

# What can we learn from synaptic weight distributions?

Boris Barbour<sup>1</sup>, Nicolas Brunel<sup>2</sup>, Vincent Hakim<sup>3</sup> and Jean-Pierre Nadal<sup>3</sup>

<sup>1</sup> Laboratoire de Neurobiologie (CNRS UMR 8544), Ecole Normale Supérieure, 46 rue d'Ulm, 75005 Paris, France

<sup>2</sup> Laboratoire de Neurophysique et Physiologie (CNRS UMR 8990), Université René Descartes, 45 rue des Saints Pères, 75006 Paris, France

<sup>3</sup> Laboratoire de Physique Statistique (CNRS UMR 8550), Ecole Normale Supérieure, 24 rue Lhomond, 75005 Paris, France

**Much research effort into synaptic plasticity has been motivated by the idea that modifications of synaptic weights (or strengths or efficacies) underlie learning and memory. Here, we examine the possibility of exploiting the statistics of experimentally measured synaptic weights to deduce information about the learning process. Analysing distributions of synaptic weights requires a theoretical framework to interpret the experimental measurements, but the results can be unexpectedly powerful, yielding strong constraints on possible learning theories as well as information that is difficult to obtain by other means, such as the information storage capacity of a cell. We review the available experimental and theoretical techniques as well as important open issues.**

## Introduction

Synaptic plasticity has been the subject of intense research since it was hypothesised to play a major role in learning and memory [1] and was demonstrated by Bliss and Lomo [2]. There have been significant efforts to link particular forms of synaptic plasticity to behavioural learning tasks [3–5]. Because synaptic plasticity alters synaptic weights within the framework of particular learning tasks, we would expect the resulting weights to depend causally upon both the learning rule and the task. Synaptic weights should, therefore, be a source of information about plasticity and learning. Ideally, one would be able to monitor the weights of multiple synapses before, during, and after learning, but such measurements are currently beyond the state of the art. However, some information about synaptic weights can be gathered using today's methodology. In particular, statistical information about the distribution of synaptic weights is available for several connection types.

Here, we review what can be deduced from measurements of synaptic weight distributions. First, we shall describe measured distributions at different types of synapses in various brain areas. We emphasise two striking observations: (i) the strong similarity of these distributions, and (ii) the small fraction of detectable synapses, compared to the expectation from the geometry of axons and dendritic arbors. We shall then examine two types of theoretical framework. The first approach is to compare the experimental data to the distribution produced by a

specific synaptic plasticity rule, given some statistical assumptions about pre- and postsynaptic activity. A second approach is to link synaptic weights directly to the learning task, bypassing the particular learning rule. Surprisingly, this is, in some cases, possible, but a key condition is that the learning has been optimised in some way. This latter approach appears to fit the data remarkably well and has the ability to access high-level information (such as information storage capacity) that is very difficult to obtain experimentally. We end by highlighting several open issues and prospects for future progress.

## Experimental measurements of synaptic weight distributions

Available measurements of excitatory weights (see Box 1, Figure 1) have been obtained from somatic recordings. We shall therefore define synaptic weight as the peak depolarisation a synapse produces at the soma. All of the distributions of synaptic weights shown in Box 1 have very similar shapes — a monotonic decay from a peak very close to zero weight. Another similarity between all of the synapse types is the apparent probability of connection of around 10%. For the remaining 90% of apparently unconnected cell pairs, it is, in general, unclear what fraction might be connected by synapses with weights that fall below the detection threshold. Because the numbers of these synapses will be of particular theoretical interest, we shall review the evidence quantifying these “silent” synapses, which we define as lacking a detectable AMPA-receptor-mediated component (NMDA or mGluR receptors might be present).

In juvenile hippocampus, “silent” synapses have been reported between Schaffer collaterals and CA1 pyramidal cells. These synapses contain NMDA receptors, but few, if any, AMPA receptors [6,7]. The numbers of these synapses are unknown; however, in immunochemical experiments, ~20% of Schaffer collateral synapses in CA1 had no detectable AMPA receptors in the adult [8]. Moreover, a much larger fraction of synapses displayed very weak labelling, indicating that many synapses might have relatively weak responses. As for all of the connections presented in Box 1, the study of Sayer *et al.* [9] found that ~10% of recordings of CA3-CA1 pyramidal cell pairs yielded detectable responses, which would certainly leave room for many undetected synapses.

At the cerebellar granule cell-Purkinje cell synapse, a large discrepancy between the connection rate predicted on

Corresponding author: Barbour, B. (barbour@ens.fr).

Available online 5 November 2007.

### Box 1. Experimentally determined weight distributions

The majority of published synaptic weight distributions concern recurrent connections between pyramidal cells. One of the first such studies involved the recurrent connections between layer 2/3 pyramidal cells [56]. Typical pre- and postsynaptic responses from that work are shown in the inset of Figure 1a. The distribution obtained is shown in Figure 1a with those from two more recent studies [57,58]. The overall shape is one of a monotonic decay from a peak (mode) near zero. Similar studies of the recurrent connections between layer 5 pyramidal cells [38,59–61] have revealed a remarkably similar distribution (Figure 1b). By far the largest data set is that of Sjöström *et al.*, which contains 1004 connections. The mean response amplitude of these connections was  $0.85 \pm 0.93$  mV (mean  $\pm$  SD), whereas the probability of connection was 11%.

Distributions obtained in other structures also have very similar shapes, but different amplitude scales. In hippocampus, the connection between CA3 and CA1 pyramidal cells (Figure 1c) displayed a mean of  $131 \mu\text{V}$  and a connection probability of 6% [9]. In cerebellum, the granule cell–Purkinje cell synapse [11,12] exhibited a mean amplitude of  $72 \pm 63 \mu\text{V}$  and a connection probability of 7% (Figure 1d).

Weight distributions are also available for other synapse types, including synapses involving neocortical interneurons [57], hippocampal interneurons [62], and layer 4 spiny stellate cells in the neocortex [63]. Remarkably, these distributions all have similar shapes to those presented above, with one significant difference—pyramidal cell–interneuron connections (in both directions) had much higher connection probabilities (>50%) [57].

anatomical grounds [10] and that observed experimentally [11] strongly suggests the existence of a large majority of undetectable/silent synapses [11,12]. This notion is supported by the very restricted simple-spike receptive fields of Purkinje cells *in vivo* [13] and their huge expansion after the induction of synaptic plasticity [14].

Similar anatomical arguments for a high likelihood of a connection have been made for recurrent pyramidal-cell connections in neocortex [15]. The current consensus, however, is that connectivity between neocortical pyramidal cells is highly diluted, with close axon-dendrite approaches forming a large pool of “latent” or “potential” synapses [16] in dynamic equilibrium with genuine synaptic connections [17–20]. Reconstructions have shown [18,21,22] that a pyramidal neurone axon shares potential synapses with the dendrites of all neighbouring neurones. However, detectable connections were found only for a small fraction of pairs of pyramidal cells (see Box 1) and appeared to be correlated with the existence of synaptic boutons [18].

To summarise: there is evidence in cortex, hippocampus, and cerebellum suggesting the existence of a large fraction of potential or silent synapses. Furthermore, detectable synaptic weights in these structures have strikingly similar distribution shapes (but different scales). We now consider the theoretical frameworks that can be used to interpret these distributions.

### Theories of synaptic weight distributions

#### Linking synaptic plasticity and weight distributions

There have been several attempts to link weight distributions and synaptic plasticity rules; spike-timing-dependent plasticity (STDP) rules [23] have received the most attention. Besides the choice of a precise plasticity rule, determining synaptic weights also requires knowledge of the activity of the pre- and postsynaptic

neurones. This is usually supplied in a statistical way, and Poisson spike trains are a frequent simplifying assumption. Correlations between pre- and postsynaptic activity obviously play an important role. This correlation can result from the explicit dynamics of the chosen (model) neurone subjected to presynaptic activity at its many synapses [24] or be imposed ad hoc to assess its effect on the weight distribution [25].

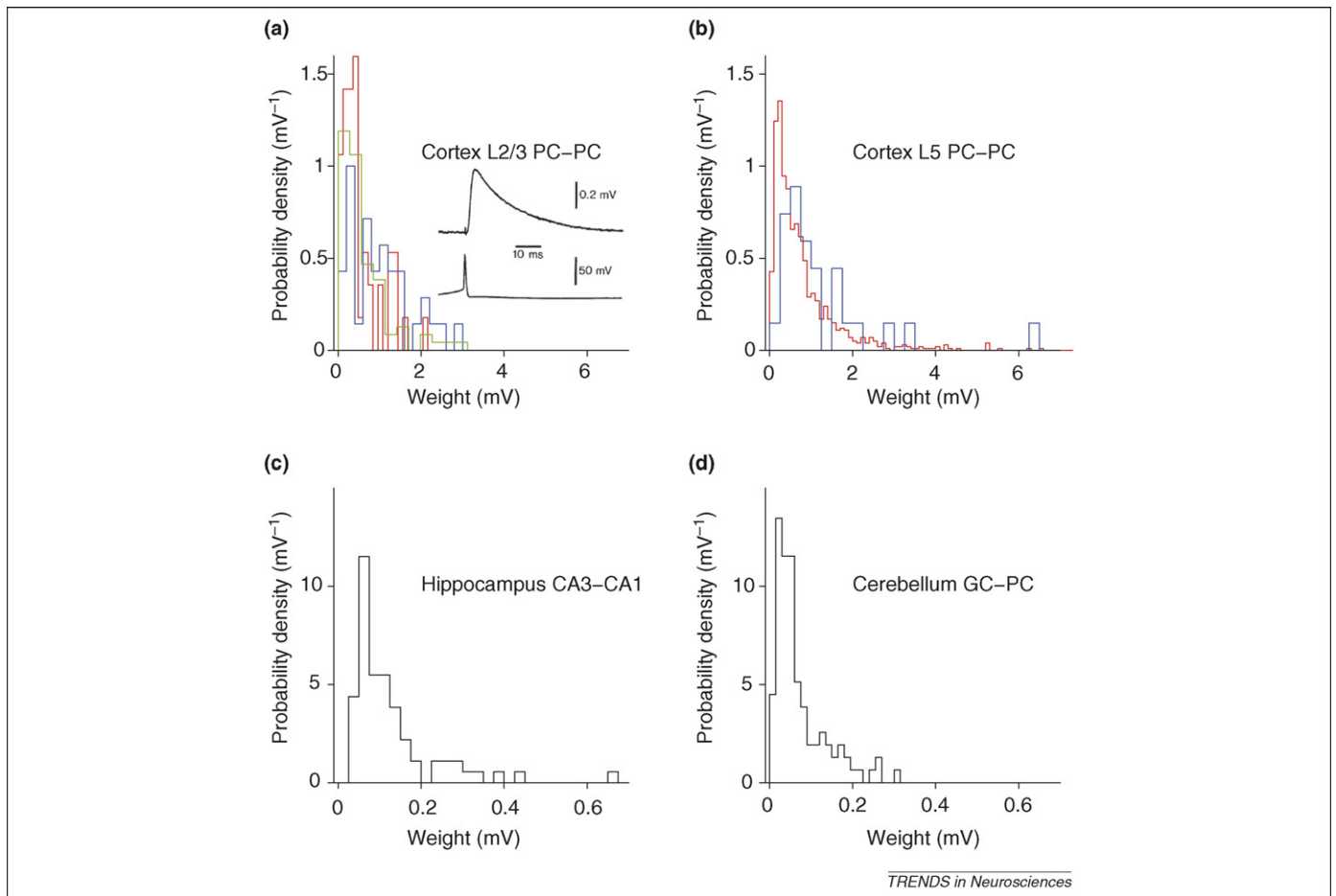
Using a simple, additive rule, according to which the weight change is independent of the synapse strength, STDP generally produces a bimodal distribution [24,26], with synaptic strengths clustering both around zero and at the maximum synaptic weight that must be imposed in this case to avoid unphysiologically strong synapses (see Figure 2). This can be understood rather simply: for uncorrelated synaptic activity, the overall effect of the plasticity rule needs to be a weakening rather than a strengthening of synaptic strength to avoid uncontrolled growth of synaptic strength. This produces the mode at zero. However, postsynaptic activity tends to be correlated with and to follow presynaptic activity at stronger synapses. This positive feedback for sufficiently strong synapses leads to the other mode at the maximum synaptic strength.

The bimodal distribution resulting from an additive rule appears to be in conflict with existing data, in which no such bimodality can be detected. It could however potentially be reconciled with experimental data if maximum synaptic weights themselves have a wide distribution, as can be expected for synapses distributed uniformly along the dendritic tree, because of electrotonic filtering effects. The nature of the distribution is also sensitive to details of the rule. Multiplicative STDP rules, in which a synaptic change depends on the strength of the synapse, can lead to unimodal distributions, in particular when depression increases with synaptic strength (see Figure 2). Such distributions have been adjusted to fit amplitude distributions of miniature EPSCs [26].

In summary, the link between synaptic plasticity and weight distributions depends upon a difficult-to-determine parameter, namely, the activity history of pre- and postsynaptic neurones. Moreover, subtle changes in the plasticity rule radically alter the relationship between the plasticity rule and the weight distribution.

#### Distributions determined by optimality principles

An alternative approach to understanding weight distributions comes from the hypothesis that weights are altered by learning with the aim of performing a task. Knowledge of that task should therefore provide strong and direct constraints upon the weights, independently of the exact manner of learning (and thus independently of the precise plasticity rule). Our own work on the distribution of synaptic weights at the cerebellar parallel fibre–Purkinje cell (PF-PC) synapse [12] serves as an illustration of this approach. We reconsidered a classic model of the Purkinje cell introduced by Marr [27], in which each cerebellar Purkinje cell acts as a pattern classifier—whether it is active or not depends upon which parallel fibre inputs are active (Box 2, Figure 3). Marr proposed that the correct association between an input pattern and the Purkinje cell output could be realised by a suitable



**Figure 1.** Experimental distributions of synaptic weights. (a) Cortical layer 2/3 pyramidal-pyramidal synapses. Red, [56]; green, [57]; blue, [58]. Inset: average presynaptic action potential (bottom) and excitatory postsynaptic potential (top) of a connected pair of layer 2/3 pyramidal cells; adapted from [56], copyright 1991 by the Society for Neuroscience. (b) Cortical layer 5 pyramidal-pyramidal synapses. Red, Sjöström *et al.* data set [38,59]; blue, [60]. (c) Hippocampal CA3-CA1 synapses [9]. (d) Cerebellar granule cell-Purkinje cell synapses [11,12].

adjustment of the weights of PF-PC synapses, by using an error signal provided by the climbing fibre (a form of “supervised learning”), a hypothesis subsequently refined by Albus [28]. Experimental confirmation of the plasticity of these weights came ~10 years later [29].

If we assume therefore that the task of the Purkinje cell is to associate given outputs to specific input patterns, can anything be deduced about the quantitative distribution of synaptic weights at the PF-PC synapse without specifying the plasticity rule at this synapse? The answer is yes, particularly if it is also assumed that the performance of the task has been optimised in some way (see Box 2). The appeal to optimality is justified in part by a good fit to experimental data (see below), but it is also reasonable to consider that evolution would favour effective and efficient computation.

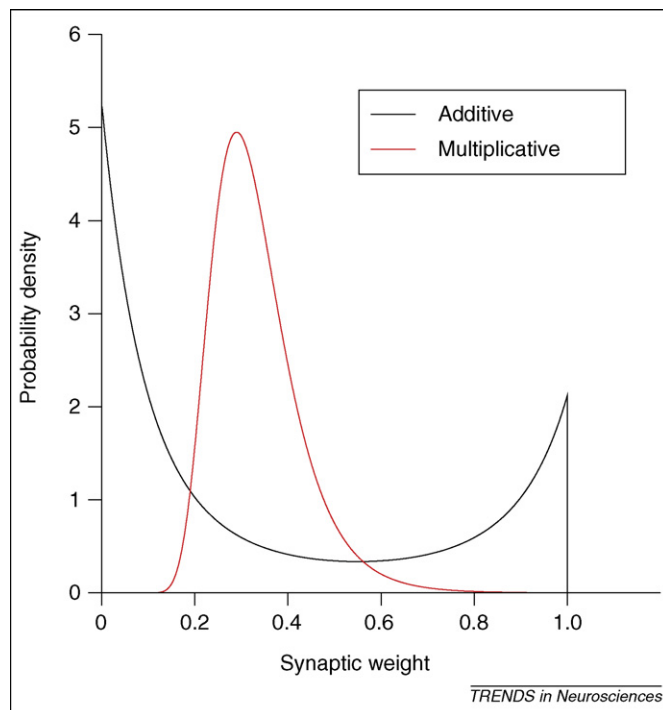
The optimal weight distribution for a perceptron with excitatory synapses that best fits the granule cell-Purkinje cell weight distribution is shown in Figure 4. The shape of the distribution is determined by the requirement of maximum storage, together with the key constraint that the synaptic weights are positive. From a maximum at zero weight, the distribution decays monotonically as a truncated Gaussian. In addition, a large fraction of zero weight (i.e. silent, or potential) synapses is present. The distribution depends only on two parameters: the mean synaptic

weight, which fixes the normalisation of the weight axis, and a parameter,  $\kappa$ , which quantifies how robust the classification is to noise perturbations.

We have described optimal learning as maximising the number of associations for a fixed resistance to noise  $\kappa$ . An alternative interpretation is that the resistance to noise is maximised for a given number of learned associations. This leads to the same optimal weight distribution, and it might be biologically more plausible, because the neural network probably does not have control over the number of associations it is presented with.

Strikingly, this analysis also offers a natural explanation for the existence of a large proportion of silent synapses: they are a necessary byproduct of optimal learning in the presence of the constraint that weights are non-negative (i.e. at excitatory synapses). As learning proceeds, some active synapses need to be depressed to avoid erroneous spike outputs. Repeated depression causes a finite fraction of synaptic weights to accumulate at zero. For non-zero  $\kappa$  (noise resistance), even more synapses must be silenced.

Fitting the theoretical distribution to the experimental one yields values for several parameters (including activity levels and resistance to noise). These values can be used to obtain precise estimates for the storage capacity (~5 kilobytes of information per Purkinje cell in the form of 40,000 input-output associations).



**Figure 2.** Distributions of synaptic weights produced by STDP rules when synapses are subjected to random Poisson inputs. For additive rules (black curve), the weight change is independent of the current weight. The resulting distribution is bimodal, with peaks close to zero weight and the maximum allowed weight. For multiplicative rules (red curve), the synaptic change depends on the weight, producing a positive equilibrium weight; the distribution obtained is unimodal, with a peak around the equilibrium weight. Adapted from [26].

Thus, interpretation of the weight distribution for the cerebellar granule cell-Purkinje cell synapse by using the perceptron model explains the existence of a large fraction of silent synapses and can yield information regarding several cellular and network properties that would be difficult to address by direct experimental methods.

#### Optimality principles for recurrent networks

Remarkably, this distribution of weights derives directly from the basic assumptions of optimal learning with excitatory synapses and is independent of the details of the learning rule—as long as it is able to attain the maximum capacity. There might be many other types of excitatory synapse involved in optimal learning in the brain; thus, it is natural to ask whether the optimal weight distribution obtained for the feedforward architecture of the perceptron could apply to other network topologies, resembling, for instance, the recurrent excitatory synapses of the neocortex. This might indeed be the case if neocortex is governed by attractor dynamics in which attractor states are stored via synaptic plasticity mechanisms [30,31]. The idea of attractor dynamics is consistent with several experimental findings on persistent activity in various areas of the cortex both *in vivo* [32–34] and *in vitro* [35,36]. If the neurones in a given cortical network and a given attractor state are simply described as being either “active” or “inactive,” a prescribed configuration of neuronal states persists if each neurone achieves its prescribed state given the synaptic inputs coming from the other neurones in the network. In other words, imposing a given number of attractors in a fully connected recurrent network of binary

#### Box 2. Obtaining weight distributions from optimality principles

Finding a set of synaptic weights that optimises performance of a task by a neural network is, in general, a difficult problem. One tractable case consists of a single binary neurone receiving a large number,  $N$ , of binary inputs (“binary” refers to the absence or presence of an action potential; synaptic weights can be continuous). This architecture is known as a *perceptron*, and a well-studied task is to learn random input-output associations by modifying synaptic weights (see Figure 3a). The *perceptron learning algorithm* (see e.g. [46]) is a supervised-learning rule that will find a set of synaptic weights solving this task—provided at least one such set exists.

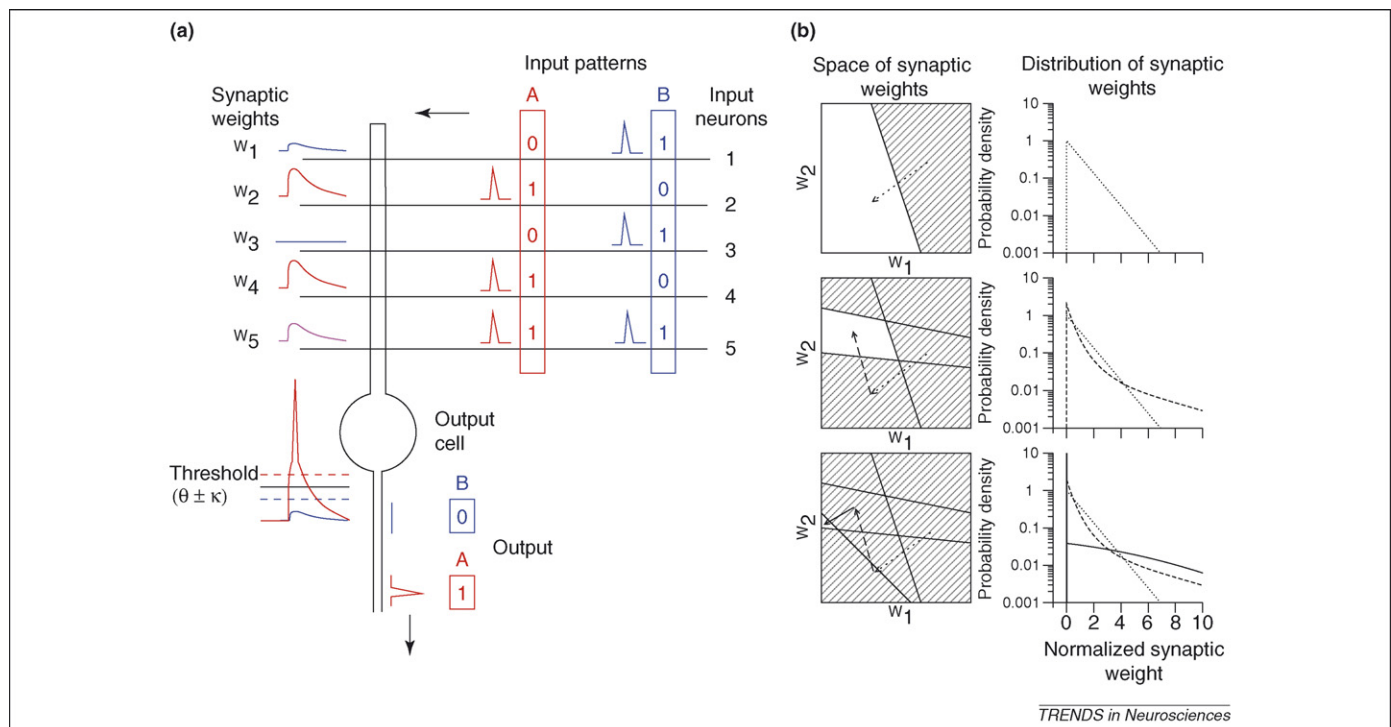
A powerful method for analysing the properties of solutions for the perceptron random association learning problem when the number of synapses,  $N$ , is large was developed by Gardner (1988), who considered the  $N$ -dimensional space representing all possible configurations of the synaptic weights of the neurone. Typically, only a finite subspace of this space will satisfy all of the input-output associations (Figure 3b). As the number of associations to be satisfied increases, this subspace decreases. The minimum number of associations causing the subspace to vanish is called the maximum (or *critical*) capacity of the perceptron. The power of Gardner’s approach arises from its adaptability: it has been extended to incorporate resistance to noise, varying levels of input and output activity, constrained synaptic weights, and even discrete synaptic weights [46].

In addition, this approach allows study of generic properties of synapses of a neurone that has achieved a given storage level, with a given robustness. In particular, it is possible to compute the distribution of weights for a system below or at maximum capacity [12,64]. For synapses with non-negative weights, the distribution at maximum capacity has two components: a fraction of at least 50% of the synapses have zero weight – the “silent” or potential synapses—the exact fraction of which increases with the robustness,  $\kappa$ . The remaining synapses (those with positive weights) are distributed according to the tail of a Gaussian distribution. How the distribution evolves as a function of the number of stored associations is shown in Figure 3b.

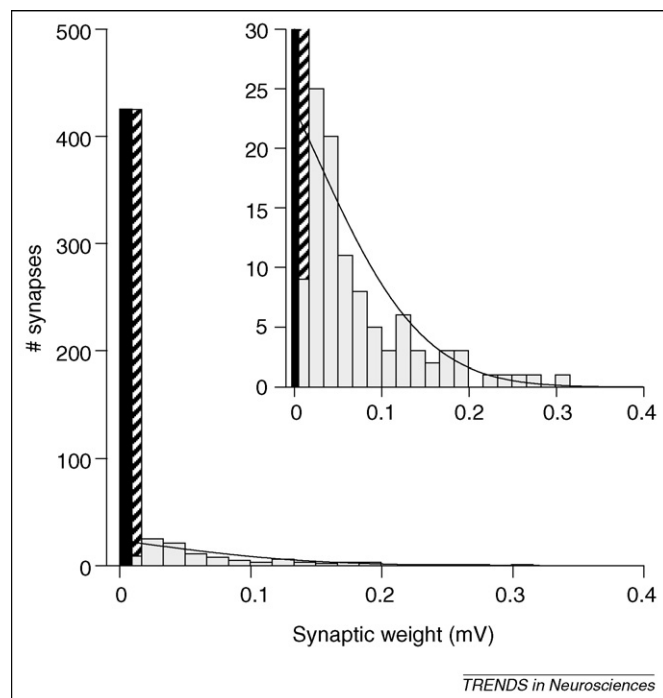
neurones is equivalent to independently solving a perceptron problem for each neurone of the network [37]. When the number of stored attractors, or their resistance to noise (i.e. their basin of attraction), is maximised, the distribution of synaptic weights is identical to that of the perceptron at maximum capacity.

The suggestion that synaptic weight distributions for excitatory synapses between pyramidal neurones could be similar to the perceptron optimal distribution implies that an important fraction of the synapses have zero weight. This is entirely consistent with the reported low probability ( $\sim 10\%$ ) of detecting a connection between nearby pyramidal cells. Note that this observation is compatible with either of the following scenarios: (i) the existence of a large number of “potential synapses,” i.e. close axonal-dendritic synapses where synapses could be created by structural plasticity; or (ii) the existence of a large fraction of undetectable (silent) synapses. This theory would also predict that the positive weights are distributed according to the tail of a Gaussian. This seems to be in contradiction with the results of Song *et al.* [38], who have fitted by a lognormal distribution the weight distribution of recurrent connections between layer 5 pyramidal cells. The lognormal distribution decays more slowly than the truncated Gaussian of the perceptron distribution. However, a slower-than-Gaussian decay could be reconciled with the





**Figure 3.** Synaptic weights in the perceptron model. (a) Perceptron model. Input patterns of action potentials elicit a compound depolarisation equal to the sum of active synaptic weights. If this compound depolarisation is greater than the threshold, an action potential is emitted (output active). For some patterns (A), the output neurone should be active; for others (B), it should remain inactive. Ensuring appropriate outputs for all input patterns requires adjusting weights accordingly. To impart resistance to noise, learned inputs should never sum closer to threshold than  $\kappa$ . Adapted from [12]. (b) Left: schematic two-dimensional representation of the N-dimensional space of synaptic weight configurations showing how learning associations restrict the space that satisfies all of the input-output associations (white region). Arrows indicate a possible trajectory of the set of weights subjected to a learning rule that tries to satisfy the constraints imposed by increasing numbers of associations. Right: corresponding synaptic weight distributions. When few associations have been learned, the distribution is exponential (top, dotted curve); middle, as the number of associations increases, the distribution stretches, with more synapses assuming very small and very large values (dashed curve); bottom, at maximum capacity, the available space of weights vanishes, and the distribution is composed of a delta function at zero (marked schematically as the thick, black line) and a truncated Gaussian (full line). For further details, see [12].



**Figure 4.** Comparing experimental and theoretical weight distributions. Granule cell-Purkinje cell synaptic weight distribution (grey histogram); the shaded bin close to zero shows the estimated number of silent synapses. The inset shows in more detail the parts of the distributions with non-zero weight. The solid line represents the best-fit optimal distribution of weights for a binary perceptron model that learns to classify random inputs into two classes. Adapted from [12].

above-described theoretical framework by a suitably nonlinear summation of inputs, or if the network stores a number of attractors that is lower than its theoretical limit [12].

Optimality criteria other than attractor maximisation have also been invoked to rationalise the large proportion of potential synapses in cortex. Assuming that synapses are binary and can take only two values (zero and some positive arbitrary value), the “information capacity” of a circuit (defined as the logarithm of the number of possible realisable circuits) can be easily computed as a function of the “filling fraction,”  $c$  (the ratio of the number of actual synapses to the total number of potential and actual synapses) [16]. This capacity is maximised when the filling fraction is  $c = 0.5$ . However, the information capacity per synapse increases monotonically when  $c$  decreases, indicating that, when the cost of maintaining positive synapses is taken into account, the optimal filling fraction should be small. In addition, synapse cost, together with optimal storage in the presence of noise, has recently been proposed as a justification for the wide distribution of synaptic weights in neocortex [39].

To summarise, theoretical analysis has shown how various features of the weight distributions can be understood through optimality principles. These theories can be applied to both the feedforward circuitry of the cerebellum and recurrent networks resembling the neocortex.

However, further work undoubtedly needs to be done to determine which quantity, if any, is optimised in these areas.

### Open issues

These initial analyses of the link between weight distributions and learning tasks open several avenues for future research; we discuss some of them here.

#### *Experimental testing of the link between optimal learning and the distribution of synaptic weights*

Analysis of synaptic weight distributions can test learning theories and offer access to difficult-to-obtain information, such as the storage capacity of a neurone. However, these analyses would obviously be strengthened by a direct demonstration that the distribution shape was indeed linked to (optimal) learning. The most promising approach would be to compare distributions when different quantities of information have been stored. One obvious idea is to compare distributions from immature and mature animals, although this might be confounded by concurrent developmental processes. Another possibility would be to compare distributions from animals raised in feature-poor and -enriched environments; presumably, the latter would have learned more. Finally, chronic pharmacological or genetic interventions might allow manipulation of specific model parameters (e.g. activity or noise levels) and testing of their expected effects on distribution shape.

#### *Discrete versus analogue synapses*

It is generally implicitly assumed that single synapses can be characterised by a continuous weight variable, which seems consistent with the continuous distributions reported in [Box 1](#). However, the distributions are also consistent with a scenario in which each synapse has only a discrete set of possible values, but with these values differing between synapses. Some experiments actually suggest that single synapses could have discrete weights or even be binary [\[40–42\]](#). However, this issue is far from settled experimentally and will require further studies of plasticity at unitary synapses.

#### *Mode of synaptic integration*

The theoretical analyses of optimal synaptic weight distributions have assumed a relatively simple mode of synaptic integration of the cell—linear summation of synaptic inputs, followed by a threshold for spike emission. The complex geometry of dendritic trees of pyramidal cells, together with the potential nonlinearities inherent in the dynamics of various intrinsic and synaptic currents, have led investigators to propose a radically different view, in which linear summation is performed only locally, at the dendritic level [\[43,44\]](#). A threshold is then applied to these local compound synaptic inputs. Finally, the resulting dendritic outputs are summed linearly at the soma, and a second threshold applied. Interestingly, such a model bears similarities with multilayer perceptrons, which are known to have more powerful learning abilities than single-layer perceptrons [\[45\]](#). The optimal distribution of weights for such single-neurone models remains to be determined.

More precise information on synaptic integration in different cell types would also be desirable.

#### *Learning rules leading to optimal performance*

We have discussed above how one can obtain the maximum capacity of a simple neuronal model as well as the resulting distribution of weights at maximum capacity. But, which learning rules are able to reach this maximum capacity? An answer to this question would connect optimality conditions to the more mechanistic approach relying on a particular synaptic plasticity rule. Learning rules that can reach the optimum are known to exist when weights are continuous. In particular, the “perceptron algorithm” [\[46\]](#) converges in a finite number of learning steps if a solution exists. This algorithm implements supervised learning by combining local information, available at the synapse, with global information provided by a “teacher.” A learning step is performed if, and only if, the cell output is different from the desired output. How this error signal might be provided to real cells is generally not clear. In the case of Purkinje cells, climbing fibres are supposed to carry the error information [\[29\]](#), but this issue remains controversial [\[47\]](#). Determining the degree and source of any supervision thus remains an important issue for future study.

The perceptron learning algorithm can be adapted to the case of sign-constrained synapses [\[48\]](#). In the case of discrete synapses, however, reaching the maximum capacity is much harder. A recently proposed algorithm comes close to this maximum capacity [\[49\]](#). Finally, whereas algorithms exist for which learning is guaranteed below the maximum storage capacity, they do not perform well when that capacity is exceeded. Some learning rules forget old memories while storing new information [\[50–52\]](#). However, no such algorithm is known to lead to optimal performance. Theories of graceful and optimal forgetting therefore also require further study.

#### *Inhibitory synapses*

We have focused here on excitatory synapses. A fundamental and still-open question is related to the asymmetric functions of excitation and inhibition. Inhibition is known to play an important role in the organisation of activity [\[53\]](#). In addition, plasticity of inhibitory circuits [\[14,54,55\]](#) suggests that they are also involved in learning. One could expect that if inhibitory synapses are also plastic, the same type of distribution would emerge for inhibitory synapses at maximum capacity (large fraction of silent/potential synapses, monotonically decaying distribution). However, this issue remains to be clarified.

### Conclusion

We have discussed here how various approaches lead to specific predictions for the distribution of synaptic weights. In our view, the most powerful approaches are those that consider optimal performance of the considered system, because no hypothesis needs to be made about the details of the plasticity rule, apart from the fact that it is able to reach the optimum. Such approaches also have the advantage of providing access to quantities such as storage capacity, which are currently beyond experimental reach.

Important challenges remain, including elucidating and experimentally testing what is optimised in a given task, as well as determining learning rules leading to optimal performance, given biophysical constraints acting at the level of single synapses. Despite such open questions, the analysis of weight distributions has the potential to become a powerful tool for studying the mechanisms of learning in the brain.

### Acknowledgements

We thank Jesper Sjöström for sharing his data and commenting on the manuscript; we also thank Dirk Feldmeyer and Yuri Zilberter for helpful discussion. This work was supported by the Action Concertée Incitative (ACI) "Neurosciences intégratives et computationnelles" and the Agence Nationale pour la Recherche (ANR).

### References

- Hebb, D.O. (1949) *Organization of Behavior*, Wiley
- Bliss, T.V.P. and Lomo, T. (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.* 232, 331–356
- De Zeeuw, C.I. *et al.* (1998) Expression of a protein kinase c inhibitor in purkinje cells blocks cerebellar ltd and adaptation of the vestibulo-ocular reflex. *Neuron* 20, 495–508
- Lynch, M.A. (2004) Long-term potentiation and memory. *Physiol. Rev.* 84, 87–136
- Whitlock, J.R. *et al.* (2006) Learning induces long-term potentiation in the hippocampus. *Science* 313, 1093–1097
- Liao, D. *et al.* (1995) Activation of postsynaptically silent synapses during pairing-induced ltp in ca1 region of hippocampal slice. *Nature* 375, 400–404
- Isaac, J.T. *et al.* (1995) Evidence for silent synapses: implications for the expression of ltp. *Neuron* 15, 427–434
- Nusser, Z. *et al.* (1998) Cell type and pathway dependence of synaptic ampa receptor number and variability in the hippocampus. *Neuron* 21, 545–559
- Sayer, R.J. *et al.* (1990) The time course and amplitude of EPSPs evoked at synapses between pairs of hippocampal CA3/CA1 neurons in the hippocampal slice. *J. Neurosci.* 10, 826–836
- Harvey, R.J. and Napper, R.M. (1991) Quantitative studies on the mammalian cerebellum. *Prog. Neurobiol.* 36, 437–463
- Isope, P. and Barbour, B. (2002) Properties of unitary granule cell–>purkinje cell synapses in adult rat cerebellar slices. *J. Neurosci.* 22, 9668–9678
- Brunel, N. *et al.* (2004) Optimal information storage and the distribution of synaptic weights: perceptron versus purkinje cell. *Neuron* 43, 745–757
- Ekerot, C.F. and Jörntell, H. (2001) Parallel fibre receptive fields of purkinje cells and interneurons are climbing fibre-specific. *Eur. J. Neurosci.* 13, 1303–1310
- Jörntell, H. and Ekerot, C-F. (2002) Reciprocal bidirectional plasticity of parallel fiber receptive fields in cerebellar purkinje cells and their afferent interneurons. *Neuron* 34, 797–806
- Hellwig, B. (2000) A quantitative analysis of the local connectivity between pyramidal neurons in layers 2/3 of the rat visual cortex. *Biol. Cybern.* 82, 111–121
- Stepanyants, A. *et al.* (2002) Geometry and structural plasticity of synaptic connectivity. *Neuron* 34, 275–288
- Le Bé, J.V. and Markram, H. (2006) Spontaneous and evoked synaptic rewiring in the neonatal neocortex. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13214–13219
- Kalisman, N. *et al.* (2005) The neocortical microcircuit as a tabula rasa. *Proc. Natl. Acad. Sci. U. S. A.* 102, 880–885
- Stepanyants, A. and Chklovskii, D.B. (2005) Neurogeometry and potential synaptic connectivity. *Trends Neurosci.* 28, 387–394
- Holtmaat, A.J. *et al.* (2005) Transient and persistent dendritic spines in the neocortex *in vivo*. *Neuron* 45, 279–291
- Stepanyants, A. *et al.* (2004) Class-specific features of neuronal wiring. *Neuron* 43, 251–259
- Chklovskii, D.B. *et al.* (2004) Cortical rewiring and information storage. *Nature* 431, 782–788
- Dan, Y. and Poo, M.M. (2006) Spike timing-dependent plasticity: from synapse to perception. *Physiol. Rev.* 86, 1033–1048
- Song, S. *et al.* (2000) Competitive hebbian learning through spike-time-dependent synaptic plasticity. *Nat. Neurosci.* 3, 919–926
- Rubin, J. *et al.* (2001) Equilibrium properties of temporally asymmetric hebbian plasticity. *Phys. Rev. Lett.* 86, 364–367
- van Rossum, M.C. *et al.* (2000) Stable hebbian learning from spike timing-dependent plasticity. *J. Neurosci.* 20, 8812–8821
- Marr, D. (1969) A theory of cerebellar cortex. *J. Physiol.* 202, 437–470
- Albus, J.S. (1971) A theory of cerebellar function. *Math. Biosci.* 10, 26–51
- Ito, M. (2006) Cerebellar circuitry as a neuronal machine. *Prog. Neurobiol.* 78, 272–303
- Hopfield, J.J. (1982) Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U. S. A.* 79, 2554–2558
- Amit, D.J. (1995) The hebbian paradigm reintegrated: local reverberations as internal representations. *Behav. Brain Sci.* 18, 617–626
- Fuster, J.M. and Alexander, G. (1971) Neuron activity related to short-term memory. *Science* 173, 652–654
- Miyashita, Y. (1988) Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* 335, 817–820
- Goldman-Rakic, P.S. (1995) Cellular basis of working memory. *Neuron* 14, 477–485
- Cossart, R. *et al.* (2003) Attractor dynamics of network up states in the neocortex. *Nature* 423, 283–288
- Ikegaya, Y. *et al.* (2004) Synfire chains and cortical songs: temporal modules of cortical activity. *Science* 304, 559–564
- Gardner, E. (1988) The phase space of interactions in neural network models. *J. Phys. A: Math. Gen.* 21, 257–270
- Song, S. *et al.* (2005) Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* 3, e68
- Varshney, L.R. *et al.* (2006) Optimal information storage in noisy synapses under resource constraints. *Neuron* 52, 409–423
- Petersen, C.C. *et al.* (1998) All-or-none potentiation at ca3-ca1 synapses. *Proc. Natl. Acad. Sci. U. S. A.* 95, 4732–4737
- O'Connor, D.H. *et al.* (2005) Dissection of bidirectional synaptic plasticity into saturable unidirectional processes. *J. Neurophysiol.* 94, 1565–1573
- O'Connor, D.H. *et al.* (2005) Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9679–9684
- Poirazi, P. *et al.* (2003) Arithmetic of subthreshold synaptic summation in a model ca1 pyramidal cell. *Neuron* 37, 977–987
- Poirazi, P. *et al.* (2003) Pyramidal neuron as two-layer neural network. *Neuron* 37, 989–999
- Vapnik, V.N. (1995) *The Nature of Statistical Learning Theory*, Springer
- Engel, A. and van den Broeck, C. (2001) *Statistical Mechanics of Learning*, Cambridge University Press
- Catz, N. *et al.* (2005) Cerebellar complex spikes is suitable to induce as well as to stabilize motor learning. *Curr. Biol.* 15, 2179–2189
- Amit, D.J. *et al.* (1989) The interaction space of neural networks with sign-constrained synapses. *J. Phys. A: Math. Gen.* 22, 4687–4693
- Baldassi, C. *et al.* (2007) Efficient supervised learning in networks with binary synapses. *Proc. Natl. Acad. Sci. U. S. A.* 104, 11079–11084
- Nadal, J-P. *et al.* (1986) Networks of formal neurons and memory palimpsests. *Europhys. Lett.* 1, 535–542
- Parisi, G. (1986) A memory which forgets. *J. Phys. A: Math. Gen.* 19, L617–L620
- Amit, D.J. and Fusi, S. (1994) Dynamic learning in neural networks with material synapses. *Neural Comput.* 6, 957–982
- Bartos, M. *et al.* (2007) Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat. Rev. Neurosci.* 8, 45–56
- Pelkey, K.A. *et al.* (2005) mglur7 is a metaplastic switch controlling bidirectional plasticity of feedforward inhibition. *Neuron* 46, 89–102
- Soler-Llavina, G.J. and Sabatini, B.L. (2006) Synapse-specific plasticity and compartmentalized signaling in cerebellar stellate cells. *Nat. Neurosci.* 9, 798–806
- Mason, A. *et al.* (1991) Synaptic transmission between individual pyramidal neurons of the rat visual cortex *in vitro*. *J. Neurosci.* 11, 72–84

- 57 Holmgren, C. *et al.* (2003) Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *J. Physiol.* 551, 139–153
- 58 Feldmeyer, D. *et al.* (2006) Efficacy and connectivity of intracolumnar pairs of layer 2/3 pyramidal cells in the barrel cortex of juvenile rats. *J. Physiol.* 575, 583–602
- 59 Sjöström, P.J. *et al.* (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 32, 1149–1164
- 60 Frick, A. *et al.* (2007) Monosynaptic connections between pairs of l5a pyramidal neurons in columns of juvenile rat somatosensory cortex. *Cereb. Cortex*
- 61 Markram, H. *et al.* (1997) Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *J. Physiol.* 500, 409–440
- 62 Miles, R. (1990) Variation in strength of inhibitory synapses in the ca3 region of guinea-pig hippocampus *in vitro*. *J. Physiol.* 431, 659–676
- 63 Feldmeyer, D. *et al.* (1999) Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single 'barrel' of developing rat somatosensory cortex. *J. Physiol.* 521, 169–190
- 64 Kohler, H.M. and Widmaier, D. (1991) Sign-constrained linear learning and diluting in neural networks. *J. Phys. A: Math. Gen.* 24, L495–L502

## Elsevier celebrates two anniversaries with a gift to university libraries in the developing world

In 1580, the Elzevir family began their printing and bookselling business in the Netherlands, publishing works by scholars such as John Locke, Galileo Galilei and Hugo Grotius. On 4 March 1880, Jacobus George Robbers founded the modern Elsevier company intending, just like the original Elzevir family, to reproduce fine editions of literary classics for the edification of others who shared his passion, other 'Elzevirians'. Robbers co-opted the Elzevir family printer's mark, stamping the new Elsevier products with a classic symbol of the symbiotic relationship between publisher and scholar. Elsevier has since become a leader in the dissemination of scientific, technical and medical (STM) information, building a reputation for excellence in publishing, new product innovation and commitment to its STM communities.

In celebration of the House of Elzevir's 425th anniversary and the 125th anniversary of the modern Elsevier company, Elsevier donated books to ten university libraries in the developing world. Entitled 'A Book in Your Name', each of the 6700 Elsevier employees worldwide was invited to select one of the chosen libraries to receive a book donated by Elsevier. The core gift collection contains the company's most important and widely used STM publications, including *Gray's Anatomy*, *Dorland's Illustrated Medical Dictionary*, *Essential Medical Physiology*, *Cecil Essentials of Medicine*, *Mosby's Medical, Nursing and Allied Health Dictionary*, *The Vaccine Book*, *Fundamentals of Neuroscience*, and *Myles Textbook for Midwives*.

The ten beneficiary libraries are located in Africa, South America and Asia. They include the Library of the Sciences of the University of Sierra Leone; the library of the Muhimbili University College of Health Sciences of the University of Dar es Salaam, Tanzania; the library of the College of Medicine of the University of Malawi; and the University of Zambia; Universite du Mali; Universidade Eduardo Mondlane, Mozambique; Makerere University, Uganda; Universidad San Francisco de Quito, Ecuador; Universidad Francisco Marroquin, Guatemala; and the National Centre for Scientific and Technological Information (NACESTI), Vietnam.

Through 'A Book in Your Name', these libraries received books with a total retail value of approximately one million US dollars.

**For more information, visit [www.elsevier.com](http://www.elsevier.com)**