

Redistribution of Synaptic Efficacy Supports Stable Pattern Learning in Neural Networks

Gail A. Carpenter

gail@cns.bu.edu

Boriana L. Milenova

boriana.milenova@oracle.com

*Department of Cognitive and Neural Systems, Boston University,
Boston, Massachusetts 02215, U.S.A.*

Markram and Tsodyks, by showing that the elevated synaptic efficacy observed with single-pulse long-term potentiation (LTP) measurements disappears with higher-frequency test pulses, have critically challenged the conventional assumption that LTP reflects a general gain increase. This observed change in frequency dependence during synaptic potentiation is called redistribution of synaptic efficacy (RSE). RSE is here seen as the local realization of a global design principle in a neural network for pattern coding. The underlying computational model posits an adaptive threshold rather than a multiplicative weight as the elementary unit of long-term memory. A distributed instar learning law allows thresholds to increase only monotonically, but adaptation has a bidirectional effect on the model postsynaptic potential. At each synapse, threshold increases implement pattern selectivity via a frequency-dependent signal component, while a complementary frequency-independent component nonspecifically strengthens the path. This synaptic balance produces changes in frequency dependence that are robustly similar to those observed by Markram and Tsodyks. The network design therefore suggests a functional purpose for RSE, which, by helping to bound total memory change, supports a distributed coding scheme that is stable with fast as well as slow learning. Multiplicative weights have served as a cornerstone for models of physiological data and neural systems for decades. Although the model discussed here does not implement detailed physiology of synaptic transmission, its new learning laws operate in a network architecture that suggests how recently discovered synaptic computations such as RSE may help produce new network capabilities such as learning that is fast, stable, and distributed.

1 Introduction

The traditional experimental interpretation of long-term potentiation (LTP) as a model of synaptic plasticity is based on a fundamental hypothesis: “Changes in the amplitude of synaptic responses evoked by single-shock extracellular electrical stimulation of presynaptic fibres are usually considered to reflect a change in the gain of synaptic signals, and are the most frequently used measure for evaluating synaptic plasticity” (Markram & Tsodyks, 1996, p. 807). LTP experiments tested only with low-frequency presynaptic inputs implicitly assume that these observations may be extrapolated to higher frequencies. Paired action-potential experiments by Markram and Tsodyks (1996) call into question the LTP gain-change hypothesis by demonstrating that adaptive changes in synaptic efficacy can depend dramatically on the frequency of the presynaptic test pulses used to probe these changes. In that preparation, following an interval of pre- and postsynaptic pairing, neocortical pyramidal neurons are seen to exhibit LTP, with the amplitude of the post-pairing response to a single test pulse elevated to 166% of the pre-pairing response. If LTP were a manifestation of a synaptic gain increase, the response to each higher-frequency test pulse would also be 166% of the pre-pairing response to the same presynaptic frequency. Although the Markram–Tsodyks data do show an amplified response to the initial spike in a test train ($EPSP_{init}$), the degree of enhancement of the stationary response ($EPSP_{stat}$) declines steeply as test pulse frequency increases (see Figure 1). In fact, post-pairing amplification of $EPSP_{stat}$ disappears altogether for 23 Hz test trains and then, remarkably, reverses sign, with test trains of 30–40 Hz producing post-pairing stationary response amplitudes that are less than 90% the size of pre-pairing amplitudes. Pairing is thus shown to induce a redistribution rather than a uniform enhancement of synaptic efficacy.

As Markram, Pikus, Gupta, and Tsodyks (1998) point out, redistribution of synaptic efficacy has profound implications for modeling as well as experimentation: “Incorporating frequency-dependent synaptic transmission into artificial neural networks reveals that the function of synapses within neural networks is exceedingly more complex than previously imagined” (p. 497). Neural modelers have long been aware that synaptic transmission may exhibit frequency dependence (Abbott, Varela, Sen, & Nelson, 1997; Carpenter & Grossberg, 1990; Grossberg, 1968), but most network models have not so far needed this feature to achieve their functional goals. Rather, the assumption that synaptic gains, or multiplicative weights, are fixed on the timescale of synaptic transmission has served as a useful cornerstone for models of adaptive neural processes and related artificial neural network systems. Even models that hypothesize synaptic frequency dependence would still typically have predicted the constant upper dashed line in Figure 1 (see section 2.1), rather than the change in frequency dependence

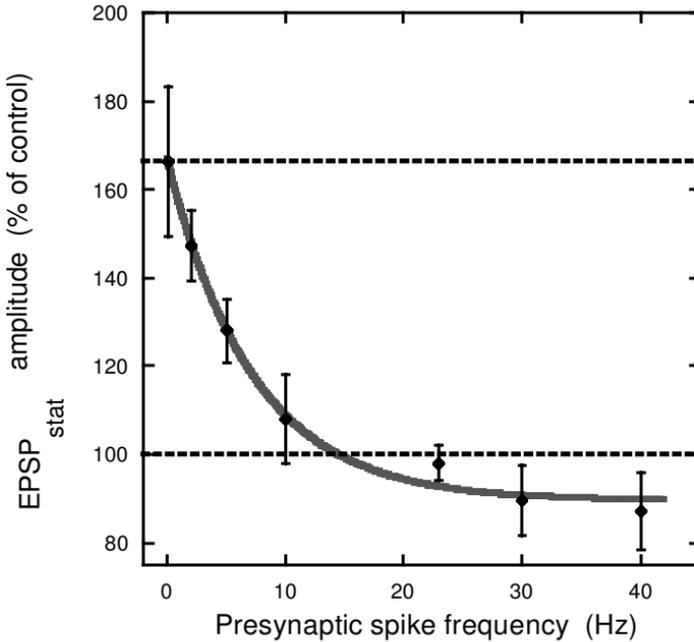


Figure 1: Relative amplitude of the stationary postsynaptic potential $EPSP_{stat}$ as a function of presynaptic spike frequency (I) (adapted from Markram & Tsodyks, 1996, Figure 3c, p. 809). In the Markram-Tsodyks pairing paradigm, sufficient current to evoke 4–8 spikes was injected, pre- and post-, for 20 msec; this procedure was repeated every 20 sec for 10 min. Data points show the $EPSP_{stat}$ after pairing as a percentage of the control $EPSP_{stat}$ before pairing, for $I = 2, 5, 10, 23, 30, 40$ Hz; plus the low-frequency “single-spike” point, shown as a weighted average of the measured data: 2×0.25 and 17×0.067 Hz. If pairing had produced no adaptation, $EPSP_{stat}$ would be a function of I that was unaffected by pairing, as represented by the lower dashed line (100% of control). If pairing had caused an increase in a gain, or multiplicative weight, then $EPSP_{stat}$ would equal the gain times a function of I , which would produce the upper dashed line (166% of control). Markram and Tsodyks fit their data with an exponential curve, approximately $(1 + 0.104[e^{-\frac{I-14.5}{7.23}}] - 1)100\%$, which crosses the neutral point at $I = 14.5$ Hz.

observed in the Markram-Tsodyks redistribution of synaptic efficacy (RSE) experiments.

A “bottom-up” modeling approach might now graft a new process, such as redistribution of synaptic efficacy, onto an existing system. While such a step would add complexity to the model’s dynamic repertoire, it may be difficult to use this approach to gain insight into the functional advantages of

the added element. Indeed, adding the Markram–Tsodyks effect to an existing network model of pattern learning would be expected to alter drastically the dynamics of input coding—but what could be the benefit of such an addition? A priori, such a modification even appears to be counterproductive, since learning in the new system would seem to reduce pattern discrimination by compressing input differences and favoring only low-frequency inputs.

A neural network model called distributed ART (dART) (Carpenter, 1996, 1997; Carpenter, Milenova, & Noeske, 1998) features RSE at the local synaptic level as a consequence of implementing system design goals at the pattern processing level. Achieving these global capabilities, not the fitting of local physiological data, was the original modeling goal. This “top-down” approach to understanding the functional role of learned changes in synaptic potentiation suggests by example how the apparently paradoxical phenomenon of RSE may actually be precisely the element needed to solve a critical pattern coding problem at a higher processing level.

The dART network seeks to combine the advantages of multilayer perceptrons, including noise tolerance and code compression, with the complementary advantages of adaptive resonance theory (ART) networks (Carpenter & Grossberg, 1987, 1993; Carpenter, Grossberg, & Reynolds, 1991; Carpenter, Grossberg, Markuzon, Reynolds, & Rosen, 1992). ART and dART models employ competitive learning schemes for code selection, and both are designed to stabilize learning. However, because ART networks use a classical steepest-descent paradigm called instar learning (Grossberg, 1972), these systems require winner-take-all coding to maintain memory stability with fast learning. A new learning law called the *distributed instar* (*dInstar*) (see section 2.1) allows dART code representations to be distributed across any number of network nodes.

The dynamic behavior of an individual dART synapse is seen in the context of its role in stabilizing distributed pattern learning rather than as a primary hypothesis. RSE here reflects a trade-off between changes in frequency-dependent and frequency-independent postsynaptic signal components, which support a trade-off between pattern selectivity and nonspecific path strengthening at the network level (see Figure 2). Models that implement distributed coding via gain adaptation alone tend to suffer catastrophic forgetting and require slow or limited learning. In dART, each increase in frequency-independent synaptic efficacy is balanced by a proportional decrease in frequency-dependent efficacy. With each frequency-dependent element assumed to be stronger than its paired frequency-independent element, the net result of learning is redistribution rather than nonspecific enhancement of synaptic efficacy. The system uses this mechanism to achieve the goal of a typical competitive learning scheme, enhancing network response to a given pattern while suppressing the response to mismatched patterns. At the same time, the dART network learning laws are designed to preserve prior codes. They do so by formally replacing the mul-

tiplicative weight with a dynamic weight (Carpenter, 1994), equal to the rectified difference between target node activation and an adaptive threshold, which embodies the long-term memory of the system. The dynamic weight permits adaptation only at the most active coding nodes, which are limited in number due to competition at the target field. Replacing the multiplicative weight with an adaptive threshold as the unit of long-term memory thus produces a coding system that may be characterized as quasi-localist (Carpenter, 2001) rather than localist (winner-take-all) or fully distributed. Adaptive thresholds, which are initially zero, become increasingly resistant to change as they become larger, a property that is essential for code stability.

Both ART and dART also employ a preprocessing step called complement coding (Carpenter, Grossberg, & Rosen, 1991), which presents to the learning system both the original external input pattern and its complement. The system thus allocates two thresholds for coding each component of the original input, a device that is analogous to on-cell/off-cell coding in the early visual system. Each threshold can only increase, and, as in the Markram-Tsodyks RSE experiments, each model neuron can learn to enhance only low-frequency signals. Nevertheless, by treating high-frequency and low-frequency component of the original input pattern symmetrically, complement coding allows the network to encode a full range of input features.

Elements of the dART network that are directly relevant to the discussion of Markram-Tsodyks RSE during pairing experiments will now be defined quantitatively.

2 Results

2.1 Distributed ART Model Equations. A simple, plausible model of synaptic transmission might hypothesize a postsynaptic depolarization T in response to a presynaptic firing rate I as $T = w^{eff} * I$, where the effective weight w^{eff} might decrease as the frequency I increases. Specifically, if:

$$T = [f(I) * w] * I,$$

where w is constant on a short timescale, then the ratio of T before versus after pairing would be independent of I . An LTP experiment that employs only single-shock test pulses relies on such a hypothesis for in vivo extrapolation and therefore implicitly predicts the upper dashed line in Figure 1. However, this synaptic computation is completely at odds with the Markram-Tsodyks measurements of adaptive change in frequency dependence.

The net postsynaptic depolarization signal T at a dART model synapse is a function of two formal components with dual computational properties: a frequency-dependent component S , which is a function of the current presynaptic input I , and a frequency-independent component Θ , which is independent of I . Both components depend on the postsynaptic voltage y

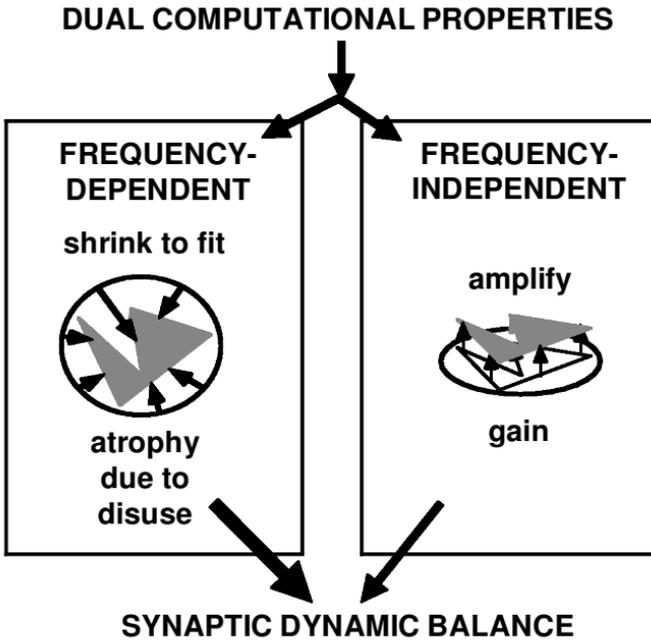


Figure 2: During dART learning, active coding nodes tend simultaneously to become more selective with respect to a specific pattern and to become more excitable with respect to all patterns. This network-level trade-off is realized by a synaptic-level dynamic balance between frequency-dependent and frequency-independent signal components. During learning, “disused” frequency-dependent elements, at synapses where the dynamic weight exceeds the input, are converted to frequency-independent elements. This conversion will strengthen the signal transmitted by the same path input (or by a smaller input), which will subsequently have the same frequency-dependent component but a larger frequency-independent component. Network dynamics also require that an active frequency-dependent (pattern-specific) component contribute more than the equivalent frequency-independent (nonspecific) component, which is realized as the hypothesis that parameter α is less than 1 in equation 2.1. This hypothesis ensures that among those coding nodes that would produce no new learning for a given input pattern, nodes with learned patterns that most closely match the input are most strongly activated.

and the adaptive threshold τ :

$$\begin{aligned}
 \text{Frequency dependent:} & \quad S = I \wedge [y - \tau]^+ \\
 \text{Frequency independent:} & \quad \Theta = y \wedge \tau \\
 \text{Total postsynaptic signal:} & \quad T = S + (1 - \alpha)\Theta = I \wedge [y - \tau]^+ \\
 & \quad \quad \quad + (1 - \alpha)y \wedge \tau.
 \end{aligned} \tag{2.1}$$

In 2.1, $a \wedge b \equiv \min\{a, b\}$ and $[a]^+ \equiv \max\{a, 0\}$. Parameter α is assumed to be between 0 and 1, corresponding to the network hypothesis that the pattern-specific component contributes more to postsynaptic activation than the nonspecific component, all other things being equal. The dynamic weight, defined formally as $[y - \tau]^+$, specifies an upper bound on the size of S ; for smaller I , the frequency-dependent component is directly proportional to I . Note that this model does not assign a specific physiological interpretation to the postsynaptic signal T . In particular, T cannot simply be proportional to the transmitted signal, since T does not equal 0 when $I = 0$.

The adaptive threshold τ , initially 0, increases monotonically during learning, according to the dInstar learning law:

$$\begin{aligned} \text{dInstar: } \quad \frac{d}{dt} \tau &= [[y - \tau]^+ - I]^+ \\ &= [y - \tau - I]^+. \end{aligned} \quad (2.2)$$

The distributed instar represents a principle of atrophy due to disuse, whereby a dynamic weight that exceeds the current input “shrinks to fit” that input (see Figure 2). When the coding node is embedded in a competitive network, the bound on total network activation across the target field causes dynamic weights to impose an inherited bound on the total learned change any given input can induce, with fast as well as slow learning. Note that τ remains constant if y is small or τ is large and that

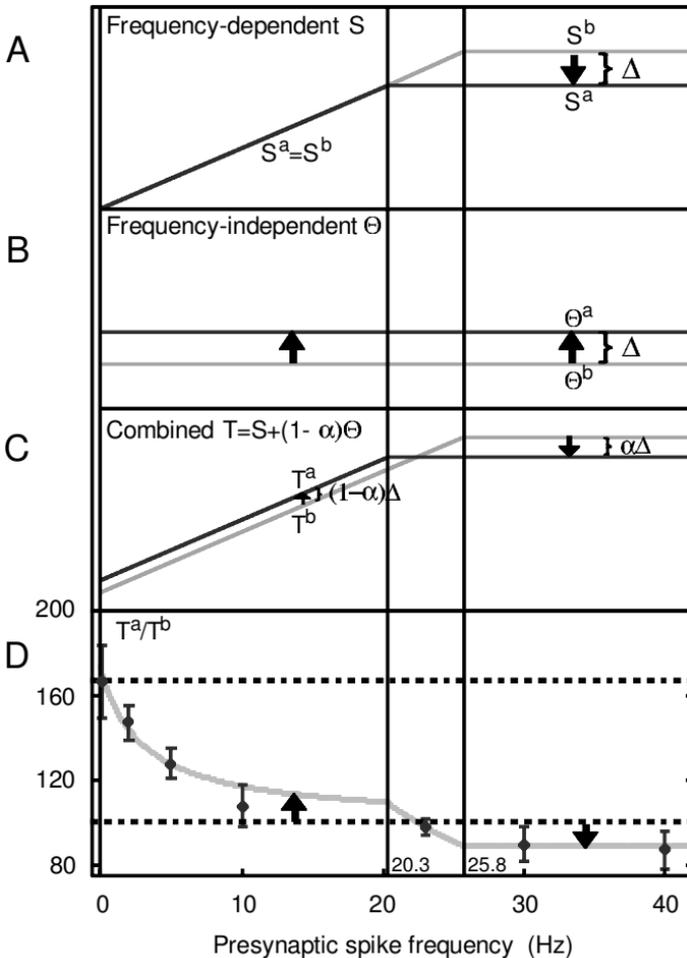
$$\begin{aligned} \frac{d}{dt} \tau &= [y - \tau]^+ - [y - \tau]^+ \wedge I \\ &= y - y \wedge \tau - [y - \tau]^+ \wedge I \\ &= y - \Theta - S. \end{aligned}$$

When a threshold increases, the frequency-independent, or nonspecific, component Θ (see equation 2.1) becomes larger for all subsequent inputs, but the input-specific component S becomes more selective. For a high-frequency input, a nonspecifically increased component is neutralized by a decreased frequency-dependent component. The net computational effect of a threshold increase (e.g., due to pairing) is an enhancement of the total signal T subsequently produced by small presynaptic inputs, but a smaller enhancement, or even a reduction, of the total signal produced by large inputs.

2.2 Distributed ART Model Predicts Redistribution of Synaptic Efficacy. Figure 3 illustrates the frequency-dependent and frequency-independent components of the postsynaptic signal T and shows how these two competing elements combine to produce the change in frequency dependence observed during pairing experiments. In this example, model elements, defined by equation 2.1, are taken to be piecewise linear, although

this choice is not unique. In fact, the general dART model allows a broad range of form factors that satisfy qualitative hypotheses. The model presented here has been chosen for minimality, including only those components needed to produce computational capabilities, and for simplicity of functional form.

Throughout, the superscript *b*(before) denotes values measured before the pairing experiment, and the superscript *a*(after) denotes values measured after the pairing experiment. The graphs show each system variable as a function of the presynaptic test frequency (*I*). Variable *I* is scaled by a factor (\bar{I} Hz), which converts the dimensionless input (see equation 2.1) to frequency in the experimental range. The dimensionless model input corresponds to the experimental test frequency divided by \bar{I} .



In the dART network, postsynaptic nodes are embedded in a field where strong competition typically holds a pattern of activation as a working memory code that is largely insensitive to fluctuations of the external inputs. When a new input arrives, an external reset signal briefly overrides internal competitive interactions, which allows the new pattern to determine its own unbiased code. This reset process is modeled by momentarily setting all postsynaptic activations $y = 1$. The resulting initial signals T then lock in the subsequent activation pattern, as a function of the internal dynamics of the competitive field. Thereafter, signal components S and Θ depend on y , which is small at most nodes due to normalization of total activation across the field. The Markram-Tsodyks experiments use isolated cells, so network

Figure 3: *Facing page.* dART model and Markram-Tsodyks data. (A) The dART postsynaptic frequency-dependent component S increases linearly with the presynaptic test spike frequency I , up to a saturation point. During pairing, the model adaptive threshold τ increases, and the saturation point of the graph of S is proportional to $(1 - \tau)$. The saturation point therefore declines as the coding node becomes more selective. Pairing does not alter the frequency-dependent response to low-frequency inputs: $S^a = S^b$ for small I . For high-frequency inputs, S^a is smaller than S^b by a quantity Δ , which is proportional to the amount by which τ has increased during pairing. (B) The dART frequency-independent component Θ , which is a constant function of the presynaptic input I , increases by Δ during pairing. (C) Combined postsynaptic signal $T = S + (1 - \alpha)\Theta$, where $0 < \alpha < 1$. At low presynaptic frequencies, pairing causes T to increase ($T^a = T^b + (1 - \alpha)\Delta$), because of the increase in the frequency-independent signal component Θ . At high presynaptic frequencies, pairing causes T to decrease ($T^a = T^b - \alpha\Delta$). (D) For presynaptic spike frequencies below the post-pairing saturation point of S^a , T^a is greater than T^b . For frequencies above the pre-pairing saturation point of S^b , T^a is less than T^b . The interval of intermediate frequencies contains the neutral point where $T^a = T^b$.

Parameters for the dART model were estimated by minimizing the chi-squared (Press, Teukolski, Vetterling, & Flannery, 1994) statistic: $\chi^2 = \sum_{i=1}^N \left(\frac{y_i - \hat{y}_i}{\sigma_i} \right)^2$, where y_i and σ_i are the mean value and standard deviation of the i th measurement point, respectively, while \hat{y}_i is the model's prediction for that point. Four parameters were used: threshold before pairing ($\tau^b = 0.225$), threshold after pairing ($\tau^a = 0.39$), a presynaptic input scale ($\bar{I} = 33.28$ Hz), and the weighting coefficient ($\alpha = 0.6$), which determines the contribution of the frequency-dependent component S relative to the frequency-independent component Θ . The components of the dimensionless postsynaptic signal $T = S + (1 - \alpha)\Theta$, for a system with a single node in the target field ($y = 1$), are $S = (I/\bar{I}) \wedge (1 - \tau)$ and $\Theta = \tau$. The dART model provides a good fit of the experimental data on changes in synaptic frequency dependence due to pairing ($\chi^2(3) = 1.085$, $p = 0.78$).

properties are not tested, and Figures 3 and 4 plot dART model equations 2.1 with $\gamma = 1$.

Figure 3A shows the frequency-dependent component of the postsynaptic signal before pairing (S^b) and after pairing (S^a). The frequency-dependent component is directly proportional to I , up to a saturation point, which is proportional to $(1 - \tau)$. Tsodyks and Markram (1997) have observed a similar phenomenon: "The limiting frequencies were between 10 and 25 Hz. ... Above the limiting frequency the average postsynaptic depolarization from resting membrane potential saturates as presynaptic firing rates increase" (p. 720). The existence of such a limiting frequency confirms a prediction of the phenomenological model of synaptic transmission proposed by Tsodyks and Markram (1997), as well as the prediction of distributed ART (Carpenter, 1996, 1997). The dART model also predicts that pairing lowers the saturation point as the frequency-dependent component becomes more selective.

Figure 3B illustrates that the frequency-independent component in the dART model is independent of I and that it increases during training. Moreover, the increase in this component ($\Delta \equiv \Theta^a - \Theta^b = \tau^a - \tau^b$) balances the decrease in the frequency-dependent component at large I , where $S^b - S^a = \Delta$.

Figure 3C shows how the frequency-dependent and frequency-independent components combine in the dART model to form the net postsynaptic signal T . Using the simplest form factor, the model synaptic signal is taken to be a linear combination of the two components: $T = S + (1 - \alpha)\Theta$ (see equation 2.1). For small I (below the post-pairing saturation point of S^a), pairing causes T to increase, since S remains constant and Θ increases. For large I (above the pre-pairing saturation point of S^b), pairing causes T to decrease: because $(1 - \alpha) < 1$, the frequency-independent increase is more than offset by the frequency-dependent decrease. The neutral frequency, at which the test pulse I produces the same postsynaptic depolarization before and after pairing, lies between these two intervals.

Figure 3D combines the graphs in Figure 3C to replicate the Markram-Tsodyks data on changes in frequency dependence, which are redrawn on this plot. The graph of T^a/T^b is divided into three intervals, determined by the saturation points of S before pairing ($I = \bar{I}(1 - \tau^b) = 25.8$ Hz) and after pairing ($I = \bar{I}(1 - \tau^a) = 20.3$ Hz) (see Figure 3A). The neutral frequency lies between these two values.

System parameters of the dART model were chosen, in Figure 3, to obtain a quantitative fit to the Markram-Tsodyks (1996) results concerning changes in synaptic potentiation, before pairing versus after pairing. In that preparation, the data exhibit the reversal phenomenon where, for high-frequency test pulses, post-pairing synaptic efficacy falls below its pre-pairing value. Note that dART system parameters could also be chosen to fit data that might show a reduction, but not a reversal, of synaptic efficacy. This might occur, for example, if the test pulse frequency of the theoretical reversal point

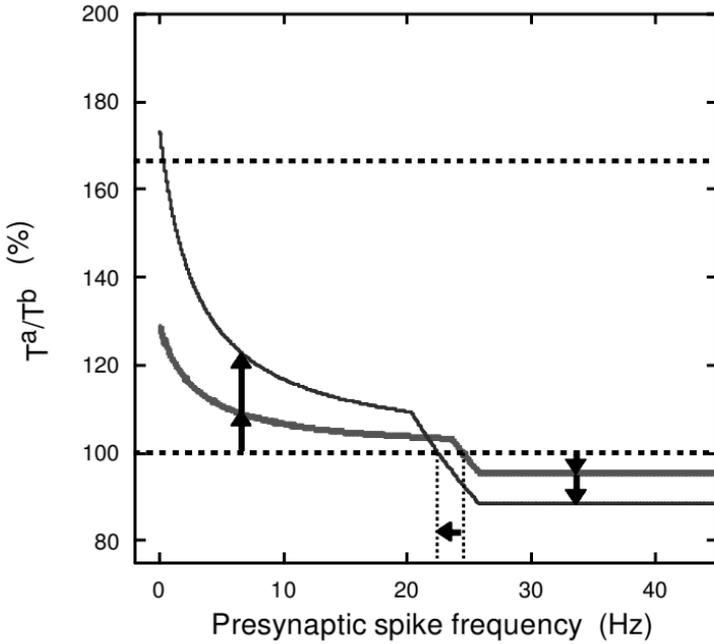


Figure 4: Transitional RSE ratios. The dART model predicts that if postsynaptic responses were measured at intermediate numbers of pairing intervals, the location of the neutral point, where pairing leaves the ratio T^a/T^b unchanged, would move to the left on the graph. That is, the cross-over point would occur at lower frequencies I .

were beyond the physiological range. Across a wide parameter range, the qualitative properties illustrated here are robust and intrinsic to the internal mechanisms of the dART model.

Analysis of the function $T = S + (1 - \alpha)\Theta$ suggests how this postsynaptic signal would vary with presynaptic spike frequency if responses to test pulses were measured at transitional points in the adaptation process (see Figure 4), after fewer than the 30 pairing intervals used to produce the original data. In particular, the saturation point where the curve modeling T^a/T^b flattens out at high presynaptic spike frequency depends on only the state of the system before pairing, so this location remains constant as adaptation proceeds. On the other hand, as the number of pairing intervals increases, the dART model predicts that the neutral point, where the curve crosses the 100% line and $T^a = T^b$, moves progressively to the left. That is, as the degree of LTP amplification of low-frequency inputs grows, the set of presynaptic frequencies that produce any increased synaptic efficacy shrinks.

3 Discussion

3.1 Redistribution of Synaptic Efficacy Supports Stable Pattern Learning. Markram and Tsodyks (1996) report measurements of the initial, transient, and stationary components of the excitatory postsynaptic potential in neocortical pyramidal neurons, bringing to a traditional LTP pairing paradigm a set of nontraditional test stimuli that measure postsynaptic responses at various presynaptic input frequencies. The dART model analysis of these experiments focuses on how the stationary component of the postsynaptic response is modified by learning. This analysis places aspects of the single-cell observations in the context of a large-scale neural network for stable pattern learning.

While classical multiplicative models are considered highly plausible, having succeeded in organizing and promoting the understanding of volumes of physiological data, nearly all such models failed to predict adaptive changes in frequency dependence. Learning laws in the dART model operate on a principle of atrophy due to disuse, which allows the network to mold parallel distributed pattern representations while protecting stored memories. The dynamic balance of competing postsynaptic computational components at each synapse dynamically limits memory change, enabling stable fast learning with distributed code representations in a real-time neural model. To date, other competitive learning systems have not realized this combination of computational capabilities.

Although dART model elements do not attempt to fit detailed physiological measurements of synaptic signal components, RSE is the computational element that sustains stable distributed pattern coding in the network. As described in section 2.2, the network synapse balances an adaptive increase in a frequency-independent component of the postsynaptic signal against a corresponding frequency-dependent decrease. Local models of synaptic transmission designed to fit the Markram-Tsodyks data are reviewed in section 3.2. These models do not show, however, how adaptive changes in frequency dependence might be implemented in a network with useful computational functions.

In the dART model, the synaptic location of a frequency-independent bias term, realized as an adaptive threshold, leads to dual postsynaptic computations that mimic observed changes in postsynaptic frequency dependence, before versus after pairing. However, producing this effect was not a primary design goal; in fact, model specification preceded the data report. Rather, replication of certain aspects of the Markram-Tsodyks experiments was a secondary result of seeking to design a distributed neural network that does not suffer catastrophic forgetting. The dInstar learning law (see equation 2.2) allows thresholds to change only at highly active coding nodes. This rule stabilizes memory because total activation across the target field is assumed to be bounded, so most of the system's memory traces remain constant in response to a typical input pattern. Defining

long-term change in terms of dynamic weights thus allows significant new information to be encoded quickly at any future time, but also protects the network's previous memories at any given time. In contrast, most neural networks with distributed codes suffer unselective forgetting unless they operate with restrictions such as slow learning.

The first goal of the dART network is the coding process itself. In particular, as in a typical coding system, two functionally distinct input patterns need to be able to activate distinct patterns at the coding field. The network accomplishes this by shrinking large dynamic weights just enough to fit the current pattern (see Figure 2). Increased thresholds enhance the net excitatory signal transmitted by this input pattern to currently active coding nodes because learning leaves all frequency-dependent responses to this input unchanged while causing frequency-independent components to increase wherever thresholds increase. On the other hand, increased thresholds can depress the postsynaptic signal produced by a different input pattern, since a higher threshold in a high-frequency path would now cause the frequency-dependent component to be depressed relative to its previous size. If this depression is great enough, it can outweigh the nonspecific enhancement of the frequency-independent component. Local RSE, as illustrated in Figure 3, is an epiphenomenon of this global pattern learning dynamic.

A learning process represented as a simple gain increase would only enhance network responses. Recognizing the need for balance, models dating back at least to the McCulloch-Pitts neuron (McCulloch & Pitts, 1943) have included a nodal bias term. In multilayer perceptrons such as back-propagation (Rosenblatt, 1958, 1962; Werbos, 1974; Rumelhart, Hinton, & Williams, 1986), a single bias weight is trained along with all the pattern-specific weights converging on a network node. The dART model differs from these systems in that each synapse includes both frequency-dependent (pattern-specific) and frequency-independent (nonspecific bias) processes. All synapses then contribute to a net nodal bias. The total increased frequency-independent bias is locally tied to increased pattern selectivity. Although the adaptation process is unidirectional, complement coding, by representing both the original input pattern and its complement, provides a full dynamic range of coding computations.

3.2 Local Models of the Markram-Tsodyks Data. During dInstar learning, the decrease in the frequency-dependent postsynaptic component S balances the increase in the frequency-independent component Θ . These qualitative properties subserve necessary network computations. However, model perturbations may have similar computational properties, and system components do not uniquely imply a physical model. Models that focus more on the Markram-Tsodyks paradigm with respect to the detailed biophysics of the local synapse, including transient dynamics, are now reviewed.

In the Tsodyks-Markram (1997) model, the limiting frequency, beyond which $EPSP_{stat}$ saturates, decreases as a depletion rate parameter U_{SE} (utilization of synaptic efficacy) increases. In this model, as in dART, pairing lowers the saturation point (see Figure 3C). Tsodyks and Markram discuss changes in presynaptic release probabilities as one possible interpretation of system parameters such as U_{SE} .

Abbott et al. (1997) also model some of the same experimental phenomena discussed by Tsodyks and Markram, focusing on short-term synaptic depression. In other model analyses of synaptic efficacy, Markram, Pikus, Gupta, & Tsodyks (1998) and Markram, Wang, and Tsodyks (1998) add a facilitating term to their 1997 model in order to investigate differential signaling arising from a single axonal source. Tsodyks, Pawelzik, and Markram (1998) investigate the implications of these synaptic model variations for a large-scale neural network. Using a mean-field approximation, they “show that the dynamics of synaptic transmission results in complex sets of regular and irregular regimes of network activity” (p. 821). However, their network is not constructed to carry out any specified function; neither is it adaptive. Tsodyks et al. (1998) conclude “An important challenge for the proposed formulation remains in analyzing the influence of the synaptic dynamics on the performance of other, computationally more instructive neural network models. Work in this direction is in progress” (pp. 831–832). Because the Markram-Tsodyks RSE data follow from the intrinsic functional design goals of a complete system, the dART neural network model begins to meet this challenge.

Acknowledgments

This research was supported by grants from the Air Force Office of Scientific Research (AFOSR F49620-01-1-0397), the Office of Naval Research and the Defense Advanced Research Projects Agency (ONR N00014-95-1-0409 and ONR N00014-1-95-0657), and the National Institutes of Health (NIH 20-316-4304-5).

References

- Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275*, 220–224.
- Carpenter, G. A. (1994). A distributed outstar network for spatial pattern learning. *Neural Networks*, *7*, 159–168.
- Carpenter, G. A. (1996). Distributed activation, search, and learning by ART and ARTMAP neural networks. In *Proceedings of the International Conference on Neural Networks (ICNN'96): Plenary, Panel and Special Sessions* (pp. 244–249). Piscataway, NJ: IEEE Press.
- Carpenter, G. A. (1997). Distributed learning, recognition, and prediction by ART and ARTMAP neural networks. *Neural Networks*, *10*, 1473–1494.

- Carpenter, G. A. (2001). Neural network models of learning and memory: Leading questions and an emerging framework. *Trends in Cognitive Sciences*, *5*, 114–118.
- Carpenter, G. A., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, *37*, 54–115.
- Carpenter, G. A., & Grossberg, S. (1990). ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks*, *3*, 129–152.
- Carpenter, G. A., & Grossberg, S. (1993). Normal and amnesic learning, recognition, and memory by a neural model of cortico-hippocampal interactions. *Trends in Neuroscience*, *16*, 131–137.
- Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J. H., & Rosen, D. B. (1992). Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, *3*, 698–713.
- Carpenter, G. A., Grossberg, S., & Reynolds, J. H. (1991). ARTMAP: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, *4*, 565–588.
- Carpenter, G. A., Grossberg, S., & Rosen, D. B. (1991). Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks*, *4*, 759–771.
- Carpenter, G. A., Milenova, B. L., & Noeske, B. W. (1998). Distributed ARTMAP: A neural network for fast distributed supervised learning. *Neural Networks*, *11*, 793–813.
- Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proc. Natl. Acad. Sci. USA*, *60*, 758–765.
- Grossberg, S. (1972). Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik*, *10*, 49–57.
- Markram, H., Pikus, D., Gupta, A., & Tsodyks, M. (1998). Potential for multiple mechanisms, phenomena and algorithms for synaptic plasticity at single synapses. *Neuropharmacology*, *37*, 489–500.
- Markram, H., & Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*, *382*, 807–810.
- Markram, H., Wang, Y., & Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc. Natl. Acad. Sci. USA*, *95*, 5323–5328.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, *9*, 127–147.
- Press, W. H., Teukolski, S. A., Vetterling, W. T., & Flannery, B. P. (1994). *Numerical recipes in C: The art of scientific computing* (2nd ed.). Cambridge: Cambridge University Press.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*, 386–408.
- Rosenblatt, F. (1962). *Principles of neurodynamics*. Washington, DC: Spartan Books.

- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructures of cognitions* (Vol. 1, pp. 318–362). Cambridge, MA: MIT Press.
- Tsodyks, M., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA*, *94*, 719–723.
- Tsodyks, M., Pawelzik, K., & Markram, H. (1998). Neural networks with dynamic synapses. *Neural Computation*, *10*, 821–835.
- Werbos, P. J. (1974). *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. Unpublished doctoral dissertation, Harvard University.

Received July 28, 1999; accepted July 5, 2001.