Characterizing Computation in Artificial Neural Networks by their Diclique Covers and Forman-Ricci Curvatures

Allen D. Parks

Abstract—The relationships between the structural topology of artificial neural networks, their computational flow, and their performance is not well understood. Consequently, a unifying mathematical framework that describes computational performance in terms of their underlying structure does not exist. This paper makes a modest contribution to understanding the structure-computational flow relationship in artificial neural networks from the perspective of the dicliques that cover the structure of an artificial neural network and the Forman-Ricci curvature of an artificial neural network’s connections. Special diclique cover digraph representations of artificial neural networks useful for network analysis are introduced and it is shown that such covers generate semigroups that provide algebraic representations of neural network connectivity.

Index Terms—Neural Networks, Band Semigroup, Computational Divergence, Computational Flow, Dicliques, Digraph, Forman-Ricci Curvature.

I. INTRODUCTION

In recent years there has been an exponential growth in research focused upon machine learning and its application to those classes of real world problems that cannot be easily addressed using traditional computer algorithms. Such classes include image analysis, e.g. [1], web searches, e.g. [2], content filtering and matching, e.g. [3], and speech recognition, e.g. [4]. Due to its robustness and adaptability, the artificial neural network (ANN) has emerged as a computational tool of choice for solving many problems in these domains. The especially powerful deep neural networks and their biologically inspired learning algorithms, e.g. [5], have consistently demonstrated state of the art performance when applied to real world tasks such as these.

The relationship between the structural topology of an ANN, its computational flow, and its performance is not well understood. Although graph theory and algebraic topology have been used with some success to analyze local and global network properties, e.g. [6,7], a unifying mathematical framework that describes the computational performance of an ANN in terms of its underlying structure does not yet exist. This paper makes a modest contribution to understanding the structure-computational flow relationship in ANNs from the perspective of the dicliques, e.g. [8], that cover the structure of an ANN and the Forman-Ricci (FR) curvature, e.g. [9,10], of an ANN’s connections.

These diclique covers provide digraphs useful for the analysis of ANNs and are also shown to generate band semigroups [11] in which the natural order of their idempotent elements provides lower semi-lattice representations [12] related to ANN connectivity.

Each connection in an ANN has an easily calculated FR curvature that is determined by the associated node and connection weights. FR curvature provides a measure of the total amount of “computational flow” – or more simply flow - through a network node and quantifies the divergence of the flow emerging from a connection – i.e., the more negative the curvature, the more the divergence of the emergent flow, and, intuitively, the more wide spread the influence the connection has on the network. Thus, both curvature and flow – along with the underlying connection topology – provide a characterization of the relationship between computation and the structure of an ANN.

When an ANN’s node weight, transfer function, and connection weight labels are ignored, the associated underlying connection topology of the ANN is assumed here to be that of an irreflexive binary relation \( R \subset N \times N \) on the ANN’s set \( N \) of nodes. Even though the numbers of nodes and connections in \( R \) can number in the tens of thousands, \( R \) can be algorithmically decomposed into maximally related subnetworks of the ANN called dicliques. Collections of dicliques which cover \( R \) provide a simpler system-like wiring diagram representation of the ANN’s underlying structure than the digraph representation of \( R \). An even simpler representation of the ANN’s structure is its diclique digraph which is easily obtained from its wiring diagram representation. Judicious assignments of FR curvatures and flows to the cover’s dicliques and wiring-diagram connections – or to the diclique digraph’s vertices and arcs – provide high level insights into the general importance of subnetwork and node contributions to an ANN computation. Furthermore, although not addressed here, time series of such diclique representations obtained as computations propagate through the layers of an ANN can provide additional insights into the relative importance of maximal subnetworks and nodes to ANN computation.

II. ARTIFICIAL NEURAL NETWORKS

An ANN is defined as an interconnected collection of nodes – i.e., artificial neurons. A connection (directed edge) between neurons – i.e., an artificial synapse - transmits a signal from the pre-synaptic neuron to the post-synaptic neuron which – in turn – aggregates and processes the incoming signals and then sends signals to post-synaptic neurons connected to it. The neurons and synapses of an ANN which has been configured to perform a class of tasks have assigned weights determined by the algorithm and

Published on October 30, 2016.

This project was supported by a short term summer Grant from the Naval Surface Warfare Center Dahlgren Division Naval Innovative Science and Engineering Program. A. D. Parks is with the Electromagnetic and Sensor Systems Department, 18444 Frontage Road Suite 327 Naval Surface Warfare Center Dahlgren Division, Dahlgren, VA 22448-5161 USA (allen.parks@navy.mil).

DOI: http://dx.doi.org/10.24018/ejers.2020.5.2.1689
input cases used to train it. These weights increase or decrease transmitted signal strengths. A neuron also has a state value determined by the aggregate pre-synaptic signals it receives and its transfer function. Application of a threshold value to a neuron’s aggregate input signal can be used to determine if the neuron sends a transfer function valued signal to post-synaptic neurons connected to it.

Neurons are typically organized into layers, each of which can perform different transformations on their input signals. The first layer is the input layer, the last layer is the output layer, and layers between the first and last are hidden layers. Input signals are transmitted from the first layer through the hidden layers and emerge as an output signal in the last layer. To be more precise, let the set $I = \{1, 2, \ldots, \ell\}$ index the layers and the set $J_1 = \{1, 2, \ldots, j_1\}, i \in I$, index the neurons within the $i$th layer. Thus, $1 \in I$ corresponds to the input layer, $\ell \in I$ corresponds to the output layer, and $I - \{1, \ell\}$ indexes the hidden layers. The collection of neurons in an ANN is the set $N = \{n_j; i \in I, j \in J_1\}$, where $n_j^i$ is the $j^{th}$ neuron in the $i^{th}$ layer. The associated collection of synapses are the ordered pairs $s_{jk}^i \equiv (n_j^i, n_k^{i+1})$ of an irreflexive binary relation $R \subset N \times N$, where $i \in I - \{\ell\}, j \in J_1$ and $k \in J_{i+1}$ (here only feed forward ANNs are addressed and only consecutive layers are connected). An ANN is fully connected if for each $i \in I - \{\ell\}$ there is an $s_{jk}^i$ for every $j \in J_1$ and $k \in J_{i+1}$.

Here an ANN computation is initiated when signals $x_j^i, k \in J_1$, are input simultaneously at layer 1. The computation proceeds though the network one layer at a time according to (1) and the result of the computation emerges at the output layer $i = \ell$ where each neuron $n_j^\ell$ manifests the output signal $y_j^\ell, j \in J_\ell$. As previously mentioned, a threshold value $\theta_j^1$ can also be applied to neuron $n_j^1$’s aggregate signal $y_j^1$ in order to determine if $n_j^1$ sends a signal to post-synaptic neurons connected to it.

### III. DICLIQUE REPRESENTATIONS OF ARTIFICIAL NEURAL NETWORKS

Recall that the underlying network structure of an ANN is assumed here to be an irreflexive binary relation $R$ on the ANN’s set of neurons. For realistically useful ANN architectures, $R$ will generally consist of extremely large numbers of neurons and synapses so that a diagrammatic representation of $R$ is typically not easily understood. Consequently, information about the global topology of an ANN; the importance of various regions within an ANN; and relationships between various internal regions of an ANN are difficult – if not impossible – to extract from these complicated representations. However, instead of attempting to analyze diagrammatic representations of $R$, issues such as these can be largely alleviated by analyzing diclique cover representations of $R$ obtained from a diclique decomposition of $R$.

A diclique of $R$ is a pair $(X, Y) \equiv D, X \subseteq N, Y \subseteq N$, such that $X \times Y \subseteq R$ and $X \subseteq X', Y \subseteq Y'$ implies $(X, Y) = (X', Y')$ [8]. Within the context of ANNs, a diclique defines the set of synapses associated with a maximally connected (bipartite) sub-network of $R$, where the first and second diclique entries $X$ and $Y$ are sets of pre- and post-synaptic neurons, respectively. Let $\mathcal{D}$ be the set of dicliques of $R$. A diclique cover of $R$ is a smallest subset $C \subseteq \mathcal{D}$ which satisfies the equality

$$R = \bigcup_{(X, Y) \in C} X \times Y, \quad (2)$$

In which case $C$ is said to cover the ANN associated with $R$ (note that in general an ANN can have many covers). The dicliques of an ANN cover provide a system level representation of an ANN that is much easier to analyze and understand than the standard diagrammatic representation of $R$. In such a system-level representation: (i) each element in $C$ corresponds to a maximally connected sub-network (or sub-system) of neurons in the associated ANN (or total system); and (ii) if $(X, Y) \in C, (U, V) \in C$, then for each $n \in Y \cap U$ there is a directed connection from $(X, Y)$ to $(U, V)$ which corresponds to a fundamental relationship between the sub-networks $(X, Y)$ and $(U, V)$ established by the neuron $n$. Thus, the essence of the connectivity occurring in an ANN is given by a simplified (relative to its $R$ diagram) system level diagram, or multi-digraph, derived from its diclique cover. The order of such a representation is $|C|$ and its size is the number of connections between dicliques.

As an example, observe from Fig. 1 that the set $\mathcal{D} = \{D_1, D_2, D_3, D_4, D_5\}$, where $D_1 = \{(n_2^1, n_1^1), (n_2^1, n_3^3), (n_4^1, n_3^3)\}$,
$D_2 = \{(n_1^2, n_3^2), (n_3^2, n_1^3)\}$,
$D_3 = \{(n_1^3, n_2^3, n_3^3), (n_1^3, n_2^3)\}$,
$D_4 = \{(n_2^3, n_3^3), (n_1^3, n_2^3)\}$, and
$D_5 = \{(n_1^3, n_2^3, n_3^3), (n_1^3, n_2^3)\}$, is a diclique decomposition of the $R$ underlying the simple ANN shown there. Given that $D$ is a smallest set such that (2) is satisfied when $C = D$, then $D$ covers the ANN. Its associated system level representation is given in Fig. 2. The boxes in Fig. 2 are labelled with the corresponding dicliques of $D$ and the directed connections between boxes are labelled with the neurons common to each diclique. Note the simplification introduced by using a diclique cover representation: whereas the $R$ for the ANN in Fig. 1 has 12 neurons and 24 synapses, its diclique cover representation in Fig. 2 has 5 dicliques (i.e., is order 5) and 9 connections (i.e., is size 9) and effectively contains the same connectivity and direction of computational flow as $R$ (line segments joining connections indicate that the associated neuron is shared by two dicliques).

![Fig. 2. A diclique cover representation of the ANN in Fig. 1](image)

It is interesting to consider the case where the ANN in Fig. 1 is fully connected such that the dicliques are

$D_1 = \{(n_1^2, n_3^2, n_1^3), (n_3^2, n_1^3, n_3^3)\}$,
$D_2 = \{(n_1^3, n_2^3, n_3^3, n_2^3), (n_1^3, n_2^3)\}$,
and
$D_3 = \{(n_1^3, n_2^3, n_3^3), (n_1^3, n_2^3)\}$

It is clear that these dicliques uniquely cover the associated underlying relational structure and yield the greatly simplified order 3 and size 7 diclique cover representation in Fig. 3.

![Fig. 3. Diclique cover representation of the ANN in Fig. 1 when it is fully connected](image)

This result can obviously be generalized for an arbitrary fully connected ANN. This generalization is stated for completeness and without proof as the following somewhat trivial theorem:

*Theorem 1.* The diclique cover of a fully connected $\ell$ layer ANN is unique and its diclique representation is order $\ell - 1$ with dicliques $D_1 = \{(n_1^2; j \in J_1), \{n_1^2; k \in J_{i+1}\}, i \in I - \{\ell\}$, and size $\sum_{i=1}^{\ell-1} |J_{i+1}|$ with connections $U_{i=1}^{\ell-1} \{n_1^i; k \in J_{i+1}\}$.

An even more simplified diclique representation of an ANN is given by its diclique digraph which has as its vertices the dicliques in the cover $C$ associated with the ANN. However, unlike the diclique cover representation which can have multiple connections between dicliques, a diclique digraph representation has at most one directed edge connecting dicliques, i.e., there is a directed edge from diclique $(X, Y) \in C$ to diclique $(U, V) \in C$ only if $Y \cap U \neq \emptyset$. The neurons in the set $Y \cap U$ label the directed edge. The order of a diclique digraph is $|C|$ and its size is its number of directed connections. Thus, a diclique digraph representation significantly reduces the size of a diclique cover representation of an ANN while effectively containing the same connectivity and flow direction information as an associated diclique cover representation. The size 5 diclique digraph for the ANN in Fig. 1 is shown in Fig. 4. Note that $D_1$ and $D_2$ serve as sources and $D_5$ is a sink in the diclique digraph in Fig. 4.

![Fig. 4. The diclique digraph representation of the ANN in Fig. 1](image)

It is easy to see that the diclique digraph representation of the fully connected ANN whose diclique cover representation is given by Fig. 3 is a dipath of length 2. The generalization of this to an arbitrary fully connected ANN is obvious and stated without proof as the following theorem:

*Theorem 2.* The diclique digraph representation of a fully connected $\ell$ layer ANN is a dipath of length $\ell - 1$ whose vertices are the dicliques $D_i = \{(n_1^2; j \in J_1), \{n_1^2; k \in J_{i+1}\}, i \in I - \{\ell\}$, and whose directed edge connecting $D_i$ and $D_{i+1}, i \in I - \{\ell - 1, \ell\}$, is labeled with the $|J_{i+1}|$ elements of the set $\{n_1^i+1; k \in J_{i+1}\}$.
IV. SYNAPTIC CURVATURE AND NEURON FLOW

Although the \( R \), the \( C \) cover, and cliquish digraph representations of an ANN provide connectivity and direction of flow information about the ANN, none of them provide a quantification of the importance and influence of neurons and synapses to ANN computation. Recently, FR curvature has been used to quantify the notions of flow and its divergence in directed networks [9,10]. Here FR curvature and flow is applied to ANN computation.

The FR curvature \( \gamma(s_{jk}) \) for a synapse \( s_{jk} \) in an ANN is given by the expression

\[
\gamma(s_{jk}) = \gamma^{i+1}_k \left( 1 - \sum_{t \in \Gamma_k} \frac{w_{jk}^t}{w_{kl}^t} \right), \quad i \in I - \{\ell\}, \tag{3}
\]

Note that here: (i) the FR curvature for \( s_{jk} \) is determined using only synapses originating at \( n_{k+1} \); and (ii) \( \gamma(s_{jk}) \) quantifies the divergence of flow emerging from \( s_{jk} \) and the more negative \( \gamma(s_{jk}) \) is, the more divergent the flow. This increase in divergence with increasing negative curvature is implicit in (3), where the sum effectively adds the (square roots of the) relative (to \( w_{jk}^t \)) contributions of the weights of all synapses incident on \( n_{k+1} \). Since the terms in the sum are positive, so is the sum. The greater the cardinality of the indexing set \( T_{k+1} \) – i.e., the larger the number of synapses incident on \( n_{k+1} \) – the greater the sum and – provided the sum exceeds unit value - the more divergent the flow. The following theorem and corollary illustrate this using the special cases of unit synaptic weights and full connectivity.

**Theorem 3.** If every synaptic weight in an ANN has unit value, then

\[
\gamma(s_{jk}) = \gamma^{i+1}_k(1 - |T_{k+1}^j|), \quad i \in I - \{\ell\},
\]

where

\[
\gamma^{i+1}_k = \sum_{j \in I_k} \gamma^j(y_i^j).
\]

**Proof.** \( w_{jk}^t = w_{kl}^t = 1 \Rightarrow \frac{w_{jk}^t}{w_{kl}^t} = 1 \Rightarrow \sum_{t \in \Gamma_k} \frac{w_{jk}^t}{w_{kl}^t} = |T_{k+1}^j|, i \in I - \{\ell\} \). Also, \( w_{jk}^t = w_{kl}^t = 1 \Rightarrow \gamma_{jk}^{i+1} = \sum_{j \in I_k} \gamma^j(y_i^j) = \sum_{j \in I_k} \gamma^j(y_i^j), i \in I - \{\ell\} \). 

**Corollary 4.** If the ANN is also fully connected, then

\[
\gamma(s_{jk}) = \gamma^{i+1}_k(1 - |I_{k+2}|), \quad i \in I - \{\ell\}.
\]

**Proof.** This result follows from the fact that for a fully connected ANN, \( T_{k+1}^j = I_{k+2} \).

FR curvature also quantifies the total computational flow \( \phi(n_j) \) through hidden layer neurons according to

\[
\phi(n_j) = \phi_{in}(n_j) - \phi_{out}(n_j), \quad i \in I - \{1, \ell\} \tag{4}
\]

where

\[
\phi_{in}(n_j^i) = \sum_{k \in E_{j+1}} \gamma(s_{kj}^{i-1})
\]

is the computational flow entering \( n_j^i \) and

\[
\phi_{out}(n_j^i) = \sum_{k \in E_{j+1}} \gamma(s_{jk}^{i+1})
\]

is the computational flow leaving \( n_j^i \).

The typical situation that occurs in highly connected ANNs is \( \phi_{in}(n_j^i), \phi_{out}(n_j^i) < 0 \). This has the interpretation that \( n_j^i \) enhances (reduces) the divergence of flow and hence enhances (reduces) the computational influence of the \( s_{kj}^{i-1} \) synapses since \( \phi(n_j^i) > 0 \) \( (\phi(n_j^i) < 0) \) when \( \phi_{out}(n_j^i) < 0 \) \( (\phi_{out}(n_j^i) > 0) \). From the perspective of curvature, this latter case follows from the facts that: (i) the more negative the total curvature (or the greater the total divergence of computational flow) of the synapses directed away from \( n_j^i \) relative to that associated with the synapses directed into (or convergent upon) \( n_j^i \), the more positive – or the greater - the total computational flow through \( n_j^i \); and (ii) the more negative the total curvature (or the greater the total divergence of computational flow) of the synapses convergent upon \( n_j^i \) relative to that associated with the synapses directed away from \( n_j^i \), the more negative – or the less - the total computational flow is through \( n_j^i \).

The following theorem illustrates the notion of computational flow for fully connected ANNs with unit synaptic weights.

**Theorem 5.** The computational flow through the hidden neuron \( n_j^i \) in a fully connected ANN in which every synapse has unit weight is

\[
\phi(n_j^i) = |I_{i-1}^j(1 - |I_{i+1}^j|) y_j^i - (1 - |I_{i+2}^j|) \sum_{k \in E_{j+1}} y_k^{i+1}.
\]

**Proof.** The result follows from the facts that (Corollary 4),

\[
\phi_{in}(n_j^i) = \sum_{k \in E_{j+1}} \gamma(s_{kj}^{i-1}) = \sum_{k \in E_{j+1}} y_k^{i+1} = |I_{i-1}^j(1 - |I_{i+1}^j|) y_j^i - (1 - |I_{i+2}^j|) \sum_{k \in E_{j+1}} y_k^{i+1}
\]

and

\[
\phi_{out}(n_j^i) = \sum_{k \in E_{j+1}} \gamma(s_{jk}^{i+1}) = (1 - |I_{i+2}^j|) \sum_{k \in E_{j+1}} y_k^{i+1}.
\]

**Corollary 6.** The computational flow through the hidden neuron \( n_j^i \) in a fully connected ANN in which every synapse has unit weight is positive when

\[
y_j^i > \frac{(1 - |I_{i+1}^j|)}{|I_{i-1}^j(1 - |I_{i+1}^j|)} \sum_{k \in E_{j+1}} y_k^{i+1},
\]

**Proof.** Obvious.

Note from this theorem that the typical case \( \phi_{out}(n_j^i) < \phi_{in}(n_j^i) < 0 \) for which \( \phi(n_j^i) > 0 \) occurs when \( |I_{i+1}^j|, |I_{i+2}^j| > 1 \) and

\[
|I_{i+2}^j| > 1 - |I_{i-1}^j(1 - |I_{i+1}^j|)} \sum_{k \in E_{j+1}} y_k^{i+1} > 1
\]

DOI: http://dx.doi.org/10.24018/ejers.2020.5.2.1689
V. THE FR GRAPHS OF AN ARTIFICIAL NEURAL NETWORK

The synaptic curvatures and neuron flows of an ANN can be used to generate three general types of - what will be called here - FR graphs of an ANN. The first type - i.e., Type 1 – is the labeled digraph representation of the underlying binary relation \( R \) associated with an ANN. Each neuron in \( R \) is labeled with the value of the flow through it and each synapse in \( R \) is labeled with the value of its FR curvature. Although Type 1 FR graphs provide very detailed quantitative curvature-flow-topology descriptions of ANN’s, they are complicated to interpret and their analytical value is likely limited to the simplest ANNs.

A Type 2 FR graph of an ANN is its diclique cover representation in which each diclique is labeled with a single curvature value derived from the curvatures of the synapses contained within the diclique (discussed below) and each connection between dicliques is labeled with the value of the flow through each neuron associated with the connection. Such graphs are useful for analyzing which sub-networks (dicliques) are most effective for dispersing flow in an ANN and for identifying individual neurons which are important to the flow between subnetworks.

A Type 3 FR graph is a labeled diclique digraph representation of an ANN. Its dicliques are labeled with single curvature values obtained from the curvature values of the synapses contained in the dicliques (discussed below) and its connections between dicliques are labeled with single flow values obtained from the values of the flows through each neuron associated with the connection (discussed below). The advantages of Type 3 FR graphs are: (i) although they somewhat dilute flow information of individual neurons connecting dicliques - they provide for easier analyses of the overall flow and divergence properties of ANNs than do Type 1 or 2 FR graphs; and (ii) they lend themselves to analyses using standard digraph algorithms (e.g., shortest dipaths, etc.).

Although a variety of approaches for determining the numerical values of diclique curvature could probably be developed, the obvious are the curvature sum and the mean curvature. Consider an ANN diclique \((X, Y)\) and let the sets \( K = \{1, 2, 3, \ldots, |X|\} \) and \( L = \{1, 2, 3, \ldots, |Y|\} \) index the neurons in \( X \) and \( Y \). If the neurons in \( X \) and \( Y \) are in layers \( i \) and \( i + 1 \), respectively, then the curvature sum \( S(X, Y) \) of \((X, Y)\) is defined by

\[
S(X, Y) = \sum_{j \in K, k \in L} \gamma(s_{jk})
\]

and the mean curvature \( \mu(X, Y) \) by

\[
\mu(X, Y) = \frac{S(X, Y)}{|X||Y|}
\]

Similarly, flow sum and mean flow can be used to provide numerical values for the connections between dicliques in the diclique digraph representation of an ANN. Let \( A = Y \cap U \neq \emptyset \) be the set of neurons which label the directed edge between \((X, Y)\) and \((U, V)\). The associated flow sum \( \Phi(X, Y), (U, V) \) and mean flow \( \mathcal{M}(X, Y), (U, V) \) are then given by

\[
\Phi(X, Y), (U, V) = \sum_{n \in A} \phi(n)
\]

and

\[
\mathcal{M}(X, Y), (U, V) = \frac{\Phi(X, Y), (U, V)}{|A|}
\]

Of course, standard deviations, etc. can also be associated with both \( \mu(X, Y) \) and \( \mathcal{M}(X, Y), (U, V) \).

As an illustration of an FR graph and an interpretation of a portion of it, consider the Type 3 FR graph derived from Fig. 4 and shown in Fig. 5 – where fictitious mean curvature and mean flow values are assigned to its dicliques and connections. Since \( \mu(D_2) < \mu(D_4) \), then \( D_2 \) influences early computational flow more than \( D_4 \) since it “spreads” the computational flow more effectively. In addition, more of the \( D_2 \) flow is sent to \( D_4 \) than to \( D_3 \) because \( \mathcal{M}(D_2, D_4) = +3 > +1 = \mathcal{M}(D_2, D_3) \). Note that the longest dipath in the graph is \( D_2 \rightarrow D_3 \rightarrow D_2 \) (length = 3 + 2 = 5) and the shortest dipath is \( D_2 \rightarrow D_3 \rightarrow D_5 \) (length = 1 + 2 = 3). It can be concluded from this that more (less) computational flow to \( D_5 \) occurs along dipath \( D_2 \rightarrow D_3 \rightarrow D_5 \) than any other dipath in the ANN. By inspection of Fig. 4 (or from the eigenvalue spectrum 2.1358, 0.6622, 0, -0.6622, -2.1358 of the underlying graph of the digraph) it is easily seen that the chromatic number for this digraph is 2 – i.e., it is bipartite - so that its dicliques can be partitioned into the two sets \( \pi_1 = \{D_1, D_2, D_3\} \) and \( \pi_2 = \{D_4, D_5\} \). This means that there is no direct computational flow between the dicliques in \( \pi_1 \) and between the dicliques in \( \pi_2 \). Some sense of the resistance of the ANN to computational disruption is obtained by noting that for example: (i) neither \( D_3 \) nor \( D_4 \) are cut-dicliques - i.e., the removal of either one of them does not totally stop computational flow (Fig. 5); and (ii) removal of \( n_1^2, n_2^2 \), and \( n_3^2 \) from the ANN does not completely stop computational flow (Figs. 1, 4, and 5).

VI. THE SEMIGROUP OF AN ANN: AN ALGEBRAIC DESCRIPTION OF DICLIQUE CONNECTION TOPOLOGY

Semigroups can be associated with ANNs by defining the multiplication "·" of \( R \) dicliques \( D = (X, Y) \) and \( D' = (X', Y') \) as
(U, V)

as the associative binary operation [8]

\[ D \cdot D' = (X \cup U_{x \in X \cap D} R(x)) , \]

where \( R(x) = \{ (x, y) \in R \} \). Observe that: (i) \( \cdot \) is commutative; (ii) \( D \cdot D' \) is an \( R \) diclique; and (iii) \( D \) is an idempotent since \( D^2 = D \cdot D = D \). The semigroup of an ANN is the band \([11] B_C = (C, \cdot ) \) of dicliques generated by the cover \( C \) of an ANN under the operation \( \cdot \). Because there can be multiple covers for an ANN, \( B_C \) is not necessarily unique.

Although \( B_C \) contains no curvature information, its algebraic structure can provide some insight about the connectivity of the cover of the associated ANN. For example, consider the ANN of Fig. 1 and the diclique cover \( C = D = \{ D_1, D_2, D_3, D_4, D_5, D_6 \} \) defined in section III. The Cayley table for this \( B_C \) is given by:

**TABLE I: THE CAYLEY MULTIPLICATION TABLE FOR THE BAND \( B_C \) ASSOCIATED WITH THE ANN IN FIG. 1**

|    | \( D_1 \) | \( D_2 \) | \( D_3 \) | \( D_4 \) | \( D_5 \) | \( D_6 \) | \( \emptyset \) |
|----|-----|-----|-----|-----|-----|-----|-----|
| \( D_1 \) | \( D_1 \) | \( D_2 \) | \( D_3 \) | \( D_4 \) | \( D_5 \) | \( D_6 \) | \( \emptyset \) |
| \( D_2 \) | \( D_2 \) | \( D_4 \) | \( D_5 \) | \( D_6 \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) |
| \( D_3 \) | \( D_3 \) | \( D_5 \) | \( D_6 \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) |
| \( D_4 \) | \( D_4 \) | \( D_6 \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) |
| \( D_5 \) | \( D_5 \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) |
| \( D_6 \) | \( D_6 \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) |
| \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) |

Here \( D_6 = ((n_2^2), (n_1^2, n_2^2, n_3^2, n_4^2)) \notin C \) and \( D_7 = ((n_3^2), (n_2^2, n_3^2, n_4^2)) \notin C \) are \( R \) dicliques and \( \emptyset \equiv (\emptyset, \emptyset) \notin C \) is the empty diclique generated by \( C \) under \( \cdot \). Note that \( \emptyset \) is a zero element in \( B_C \) and that the sets \( \{ D_1, D_2, D_6, \emptyset \}, \{ D_3, D_4, D_5, \emptyset \}, \) and \( \{ D_5, \emptyset \} \) are ideals and sub-semigroups in \( B_C \).

Contrast this Cayley table with the following Cayley table for the band generated by the cover of the fully connected ANN in Fig. 3.

**TABLE II: THE CAYLEY MULTIPLICATION TABLE FOR THE BAND \( B_C \) ASSOCIATED WITH THE FULLY CONNECTED ANN IN FIG. 3**

|    | \( D_1 \) | \( D_2 \) | \( D_3 \) | \( \emptyset \) |
|----|-----|-----|-----|-----|
| \( D_1 \) | \( D_1 \) | \( D_2 \) | \( \emptyset \) | \( \emptyset \) |
| \( D_2 \) | \( D_2 \) | \( D_3 \) | \( \emptyset \) | \( \emptyset \) |
| \( D_3 \) | \( \emptyset \) | \( \emptyset \) | \( D_3 \) | \( \emptyset \) |
| \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) | \( \emptyset \) |

It is evident from Table 2 that - other than the empty band - the cover for the fully connected ANN does not generate dicliques that are not in the cover. This observation leads to the following general theorem:

**Theorem 7.** If \( C \) is the unique diclique cover for a fully connected ANN, then \( C \cup \{ \emptyset \} \) are the only elements of \( (C, \cdot ) \).

**Proof.** This follows from the fact that the first entries for any two dicliques in the cover of a fully connected ANN have no neurons in common. More specifically, let \( C \) be the unique cover for a fully connected ANN and \( D_i = (X_i, Y_i) \in C, D_j = (X_j, Y_j) \in C \). Since \( X_i \cap X_j = X_i \) and \( X_i \cap X_j = \emptyset, i \neq j, \) then

\[ D_i \cdot D_j = (X_i \cap X_j \cup Y_i \cap Y_j, R(x)) = (X_i, Y_i) = D_i \]

and

\[ D_i \cdot D_j = (X_i \cap X_j \cup Y_i \cap Y_j, R(x)) = (X_i, Y_i) = D_i \]

Thus, \( C \cup \{ \emptyset \} \) are the only elements in \( (C, \cdot ) \).

Because \( B_C \) is a band, it is also a lower semi-lattice under the natural order \( \leq \) of its idempotents \([12] \). In particular, if \( e \) and \( f \) are idempotents in a band, then \( e \leq f \) means \( ef = fe = e \) and \( e \) is said to be under \( f \). For the band with the Cayley table given by Table 1, it is easy to see that \( D_6 \leq D_1, D_6 \leq D_2, D_7 \leq D_3, D_7 \leq D_4, \) and \( \emptyset \leq D_i, i = 1,2,\ldots 7. \) This natural order provides the following lower semi-lattice for the band associated with the ANN of Fig. 1:

**Fig. 6. The lower semi-lattice of dicliques for the ANN in Fig. 1**

Note that \( D_5, D_6, \) and \( D_7 \) are primitive idempotents, i.e. they satisfy the condition that the only idempotents under each \( D_i \) are \( D_i \) and \( \emptyset \) (recall that \( \emptyset \) is the zero element of the band) – whereas \( D_1, D_2, D_3, D_4 \) are not.

By definition, the meet of two dicliques \( D_i \land D_j \) in a diclique semi-lattice for an ANN is their semigroup product \( D_i \cdot D_j \) and corresponds to a diclique in \( R \) which has as its first entry the neurons common to the first entries of \( D_i \) and \( D_j \). Thus, such meets can be useful for identifying other maximally connected subnetworks in ANNs. For example, refer to Fig. 1 to see that \( D_6 \) and \( D_7 \) are maximally connected subnetworks.

**Fig. 7. The semi-lattice of dicliques for the ANN in Fig. 3**

The semi-lattice associated with the band for the fully connected ANN of Fig. 3 is given by Fig. 7. There the natural order for the idempotents is simply \( \emptyset \leq D_i, i = 1,2,3. \) Thus, \( D_1, D_2, \) and \( D_3 \) are primitive idempotents. This leads to the following result:

**Theorem 8.** All dicliques in the band associated with a fully connected ANN are primitive idempotents.

**Proof.** Since the first entries for every two dicliques in the unique cover of a fully connected ANN have empty intersection, then the only diclique generated by the cover that is not in the cover is the empty diclique. Consequently, the only idempotents under any diclique \( D_i \in (C, \cdot ) = (C \cup \{ \emptyset \}, ) \) is \( D_i \) and \( \emptyset \).
This theorem therefore identifies an algebraic signature for fully connected ANNs.

VII. CONCLUDING REMARKS

Dicliques have been shown to provide simplified graph representations of ANN connection topology. When combined with FR curvatures, such representations can provide insights into the flow of computation within ANNs. It was also shown that a band semigroup of an ANN’s dicliques can be associated with the ANN and that the natural idempotent order of the dicliques in the band defines a lower semi-lattice whose meets identify other fully connected subnetworks in the ANN. As a final result, it was shown that the dicliques in the band associated with a fully connected ANN are all primitive idempotents.

Future efforts should involve testing the theory developed above using real functioning ANNs; using the theory to analyze the flow of computations through ANNs; and investigating other algebraic properties of ANNs.

REFERENCES

[1] W. Rawat, and Z. Wang, “Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review”, Neural Computation, vol. 29, pp. 2352-2449, 2017. Doi:10.1162/NECO_a_00990.

[2] M. Chau, and H. Chen, “Incorporating Web Analysis into Neural Networks: An Example in Hopfield Net Searching”, IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, vol. 37, no. 3, pp. 352-358, 2007. Doi:10.1109/TRMCC.2007.893277.

[3] P. Lee, S. Hua, and A. Fong, “Neural Networks for Web Content Filtering”, IEEE Intelligent Systems, pp. 48-57, 2002. Doi:10.1109/MIS.2002.1039832.

[4] M. Mustafa, T. Allen, and K. Appiah, “A comparative review of dynamic neural networks and hidden Markov model methods for mobile on-device speech recognition”, Neural Comput & Applic, vol. 31 (Suppl 2), pp. S891-S899, 2019. Doi:10.1007/s00521-017-3028-2.

[5] D. Krotov, and J. Hopfield, “Unsupervised learning by competing hidden units”, PNAS, vol. 116, no. 16, pp. 7723-7731, 2019. Doi:10.1073/pnas.1820458116.

[6] S. Xie, A. Kirillov, R. Girshick, and H. He, “Exploring Randomly Wired Neural Networks for Image Recognition”, arXiv preprint arXiv:1904.01569v2 [cs.CV], 2019.

[7] W. Guss, and R. Salakhutdinov, “On Characterizing the Capacity of Neural Networks using Algebraic Topology”, arXiv preprint arXiv:1802.04443v1 [cs.LG], 2018.

[8] R. Haralick, “The Diclique Representation and Decomposition of Binary Relations”, Journal of the Association for Computing Machinery, vol. 24, no. 4, pp. 356-366, 1974.

[9] M. Weber, E. Saucan, and J. Jost, “Characterizing complex networks with Forman-Ricci curvature and associated geometric flows”, Journal of Complex Networks, Vol. 5, No. 4, pp. 527-550, 2017. Doi:10.1093/comnet/cnw030.

[10] R. Sreejith, J. Jost, E. Saucan, and A. Samal, “Forman curvature for directed networks”, arXiv preprint arXiv:1605.04662v2 [q-bio.MN], 2017.

[11] J. Howie, An Introduction to Semigroup Theory, Academic Press: London, 1976, pp. 95-103.

[12] A. Clifford, and G. Preston, The Algebraic Theory of Semigroups, Volume I. American Mathematical Society: Providence, Rhode Island, 1961, pp. 23-26.

Allen D. Parks was born in the USA in 1945. He received a BA in Mathematics from Marshall University in 1969 and a Ph.D. in Physics/Astronomy with a dissertation in electron impact phenomena from the Pennsylvania State University in 1975. Since then he has been employed by FMC Corp., General Dynamics Corp., Computer Sciences Corp, and the General Electric Company in a variety of space flight engineering related capacities. Dr. Parks has been a research scientist at the Naval Surface Warfare since 1981 and has served there as a Principal Research Scientist for the last 15 years. In addition to neural networks, Dr. Parks’ research interests include foundational quantum physics, celestial mechanics, and applications of advanced abstract mathematics to data science, networks, and complex systems.