

Contrasting Rules for Synaptogenesis, Modification of Existing Synapses, and Synaptic Removal as a Function of Neuronal Computation

William B Levy *

Dept. of Neurosurgery, University of Virginia, Charlottesville, VA 22908; *Corresponding author

Keywords: shedding, pruning, prediction, energy-efficient, sparse connectivity

Abstract. There are three easily distinguishable forms of synaptic plasticity: synaptogenesis, modification of existing synapses, and synapse removal. For any particular synaptic class, only a compatible set of plasticity rules makes sense for these three distinguishable forms. Here we point out two such compatible sets, one for each of two classes of excitatory synapses. In both cases, synaptogenesis is local and anti-correlational while modification of existing synapses encodes mostly positive correlations as adjusted by a particular negative correlation. Connections that offer little or nothing to the local, postsynaptic information processing is an inefficiency and leads to synapse removal.

The theme of this presentation brings together two earlier lines of research: (a) an integrated theory of synaptic modification and (b) the minimal functional aspects of prediction/classifying/decoding neurons that imply a minimal but necessary physiology and anatomy of such neurons.

In a series of thirteen papers (see LevyLab website for details) beginning with Levy and Desmond (1985), we defined and studied a tripartite set of synaptic modification rules including synaptogenesis, modification of existing synapses, and removal of existing synapses. Many of these papers are devoted to infomax-like transformations (Barlow, 1959), and there is a key insight making the rules useful for such a transformation – neurons are most useful when carrying a genetically predefined amount of information and innervated in a biased random manner that produces a relatively small fractional connectivity to each neuron. The appropriate synaptogenesis rule can guarantee this information constraint and, as well, a connectivity in which each neuron only receives a small random fraction of the total available inputs

(Levy et al., 1999). Because activity levels are also being guaranteed, these rules can also produce neurons that fire at an energy-efficient information rate (Levy & Baxter, 1996).

In parallel with this work on synaptogenesis, there was also ongoing work describing certain neuronal computations by CA1 hippocampal pyramidal cells (Levy, 1989; 1990; Levy et al., 1990). In particular, a cell much like a CA1 pyramid of the hippocampus (and as well, like a neocortical pyramid that has a main apical dendrite and which receives different classes of inputs distally and proximally on the apical dendritic branches) is a candidate for performing as a prediction/classifying neuron. (A prediction neuron essentially forecasts a classification prior to the event occurring and a classification neuron is essentially a retrodicting neuron where a feature set is used to decide on the existence or not of a named category. Such a neuron can also be a decoder.) Figure 1 illustrates, with the two distinct apical dendritic compartments, the anatomical description of such a pyramidal-type neuron. Compartment 1 (C1) is the most distal compartment. Functionally it is an information transforming compartment, and at this functional level, it resembles an entire information transforming neuron (illustrated as a dentate gyrus granule cell to the right of C1 in Fig. 1). The inputs that excite C1 essentially define the set of events that the pyramidal neuron is constructed (self-constructed) to recognize, classify, decode, or predict.

Compartment 2 (C2) is devoted to predicting something about the excitation of C1. That is, we hypothesize that the synaptic excitation of C2 generates predictions or classifications in the absence of the C1 event being predicted or classified. Thus, C2 combines feature information to predict about the “named category” or event defined by compartment 1. A critical characteristic that distinguishes the two compartments is the associative synaptic modification rule of existing synapses. In C1 the rule depends only on events produced by the inputs to C1. Thus, the modification is unsupervised or self-supervised. On the other hand, the associative synaptic modification rule for an existing synapse in C2 depends upon the coactivity between this synapse and a postsynaptic excitation arising out of compartment 1. Thus, the synaptic modification here is supervised.

(FIGURE 1 ABOUT HERE.)

Although not absolutely critical, we will assume that the C1 excitation is distally thresholded and call this $Z_j \in \{0,1\}$. Likewise thresholding the excitation of C2 produces $\bar{Z}_j \in \{0,1\}$ where this variable is the predicted value of Z_j .

Here we now consider what forms of synaptogenesis and synapse removal are compatible with such a synapse and such function. At all times biological plausibility is our utmost concern (see Levy et al., 1990; Levy & Desmond, 1985 for details illustrating the relevant biology). The most interesting idea is that synaptogenesis is a decreasing function of postsynaptic excitation.

| Table 1 General Principles for Excitatory Synaptic Plasticity in Neocortex | | |
|--|--|--|
| <i>Synaptogenesis</i> | <i>Modification of Existing Synapses</i> | <i>Synapse Removal</i> |
| stochastic process anti-correlational | deterministic process correlational and anti-correlational | stochastic process |
| receptivity for new innervation is negatively correlated with a certain aspect of postsynaptic excitation | stores statistic(s) in synaptic weights, w_{ij} , for presynaptic input i and postsynaptic neuron j | excitatory synapses falling below a minimum value are wasteful and should be removed |
| independent contributions by pre- and post (pre = i , post = j) A_i - presynaptic avidity R_j - postsynaptic receptivity Prob(of new synapse ij) \propto $A_i(t) * R_j(t)$ | $\Delta w(t+1) \propto$ $f(post_j(t)) * g(pre_i(t), w_{ij}(t));$ $w_{ij} \geq 0;$ $f(post_i) \geq 0;$ $g(\quad)$, over a restricted range, can be positive, negative, or zero | Prob(removal ij) $= \begin{cases} 0 & \text{if } w_{ij}(t) + \Delta w_{ij}(t+1) > c \\ > 0 & \text{otherwise} \end{cases}$ (c is some nonnegative constant) Δw can take on negative or positive rules |

Table 1 emphasizes the similarities between the two sets of synaptic plasticity rules. In particular, we hypothesize that synaptogenesis is, fundamentally, anti-correlational in a way that complements the positive correlational encodings produced by modification of the synapse once formed. Synaptogenesis is invigorated when the postsynaptic neuron is failing to do something it was designed to do. In one case (C1), the hypothesized function is carrying enough information, in the other case (C2), the function is successful prediction.

A second aspect of synaptogenesis is the local principle. It is not possible to maintain specificity between pre- and postsynaptic elements if temporally coactive elements are too widely (more than one or

two microns) separated in space. That is, if there is a multiplicity of presynaptic elements and a multiplicity of postsynaptic elements as actually exist in the brain, activity-based, chemically diffusing signals being passed between pre- and postsynaptic structures will inevitably be confused once the distances between pre- and postsynaptic structures are too large and the times between presynaptic and postsynaptic coactivation is too long. Thus, although we posit correlational statistics are encoded at synapses, this only occurs after synaptogenesis.

The general principles of synaptogenesis noted in Table 1 reflect this local principle (that is, avoid action at a distance). As a result the presynaptic rule depends on something about the presynaptic structure and the postsynaptic rule depends on something local to the postsynaptic structure. These separate events would each control distinct biological growth processes that, when multiplied together, can be interpreted as a probability (e.g., surface area, diffusable growth factor, binding sites, or random growth). Thus, physical separateness produces an independence-like idea which is reflected in the multiplication (see Table 1).

Again referring to Table 1 and its generic descriptions, synapse removal is an efficiency mechanism. That is, a synapse on postsynaptic neuron j that makes little or no contribution to the computation performed by j is inefficient; it wastes space and energy, and in Nature such inefficiency is a demand for synapse removal. As we are limiting the discussion here to excitatory synapses (and obviously inhibitory synapses do carry information and would require a very different set of synaptic plasticity rules), we can see one clear requirement for synapse removal: excitatory synapses that are driven to a value of zero or less should be removed. In fact it might even be argued, they should be removed when they are at a value proportional to uncorrelated or less. Surely once they are driven to a value below uncorrelated, they should no longer be excitatory synapses and should, if anything, make some contribution to the inhibition of the cell j .

Tables 2 and 3 detail some of the specific differences that distinguish these synaptic plasticity rules as a function of whether the compartment is designed for efficient representational information capacity, i.e., C1, or whether the function is prediction/classification as in C2.

| Table 2a Comparing Synaptic Weights | | |
|---|-----|--|
| <i>Compartment 1 (C1) of Prediction Neuron</i> (\equiv Information Transforming Neuron) | vs. | <i>Prediction Compartment (C2)</i> |
| $w \propto e_1$ where for a particular input set $i \in \{1, \dots, n\}$ to C1 with random vector values $X_i(t)$, $(E[XX^T] - E(X)E[X^T])e_1 = \lambda e_1$; e_1 is the dominant eigenvector | | $v_i = \log \frac{w_{ij}^+ (1 - w_{ij}^-)}{w_{ij}^- (1 - w_{ij}^+)}; c_i = \log \frac{(1 - w_{ij}^+)}{(1 - w_{ij}^-)}$ $w_{ij}^+ = E[X_i Z_j = 1]$ $w_{ij}^- = E[X_i Z_j = 0]$ |
| Table 2b Internal Excitation | | |
| <i>C1</i> | vs. | <i>C2</i> |
| $Y(t) = \sum_i X_i(t) w_{ij}$ $Z_j(t) = \theta(Y(t))$ | | $U(t) = \sum_i X_i v_i + \sum_i c_i$ $\bar{Z}_j(t) = \bar{\theta}(U(t), E[Z_j])$ |
| $\theta(\)$ and $\bar{\theta}(\)$ are threshold functions | | |

Table 2 summarizes the hypothesized computations. For information transforming compartments, a self-supervised synaptic modification rule can maximize information transmission in an energy-efficient manner if the distributional form of the excitation is nearly gaussian. Specifically, the set of synaptic weights are hypothesized to take on values proportional to the dominant eigenvector of the covariance matrix of these inputs themselves. This dominant eigenvector is the best linear filter in the sense that it maximizes information throughput. The standard synaptic modification rule of Oja (1982) and the one that we have suggested (Levy, 1982; Levy & Geman, 1982) are two of many variants that encode synaptic weights proportional to the dominant eigenvector. Moreover, if the mean and variance of this excitation are available to the neuron, the representational firing rate, $E[Z_j]$, can be set to its energy-efficient rate so long as a central limit theorem is well approximated by the net excitation of C1 (see Tyrcha & Levy, 2003).

This “self-supervised” rule stands in contrast to the synaptic modification rule in C2 where simple conditional statistics are postulated as being encoded (Levy et al., 1990). A pair of correlational learning

rules are postulated. Together we call this pair the four quadrant rule (in reference to a 2 x 2 contingency table and with $E[Z_j]$ known to the postsynaptic neuron j). For input i , the four quadrant rule encodes

$\left\{ w_{ij}^+ = E[X_i | Z_j = 1] \text{ and } w_{ij}^- = E[X_i | Z_j = 0] \right\}$. The contribution of each synapse to a prediction is

$X_i \log \frac{w^+ (1 - w^-)}{w^- (1 - w^+)} + \log \frac{(1 - w^+)}{(1 - w^-)}$. As the second logarithmic term is not multiplied by X_i , it is a

constant and can be thought of as a resting conductance (perhaps implemented by the additional spine surface area of each synapse). This synaptic form allows the generation of \bar{Z}_j as a Bayesian-like prediction of Z_j (see Levy et al., 1990 for details).

Now that we have stated the synaptic modification rule for existing synapses, one can think sensibly about synapse removal. In the case of C1 and the information transforming neuron, any input that is being forced to a negative value by the Δw_{ij} rule will tend to be discarded. Such negative values will occur when the row sum associated with this input of covariance matrix is less than a certain value.

Specifically, $E \left[X_k, \sum_{i \neq k} X_i w_{ij} \right] - E[X_k] E \left[\sum_{i \neq k} X_i w_{ij} \right] < 0$. That is, on average negative correlation is

cause for removal of the k^{th} input into C1. On the other hand, removal from the C2, prediction

compartment, hinges on whether activity produces a positive contribution to the \bar{Z}_j prediction.

Specifically, if the active state of input i makes the event $Z_j = 1$ less likely, then the synapse should be discarded. Specific conditions for removal are $\log (E[X_i | Z_j = 1] / E[X_i | Z_j = 0]) < 0$ or this same logarithm less than the prior log odds of Z_j .

Lastly, we come to the anti-correlational synaptogenesis rules (Table 3). Synaptogenesis should serve up a reasonable set of potential synapses. This “reasonable set,” will be better than an unbiased random selection; surely appropriately biased synaptogenesis saves time. In either case (C1 or C2), the synaptogenesis rule is broken down into two parts as required by the locality principle. Postsynaptically,

the receptivity for new innervation (R_j) in the information transforming compartment should help bring the neuron up to the desired information transmission levels. This can be done by producing appropriate retrograde chemical signals and attachment sites that depend either on the variance of postsynaptic

| Table 3 | | Comparing the Synaptogenesis Rules | |
|---|-----|---|--|
| <i>Compartment 1 (C1) of Prediction Neuron</i> | | vs. | <i>Prediction Compartment (C2)</i> |
| <i>(≡ Information Transforming Neuron)</i> | | | |
| an axon's avidity decreases as the synaptic territory of this axon increases | | | avidity goes up when active |
| $A_i \propto f_A \left(\sum_j w_{ij} \right)$ | vs. | | $A_i(t) = \begin{cases} 0 & \text{if } X_i(t) = 0 \\ > 0 & \text{otherwise} \end{cases}$ |
| receptivity goes down as average postsynaptic excitation goes up | | | receptivity goes up when a positive Z_j event is missed by the predictor \bar{Z}_j |
| $R_j \propto f_R(E[Z_j])$ | or | | |
| $R_j \propto f_R \left(\text{var} \left(\sum_i X_i w_{ij} \right) \right)$ | vs. | | $R_j(t) = \begin{cases} > 0 & \text{if } \{Z_j(t) = 1, \bar{Z}_j(t) = 0\} \\ 0 & \text{otherwise} \end{cases}$ |
| $X_i \in \{0,1\}; Z_j \in \{0,1\}; \bar{Z}_j \in \{0,1\}; Z_j$ is the event being predicted; \bar{Z}_j is the prediction. | | | |
| The f 's are monotonic nonincreasing functions that range over $[0,1]$. | | | |

excitation $\left(\text{var} \left(\sum_i X_i w_{ij} \right) \right)$ or upon the firing rate Z_j of the neuron. That is to say, R_j should be

inversely related to these variables. In regard to C1, the less a neuron is firing below its desired level, the more receptive to new innervation the neuron should be. Each neuron must reach its genetically programmed, energy-efficient information rate or it is eliminated. Once this rate is obtained,

synaptogenesis should stop; that is, R_j goes to zero. On the other hand, since the job of the prediction compartment is successful prediction, synaptogenesis should hinge upon unsuccessful prediction.

Specifically, new excitatory inputs are needed when $Z_j = 1$ while the prediction is $\bar{Z}_j = 0$. The more often this happens, the greater the receptivity should be for innervation in the prediction compartment.

Thus for C2, R_j is negatively correlated with \bar{Z}_j in a conditional sense (that is, if $(1 - \bar{Z}_j) Z_j$ is positive

R_j should increase). Whether a suitable input exists to lower such errors will be determined by the synaptic modification rule in concert with the synapse removal term.

Presynaptically, things are a little different when we compare C1 vs. C2 (see Table 3). In the case of information transforming systems, that is for C1, the avidity of random growth by an input i depends on this input's entire set of synapses w_{ij} . Note that the summation of the w_{ij} 's is over j . That is to say, a single input i will innervate many postsynaptic cells j , and this summation over j will express the total excitatory territory a presynaptic input has over a single class of postsynaptic cells. With the infomax perspective, it is desirable that no single input goes to too many j 's. If this were to happen, the positive correlation between postsynaptic neurons would rise. The avidity function reduces correlations between postsynaptic neurons and raises the information capacity of the entire system (Adelsberger-Mangan & Levy, 1994). Again, there is a contrast with the C2. In this case, inputs to the prediction compartment conform more to the conventional idea that presynaptic activity should encourage synaptogenesis. However, just being active is not enough because of the multiplication of avidity times receptivity. That is, if a postsynaptic cell j is not receptive to new innervation ($R_j = 0$), a large A_i will be ignored by the stochastic synaptogenesis mechanism.

In summary then, we see that the predicted biology, that of late development and of adult synaptic plasticity, should be quite different in adjacent compartments of a pyramidal cell. Thus, we have another illustration where computational theory can guide the hypotheses of experimental neuroscience.

ACKNOWLEDGEMENTS

This work was supported by NIH MH48161, MH63855 and RR15205 to WBL.

REFERENCES

Adelsberger-Mangan, D. M. and Levy, W. B The influence of limited presynaptic growth and synapse removal on adaptive synaptogenesis. Biol. Cybern. 71, 1994, 461-468.

- Barlow, H. B. Sensory mechanisms, the reduction of redundancy and intelligence. (National Physical Laboratory Symposium no. 10. The mechanisation of thought processes.) Her Majesty's Stationary Office, London, pp. 537-559.
- Levy, W. B. Associative encoding at synapses. Proceedings of the Fourth Annual Conference of Cognitive Science Society, 1982, 135-136.
- Levy, W. B. A computational approach to hippocampal function. In: Computational Models of Learning in Simple Neural Systems. (R. D. Hawkins and G. H. Bower, Eds.), New York: Academic Press, pp. 243-305, 1989.
- Levy, W. B. Maximum entropy prediction in neural networks. International Joint Conference on Neural Networks, 1990, I-7—I-10.
- Levy, W. B. and Baxter, R. A. Energy efficient neural codes. Neural Comp. 8, 1996, 531-543.
- Levy, W. B., Colbert, C. M., and Desmond, N. L. Elemental adaptive processes of neurons and synapses: A statistical/computational perspective. In: Neuroscience and Connectionist Models. (M. A. Gluck and D. E. Rumelhart, Eds.), Hillsdale, NJ: Lawrence Erlbaum, pp. 187-235, 1990.
- Levy, W. B., Delic H., and Adelsberger-Mangan, D. M. The statistical relationship between connectivity and neural activity in fractionally connected feed-forward networks. Biol. Cybern., 80, 1999, 131-139.
- Levy, W. B. and Desmond, N. L. Associative potentiation/depression in the hippocampal dentate gyrus. In: Electrical Activity of the Archicortex. (G. Buzsaki & C.H. Vanderwolf, Eds.), Budapest, Hungary: Akademiai Kiado, pp. 359-373, 1985.
- Levy, W. B. and Geman, S. Limit behavior of experimentally derived synaptic modification rules. Technical Report, Div. of Applied Mathematics, Brown University, Providence, RI, Reports in Pattern Analysis No. 121, 1982, 12 pages.
- Oja, E. A simplified neuron model as a principal component analyzer. J. Math. Biol., 15, 1982, 267-273.
- Tyrcha, J. and Levy, W. B. Another contribution by synaptic failures to energy efficient processing by neurons, unpublished (CNS03 submission).

FIGURE LEGENDS

Figure 1. The computational distinction between a prediction neuron and an information transforming neuron has morphological correlates. The distal tuft of the prediction neuron corresponds to the entire dendritic field of the information transforming neuron.

A prediction neuron has two dendritic compartments (C).
 C1 defines what the neuron predicts or classifies.
 C2 generates the prediction/ classification.

An information transforming neuron is like compartment one.

