Computational Modeling of Neural Plasticity for Self-Organization of Neural Networks

Joseph Chrol-Cannon and Yaochu Jin*

Department of Computing, University of Surrey, Guildford, GU2 7XH, United Kingdom

Abstract

Self-organization in biological nervous systems during the lifetime is known to largely occur through a process of plasticity that is dependent upon the spike-timing activity in connected neurons. In the field of computational neuroscience, much effort has been dedicated to building up computational models of neural plasticity to replicate experimental data. Most recently, increasing attention has been paid to understanding the role of neural plasticity in functional and structural neural self-organization, as well as its influence on the learning performance of neural networks for accomplishing machine learning tasks such as classification and regression. Although many ideas and hypothesis have been suggested, the relationship between the structure, dynamics and learning performance of neural networks remains elusive. The purpose of this article is to review the most important computational models for neural plasticity and discuss various ideas about neural plasticity's role. Finally, we suggest a few promising research directions, in particular those along the line that combines findings in computational neuroscience and systems biology, and their synergetic roles in understanding learning, memory and cognition, thereby bridging the gap between computational neuroscience, systems biology and computational intelligence.

Keywords: neural plasticity, neural networks, gene regulatory networks, learning, neural self-organization

1. Introduction

Understanding the principles behind the self-organization of biological nervous systems is the key to understanding cognition. Generally speaking, neural self-organization can be studied from the evolutionary and developmental perspectives. There were a number of major transitions or divergences in the evolution of nervous systems, for example, from the diffused nervous structure

*Corresponding author. Tel: +44 1483686037. Fax: +44 1483686051. Email: yaochu.jin@surrey.ac.uk.

in cnidaria to the bilaterally symmetric one in flatworm [1]. Computational models have been built up for co-evolving the development of the neural system and body plan of an animate based on primitive organisms such as hydra and flatworm and the results suggest that energy efficiency might be the most important constraint in neural self-organization [2, 3]. In addition, a strong coupling between the evolution of neural systems and body plan is also revealed [4, 5].

Meanwhile, increasing evidence has shown that adult brains undergo intensive rewiring [6], which involves neural plasticity including the strengthening or weakening of existing connections, or even formation of new synapses and elimination existing ones. Seminal studies by Merzenich and Kaas [7, 8] demonstrated that once sensory nerves are severed, the cortical maps to which they projected are subsequently reorganized to accept nerves from surrounding nerves. This topographic adaptation can only be realized through neural plasticity and indicates the experience-dependent nature of plasticity and its central role in forming the basis of continual learning.

There has been a number of trend changes in the investigation of plasticity models. Initially, the focus was to provide a stable, self-regulated formulation of Hebbian learning [9, 10, 11]. Then, a shift towards spiking neural networks had models of plasticity emerge that depended on the precise timing of spikes between connected neurons [12]. More recently, all of these models have been recognized as phenomenological approaches [13], and more biological, molecular bases are being sought [14, 15]. Also, neuro-modulators are being included in spike-timing models that add reinforcement capabilities on top of the purely associative [16, 17].

While the high level functions of neuroplasticity – learning and memory – are taken for granted, the suggested roles of plasticity in formally defined neural network models are varied and often contradictory. In some cases, simply applying models of plasticity to existing paradigms, such as reservoir computing, has yielded improved results [18, 19, 20, 21]. Other studies [22, 23, 24, 25, 20, 26] link the role of plasticity with increasing the mutual information in the signals between connected neurons. Some claim that Hebbian plasticity thus increases the correlation between neurons in a reservoir [27], while others suggest that the neural activity is decorrelated and that this is, in fact, a desirable property [28, 29]. All of this is in addition to the classically proposed purpose of Hebbian learning as associative. Of course, there could be multiple roles that plasticity has to play in actual Human learning, each emerging in certain situations. Here we do not argue for one functional role in particular, but present a number of viewpoints.

The increasingly complex and self-regulated biological models of plasticity present a qualitatively different approach to the statistical optimization methods in machine learning. However, the success of these machine learning methods, particularly the recent advances made in deep learning [30], cannot be ignored. Somehow, the new, biologically inspired findings in neuroscience must be systematically incorporated into applied methods in order to realize more advanced capabilities that it is clear many living beings possess.

This review focuses on the role of neural plasticity in dynamics, structure and

functions rather than a detailed review of research on computational modeling of neural plasticity only. Related reviews can be found of spike-timing dependent plasticity [31] and plasticity dynamics in recurrent networks [32]. Reviews of the reservoir computing paradigm [33, 34] are also relevant to much of the current practical application of computationally modeled plasticity.

The rest of the paper is organized as follows. Section 2 describes reservoir computing neural network models that have benefited from the application of neuro-plasticty, and deep neural networks that have the potential to. Section 3 outlines the early progression of formally defined and naturally inspired plasticity models. Section 4 focuses on some recent developments in plasticity that capture more details observed in more recent biological experiments. Section 5 explores the functional roles that have been suggested for plasticity models. Some important challenges for future research are raised and promising areas of potential in the field are discussed in Section 6.

2. Neural Network Models

Two recent neural network models are described in this section. In different ways, they take inspiration from neural structures observed in the mammalian cortex. However, while biologically motivated, both are also designed to work algorithmically with machine learning principles on data classification and prediction tasks. We propose in this review, that these models are prime candidates for being augmented with neural plasticity models in order to improve their performance.

2.1. Recurrent Reservoir

Reservoir computing [35, 36] is a random projection paradigm in which a randomly connected recurrent neural network transforms an input signal into temporal, dynamic activity from which states are periodically 'read-out' and used with standard regression to learn and predict supervised outputs. The reservoir computing framework is illustrated in Figure 1.

There are two main flavors of the reservoir computing paradigm.

2.1.1. Echo State Networks

Echo State Networks (ESNs) [35] consist of artificial neurons, typically with sigmoid or tanh activation functions. There are no variable time delays at the neurons connections and the reservoir state is simply taken as the population of neuron activation values within a single time-step.

2.1.2. Liquid State Machines

Liquid State Machines (LSMs) [36] consist of spiking neurons, in which an excitable membrane is modeled and produces a sequence of timed spike activations in response to input stimuli. When taking the reservoir state, the spike sequences must be converted into scalar values, typically by using a low-pass filtering method. The connections are given varying delays to incorporate fast and

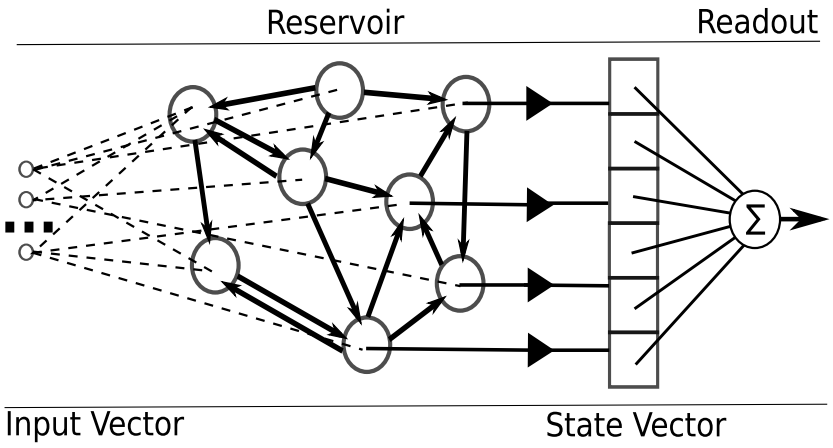

Figure 1: Structural depiction of the Reservoir Computing architecture. The ***reservoir*** consists of a randomly connected set of artificial or spiking neurons. The ***input vector*** is fully or partially connected to a typically random subset of the reservoir nodes. The ***state vector*** consists of scalar values and has the same dimension as the number of neurons in the reservoir. If the reservoir consists of spiking neurons, the spike-train activity is typically transformed to scalar values to produce the state vector. The ***readout*** is a single layer perceptron, trained using gradient descent to produce a desired output.

longer-term dynamics to the recurrent activity. Synapses often use a dynamic model of transmission between neurons, that further increases the long-term dynamics of past activity in the network.

The potential for incorporating plasticity is greater for LSMs than it is for ESNs due to the former include information in activation timings between neurons as well as the strength of activation that ESNs rely on. Also, the neural and synaptic models are far richer, with more parameters to affect plasticity in activity-dependent ways.

2.2. Deep Belief Network

A deep belief network is a probabilistic, generative model that represents its training data at multiple levels of abstraction. Sampled hidden layers can be used to infer values of input and vice versa. In this way, a generative associative memory exists that can be used to predict abstract features or input patterns. To train weights, an efficient learning algorithm is presented [30]. Figure 2 illustrates the structure and process. A contrastive divergence [37] procedure is used to learn the weights layer by layer from the input to each additional hidden layer. This learning procedure between pair-wise layers is termed a Restricted Boltzmann Machine (RBM). Then, a fine-tuning of the learned weights will adjust them in a downward pass by starting with a high level state and adjusting the bottom-up weights to maximize the likelihood of generating that state. So far, this process is entirely unsupervised. It is possible to introduce supervised learning by having target labels appended to the input of the highest hidden

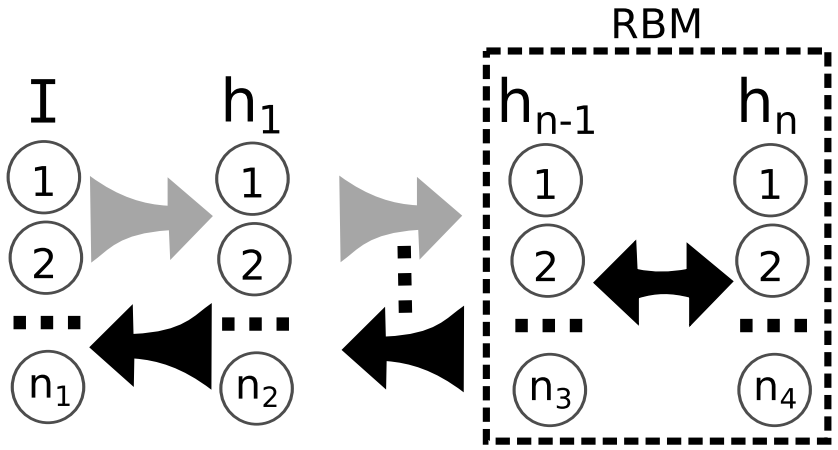

Figure 2: Deep Belief Network (DBN) illustration. The input vector, $\boldsymbol{I}$, is the lowest layer, followed by a number of hidden layers, $\boldsymbol{h}$. Training is performed bottom-up, between pair-wise layers, with the previously trained hidden layer becoming the visible input states to learn a new layer of hidden causes. The bi-directional arrow in the Restricted Boltzmann Machine (RBM) training procedure signifies supervised gradient descent that iteratively maximizes the likelihood of the hidden probability distributions to infer sampled input states.

layer. It is also possible to use the above described method to initialize a feed-forward network that can be further trained with back-propagation.

3. Early Plasticity Models

3.1. Types of Neural Plasticity

There are various types of neural plasticity that have been observed in experiments and thus also investigated in computational modeling. Generally speaking, neural plasticity can be divided into the following three types.

- *Synaptic plasticity.* Synaptic plasticity means the strengthening or weakening of the synapses that connect neurons and facilitate transmission of electro-chemical signals [38]. The change in synaptic strength can be caused by a change in the abundance of neuro-transmitter molecules at the synapse, or by a rise or fall in conduction of post-synaptic receptors. Any models that directly modify the connection strength between neurons are examples of synaptic plasticity.

- *Intrinsic Plasticity.* Intrinsic plasticity denotes modification of a neuron's intrinsic ability to generate or propagate action potentials [39]. This process is neuron wide and not synapse specific between two neurons. Intrinsic plasticity is often taken to self-regulate a neuron's activity and to be involved in a kind of homoeostatic mechanism to keep the activity within a practical range.

- *Homeostatic plasticity.* Modeling studies have demonstrated that changes in synaptic weight based on correlations between pre- and post- synaptic activity can be inherently unstable, as increases or decreases in synaptic weights cause increases or decreases in post- synaptic firing rate that subsequently generate further increases or decreases in synaptic weights in a positive feedback loop. Hence, homeostatic processes that regulate the total synaptic drive to a neuron and / or the long-term average firing rate of a neuron are critical for the stable operation of neural networks.

  Two principal homeostatic mechanisms are synaptic scaling and metaplasticity. The former refers to a uniform, cell-wide modulation of synaptic efficacies [40]; and the latter refers to a uniform, cell-wide modulation of the modification threshold for synaptic plasticity, each controlled by a long-term average of pre- or post- synaptic firing rate [41].

3.2. Hebb's Postulate

An important landmark in the basis of associative learning in the brain came from Hebb [9] in the form of a highly influential postulate. Essentially, it states that if cells are active at the same time, a connection will develop between them to facilitate further co-activation driven in a causal manner. In this way, coincident neural activity becomes associated activity and provides a mechanistic basis for the idea of associative learning that, until Hebb's postulation, had existed largely as an abstract notion.

There are numerous quotes from [9] that describe Hebb's postulate. The following quote is chosen for being succinct:

"Any two cells or systems of cells that are repeatedly active at the same time will tend to become 'associated', so that activity in one facilitates activity in the other."

This can be formulated in the following equation:

$$\Delta w_i = \eta x_i y \tag{1}$$

Δw_i, the change in synaptic strength, is increased in proportion to the product of the pre- (x_i) and post- (y) synaptic activity, multiplied by a learning rate, η. As can be seen from both Hebb's postulate and the formula, connection strength can only increase due to there being no depression term. This leads to an untenable model in practice, as the synapses will continue to increase indefinitely. If maximum bounds are put on the synaptic strength, then the weights will saturate to the maximum values, at which point no further learning can take place.

3.3. Homeostatic Regulation

In early simulations incorporating Hebbian learning, it became apparent that a mechanism to reduce synaptic weight was needed to allow stability in the adaptation. Von der Malsburg used a technique of synaptic scaling [42] to always keep the sum total of the synaptic weights connected to a post-synaptic neuron,

constant. This way, any increases in synaptic weight will be automatically balanced out with a decrease in the others. It is therefore ensured that further changes to all synapses are possible and always subject to the neural activity.

Synaptic re-normalization equation from [42]:

$$w_{ik} = w'_{ik} \cdot N \cdot \frac{w_{\text{avg}}}{w'_k} \tag{2}$$

The normalized synaptic weight from i to k, is w_{ik}. The previous, un-normalized weight is w'_{ik} and the sum total of all weights to a post neuron k is w'_k. The number of driving inputs is given as N while w_{avg} is the average weight value that all of the synapses take.

3.4. Self-Organizing Networks

Von der Malsburg [42] was the first to show that Hebb's plasticity could lead the network to functionally self-organize into stimuli sensitive 'feature detectors'. The simulations showed parallel with findings from biological experiments – neighboring cells responding to similar stimuli [43] (simple and complex cells in Hubel and Weisel's early work).

This work demonstrated that neural networks would organize according to activity dependent synaptic modification in addition to the genetically determined connectivity that was previously assumed.

Willshaw and Malsburg [44] show how a continuous topographic map between a layer of pre- and post-synaptic cells emerges with Hebbian plasticity acting laterally between the cell sheets. It is argued that the concept of neural mapping goes beyond the feature detecting neuron, and has a system-matching property where symmetries in one layer are preserved in another. Topographic maps are known to be particularly important in brain areas between the retina and visual cortex in order to preserve spatial patterns in images. For ordered maps to develop between cell sheets of varying dimensions, it was demonstrated that a marker model is needed to establish the initial synaptic contacts [45]. However, later work [46] proposes a model based only on activity dependent synaptic modifications that can form topographic maps. It uses noise induced transitions to determine the ordered connections.

Amari [47] formulates a complete mathematical description of a self-organizing neural field that treats the population of individual cells as a continuum, divided spatially into a finite set of parameters. A time step is also assumed so the framework is coarse-grained in both space and time. Amari shows analytically how pattern formation in the model has the ability to produce feature detectors for categorization and topographic maps. Stability conditions for these patterns are also derived. When it comes to the stability of self-organizing neural systems, some works have gone beyond the synaptic modification models. A morphogenetic model of synaptogenesis is presented in [48], that just considers the growth and removal of synapses. A number of free elements are quantified for each cell that allows it to form new connections or those of a different strength. The elements are re-distributed through the network and therefore so are the

connections. In this regulated algorithm, it leads to a form of homeostasis. The stability conditions required for ongoing operation are analyzed in another work [49].

3.5. Anti-Hebbian Learning

An inverted form of Hebbian adaptation – anti-Hebbian plasticity – is proposed [50] to be active in decorrelating sensory inputs in taste and vision, between laterally connected ganglion cells. This is to ensure that the output signals of these cells represent changes in the input signal in the most efficient manner.

Due to overlap in the cell sensitivities to the inputs, the initial outputs are correlated. After a period of anti-Hebbian training on the lateral connections between the cells, the variables are decorrelated. This is shown to produce a larger spread of output values for correlated inputs and therefore increases the sensitivity for small changes in the input which leads to a more efficient representation of the input variables.

The equation for anti-Hebbian adaptation takes the same form as Equation 1 above, but now with a negative learning rate.

3.6. Oja's Rule

Oja's rule [10] is a modification of the plain Hebbian interpretation that aims to address the stability problem of an exclusively potentiating mechanism. The rule is given in the following formula:

$$\Delta w_i = \eta(x_i y - y^2 w_i) \tag{3}$$

This formula is similar to Equation 1, except that it includes a depressive term, $-y^2 w_i$. The depression is proportional to both the current weight of the connection strength and to the square of the post synaptic activity. The higher the weight and the resultant post-synaptic activity is driven, the greater this depressive term will be. This constitutes a self-regulating system that balances against the purely potentiating Hebbian term.

In [10], it is shown that when the rule is applied to the incoming synapses of a single neuron, the post-synaptic response extracts the principle component of the pre-synaptic input vector. This is after the rule has been applied long enough for the synapses to converge to a stable state where the average weight change across the inputs is zero.

3.7. BCM Theory

Another regulated form of Hebbian plasticity was proposed at the same time as Oja's work. Beinenstock, Cooper and Munro [11] took a similar approach to regulating the post-synaptic activity in their model (named BCM), but this time using a sliding threshold to determine whether the weight change should be positive or negative.

The equation for the BCM weight update is as follows:

$$\Delta w_i = y(y - \theta_M)x_i - \epsilon w_i \tag{4}$$

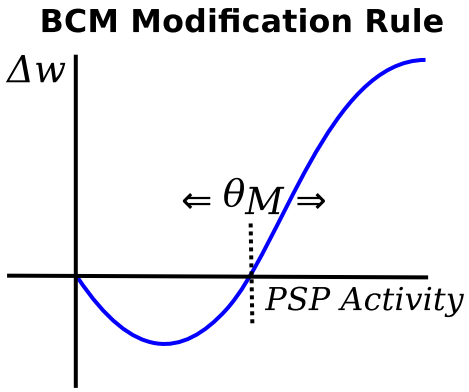

Figure 3: The Bienenstock-Cooper-Munro plasticity rule illustrated with synaptic weight change on the y-scale and post-synaptic activity on the x-scale. θ_M is the sliding modification threshold that changes based on a temporal average of post-synaptic activity.

Here, θ_M, the sliding threshold, is given as a temporal average firing rate of the post-synaptic neuron. This is given in the following formula:

$$\theta_M = E^P[(y/y_0)] \tag{5}$$

$E^P[...]$ is some function of the neural activity that constitutes a temporal averaging. y is the post-synaptic output and y_0 is a desired value that the post-synaptic output will be regulated to. The sliding threshold increases as the post-synaptic output exceeds the desired activity level. This causes the weight change to be negative, thereby providing regulation to the weight adjustment.

Like Oja's rule, the weight change is also regulated based on the current value of the connection weight. In BCM this is done in a uniform decay of all weights using a small, subtractively applied learning rate, ϵ.

In its introduction, BCM theory is claimed to explain neural selectivity observed in the receptive fields of the visual cortex. It was also claimed to provide a competitive mechanism in the context of binocular competition.

4. Recent Detailed Plasticity Models

With an increased focus on spike based neural models that model a biologically more realistic excitable membrane, new interpretations of plasticity were required to make use of the new parameters that went beyond a simple rate code. This trend, combined with increasingly detailed biological experiments, has led to a number of new directions in the development of plasticity models that we present in this section.

4.1. Spike Timing Dependent Plasticity

The previously described plasticity mechanisms all work on rate based neuron activation models, in which the activity level of a neuron is assumed to be

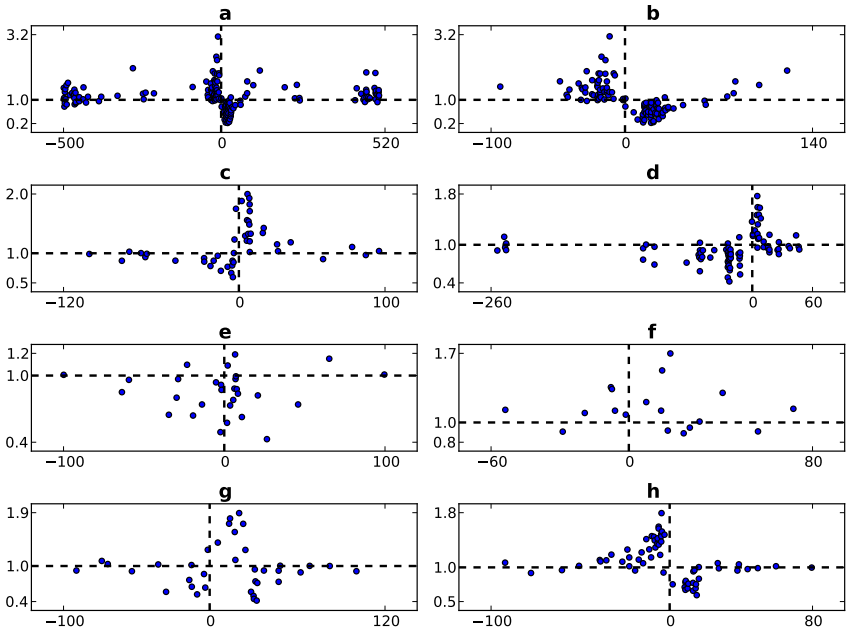

Figure 4: A collection of results from a number of STDP protocol experiments. The many, highly varying patterns that constitute the 'learning windows' shows the variety of characteristics plasticity can exhibit even when studies under the same experimental protocol. It should be noted that many of the results do come from brain slices from creatures of different species, and from different brain areas such as cortex and hippocampus. Data was extracted from plots in (a),(b) Bell 1997 [51] (c) Bi & Poo 1998 [52] (d) Feldman 2000 [54] (e),(f),(g) Wittenberg & Wang 2006 [55] (h) Zhang 1998 [53].

its instantaneous rate of generating action potentials. This means they are not directly applicable to more realistic spiking models which account for precise timing of action potentials rather than the average rate of firing.

Experimental studies in a number of works [51, 52, 53, 54, 55] have shown that the amount and direction of synaptic plasticity is dependent on both the order of pre- and post- synaptic spikes and the time delay between them. After repeated stimulation of brain cells in culture at different spike timings, the increase in post-synaptic potential is plotted against the delay between synaptic transmission and post-synaptic action potential. The resulting patterns from these experiments are reproduced in Figure 4, and constitute the observed 'learning windows' of spike timing dependent plasticity (STDP). Theoretically, these learning windows present a temporal interpretation of Hebbian learning in which causal co-activation is reinforced and anti-causal co-activation is diminished. There is no explicit self-regulation in plain STDP, yet stability can be achieved through the presence of depressive regions in the learning window.

The following subsections describe mathematical formulations of two commonly modeled forms of STDP.

4.1.1. Bi-Phasic STDP

The original formulation of STDP as a mathematical model is made in [12]. It consists of two phases: a depressive phase in which pre- follows post-synaptic spike, and a potentiating phase in which post- follows pre-synaptic spike. In both phases, the weight change decreases in magnitude as the delay between spikes increases.

The formula is as follows:

$$\Delta w_i = \begin{cases} A_+ \exp(\Delta t_i/\tau_+) & if\ \Delta t_i < 0 \\ -A_- \exp(-\Delta t_i/\tau_-) & if\ \Delta t_i > 0 \end{cases} \tag{6}$$

A_+ and A_- are the learning rates for the potentiation and depression, respectively. Δt_i is the delay of the post-synaptic spike occurring after the transmission of the pre-synaptic spike. τ_+ and τ_- control the rates of the exponential decrease in plasticity across the learning window.

4.1.2. Tri-Phasic STDP

In light of further STDP protocol experiments in the CA3-CA1 regions in the hippocampus [55], a new pattern for the learning window emerged. A tri-phasic rule is observed in this case, with a short potentiating region surrounded on either side by two depressive regions. This can be observed in a number of sub-plots in Figure 4 and is illustrated as a plotted formula in Figure 5.

One equation to describe the tri-phasic learning window is given in [56]:

$$\Delta w_i = A \left[1 - \frac{(\Delta t_i - \alpha)^2}{\alpha^2}\right] \exp\left(\frac{-|\Delta t_i - \alpha|}{\alpha}\right) \tag{7}$$

Another is given in a recent comparative study of plasticity rules [57]:

$$\Delta w_i = A_+ \exp\left(\frac{-(\Delta t_i - 15)^2}{200}\right) - A_- \exp\left(\frac{-(\Delta t_i - 15)^2}{2000}\right) \tag{8}$$

This learning window is visualized in Figure 5, with $A_+ = 0.25$ and $A_- = 0.1$. This formula simply consists of two Gaussian functions, one narrow additive curve, within another wider subtractive Gaussian of a lower magnitude. The parameters were chosen to generally match the values observed in the data in Figure 4.

4.1.3. Reward-Modulated STDP

Plain, unsupervised STDP has been questioned as a plausible mechanism for meaningful learning to occur. Legenstein [17] proposes that a reward modulated form of STDP (RM-STDP) can provide a tenable model for supervised learning while maintaining a method based on biological plasticity. Beyond the timing of pre- and post- synaptic spikes, this rule requires a third signal with the ability to consolidate useful and meaningful changes in the network based on a reward signal. It is proposed that this extra signal would be biologically realized in the concentration on a neuromodulator. This is concurred with by Izhikevich [16],

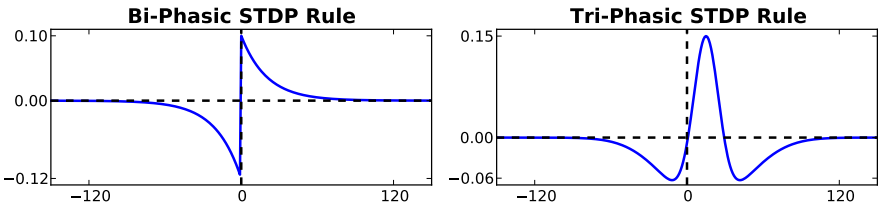

Figure 5: The two predominantly studied STDP learning windows. The bi-phasic rule is plotted from the formula in [12] and the tri-phasic rule plotted from the formula in [57]. These curves have been derived by fitting formulas to the experimental data produced from the STDP protocol. The bi-phasic rule roughly matches the observed pattern in Figure 4c, and the tri-phasic rule roughly matches Figure 4g.

in a study that demonstrates (RM-STDP) with the neuromodulator dopamine regulating the synaptic application of a globally applied reinforcement signal.

4.1.4. Reservations for Pure STDP

As a phenomenological model, STDP is particularly vulnerable to criticisms from biologically oriented studies. In particular, it has been suggested in [58] that contrary to a spike-timing model of plasticity, a post-synaptic spike is not required for a change in synaptic efficacy. The study goes further, to present experimental evidence that the backpropagating signals that are implicitly required by STDP, are neither necessary nor sufficient for plasticity *in vivo*. A follow-up work [59], reinforces this position by proposing that the protocol followed in STDP experiments is artificial. That is, the post-synaptic current injection leads to a phenomenon that is not observed when the Excitatory Post-Synaptic Potential (EPSP) causes a spike, as is the norm.

The current fixation that plasticity models have on the synapse has also recently been called into question in [60]. There, it is brought to attention that many experiments show that the intrinsic excitability of neurons change, often in accordance with the synaptic efficacy. At the extremes of potentiation/depression, the relation is observed to reverse, with a decrease/increase in excitability, respectively.

It is easy to get immersed in precise formulations of plasticity that are either derived from a class of biological experiment or inspired by a cybernetic principle of self-organization and self-regulation. Of course, there is no guarantee that the selected models for plasticity have much relevance to the processes interacting in a living brain.

In a poster report on experimental findings [61], it is stated that when under conditions of irregular, natural spiking patterns, popular models of plasticity – including STDP, voltage-dependent plasticity and a calcium controlled model – all have less influence than previously assumed. Irregular spiking reduces the level of potentiation and depression in all tested models. Furthermore, the calcium model becomes insensitive to spike correlations when there is a high average firing rate.

4.2. Voltage-Dependent Plasticity

Rather than updating a synapses weight based on the pre- and post- synaptic timing of action-potentials, a recent form of plasticity model uses the instantaneous voltage-level of the pre- and post-synaptic membrane [62]. The following equation shows this model that shares much of its form with the equation for bi-phasic STDP:

$$\Delta w_i = -A_{LTD} x_i [\bar{u} - \bar{\theta}] + A_{LTP} \bar{x}_i [u - \theta][\bar{u} - \bar{\theta}] \tag{9}$$

In this formulation; x is the pre-synaptic spike train, u is the post-synaptic voltage, and θ is a threshold to be exceeded for synaptic weight change to occur. Parameters $\bar{x}$, $\bar{u}$ and $\bar{\theta}$ are low-pass filtered values of the previous three parameters. A_{LTD} and A_{LTP} are learning rates for long-term depression and potentiation, respectively.

The notable aspect of this rule is that an action-potential is no longer required to trigger a change in synaptic efficacy, as has been observed in experiment. Under different stimulation regimes, it has been shown to fit both bi-phasic STDP and BCM learning patterns. The introductory paper used the model to explore cell selectivity and receptive field emergence that also corresponds to experimental findings.

4.3. Calcium Controlled Plasticity

STDP, as a class of models, does not consist of underlying molecular mechanisms formalized in biological terms. In fact, some of the implicit assumptions of STDP have been called into question, as discussed later in this section.

Shouval has commented that the simple viewpoint of STDP neglects that actual mechanisms that modify synapses [13]. He points out *"that synaptic plasticity is induced by a variety of receptor-generated second messengers, which in turn activate kinases, phosphatases, and other downstream targets"*. By assuming that the essential character of plasticity can be abstracted away from the biology, STDP will miss these fundamental mechanisms that may prove essential to learning and memory ability.

A new class of plasticity model that is governed by calcium concentration at the synapse is emerging as a possible underlying mechanism that is compatible with some of the empirical observations in STDP experiments. Two recent models are defined in [14] and [15]. The former has a more explicit biological grounding in that calcium concentration is regulated through kinase and phosphatase channels. The latter forgoes this detail with the benefit of having a simpler model to implement and analyse. Both models emphasize their ability in reproducing a set of commonly observed STDP learning windows. Figure 6 shows one such set produced by the model in [15].

5. Functional Role of Plasticity

5.1. Learning Input Structure and Coding

It has been demonstrated [63] that temporal information could be encoded spatially in a population of spiking neurons connected through STDP. Such a

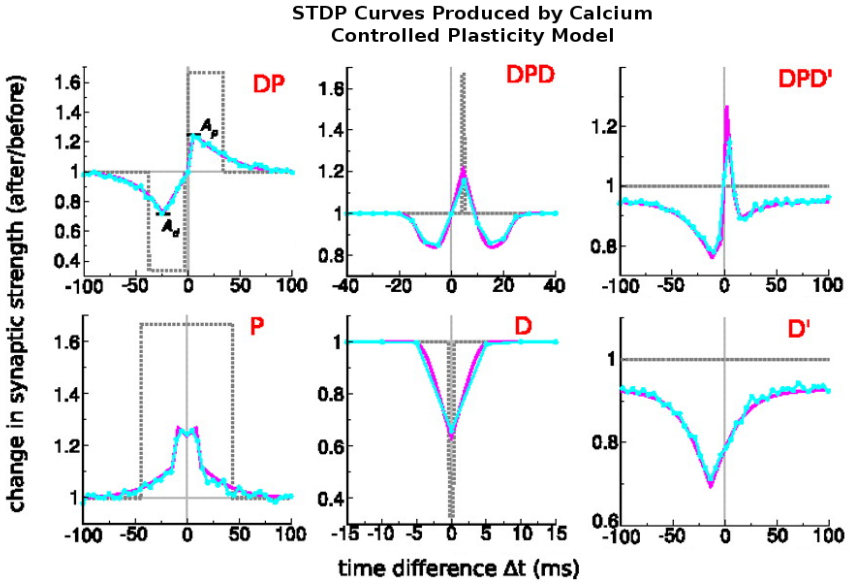

Figure 6: A variety of STDP learning windows can be predicted/explained by a unified plasticity model based on calcium concentration. If one look carefully at the experimental data presented in Figure 4, corresponding patterns can be observed for each of the learning windows. The calcium control model is [15] and these sub-plots are reproduced from there.

representation formed an auto-associative memory, where partial presentation of the temporal input signal triggered the activation of the whole spatial representation. This generalizes spatial auto-associative memories to spatio-temporal ones, with STDP taking the role of learning the temporal order of patterns through reinforcement of causal chains of activations.

Polychronous groups [64] have been used to analyze [65] recurrently connected networks trained to classify temporal spike patterns. These groups are spatio-temporal structures which were found to develop under STDP and activated in response to particular classes of input. In this way, it is evident that these groups are input specific to a certain degree and can therefore be seen as a representation of the input patterns.

The structural development of recurrent networks has been studied [62] in terms of receptive field formation. It was shown that using a voltage based model of STDP led to input specificity and structures forming that reflected the neural code. Feed-forward structure emerged temporal coded input while bi-directional structure emerged under rate coded input. This reinforces the hypothesis that the structure that develops under plasticity shares structure that is present in the input signal.

Under a number of STDP models, cell assemblies are shown to develop that can be reliably activated using either rate-based or temporal spike codes [66]. It is argued that a temporal code would be more energy efficient as it requires far

fewer spikes. Such a temporal code presented in the previously mentioned work depends on synchronized inputs in order for neurons within a cell assembly to become synchronized.

5.2. Correlate or Decorrelate Neural Activity

An interesting contradictory area in the theoretical role of plasticity is whether it is desirable to increase or decrease the correlations of spikes in a population of neurons. The standard Hebbian interpretation of associative learning would have plasticity lead to an increase in correlation for any learned associations between neurons. Work in this vein has been modeled and it was shown that STDP would select pre-synaptic activity that had higher correlation [27], thus would increase the overall level in a neural population. Conversely, it has been proposed [20, 28, 29] that decorrelation is more desirable because it would maximize the information content in the network and also improve supervised learning methods. With this aim, a form of anti-Hebbian learning is employed. It is shown in [29] that an anti-Oja rule leads to improved time-series prediction in a reservoir computing model.

5.3. Increasing Sparsity and Information Maximization

A computational model of neurons would have them perform some transformation on input patterns, such as in the case of reservoir computing mentioned above. In contrast to this view, some view neurons as signal carriers in which the maximal possible amount of information is transmitted between cells. It has been demonstrated that both BCM [22] and STDP [23] lead to maximal mutual information being retained between pre- and post-synaptic spike trains. Close to optimal transmission is shown to result from a triplet model of STDP, while the pair based model does show less improvement for information transmission [25]. A study has suggested that the purpose of STDP is to reduce the variance in a neurons response when presented with a specific spike input pattern, thus tuning selective sensitivity [24]. More recently, IP has been incorporated with a supervised learning scheme – the error-entropy minimization algorithm – as a cooperative information maximization mechanism that improves the performance on time-series prediction [26].

5.4. Improving Reservoir Computing for Learning

Reservoir computing, introduced in Section 2, does not rely on structural learning for its basic operation. However, recent studies have applied plasticity to 'shape' the reservoir and thereby improve performance in machine learning tasks such as prediction, regression and classification.

A large body of research work [19, 20, 21, 29, 67, 68] shows an improvement in the reservoirs predictive performance when plasticity is active within the reservoir in the form of an unsupervised pre-training phase. It is thought this improves the reservoir characteristics such as the fading memory of the reverberating dynamics or scaling the spectral radius of the weight matrix closer to 1, leading to an activity regime balanced on the 'edge of chaos'.

In [67], a k-winner-takes-all (kWTA) model of reservoir is shown to only improve when both IP and STDP are active together. When either plasticity rule is enabled on its own, there is a degeneration in the reservoir activity that prevents effective learning by the readout. IP alone leads to chaotic activity while STDP alone leads to time-locked spiking of initially activated neurons. These resulting patterns of activity are plotted in Figure 7. This result may be an effect unique to using the kWTA model, however.

The ESN form of reservoir computing has most notably been applied to time-series prediction data. [19, 67] are notable works that have applied IP and STDP to ESNs and that have shown improvement in regression to time-series prediction problems. NARMA and a prediction task based on Markov processes are used as the benchmark tasks. [29] applies anti-Oja plasticity to an ESN and shows an improvement in predictive performance on a time-series sun spot data set.

LSMs tend to be applied to classification of temporal sequence data, in contrast to prediction. [21, 68, 69] are works in which BCM and STDP have been used to improve results in temporal classification using LSMs. Notably, spoken digit and human behavior recognition are the applied learning tasks that consist of temporal samples in the form of sensory input. Having plasticity applied as a pre-training phase, or by having it continuously active at a slower rate than supervised learning, classification of temporal samples has been shown to improve.

6. Challenges and Potential

6.1. Relationship between Plasticity, Structure and Performance

There is a current divide in the spiking network and plasticity literature. On one side, the neuroscientific literature analyses the structural adaptation and learning of networks under the influence of plasticity [62, 71, 72, 73, 74]. In these, typically, a hierarchical network inspired by the layered structure and connectivity of the cortex is shown to develop input specific receptive fields that correspond to audio or visual stimuli. However, little/no application of these methods are made to machine learning algorithms. On the other side, some computational studies have been reported that do show that plasticity can improve applied neural network methods [21, 29, 68]. However, these works tend to have an extremely limited analysis of the input-specific structural learning that takes place within the networks. It is very unclear how or why plasticity works in these circumstances.

We have done some preliminary work on analyzing the link between plasticity, network structure and performance. After applying a number of plasticity rules to a reservoir on a speaker recognition learning task [70], the speaker specific changes to the synaptic strength were extracted as matrices of perturbed weights. These were then plotted against the average reservoir activation vector, S, and the readout weights, $Wout$, learned in supervised gradient descent. Three of the nine speaker classes are shown in Figure 8 with BCM as the plasticity rule used for this particular experiment. Each heat map is the average

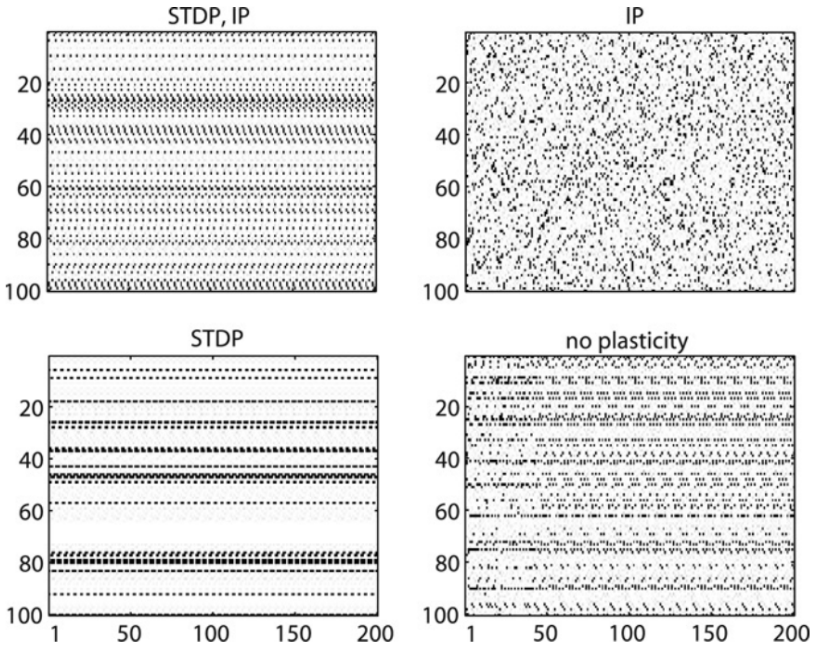

Figure 7: Spike raster plots reproduced from [67]. Each of the sub-plots show the differences in spike dynamics when Intrinsic Plasticity (IP), Spike-Timing-Dependent-Plasticity (STDP), or a combination of the two are active in a population of neurons. Only IP and STDP together lead to a spiking regime that is balanced between time-locked order and chaotic activity.

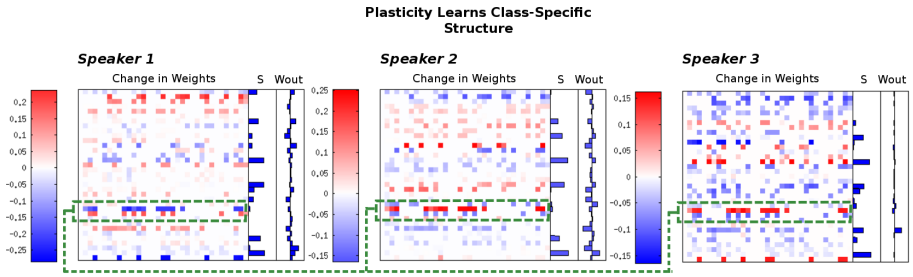

Figure 8: BCM plasticity applied to a spiking reservoir learns class-specific structure on a speaker recognition task [70]. Three out of nine total classes are shown here. The main heat maps show the change in the reservoir weight matrix after presentation of voice input data from each speaker. Blue values show a reduction in synaptic weight and red values show an increase. The bar-chart, $\boldsymbol{S}$, shows the average neuron activation for each class. The bar-chart, $\boldsymbol{Wout}$, shows the learned reservoir readout weights. The patterns highlighted in green dashes are examples of the weights being driven in opposing directions for different classes. There are many other instances that together produce class-specific weight changes over all the connections taken as a whole.

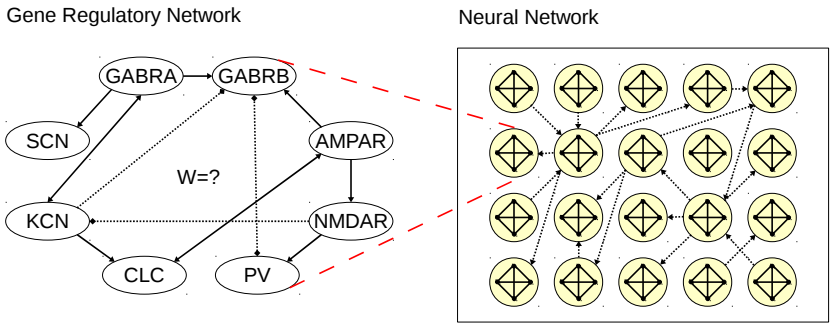

Figure 9: An illustration of a GRN regulating the structure of a spiking neural network that is optimized to produce a particular spectral output pattern. Each neuron has a GRN that determines the local connectivity of that neuron. This figure is reproduced from [75].

weight change for a specific class of input, where each point in the heat map is an increasing (red), decreasing (blue), or unchanging (white) synaptic weight. What these results show is that unsupervised plasticity does learn highly specific structural changes per class of temporal input pattern. However, by looking at the similarity of the reservoir activation vectors, this class specific structure is not retained enough for the activity of the reservoir to become as significantly sensitive to each class. This is due to the effect of catastrophic forgetting, the averaging out of these different weight directions to produce a homogeneous set that leads to minimal changes in activity for given differences in input signal. Here, plasticity has succeeded yet the reservoir computing framework is not adapted to benefit from the learned structure.

For a complete and coherent understanding of neuro-plasticity, a precise formalization of how structural adaptation contributes to specific learning tasks is required. This may be possible by applying the analysis of receptive field emergence to applied neural networks that use regression learning. However, in the process, the current network models used in machine learning tasks, like reservoirs, will have to be modified to allow the abstract, numerical feature-vector data sets to be converted into the sensory field format of input stimulus in neuroscience literature.

6.2. Systems Biology for Gene-Regulated Plasticity

Genetically driven neural development has been computationally modeled in a number of works [76, 77, 78] that demonstrate the activity-independent organization of neural networks. Of course, even after the development of the brain, the genetic factors will remain to also affect plasticity in response to activity driven from the environment.

It is becoming clear through neurobiology studies [79] that neural activity, plasticity and gene-regulatory networks are interlinked in a complex system of adaptation and regulation. A commentary [80] laments the lack of computational modeling that includes all the facets of neuronal systems biology. Some

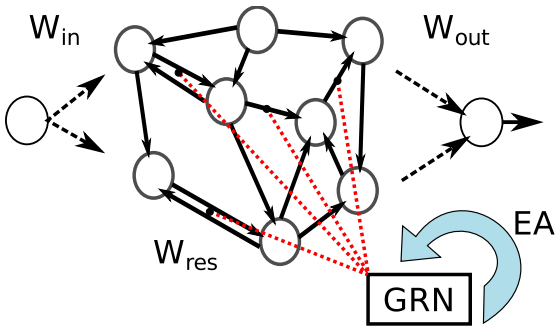

Figure 10: An illustration of a gene-regulated, evolving neural network that is recreated from [68]. A Gene Regulated Network (GRN) regulates the parameters of a plasticity rule that adapts the synapses according to the input patterns. An Evolutionary Algorithm (EA) evolves the parameters of the GRN model. There are multiple levels of regulation and structural learning that happen at different time scales in parallel.

work has been done on analyzing gene-regulated neural development [75, 81], but it has yet to be applied to learning models for data driven tasks. Figure 9 is reproduced from [75] and illustrates how a GRN can be used to regulate the structure of a neural network in order to produce particular output activity. In this case, a target spectral pattern is approximated by optimizing the GRN weights. The neuron specific GRN weights produce local dynamics that give rise to heterogeneous connectivity across the network.

Gene regulated neural plasticity in spiking neural networks have been applied to machine learning tasks. An evolutionary approach to incorporating plasticity in reservoir computing is presented in [68]. Here, a gene regulatory network (GRN) is evolved that adapts the parameters in BCM plasticity that is then applied to the reservoir. In speech and human behavior recognition tasks, this form of adaptation is shown to significantly improve classification accuracy, as well as improving regression on a time-series benchmark. The evolving process of the GRN-BCM reservoir model is illustrated in Figure 10. A similar GRN-regulated BCM rule has also been applied to adapt the weights of a feed-forward spiking neural network [82].

The models previously described in this review each consist of a relatively simplified set of formulae when compared to the overwhelming complexity found in the biological literature. Currently, calcium control theory provides one of the most comprehensive plasticity models from the systems biological perspective. In [14], calcium neuro-transmitter concentration is regulated by kinase and phosphatase channels which provides specific molecular mechanisms for a biological model of plasticity. However, there are many more ion channels, neuro-transmitter and neuro-modulating proteins that are not accounted for [83]. Furthermore, the complex and varying network of interactions between these elements is not even well understood in the biological literature.

While these approaches have the advantages of faster computability and easier comprehensibility, the diversity and variation lost from the complete ex-

perimental accounts may prove to be a weakness at the same time. The high degree of complexity and diversity in gene regulatory networks may lead to much higher tunability and robustness in the resulting epi-genetic interactions [84, 85], thus leading to unified models of neural plasticity that account for synaptic, intrinsic and homeostatic plasticity [86].

6.3. Plasticity in Deep, Semi-Supervised Learning

The current learning algorithms for deep learning, described in Section 2, share characteristics with how plasticity has been applied to reservoir computing. In both, an unsupervised iterative update of the weights are made, based on the causal inference of connected nodes. Of course, there are structural differences as well as procedural. Learning in deep networks is done by layer pairwise, while reservoirs are unstructured and randomly recurrent. Notably, contrastive divergence produces a generative set of weights, while Hebbian learning and STDP have yet to demonstrate that ability in bi-directional networks. It is not clear whether a generative model is needed for machine learning tasks, or if it is even possible in a reservoir structure. If combined, the reservoir model could provide temporal pattern learning ability to deep networks. In such a scheme, for example, the nodes of the deep network could be modeled by recurrent networks, where hierarchical and spatial features are learned through contrastive divergence and temporal features learned within the nodes through Hebbian plasticity mechanisms.

Use of plasticity rules to shape the structure of the reservoir also shares a similar philosophy in semi-supervised learning [87], where the structural information in the in features, which can be seen as sensory inputs are exploited to enhance the top-down supervised learning. The research activities in these two areas are completely separated so far, can however benefit from each other.

7. Conclusion

This article presented the prominent computational models of plasticity and their current applicability to empirical improvements as neural network adaptation mechanisms. The existing examples of plasticity in use tend to apply it to randomly connected recurrent reservoirs to learn the structural information in the inputs, achieve sparsity in neural connectivity and enhance learning performance.

We suggest that computational modeling of neural plasticity may provide a unique platform that bridges the gap between computational neuroscience, systems biology and computational intelligence. We advocate the use of gene regulatory networks in modeling neural plasticity to achieve unified plasticity models. Such models of plasticity are more likely to be self-regulated and sensitive to the network and input structure. A philosophical similarity between neural plasticity and deep learning and semi-supervised learning is pointed out. We suggest that plasticity and recurrence, when combined with deep learning or supervised learning, could yield models with benefits from each paradigm.

We hope the above suggested research will result in a gene-neuron approach to understanding cognition.

References

[1] A. Ghysen, The origin and evolution of the nervous system, International Journal of Developmental Biology 47 (2003) 555–562.

[2] Y. Jin, L. Schramm, , B. Sendhoff, A gene regulatory model for the development of primitive nervous systems, in: INNS-NNN Symposia on Modeling the Brain and Nervous Systems„ Lecture Notes in Computer Science 5506, Springer, 2009, pp. 48–55.

[3] B. Jones, Y. Jin, B. Sendhoff, X. Yao, Evolving functional symmetry in a three dimensional model of an elongated organism, in: Artificial Life IX„ 2008, pp. 305–312.

[4] B. Jones, Y. Jin, B. Sendhoff, X. Yao, Emergent distribution of computational workload in the evolution of an undulatory animat, in: The 11th International Conference on Simulation of Adaptive Behaviors (SAB' 2010, 2010, pp. 587–596.

[5] L. Schramm, Y. Jin, B. Sendhoff, Evolutionary synthesis and analysis of a gene regulatory network for dynamically stable growth and regeneration, Artificial Life 18 (4) (2012) 425–444.

[6] A. Holtmaat1, K. Svoboda, Experience-dependent structural synaptic plasticity in the mammalian brain, Nature Reviews Neuroscience 10 (2009) 647–658.

[7] M. M. Merzenich, J. H. Kaas, J. Wall, R. J. Nelson, M. Sur, D. Felleman, Topographic reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation, Neuroscience 8 (1) (1983) 33–55.

[8] M. M. Merzenich, R. J. Nelson, M. P. Stryker, M. S. Cynader, A. Schoppmann, J. M. Zook, Somatosensory cortical map changes following digit amputation in adult monkeys, J. Comp. Neurol. 224 (4) (1984) 591–605.

[9] D. Hebb, The Organization of Behavior., Wiley: New York, 1949.

[10] E. Oja, Simplified neuron model as a principal component analyzer, Journal of Mathematical Biology 15 (3) (1982) 267–273.

[11] E. L. Bienenstock, L. N. Cooper, P. W. Munro, Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex., Journal of Neuroscience 2 (1) (1982) 32–48.

[12] S. Song, K. D. Miller, L. F. Abbott, Competitive Hebbian learning through spike-timing-dependent synaptic plasticity, Nat. Neurosci. 3 (9) (2000) 919–926.

[13] H. Z. Shouval, S. S. Wang, G. M. Wittenberg, Spike timing dependent plasticity: a consequence of more fundamental learning rules, Front Comput Neurosci 1 (4) (2010) 19.

[14] D. Bush, Y. Jin, Calcium control of triphasic hippocampal stdp, Journal of Computational Neuroscience 33 (3) (2012) 495–514.

[15] M. Graupner, N. Brunel, Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location, Proceedings of the National Academy of Sciences 109 (10) (2012) 3991–3996.

[16] E. Izhikevich, Solving the distal reward problem through linkage of stdp and dopamine signaling, BMC Neuroscience 8 (Suppl 2) (2007) 1–2.

[17] R. A. Legenstein, D. Pecevski, W. Maass, A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback., PLoS Computational Biology 4 (10).

[18] J. J. Steil, Online reservoir adaptation by intrinsic plasticity for backpropagation-decorrelation and echo state learning, Neural Networks 20 (3) (2007) 353–364.

[19] B. Schrauwen, M. Wardermann, D. Verstraeten, J. J. Steil, D. Stroobandt, Improving reservoirs using intrinsic plasticity, Neurocomputing 71 (7-9) (2008) 1159–1171.

[20] P. Joshi, J. Triesch, Optimizing generic neural microcircuits through reward modulated stdp, in: C. Alippi, M. Polycarpou, C. Panayiotou, G. Ellinas (Eds.), Artificial Neural Networks ICANN 2009, Vol. 5768 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009, pp. 239–248.

[21] F. Xue, Z. Hou, X. Li, Computational capability of liquid state machines with spike-timing-dependent plasticity, Neurocomputing 122 (2013) 324–329.

[22] T. Toyoizumi, J.-P. Pfister, K. Aihara, W. Gerstner, Generalized bienenstock-cooper-munro rule for spiking neurons that maximizes information transmission, Proceedings of the National Academy of Sciences of the United States of America 102 (14) (2005) 5239–5244.

[23] T. Toyoizumi, J. P. Pfister, K. Aihara, W. Gerstner, Optimality model of unsupervised spike-timing-dependent plasticity: synaptic memory and weight distribution, Neural Comput 19 (3) (2007) 639–671.

[24] S. M. Bohte, M. C. Mozer, Reducing the Variability of Neural Responses: A Computational Theory of Spike-Timing-Dependent Plasticity, Neural Comput 19 (2007) 371–403.

[25] G. Hennequin, W. Gerstner, J.-P. Pfister, Stdp in adaptive neurons gives close-to-optimal information transmission, Frontiers in Computational Neuroscience 4 (143).

[26] Y. Li, C. Li, Synergies between intrinsic and synaptic plasticity based on information theoretic learning., PLoS ONE 8 (5).

[27] M. van Rossum, G. Turrigiano, Correlation based learning from spike timing dependent plasticity, Neurocomputing 38-40 (2001) 409–415.

[28] H. Jaeger, Reservoir riddles: suggestions for echo state network research, in: Neural Networks, 2005. IJCNN '05. Proceedings. 2005 IEEE International Joint Conference on, Vol. 3, 2005, pp. 1460–1462.

[29] S. Babinec, J. Pospichal, Improving the prediction accuracy of echo state neural networks by anti-oja's learning, in: Proceedings of the 17th international conference on Artificial neural networks, ICANN'07, Springer-Verlag, 2007, pp. 19–28.

[30] G. E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, Neural Comput. 18 (7) (2006) 1527–1554.

[31] H. Markram, W. Gerstner, P. J. Sjostrom, Spike-timing-dependent plasticity: a comprehensive overview, Frontiers in Synaptic Neuroscience 4 (2).

[32] M. Gilson, A. Burkitt, L. J. Van Hemmen, Stdp in recurrent neuronal networks, Frontiers in Computational Neuroscience 4 (23).

[33] M. Lukosevicius, H. Jaeger, Reservoir computing approaches to recurrent neural network training, Computer Science Review 3 (3) (2009) 127–149.

[34] M. Lukosevicius, H. Jaeger, B. Schrauwen, Reservoir computing trends, KI - Kuenstliche Intelligenz 26 (4) (2012) 365–371.

[35] H. Jaeger, The echo state approach to analysing and training recurrent neural networks, Tech. Rep. 148, GMD-Forschungszentrum Informationstechnik (2001).

[36] W. Maass, T. Natschlager, H. Markram, Real-time computing without stable states: A new framework for neural computation based on perturbations., Neural Computation 14 (11) (2002) 2531–2560.

[37] G. E. Hinton, Training products of experts by minimizing contrastive divergence, Neural Comput. 14 (8) (2002) 1771–1800.

[38] A. Citri, R. C. Malenka, Synaptic plasticity: Multiple forms, functions, and mechanisms, Neuropsychopharmacology 33 (1) (2007) 18–41.

[39] J. T. Brown, A. D. Randall, Activity-dependent depression of the spike after-depolarization generates long-lasting intrinsic plasticity in hippocampal CA3 pyramidal neurons, J. Physiol. (Lond.) 587 (6) (2009) 1265–1281.

[40] J. Shepherd, G. Rumbaugh, J. Wu, S. Chowdhury, N. Plath, D. Kuhl, R. Huganir, P. Worley, Arc mediates homeostatic synaptic scaling of AMPA receptors, Neuron 52 (3) (2006) 475–484.

[41] W. C. Abraham, M. F. Bear, Metaplasticity: the plasticity of synaptic plasticity, Trends in Neurosciences 19 (4) (1996) 126–130.

[42] C. Malsburg, Self-organization of orientation sensitive cells in the striate cortex, Kybernetik 14 (2) (1973) 85–100.

[43] D. H. Hubel, T. N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, J. Physiol. (Lond.) 160 (1962) 106–154.

[44] D. J. Willshaw, C. von der Malsburg, How patterned neural connections can be set up by self-organization, Proc. R. Soc. Lond., B, Biol. Sci. 194 (1117) (1976) 431–445.

[45] D. J. Willshaw, C. von der Malsburg, A marker induction mechanism for the establishment of ordered neural mappings: its application to the retinotectal problem, Philos. Trans. R. Soc. Lond., B, Biol. Sci. 287 (1021) (1979) 203–243.

[46] P. Erdi, G. Barna, Self-organizing mechanism for the formation of ordered neural mappings, Biological Cybernetics 51 (2) (1984) 93–101.

[47] S. Amari, Field theory of self-organizing neural nets, Systems, Man and Cybernetics, IEEE Transactions on SMC-13 (5) (1983) 741–748.

[48] I. Dammasch, G. Wagner, J. Wolff, Self-stabilization of neuronal networks, Biological Cybernetics 54 (4-5) (1986) 211–222.

[49] I. Dammasch, G. Wagner, J. Wolff, Self-stabilization of neuronal networks, Biological Cybernetics 58 (3) (1988) 149–158.

[50] H. Barlow, P. Foldiak, Adaptation and decorrelation in the cortex., in: I. R. Durbin, C. Miall, G. Mitchison (Eds.), The Computing Neuron, Addison-Wesley, Wokingham, England, 1989, pp. 54–72.

[51] C. C. Bell, V. Z. Han, Y. Sugawara, K. Grant, Synaptic plasticity in a cerebellum-like structure depends on temporal order, Nature 387 (6630) (1997) 278–281.

[52] G. Q. Bi, M. M. Poo, Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type, Journal of Neuroscience 18(24) (1998) 10464–72.

[53] L. I. Zhang, H. W. Tao, C. E. Holt, W. A. Harris, M.-m. Poo, A critical window for cooperation and competition among developing retinotectal synapses, Nature 395 (6697) (1998) 37–44.

[54] D. E. Feldman, Timing-based LTP and LTD at vertical inputs to layer ii/iii pyramidal cells in rat barrel cortex, Neuron 27 (1) (2000) 45–56.

[55] G. M. Wittenberg, S. S. Wang, Malleability of spike-timing-dependent plasticity at the ca3-ca1 synapse, Journal of Neuroscience 26 (2006) 6610–6617.

[56] A. Waddington, P. A. Appleby, M. De Kamps, N. Cohen, Triphasic spike-timing-dependent plasticity organizes networks to produce robust sequences of neural activity, Frontiers in Computational Neuroscience 6 (88).

[57] J. Chrol-Cannon, A. Gruning, Y. Jin, The emergence of polychronous groups under varying input patterns, plasticity rules and network connectivities., in: IJCNN, IEEE, 2012, pp. 1–6.

[58] J. Lisman, N. Spruston, Postsynaptic depolarization requirements for LTP and LTD: a critique of spike timing-dependent plasticity, Nat Neurosci 8 (7) (2005) 839–841.

[59] J. Lisman, N. Spruston, Questions about STDP as a General Model of Synaptic Plasticity, Front Synaptic Neurosci 2 (2010) 140.

[60] D. Debanne, M.-M. Poo, Spike-timing dependent plasticity beyond synapse - pre- and post-synaptic plasticity of intrinsic neuronal excitability, Frontiers in Synaptic Neuroscience 2 (21).

[61] M. Graupner, S. Ostojic, Natural firing patterns reduce sensitivity of synaptic plasticity to spike-timing, BMC Neuroscience 14 (1) (2013) 1–1.

[62] C. Clopath, L. Busing, E. Vasilaki, W. Gerstner, Connectivity reflects coding: a model of voltage-based stdp with homeostasis., Nat Neurosci 13 (3) (2010) 344–352.

[63] T. Nowotny, M. I. Rabinovich, H. D. Abarbanel, Spatial representation of temporal information through spike-timing-dependent plasticity, Phys Rev E Stat Nonlin Soft Matter Phys 68 (1) (2003) 011908.

[64] E. M. Izhikevich, Polychronization: Computation with spikes., Neural Computation 18 (2006) 245–282.

[65] H. Paugam-Moisy, R. Martinez, S. Bengio, Delay learning and polychronization for reservoir computing, Neurocomputing 71 (7–9) (2008) 1143–1158.

[66] A. Knoblauch, F. Hauser, STDP, Hebbian cell assemblies, and temporal coding by spike synchronization, BMC Neuroscience 12 (1) (2011) 1–2.

[67] A. Lazar, G. Pipa, J. Triesch, Fading memory and time series prediction in recurrent networks with different forms of plasticity., Neural Networks 20 (3) (2007) 312–322.

[68] J. Yin, Y. Meng, Y. Jin, A developmental approach to structural self-organization in reservoir computing, IEEE Transactions on on Autonomous Mental Development 4 (4) (2012) 273–289.

[69] D. Norton, D. Ventura, Preparing more effective liquid state machines using hebbian learning, in: Neural Networks, 2006. IJCNN '06. International Joint Conference on, 2006, pp. 4243–4248.

[70] M. Kudo, J. Toyama, M. Shimbo, Multidimensional curve classification using passing-through regions, Pattern Recognition Letters 20 (1999) 1103.

[71] M. Gilson, A. N. Burkitt, D. B. Grayden, D. A. Thomas, J. L. Hemmen, Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks I: Input selectivity-strengthening correlated input pathways, Biological Cybernetics 101 (2) (2009) 81–102.

[72] M. Gilson, A. N. Burkitt, D. B. Grayden, D. A. Thomas, J. L. Hemmen, Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks II: Input selectivity-symmetry breaking, Biological Cybernetics 101 (2) (2009) 103–114.

[73] M. Gilson, A. N. Burkitt, D. B. Grayden, D. A. Thomas, J. L. Hemmen, Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks III: Partially connected neurons driven by spontaneous activity, Biological Cybernetics 101 (5-6) (2009) 411–426.

[74] M. Gilson, A. N. Burkitt, D. B. Grayden, D. A. Thomas, J. L. Hemmen, Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks IV, Biological Cybernetics 101 (5-6) (2009) 427–444.

[75] L. Benuskova, N. Kasabov, Modeling brain dynamics using computational neurogenetic approach, Cogn Neurodyn 2 (4) (2008) 319–334.

[76] H. Kitano, A simple model of neurogenesis and cell differentiation based on evolutionary large-scale chaos, Artificial Life 2 (1995) 79–99.

[77] A. Van Ooyen, Modeling Neural Development, MIT Press, Cambridge, 2003.

[78] A. van Ooyen, Using theoretical models to analyse neural development, Nat. Rev. Neurosci. 12 (6) (2011) 311–326.

[79] S. W. Flavell, M. E. Greenberg, Signaling Mechanisms Linking Neuronal Activity to Gene Expression and Plasticity of the Nervous System, Annual Review of Neuroscience 31 (1) (2008) 563–590.

[80] N. Le Novere, The long journey to a systems biology of neuronal function, BMC Systems Biology 1 (1) (2007) 1–3.

[81] L. Benuskova, V. Jain, S. G. Wysoski, N. K. Kasabov, Computational neurogenetic modelling: a pathway to new discoveries in genetic neuroscience, Int J Neural Syst 16 (3) (2006) 215–226.

[82] Y. Meng, Y. Jin, J. Yin, Modeling activity-dependent plasticity for human behaviour recognition, IEEE Transactions on Neural Networks 22 (12) (2011) 1952–1966.

[83] J. Cortes-Mendoza, S. Diaz de Leon-Guerrero, G. Pedraza-Alva, L. Perez-Martinez, Shaping synaptic plasticity: the role of activity-mediated epigenetic regulation on gene transcription, Int. J. Dev. Neurosci. 31 (6) (2013) 359–369.

[84] J. Hasty, D. McMillen, F. Isaacs, J. J. Collins, Computational studies of gene regulatory networks: in numero molecular biology, Nat. Rev. Genet. 2 (4) (2001) 268–279.

[85] N. Geard, K. Willadsen, Dynamical approaches to modeling developmental gene regulatory networks, Birth Defects Res. C Embryo Today 87 (2) (2009) 131–142.

[86] D. Bush, Y. Jin, A unified computational model of the genetic regulatory networks underlying synaptic, intrinsic and homeostatic plasticity, BMC Neuroscience 12(Suppl 1):P161.

[87] O. Chapelle, B. Scholkopf, A. Zien (Eds.), Semi-supervised Learning, MIT Press, Cambridge, 2006.