Unsupervised Learning of Patterns Using Multilayer Reverberating Configurations of Polychronous Wavefront Computation

Fred Highland\textsuperscript{a*}, Corey Hart\textsuperscript{b}

\textsuperscript{a}University of Maryland Baltimore County, Baltimore, MD, USA
\textsuperscript{b}nNetworx, LLC, Philadelphia, Pennsylvania, USA

Abstract

Polychronous Wavefront Computation (PWC) is an abstraction of spiking neural networks that has been shown to be capable of basic computational functions and simple pattern recognition through multilayer configurations. The objective of this work is to apply unsupervised learning methods to multilayer PWC configurations to improve performance providing a basis for more advanced applications and deep learning. Previous work on defining multilayer PWC configurations is extended by applying biologically inspired learning methods to dynamically suppress unneeded transponders and improve configuration performance. Simple learning approaches based on concepts from spike-timing-dependent plasticity and potentiation decay models are adapted to PWC transponders and combined with training sequences to optimize the transponder configurations for recognition. Learning is further enhanced by configuring transponders in recurrent structures to activate hidden layer transponders creating reverberations that reinforce learning. A means to classify multiple input patterns into general concepts is also introduced to further enhance the recognition capabilities of the configurations. The concepts are experimentally validated and analyzed through application to a 7-segment display digit recognition problem showing that the approach can improve PWC configuration performance and reduce complexity.

© 2016 The Authors. Published by Elsevier B.V.

Keywords: polychronous wavefront computation; pattern recognition; multilayer neural networks; spiking neural networks; unsupervised learning
1. Introduction

Polychronous Wavefront Computation (PWC)\(^1\) was proposed as an abstraction of the spiking neural network paradigm\(^2\) based on temporal and spatial patterns of wavefront activity in a pulse propagating media and their interaction with transponders. It provides a potentially practical model for implementing neuromorphic computing systems because of its simple design that eliminates the need for direct connections between computational units reducing the complexity of large scale implementation.

PWC consists of a configuration of transponders (or nodes) that may emit wavefronts and sense the wavefronts from other nodes. When a node senses two (or more) wavefronts simultaneously, it emits a new wavefront. Figure 1 shows a simple example of PWC interactions. Two nodes (A and B), activated at different times, emit wavefronts (circles) that trigger a third node (C) at the intersection of those wavefronts. The possible wavefront intersections over time form a hyperbola defined by the radii of the wavefronts generated by the two nodes, which in turn are defined by the positions and relative activation times of the nodes. In this example, node A was activated before node B resulting in a larger diameter wavefront. The possible intersection points for the A and B activation are defined by the solid hyperbola. Other possible intersections, defined by different relative activation times of A and B, are shown as dashed hyperbolae.

Izhikevich and Hoppensteadt\(^1\) have defined small configurations of PWC transponders that can perform signal analysis and logical operations. The mathematical properties of simple PWC configurations have been explored in detail and some conceptual sensor configurations suggested\(^3\). It has also been suggested that numerical programming methods could be used to configure PWC transponders\(^4\). In addition, recent work has presented a method to define multilayer configurations to perform pattern recognition with PWC\(^5\) but has not shown how to optimize these configurations and improve their recognition performance. Improved performance would allow exploration of complex PWC properties and provide the basis for researching advanced recognition applications such as deep learning.

The objective of this work is to extend the prior work on multilayer PWC configurations to provide a basis for exploring the application of PWC to more complex problems. The paper presents an approach for improving the performance of multilayer PWC configurations through application of biologically inspired mechanisms. It will first overview the multilayer PWC configuration for pattern recognition as a basis for the discussion. It will then introduce two unsupervised learning methods: Activation Threshold Adjustment, which decays or enhances node activation based on frequency of stimulus, and reverberation within the configuration which enhances the effect of Activation Threshold Adjustment. It will also introduce a design for associating multiple input patterns with a common concept to enhance the recognition capability of the configurations. It then discusses the application of the approach to a 7-segment display digit recognition problem and the results obtained. It concludes with a discussion of the characteristics of the proposed method and suggestions for further research.

2. Multilayer PWC configurations

A multilayer PWC configuration\(^5\) is an analog of the Multi-Layer Perceptron to perform classification of a set of inputs. It consists of an ensemble of PWC transponders (nodes) divided into a set of input nodes, representing pattern features, hidden nodes that combine subsets of the input features and output nodes to classify sets of features. The input nodes are defined to represent all features and values of the problem and are arranged such that three wavefront intersections for all possible triples of feature values will occur. Hidden nodes are then placed at the wavefront intersections for each feature value triple. Finally, output nodes are selected based on combinations of hidden nodes that cover sets or subsets of input feature values.

![Figure 1 - PWC Two Node/Wavefront Activation](image-url)
The method creates a PWC configuration that can classify input patterns based on wavefront propagation through the configuration. In order for the method to be generally applicable, all possible input patterns, hidden nodes and outputs must be defined as the results are not always known in advance. For most problems, the number of hidden and output nodes needed is a small subset of the total possible. Elimination of these unnecessary nodes would reduce the complexity and minimize unwanted activations improving both performance and quality of the results. While this can be done by pre-analysis of the configuration with respect to desired outputs or post-analysis of performance statistics based on the results of training cases, an unsupervised learning approach that can identify and suppress unneeded nodes during training is desirable.

3. Node Suppression using Unsupervised Learning

3.1. Activation Threshold Adjustment

One approach to automatically eliminate unneeded nodes is to include a method similar to synaptic potentiation decay to suppress little used nodes. Potentiation decay reduces the weight of neural synaptic links over time from lack of stimulation. When combined with positive training instances, this can provide a mechanism to suppress connections that are not needed. Explicit nodal connections do not exist in the PWC framework but a similar decay effect on the operation of the node can be achieved by applying the concept to the activation threshold.

In typical PWC transponder models, the activation threshold defines the discrete number of wavefront intersections that cause the transponder to activate producing an output signal (wavefront). The model used here defines a continuous threshold value representing accumulated activation charge with time dependent leakage allowing incremental adjustments over time. Miller and Jin suggest implementation of potentiation decay by rescaling synaptic weights by a constant, \( \beta \), applied at fixed time increments. As this rescaling is applied to all synapses that are summed and compared against a threshold value for activation, descaling the activation threshold will have the similar effect on the sensitivity of PWC transponders. The result is applied to the activation threshold \( \theta \) over a time interval \( t \) according to:

\[
\theta_{n+t} = \theta_n / \beta^t, \quad 0 < \beta < 1
\]  

The effect of the activation threshold adjustment is to raise the threshold slowly over time. Once the threshold exceeds then number of wavefronts triggering activation, the node is no longer activated by wavefront intersections and is suppressed.

The decay of node activation must be complemented with activation enhancement based on stimulation so that short term inactivity does not permanently suppress the node. This is achieved using a modification of the approach used by Spike-Timing-Dependent Plasticity (STDP). The approach is applied to reduce the membrane threshold toward a lower limit value based on the timing between stimuli to improve sensitivity according to:

\[
\theta_{n+1} = \theta_n - A_+ e^{-\Delta t / \tau_m} (\theta_n - \theta_{\text{min}})
\]  

Where:
- \( \theta_n \) – the current activation threshold
- \( A_+ \) – max potentiation modification
- \( \Delta t \) – the time since the last transponder activation
- \( \tau_m \) – the potentiation time constant
- \( \theta_{\text{min}} \) – the threshold lower limit

As with STDP, the choice of a larger \( \tau_m \) and/or a smaller \( \Delta t \) make repeated frequent stimulations much more effective than longer terms stimulus in enhancing the sensitivity of the node.

3.2. Reverberation

The suppression of unneeded nodes by Activation Threshold Adjustment is a slow process that may require many training cycles depending on the selection of \( \beta \). By adding special reverberation nodes configured to reactivate the
hidden nodes, a higher frequency of stimulation can be achieved increasing the rate of learning from each training case. The reverberation nodes are inserted between the hidden and output nodes to provide additional stimulation while retaining the output of the recognition process.

Configuring reverberation nodes to activate hidden nodes can be done as part of a chain of node activation (similar to synfire chains\(^6\)). At each stage, three nodes can be used to activate a node at the next stage but since a single node cannot continue wavefront propagation additional hidden nodes will be required. In theory, four hidden nodes can create three combinations of three nodes that can activate three nodes at a subsequent stage to continue propagation. However, the number of features represented by the four hidden nodes must be limited to the number of input features (or less) making this approach infeasible for many problems. Using five hidden nodes as a starting point can provide sufficient flexibility but requires finding hidden node combinations that both cover the desired input pattern feature set and produce feasible node triplet intersections. In addition to the complexity of defining each stage, configuring the chain of nodes to activate the originating hidden nodes is not possible in general because of the output positions are predetermined by the input nodes.

A simpler approach to create a reverberating chain of activation is to use co-located nodes (i.e., multiple nodes at the same location) with separate wavefront activation times for each node specified by a time delay. At the node level the concept in similar to axonal conduction delays in neurons, which have been shown to be part of the learning process\(^2\) but in this case it is used to target nodes for activation. The approach leverages two properties of the multilayer PWC configurations. 1) the number of feature combinations (and therefore hidden nodes) is relatively large providing multiple alternatives to represent a given set of features and 2) co-located nodes can operate as independent nodes if they include a unique wavefront time delay for the location to prevent simultaneous wavefront emissions. Using these properties, a small set of hidden nodes may be configured to activate multiple nodes at the next level by carefully selecting the activation timings. The position of the activated nodes depends on the activating timings selected providing a means to target nodes that can be used to active existing hidden nodes and create reverberation.

Figure 2 shows the simplest form of a reverberating configuration using the activation delay approach. Co-located hidden nodes h1-h3 create three wavefronts each that propagate to reverb nodes r1-r3 as indicated by the solid (green) arrows. Wavefronts are delayed such that wavefronts from each hidden node arrive at the targeted reverb node at the same time activating that node. The co-located reverb nodes then create wavefronts that target the original hidden nodes as indicated by the dashed (light blue) curved arrows. The reverb nodes can also be used to activate output nodes by adding co-located nodes targeting the appropriate output nodes.

The time delay required at each hidden node is defined from the initial activation time of the hidden node and the wavefront propagation time to the reverb nodes. The wavefront propagation time is dependent on the reverb node position. Since the activation delay approach provides significant flexibility in target (reverb) node position, this can be determined arbitrarily (a simple regular method such as arrangement in a circle is preferred) and is assumed to be predefined. Then general form of this relationship is:

\[
t_{0j} = t_{0i} + d_{hi,rj} + delay_{hi,rj},
\]

for \(i = 1 \ldots 3\) and \(j = 1 \ldots 3\)

Where
- \(t_{0i}\) is the first activation time for the hidden node \(hi\)
- \(t_{0j}\) is the activation time for the reverb node \(rj\)
- \(d_{hi,rj}\) is the distance from hidden node \(hi\) to reverb node \(rj\)
- \(delay_{hi,rj}\) is the delay time at node \(hi\) to stimulate node \(rj\)

For a given reverberation node \(rj\), \(t_{0j}\) must be the same for all \(i\) since the wavefronts much coincide. There are many possible values for \(delay_{hi,rj}\) that meet this constraint but the smallest delay values result from using the latest arrival time at \(rj\) which occurs at \(max(t_{0hi} + d_{hi,rj})\). For a given node \(hi\) this gives:

![Figure 2 - Simple Reverberating Configuration](image-url)
\[
\text{delay}_{hn,rj} = \max(t_0_{hi} + d_{hj,rj}) - (t_0_{hn} + d_{hn,rj})
\] (4)

In order for reverberation to occur, the time between activations of each hidden node resulting from the wavefront propagation through the reverberation node, must be the same for all hidden nodes. That is:

\[
\Delta t = t_1_{hi} - t_0_{hi} = t_0_{hi} + d_{hj,rj} + t_0_{hj} + d_{hj,rj} + \text{delay}_{rj,hi} - t_0_{hi} = 2d_{hj,rj} + \text{delay}_{hj,rf} + \text{delay}_{rj,hi}, \quad \text{for } i = 1 \ldots 3 \text{ and } j = 1 \ldots 3
\] (5)

The minimum and optimal value of \(\Delta t\) is the maximum traversal time for a wavefront for all nodes:

\[
\Delta t = \max(2d_{hj,rj} + \text{delay}_{hj,rj})
\] (6)

Making the delay calculation for each reverberation node:

\[
\text{delay}_{rj,hi} = \Delta t - (2d_{hj,rj} + \text{delay}_{hj,rj})
\] (7)

The equations above define the delay times to configure nodes that will reverberate indefinitely, which is undesirable from a performance and behavior perspective. To limit the reverberation process while retaining independence of node semantics, an activation charge depletion mechanism similar to synaptic fatigue is introduced. Synaptic fatigue occurs when the neurotransmitters at a synapse are temporarily depleted due to rapid firing rates. The effect stops synaptic signal transmission preventing the postsynaptic neuron from firing. After a short period of time the neurotransmitters are replenished and normal function is restored. The PWC transponder model used in this work simulates this behavior by representing activation as a depletion of an activation resource by a fixed amount, \(\text{Cost}_{\text{Activation}}\), for each activation. The activation resource is recharged, when not at its full capacity, at an exponential rate over time. If the discharge rate (activation frequency) over time is greater than the charge rate, the activation resource will eventually be insufficient to support activation until recharged. The resulting activation resource charging occurs at an exponential rate typical of chemical processes according to:

\[
\text{Charge}_{\text{Activation}} = \text{Charge}_{\text{Activation}} + (1 - \text{Charge}_{\text{Activation}})(1 - e^{-t/t_{\text{Charge}}})
\] (8)

Where
\[
\text{Charge}_{\text{Activation}} \quad \text{is the current activation resource level}
\]
\[
t \quad \text{is the time since last charge}
\]
\[
t_{\text{Charge}} \quad \text{is the activation charge time constant}
\]

3.3. Concept Learning

The previous section defined how to implement a simple configuration for reverberation that enhances node suppression to improve network performance. However, the simple multilayer PWC structure only provides an a priori mapping of a given input pattern to an output node identifying that pattern. No structure is provided for dynamically determining which output should be associated with an input pattern. Outputs which are associated with multiple patterns or pattern subsets will be referred to as concept outputs in the following discussion.

The idea behind a concept representation is to allow multiple input patterns to be associated with a common output that can be reinforced by training examples. This requires a mechanism similar to a logical or that can combine multiple alternative inputs into an output. To accomplish this, two additions are made to the pattern recognition structures already defined. First, a concept feature is added to indicate that a pattern is associated with a specific concept – effectively a positive reinforcement input. Second, the concepts are linked to the patterns using nodes with an activation threshold initially greater than three. The nodes will require more than three wavefronts for activation but since Activation Threshold Adjustment will reduce the threshold value with repeated stimulation, the activation levels will eventually be reduced to support a three node activation. This allows stimulation from the concept feature and the input pattern to be used to learn the pattern and concept association which will persist when the concept feature is removed.

The addition of the concept feature takes a form similar to other features. Each possible concept feature is combined with a pair of input features to form a concept hidden node that is triggered by three input features. There will be a
relatively large number of these \( (n_{\text{concepts}} \times \binom{n_{\text{features}}}{2}) \) but most of them will be unused and can be eliminated during configuration of the nodes or during training.

The wavefronts of concept hidden nodes are combined with the wavefronts of other hidden nodes for the input pattern to activate a concept association node (referred to as a concept node). The concept association node is co-located with the reverberation nodes but has an initial activation threshold \( \theta < 3 \) to force activation only by a four wavefront intersection. Computing the location of the concept node involves the solution of a four variable constraint equation. The formulation used defines the variables as the distance \( r \) from the hidden concept node and the three time delays from the other hidden nodes. The distance variable is fixed to roughly locate the node and the delays are determined by gradient descent. Three separate concept nodes are created for each possible output concept at different locations to allow them to be independently modified for each concept. After training, the unmodified nodes are suppressed and can be removed to optimize performance.

As mentioned above, the location of the reverberation nodes is unconstrained in the simple reverberation configuration. With the introduction of the concept nodes and their locations are changed to match the location for concept nodes determined by the four variable intersection solution. This is also necessary as the reverberation nodes need to support a different set of hidden nodes for each concept. The three input feature hidden nodes are common but the fourth concept hidden node is unique for each concept. Since the concept and reverberation nodes are co-located, reverberation and output activation will occur repeatedly each time a concept is detected. To prevent this from occurring, reverberation nodes are also defined with an initial activation threshold \( \theta < 3 \) but have a minimum activation threshold of 3 to prevent Activation Threshold Adjustment from reducing the \( \theta \) below 3. This allows reverberation to occur only when the concept hidden node is active.

During training, the concept nodes are activated by combining the hidden concept input and the other hidden nodes for the pattern. The activation, using Activation Threshold Adjustment enhanced by reverberation, reduces the concept node threshold such that the concept nodes can be activated by only the input hidden nodes.

The need to have multiple reverberation/concept nodes per pattern (one set for each concept) creates a potential problem with scalability. Using the Simple Reverberating Configuration of Figure 2, each input hidden node location requires \( 3 \times n_{\text{concepts}} \) co-located nodes. Since these nodes are activated by any occurrence of the three input features including patterns that are not of interest or associated with the concept, this can generate a large number of unneeded wavefronts impacting performance. To eliminate this problem, a layer of pattern nodes is inserted between the input hidden nodes and the reverberation/concept nodes. The allows the input hidden nodes to identify valid patterns before the pattern nodes generate a large number of wavefronts to match possible patterns.

The revised configuration is show in Figure 3. Co-located hidden nodes at locations h1-h3 propagate wavefronts to fixed location pattern nodes p1-p3 indicated by the solid (green) arrows. Nodes p1-p3 are activated only when the patterns match the features of h1, h2 and h3. Nodes p1-p3 propagate wavefronts to reverber nodes r1-r3 and concept nodes c1-c3 indicated by the solid (light green) arrows. Nodes r1-r3 and c1-c3 have unique locations per concept. The reverber nodes r1-r3 activate only when the pattern nodes and the appropriate hidden concept node (small dashed (orange) arrows) are activated. The reverber nodes then create wavefronts that target the original hidden nodes as indicated by the dashed (light blue) curved arrows. The concept nodes activate only when the pattern nodes and appropriate hidden concept node are activated reducing their threshold below 3 through repeated stimulation. The concept nodes create wavefronts that target the output.
nodes as indicated by the dotted (red) arrows. After sufficient training, the inputs stimulating only hidden nodes h1-h3 (and not the hidden concept) propagate through the pattern nodes to activate the output concept output because the concept node threshold has been reduced by Activation Threshold Adjustment.

4. Results

To test the methods described above, a simple 7-segment display digit recognition problem was implemented. 7-segment display devices have been used to display decimal digits on electronic devices since about 1910. They are driven by decoders that convert binary numbers to a decimal representation of seven illuminated segments. The problem used here is to input the seven segments (the standard notation is shown in Figure 4) and categorize them as one of ten outputs labeled 0-9. This problem represents a simple version of a multi-feature classification problem providing a practical basis for detailed experimentation and analysis.

A spreadsheet model was created to define the node configuration and analyze the solution. This provided for layout of the input, pattern and output nodes, computation of node locations, evaluating location node alternatives, computing activation delays and analysis of the problem.

The inputs were encoded as seven features of two mutually exclusive values each (segment “on” or “off”). Input feature and concept feature nodes were arranged in a circle of radius 100 units distributed evenly around the circle with a 6 unit time delay between each feature. The “off” value units are positioned halfway between the “on” value units and were activated at the same time as the “on” features since the values are mutually exclusive. The “on” value nodes were activated in a clockwise sequence and the “off” value nodes in a counterclockwise sequence to minimize position conflicts. Pattern nodes were arranged in a circle with a radius of 60-90 units and outputs in a circle of 170 units. Reverberation nodes and concept nodes positioned based at approximately 100-130 units from the center by adjusting the concept hidden node radius in the gradient descent formulation.

The resulting node configuration was run on a PWC Simulator built with NetLogo. The simulator implements PWC transponders in a dynamic agent-based discrete-event simulation that allows the wavefront and transponder actions to be visualized. The simulator supports Leaky Integrate-then-Fire semantics on the nodes allowing the activations to be tuned to adjust for node distances. The PWC simulator also includes extensions to support dynamic Activation Threshold Adjustment, node activation delays, and activation charge depletion to control reverberation.

Tests were run with a series of patterns to determine the effectiveness of learning methods. The test cases consisted of repeated series of positive test examples representing the digits 0-9 with correct concept identification, followed by negative (non-digit pattern cases) and recognition cases consisting of valid digit patterns with no concept inputs. Each digit pattern case was separated by 5000 time units to allow time for wavefront propagation and reverberation to complete. Training cases were run repeatedly until the outputs stabilized. After each series, parameter adjustments were made to improve the results. The parameters to define reverberation ($t_{charge}$ from (8) and $Cost_{Activation}$) were adjusted to provide 3-7 cycles of reverberation between hidden nodes and the corresponding reverberation nodes for each training case. The actual number of reverberations vary per pattern based on the timing of hidden node stimulation and the distance between hidden and reverberation nodes.

Figure 5 shows the number of active nodes and average activation threshold values obtained during the training process for the two different learning approaches. Only Hidden, Reverberation and Concept node data is shown as the other node types (Input, Pattern and Output) are not affected by the learning process. Figure 5a show the results from applying Activation Threshold Adjustment only and Figure 5b are the results from applying Activation Threshold Adjustment with reverberation. Both tests used the same data with parameters optimized for the each of the methods. The training datasets contained 11 samples that are presented over a period of 60,000 time units. The timescales and node count ranges are kept the same for purposes of comparison. The solid and heavy dashed lines represent the number of active nodes during the test case while the dotted and small dashed lines represent the average node threshold values. Since the
node threshold values are averages, their values change during the test and dip significantly when node suppression takes place (suppressed nodes are not included in the average causing the dip). Suppressed nodes are defined as any node with $\beta > 3$ except for Reverb nodes which only activate with four wavefronts and are suppressed when $\beta > 4$. It should also be noted that Concept nodes have an initial $\beta > 3$ and therefore are suppressed initially. Concept nodes that are activated during training have their $\beta$ values reduced by Threshold Adjustment and become active as seen in the dashed (red) lines at the bottom of the graphs.

Both learning approaches result in a suppression of unneeded nodes to improve performance but the learning process have different performance attributes. Threshold Adjustment without reverberation requires a rescaling parameter ($\beta$) closer to 1 which adjusts the node threshold more slowly requiring a longer time period to be effective. As shown in the figure, Threshold Adjustment alone requires twice as long (24,000 time units - 4 repetitions of the training data) while Threshold Adjustment with reverberation requires half the time (12,000 time units - 2 repetitions). In addition to being faster, the use of reverberation along with the lower rescaling value allows learning to be slightly more effective. As shown in Table 1, the approach with reverberation suppressed additional hidden nodes that did not directly contribute to the recognition of the pattern (they represented features that were already included in other hidden nodes participating in the recognition). The use of reverberation incurs an overhead of additional nodes to effect the feedback required. However, as seen in Figure 5b, the approach is effective in suppressing unneeded reverberation nodes.

Only positive test data is shown in this figure for clarity. Negative test cases were also run and show that a significant portion of the hidden nodes are rarely used and can be removed. The exact amount is problem dependent but for this application it is ~60% of the possible hidden nodes. Since all hidden nodes for a pattern must activate to stimulate the reverb and concept nodes for the pattern, negative patterns are not activated and will be suppressed by Activation Threshold Adjustment making significant reductions in the network.

Table 1 shows the summary statistics for the positive training cases for hidden, reverb and concept nodes. The data is further subcategorized into inactive and suppressed, Activation Threshold Adjustment with and without reverberation. Activated nodes were activated by the training data.

### Table 1: Effects of Learning Approaches on Positive Training Cases

<table>
<thead>
<tr>
<th>Node Type</th>
<th>Count</th>
<th>Sub-category</th>
<th>Active After Training (Activation Adjustment no Reverb)</th>
<th>Active After Training (Activation Adjustment with Reverb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hidden</td>
<td>158</td>
<td>Inactive</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suppressed</td>
<td>45</td>
<td>61</td>
</tr>
<tr>
<td>Reverb</td>
<td>1167</td>
<td>Inactive</td>
<td>1053</td>
<td>1053</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suppressed</td>
<td>1053</td>
<td>1053</td>
</tr>
<tr>
<td>Concept</td>
<td>330</td>
<td>Inactive</td>
<td>297</td>
<td>297</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suppressed</td>
<td>297</td>
<td>297</td>
</tr>
<tr>
<td>Totals</td>
<td>1655</td>
<td>Inactive</td>
<td>1395</td>
<td>1395</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suppressed</td>
<td>1395</td>
<td>1411</td>
</tr>
</tbody>
</table>

![Figure 5 - Active Node Counts and Average Activation Threshold During Training without and with Reverberation](image-url)
and have an activation count > 0 at the end of the test. Suppressed nodes those with $\theta > 3$ ($\theta > 4$ for Reverb nodes) at the end of the test.

The last two columns of Table 1 show the node counts resulting from applying the two learning methods. It can be seen that both methods are effective in identifying and suppressing unneeded nodes but the use of reverberation allows identification of little used nodes and works faster than Activation Threshold Adjustment alone.

What the statistics do not show is the qualitative effect to recognition resulting from the learning processes. Overall, suppressing unneeded nodes reduces the number of spurious activations that do not represent the input patterns reducing noise in the output and improving the overall recognition quality. It also reduces the overall activity in the configuration reducing computation needs.

5. Discussion

The sharpness of the learning activity (significant drops in measurements in the Figure 5 graphs) is largely controlled by the value of $\tau_m$. Smaller values of $\tau_m$ increase the sharpness to a point while larger values create more gradual convergence. The value for $\tau_m$ must be adjusted based on the time cycles between stimulations. Smaller values of $\tau_m$ produce higher quality results, removing more unneeded nodes while maintaining needed ones, while larger values with may remove less nodes and can remove needed ones degrading recognition quality. Reverberation improves performance by creating short stimulation cycles that have a larger effect on improving node sensitivity.

The use of multiple co-located nodes in the configuration would seem to be an impediment to the physical instantiation of this work. However, it’s representation as multiple nodes is only a convenience for simulation and testing. It can also be implemented as a single physical node with multiple time-delayed output behavior activated by the same initiating wavefronts.

The work to date has focused on positive training examples and positive stimulation to enhance the recognition capabilities of the configurations. It is well known that negative examples and inhibitory signals are an important part of the learning process. The concept of inhibitory wavefronts has been proposed but not incorporated into the research at this time.

The use of three wavefront intersections along with reverberation has some unexpected impacts on the resilience of the configurations. Although three wavefront intersection criteria has made the configurations much less sensitive to unintended activations and allowed higher node density, extraneous activations still occur. However, since node activations require three wavefront intersections, it is unlikely that a stray activation, even combining with other wavefronts, will intersect at the right time and location to cause a downstream activation. In addition, the use of reverberation establishes a chain of activity that is reinforced. Stray activations are not part of this chain and are omitted from subsequent reverberation cycles making them irrelevant. These phenomena have been observed in testing and usually causes some extra activations but has little to no effect on the outputs.

The use of reverberation and concept identification add significant capabilities to the configuration of PWC nodes but at a cost of significantly more nodes. The learning methods are effective in suppressing unneeded nodes but the extra nodes add complexity to the overall approach. This increases the effort to define the configuration since a four variable gradient descent must be performed each pattern-concept combination and increases the size of the networks. Simpler node configurations that utilize less nodes would make the use of PWC more practical in this application.

6. Summary and Directions for Further Work

The work defines an approach to creating recurrent behavior in multilayer PWC configurations that can be used to improve performance when combined with STDP and other biologically inspired behaviors. Through the implementation of Activation Threshold Adjustment combined with activation fatigue and variable activation delays, a condition of controlled reverberation is induced that can be used to dynamically change the sensitivity of the nodes to stimuli. Test results have shown that it is effective in suppressing unnecessary nodes and improving the recognition performance of the configuration. A mechanism is also introduced to allow the association of multiple patterns with a single output concept. This has been incorporated in the node configurations to more closely emulate multi-layer perceptron behavior with PWC nodes.
The approach used here to incorporate reverberation in PWC configurations creates a form of recurrent behavior. The looping mechanism between hidden and reverberation nodes is a basic behavior required for recurrent designs enabling implementation other recurrent architectures such as Long Short Term Memory networks\textsuperscript{13}.

The work provides a basis to explore more advanced research topics in the area of PWC configurations including:

- Development of learning approaches to determine output and even hidden node locations
- Development of simplified configurations to reduce the number of nodes required for recognition
- Explorations of the use of inhibitory signals with PWC models
- Construction of more complex configurations such as recurrent structures
- Application to other problems such as approximate classification and deep learning approaches.

The work presented extends previous work that defines multilayer PWC configurations by adding unsupervised learning approaches to improve overall performance and defining a basic design that can create recurrent behavior and represent simple abstract concepts. This combination of capabilities provides a basis for the development of additional computational capabilities supporting further exploration of the use of the PWC framework in more complex problems.

References