complementary codes. *Neuron* 52, 1073-1083.
- Berkes, P., Orbán, G., Lengyel, M., and Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* 331, 83-87.
- Bertschinger, N., and Natschläger, T. (2004). Real-time computation at the edge of chaos in recurrent neural networks. *Neural Comput.* 16, 1413-1436.
- Bosking, W.H., Zhang, Y., Schofield, B., and Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* 17, 2112-2127.
- Bruno, R.M., and Sakmann, B. (2006). Cortex is driven by weak but synchronously active thalamocortical synapses. *Science* 312, 1622-1627.
- Buonomano, D.V., and Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nature Rev. Neurosci.* 10, 113-125.
- Buzsáki, G., Logothetis, N., and Singer, W. (2013). Scaling brain size, keeping timing: Evolutionary preservation of brain rhythms. *Neuron* 80, 751-764.
- Canolty, R.T., Ganguly, K., Kennerley, S.W., Cadieu, C.F., Koepsell, K., Wallis, J.D., and Carmena, J.M. (2010). Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proc. Natl. Acad. Sci. USA* 107, 17356-17361.
- Carvalho, T.P., and Buonomano, D.V. (2011). A novel learning rule for long-term plasticity of short-term synaptic plasticity enhances temporal processing. *Frontiers Integr. Neurosci.* 5, 20: 1-11.
- Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., et al. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neurosci.* 13, 369-378.
- Deco, G., and Kringelbach, M.L. (2016). Metastability and coherence: Extending the communication through coherence hypothesis using a whole-brain computational perspective. *Trends Neurosci.* 39, 125-135.

- DiCarlo, J.J., and Cox, D.D. (2007). Untangling invariant object recognition. *Trends Cogn. Sci.* 11, 333-341.
- Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neurosci.* 19, 158-164.
- Engel, A.K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nature Rev. Neurosci.* 2, 704-716.
- Eysel, U.T., Eyding, D., and Schweigart, G. (1998). Repetitive optical stimulation elicits fast receptive field changes in mature visual cortex. *NeuroReport* 9, 949-954.
- Feldman, D.E. (2012). The spike-time dependence of plasticity. *Neuron* 75, 556-571.
- Fell, J., Ludowig, E., Staesina, B.P., Wagner, T., Kranz, T., Elger, C.E., and Axmacher, N. (2011). Medial temporal theta/alpha power enhancement precedes successful memory encoding: Evidence based on intracranial EEG. *J. Neurosci.* 31, 5392-5397.
- Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual Rev. Neurosci.* 32, 209-224.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474-480.
- Fries, P., Reynolds, J.H., Rorie, A.E., and Desimone, R. (2001a). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291, 1560-1563.
- Fries, P., Neuenschwander, S., Engel, A.K., Goebel, R., and Singer, W. (2001b). Rapid feature selective neuronal synchronization through correlated latency shifting. *Nature Neurosci.* 4, 194-200.
- Fries, P., Roelfsema, P.R., Engel, A.K., König, P., and Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proc. Natl. Acad. Sci. USA* 94, 12699-12704.
- Gilbert, C.D., and Wiesel, T.N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci.* 9, 2432-2442.
- Gilbert, C.D., Li, W., and Piech, V. (2009). Perceptual learning and adult cortical plasticity. *J. Physiol.* 587, 2743-2751.
- Gray, C.M., and Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA* 86, 1698-1702.
- Gray, C.M., Engel, A.K., König, P., and Singer, W. (1990). Stimulus-dependent neuronal oscillations in cat visual cortex: Receptive field properties and feature dependence. *Eur. J. Neurosci.* 2, 607-619.
- Gray, C.M., König, P., Engel, A.K., and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338, 334-337.
- Grienberger, C., Chen, X., and Konnerth, A. (2015). Dendritic function in vivo. *Trends Neurosci.* 38, 45-54.
- Hebb, D.O. (1949). *The Organization of Behavior* (New York, John Wiley & Sons).
- Holthoff, K., Kovalchuk, Y., and Konnerth, A. (2006). Dendritic spikes and activity-dependent synaptic plasticity. *Cell Tissue Res.* 326, 369-377.
- Hopfield, J.J. (1987). Learning algorithms and probability distributions in feed-forward and feed-back networks. *Proc. Natl. Acad. Sci. USA* 84, 8429-8433.
- Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol. (Lond.)* 160, 106-154.
- Kayser, C., Salazar, R.F., and König, P. (2003). Responses to natural scenes in cat V1. *J. Neurophysiol.* 90, 1910-1920.
- Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., and Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature* 425, 954-956.
- Kundu, B., Sutterer, D.W., Emrich, S.M., and Postle, B.R. (2013). Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *J. Neurosci.* 33, 8705-8715.
- Kuramoto, Y., Aoyagi, T., Nishikawa, I., Chawanya, T., and Okuda, K. (1992). Neural network model carrying phase information with application to collective dynamics. *Prog. Theor. Phys.* 87, 1119-1126.

- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436-444.
- Lewis, C.M., Baldassarre, A., Comitteri, G., Romani, G.L., and Corbetta, M. (2009). Learning sculpts the spontaneous activity of the resting human brain. *Proc. Natl. Acad. Sci. USA* 106, 17558-17563.
- Lima, B., Singer, W., and Neuenschwander, S. (2011). Gamma responses correlate with temporal expectation in monkey primary visual cortex. *J. Neurosci.* 31, 15919-15931.
- Löwel, S., and Singer, W. (1992). Selection of intrinsic horizontal connections in the visual cortex by correlated neuronal activity. *Science* 255, 209-212.
- Lukoševičius, M., and Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Computer Sci. Rev.* 3, 127-149.
- Lundqvist, M., Rose, J., Herman, P., Brincat, S.L., Buschman, T.J., and Miller, E.K. (2016). Gamma and beta bursts underlie working memory. *Neuron* 90, 152-164.
- Maris, E., Fries, P., and van Ede, F. (2016). Diverse phase relations among neuronal rhythms and their potential function. *Trends Neurosci.* 39, 86-99.
- Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marini, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D.C., and Kennedy, H. (2014). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cerebral Cortex* 24, 17-36.
- Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275, 213-215.
- Nikolic, D., Häusler, S., Singer, W., and Maass, W. (2009). Distributed fading memory for stimulus properties in the primary visual cortex. *PLoS Biol.* 7: e1000260, 1-19.
- Pecka, M., Han, Y., Sader, E., and Mrsic-Flogel, T.D. (2014). Experience-dependent specialization of receptive field surround for selective coding of natural scenes. *Neuron* 84, 457-469.
- Pipa, G., and Munk, M.H.J. (2011). Higher order spike synchrony in prefrontal cortex during visual memory. *Frontiers Comput. Neurosci.* 5: 23, 1-13.
- Rauschecker, J.P., and Singer, W. (1981). The effects of early visual experience on the cat's visual cortex and their possible explanation by Hebb synapses. *J. Physiol. (Lond.)* 310, 215-239.
- Ray, S., and Maunsell, J.H.R. (2015). Do gamma oscillations play a role in cerebral cortex? *Trends Cogn. Sci.* 19, 78-85.
- Roberts, M.J., Lowet, E., Brunet, N.M., Ter Wal, M., Tiesinga, P., Fries, P., and De Weerd, P. (2013). Robust gamma coherence between macaque V1 and V2 by dynamic frequency matching. *Neuron* 78, 523-536.
- Rosenblatt, F. (1958). The perceptron. A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* 65, 386-408.
- Siapas, A.G., Lubenov, E.V., and Wilson, M.A. (2005). Prefrontal phase locking to hippocampal theta oscillations. *Neuron* 46, 141-151.
- Siegel, M., Buschman, T.J., and Miller, E.K. (2015). Cortical information flow during flexible sensorimotor decisions. *Science* 348, 1352-1355.
- Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron* 24, 49-65.
- Singer, W. (2016). Synchronous oscillations and memory formation. In *Comprehensive Neurobiology of Learning and Memory*. Byrne, J.H., Menzel, R. (Eds.). Elsevier. New Edition (in print).
- Singer, W., and Treutner, F. (1976). Unusually large receptive fields in cats with restricted visual experience. *Exp. Brain Res.* 26, 171-184.
- Smith, G.B., Sederberg, A., Elyada, Y.M., Van Hooser, S.D., Kaschube, M., and Fitzpatrick, D. (2015). The development of cortical circuits for motion discrimination. *Nature Neurosci.* 18, 252-261.
- Stettler, D.D., Das, A., and Bennett, J.G., C.D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron* 36, 739-750.
- Stiefel, K.M., Tennigkeit, F., and Singer, W. (2005). Synaptic plasticity in the absence of backpropagating spikes of layer II inputs to layer V pyramidal cells in rat visual cortex. *Europ. J. Neurosci.* 21, 2605-2610.

Uhlhaas, P.J., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolic, D., and Singer, W. (2009). Neural synchrony in cortical networks: history, concept and current status. *Frontiers Integrat. Neurosci.* 3, 1-19.

von Helmholtz, H. (1867). *Handbuch der Physiologischen Optik* (Leipzig, Leopold Voss Verlag).

Yamamoto, J., Suh, J., Takeuchi, D., and Tonegawa, S. (2014). Successful execution of working memory linked to synchronized high-frequency gamma oscillations. *Cell* 157, 845-857.