A GRAPH THEORETIC APPROACH TO CELL ASSEMBLIES

Shaunak Das

1. INTRODUCTION

In the last century, one of the more interesting ideas put forth to explain neurological phenomenon such as learning and memory has been the cell assembly, a coalition of neurons which contribute more excitation to each other than to the average neuron in the network; modern physiology (Buzsaki, 2003) has found evidence in support of such entities. Though conceptually very attractive, the notion of cell assembly had been one that was difficult to pin down concretely.

Palm (1981) viewed networks of neurons as graphs, in which vertices represented neurons, and edges corresponded to synaptic connections between them; he then was able to make precise the definition of cell assembly. In this study, we explore Palm's definition using tools from graph theory. The main objective is to better understand the structure and architecture (rather than dynamics) of these assembly networks, with the hope of being able to identify all of the cell assemblies for a given network. Such an ambitious goal requires the development of efficient algorithms, which in turn requires a better understanding of the underlying mathematics; the main focus of this work is on the latter.

2. DEFINITIONS

Given a graph $G = (V, E)$, where V denotes the set of vertices, E the set of edges, let $c(u, v)$ be the weight of the edge going from vertex u to v.

Define the function $e : \mathrm{P}(V) \times \mathrm{R} \longrightarrow \mathrm{P}(V)$ by

$$e(S, k) = \left\{ v \in V : \sum_{u \in S} c(u, v) \geq k \right\}$$

We will work with a fixed k, so that $e(S, k) = e(S)$ from henceforth. Also, we adopt the shorthand

$$e^n(S) = \underbrace{e \circ ... \circ e(S)}_{n}$$

to represent composition of the e function n times, with $e^0(S) = S$.

The operative part of this definition posed by Palm is the notion of threshold gating, a property of biological neural networks central to their functionality. Given a set S of currently active neurons, any other neuron (vertex) v can be excited, $v \in e(S)$, only if it receives at least k amount of stimulus (edge weight) from those in S.

DEFINITION 1. A set S' is ignited by S if $\exists\ n \in$ such that $S' \subset e^n(S)$.

DEFINITION 2. if $S \subset e(S)$, then S is a k-persistent (or persistent) set. Every persistent set ignites itself.

DEFINITION 3. if $S = e(S)$, then S is an invariant set. Every invariant set *only* ignites itself.

DEFINITION 4. if $\exists\ n \in \natural$ such that $e^n(S) = \emptyset$, then S is a weak set.

DEFINITION 5. A sequence $(S, e(S), ..., e^{l-1}(S))$ such that $e^l(S) = S$ is called a period of length l, $l > 1$.

Only one of the following is true of $S \subset V$:

1. S is weak
2. S ignites a period
3. S ignites an invariant set

It can be easily shown that a persistent set S must ignite an invariant set.

DEFINITION 6. The invariant set ignited by a persistent set S is called the closure of S, $\bar{S}$.

3. K-PERSISTENT SETS AND K-CORES

From henceforth, we will be working with finite simple graphs. For fixed k, e is simply a self-map of $P(V)$. Thus, e will surject onto $P(V)$ only if it is injective; however, this will hardly ever be the case (consider for $k > 1$, any two singletons both get mapped to the empty set). It appears that e lacks any useful properties that could potentially give us information about our graph $G = (V, E)$. We must look in another direction in order to gain better insight into our problem.

Recall that a set $S \subset V$ is k-persistent if $S \subset e(S)$. In the case of unweighted graphs, $c(u, v) \in \{0, 1\}\ \forall u, v \in V$, we can define $e(S)$ simply as

$$e(S) = \{v \in V : |N(v) \cap S| \geq k\}$$

where $N(v)$ is the set of vertices that v is adjacent to, its neighbors. Though very simple, this definition is crucial to everything else which follows in this study.

We now introduce the concept of a k-core. In the current literature, a set $S \subset V$ is said to induce a k-core if 1) each vertex in S has k or more neighbors in S (i.e. has degree greater than or equal to k), and 2) the induced graph $G[S]$ is maximal, i.e. not properly contained within another k-core. However, maximality is not necessary for our purposes, and so we neglect this condition in our definition.

DEFINITION 7. Given a graph $G = (V, E)$ and $S \subset V$, S induces a k-core $G[S]$ if $\delta(G[S]) \geq k$, where $\delta(G[S])$ is the minimum degree of the graph induced by S.

Though k-core refers to the induced graph $G[S]$, we may often call the vertex set S such, with the understanding that S in fact induces a k-core. We are now ready to introduce the first, and certainly the most important, result of this study.

THEOREM 1. Given a graph $G = (V, E)$, $S \subset V$ is k-persistent if and only if it induces a k-core.

Proof. Let $S \subset V$ be k-persistent. This means $S \subset e(S)$, and thus $|N(v) \cap S| \geq k$, $\forall v \in S$. Since each $v \in S$ is adjacent to k or more other members in S, $deg_{G[S]}(v) \geq k$, $\forall v \in S$. Thus, $\delta(G[S]) \geq k$; S induces a k-core. Now let S induce a k-core; this implies each $v \in S$ is adjacent to at least k other members in S. Thus $|N(v) \cap S| \geq k$, $\forall v \in S$, and so $S \subset e(S)$; S is a k-persistent set. $\diamond$

The importance of the above theorem cannot be overstated. As seen earlier, the amorphous e mapping will hardly ever have any nice properties, thus its use does not extend beyond denoting sets. THEOREM 1 is a bridge between Palm's obscure notion of persistent sets and mainstream graph theory. It provides us with concrete techniques, namely looking at vertex degrees, that will be powerful in the analysis of persistent sets of a given network.

A characteristic of cell assembly is that any "sufficiently large" portion of it is able to excite the entire assembly; just as well, it should be a closed circle of neurons. Palm came up with the following two definitions in order to capture these properties.

DEFINITION 8. Given a graph $G = (V, E)$, let $S \subset V$ be k-persistent. S is a tight set if each of its persistent subsets, whose complement in S is not weak, is able to ignite all of S.

DEFINITION 9. A cell assembly is the closure of a tight set.

The definition of a tight set is certainly a mouthful; being able to decompose this concept serves as motivation for much of the remaining paper. The following simple proposition provides some insight into where to find cell assemblies.

PROPOSITION 2. If S induces a k-core, then $e(S)$ induces a k-core.

Proof. Since S induces a k-core, it is persistent, $S \subset e(S)$. Applying e, it follows that $S \subset e(S) \subset e^2(S) \subset ...$. Since $e(S) \subset e^2(S)$, $e(S)$ is persistent and thus must induce a k-core. $\diamond$

Inductively then, it can be shown that $e^n(S)$ will also be a k-core, $\forall n$. Since $\bar{S} = e^n(S)$, for S a tight set, it follows that **all cell assemblies must be**

k**-cores**. Thus, the set of k-cores of a given graph serves as the natural hunting ground for tight sets and cell assemblies. Determining the closures of tight sets, and thus the assemblies, is computationally very easy; what is much more difficult is the actual *identification* of the tight sets within a graph. Though our goal is to find the cell assemblies for a given network, we must first investigate k-cores/persistent sets and tight sets. The rest of this work is devoted to these entities.

4. MINIMAL K-CORES

We define a minimal k-core as one which does not contain a k-core properly within it. The reason these interest us is that they themselves are tight sets (trivially). As a result, we can immediately determine their closures without any analysis of persistent subsets; these will be the easiest tight sets to find.

However, before computationally attacking the problem of finding all of the minimal k-core of a given graph, or any k-core for that matter, we need an idea of how feasible finding just *one* is. Below we state the decision problems of K-CORE and MINIMAL K-CORE; it should be noted that since our definition of k-core differs from the mainstream one, our K-CORE problem also differs from the one already posed.

DEFINITION 10: K-CORE.
Given a graph $G = (V, E)$, positive integers k and t, can we find a k-core of size at most t in G? That is, does there exist $S \subset V$ such that $|S| \leq t$ and $\delta(G[S]) \geq k$?

DEFINITION 11: MINIMAL K-CORE.
Given a graph $G = (V, E)$, positive integers k and t, can we find a minimal k-core of size at most t in G? That is, does there exist $S \subset V$ such that $|S| \leq t$ and $\delta(G[S]) \geq k$, and for $S' \subset S$, $\delta(G[S']) < k$?

DEFINITION 12: CLIQUE.
Given a graph $G = (V, E)$ and positive integer k, can we find a clique of size k in G. That is, does there exist $S \subset V$ such that $|S| = k$ and $G[S]$ is a complete subgraph in G?

THEOREM 3. K-CORE is NP-complete.

Proof. We restrict K-CORE to CLIQUE by considering only instances in which $t = k + 1$. $\diamond$

THEOREM 4. MINIMAL K-CORE is NP-complete.

Proof. We restrict MINIMAL K-CORE to CLIQUE again by considering only instances in which $t = k + 1$. $\diamond$

The following give some sufficient conditions for when and when not a k-core is minimal.

PROPOSITION 5. Every connected k-regular graph G is a minimal k-core.

Proof. Let G be a connected k-regular graph, $\Delta(G) = \delta(G) = k$. We assume that G is not a minimal k-core, and thus $\exists\ S \subset V$ such that $G[S]$ is a k-core inside of G. Since G is connected, $\exists\ v \in S$ that is adjacent to some $w \in V - S$. However, since G is also k-regular, this implies that $deg_{G[S]}(v) < k$, and so $\delta(G[S]) < k$. This is a contradiction, so G must be a minimal k-core. $\diamond$

PROPOSITION 6. If $\delta(G) > k$, then G is not a minimal k-core.

Proof. Since $\delta(G)$ is strictly greater than k, removal of a single vertex v will at most decrease the degree of any remaining vertex by 1. Thus, $\delta(G[V - \{v\}]) \geq k$, and so $G[V - \{v\}]$ is a k-core inside of G. $\diamond$

Notice that this last proposition does not mention anything about connected k-cores. However, if a given k-core is not connected, we automatically know that it is not minimal (in fact, it is not even tight, as will be shown in the next section), since each of its connected components must be k-cores.

PROPOSITION 7. Given a k-core $G = (V, E)$, if there exists $S \subset V$ such that $min\,\{deg_G(v) : v \in S\} \geq k + |V| - |S|$, then G is not minimal.

Proof. Suppose that $S \subset V$ is such that $min\,\{deg_G(v) : v \in S\} \geq k + |V| - |S|$. Let $P = V - S$; since each $v \in S$ can have at most $|P| = |V| - |S|$ neighbors in P, it follows that $\delta(G[S]) \geq k$. Thus $G[S]$ is a k-core sitting inside of G, so G is not minimal. $\diamond$

This last proposition lends itself to a very simple algorithm; find the first $k + 1$ vertices that have largest degree in G (by computing row sums of the adjacency matrix), and let these initially constitute S. If the hypothesis is not true, from the remaining pool add the next vertex of largest degree to S, and continue until either the hypothesis is true or all of V has been exhausted. Notice, however, that this condition is not necessary, merely sufficient, for the k-core to be non-minimal.

We now describe a greedy algorithm that finds the maximum k-core within a graph, and then use it to determine whether or not a given k-core is minimal. Given $G = (V, E)$, determine a vertex v of minimum degree; if $deg_G(v) \geq k$, then G is already a k-core; if not, then we know v and its edges cannot be a part of any k-core and we remove it from the graph; we continue recursively until we get a k-core or empty set.

Now suppose we know that $S \subset V$ induces a k-core; let $S = \{v_1, ..., v_{|S|}\}$. Delete v_1 from S, and then apply the greedy algorithm to $G[S - \{v_1\}]$; if a non-empty set is returned, we may stop. Otherwise, we move on to v_2, applying the greedy algorithm to $G[S-\{v_2\}]$. We continue in this fashion, removing *only one* vertex at a time from S and implementing the greedy algorithm, until we have either arrived at a non-empty set or exhausted all of S. The former case corresponds to $G[S]$ not being a minimal k-core, the latter $G[S]$ is minimal.

5. GENERATING OTHER K-CORES

As we have seen, a minimal k-core of a graph G has the property of irreducibility, in that removal of any number of vertices from it no longer makes it a k-core; it is natural then to ask whether they are the building blocks for other k-cores in G. It would be nice if the set of minimal k-cores of G could serve as a basis from which we could generate all other k-cores in the graph; the natural procedure would be to union the vertex sets of any of these minimal k-cores, and then look at the induced graph. Unfortunately, this will not give us all of the k-cores in G, as the following example illustrates for $k = 3$:

However, we can ask under what conditions will the union of two vertex sets yield a new set that induces a k-core. It is trivial that the union of two k-cores will itself be a k-core. As the next proposition shows, we are only interested in the connected k-cores.

PROPOSITION 8. Every tight set induces a connected k-core.

Proof. This is proven by contraposition. Given $G = (V, E)$, let $S \subset V$ induce a disconnected k-core, and $S' \subset S$ induce one of its connected components. $S - S'$ must induce a k-core, thus is not weak. However, S' cannot ignite all of S, since it is not even connected to $S - S'$. Thus, S is not a tight set. $\diamond$

The next simple result lets us know precisely when the union of two k-cores is connected.

PROPOSITION 9. Given $G = (V, E)$, Let $S, S' \subset V$ both induce connected k-cores. $G[S \cup S']$ is a connected k-core $\Leftrightarrow N(S) \cap S' \neq \emptyset$.

Proof. Since S and S' both induce k-cores, then $G[S \cup S']$ is a k-core. Let $N(S) \cap S' \neq \emptyset$. This means $\exists\ u \in S'$ that is adjacent to some $w \in S$. Since $G[S]$ and $G[S']$ are both independently connected, and $\exists\ (u, w) \in E$ that connects $u \in S'$ to $w \in S$, $G[S \cup S']$ is a connected graph. $G[S \cup S']$ is a connected k-core. Now let $G[S \cup S']$ be a connected k-core. The fact that this graph is connected means that some $w \in S$ is adjacent to some $u \in S'$. As a consequence, $u \in N(S) \cap S'$, and so this intersection is not empty. $\diamond$

The following is useful in trying to build the k-cores of a graph that are not simply the union of multiple k-cores.

PROPOSITION 10. Given $G = (V, E)$, let $S, S' \subset V$, where $S \cap S' = \emptyset$. If

$$min\,\{|N(v) \cap S| : v \in S'\} \geq k - \delta(G[S'])$$

and

$$min\,\{|N(v) \cap S'| : v \in S\} \geq k - \delta(G[S])$$

then $S \cup S'$ induces a k-core.

Proof: If the hypothesis is true, then we are guaranteed that each $v \in S'$ has at least $\delta(G[S'])$ neighbors in S' plus $k - \delta(G[S'])$ neighbors in S, for a total of at least k neighbors in $S \cup S'$. Similarly for all $v \in S$, and so $\delta(G[S \cup S']) \geq k$. $\diamond$

We may also consider taking a top-down approach to generating all of the k-core within a graph. Starting with the largest k-core, we could continue to decompose it into smaller and smaller ones. The greedy algorithm described at the end of the last section allows us to do this. The first application will yield the largest k-core in our graph. We can then begin to delete vertices from this graph, and apply the above algorithm each time. Working in this fashion will eventually expose all of the k-cores. Unfortunately, this approach is not much better than a brute force search algorithm.

6. MAXIMAL K-CORES AND INVARIANT SETS

We define a maximal k-core as one that is not a proper subgraph of a connected k-core. Connectedness of the super k-core is an important condition, as the following example shows.

It turns out that maximal k-cores are induced by invariant sets. Since invariance is something that is computationally very easy to check for, this section is kept modest.

PROPOSITION 11. Given $G = (V, E)$, let $S \subset V$ induce a k-core. S is invariant if and only if $\bigcap_{i=1}^{k} F(v_i) = \emptyset$, for any set $\{v_1, ..., v_k\}$ of k vertices in S.

Proof. Let S be invariant; then there does not exist $v \in V - S$ that has k or more neighbors in S. Thus, no set of k vertices in S shares a common neighbor outside of S. $\bigcap_{i=1}^{k} F(v_i) = \emptyset$, for all sets $\{v_1, ..., v_k\}$ of k vertices in S.

Now let $\bigcap_{i=1}^{k} F(v_i) = \emptyset$, for all sets $\{v_1, ..., v_k\}$ of k vertices in S. Assume S is not invariant; then $\exists w \in V - S$ that is adjacent to k or more members in S, $\{v_1, ..., v_m\}$, $m \geq k$. This means $w \in \bigcap_{i=1}^{k} F(v_i) \subset \bigcap_{i=1}^{m} F(v_i)$. This is a contradiction, so S is invariant. $\diamond$

PROPOSITION 12. Given $G = (V, E)$, a set $S \subset V$ that induces a maximal k-core is invariant.

Proof. Let S induce a maximal k-core; since $G[S]$ is not a proper subgraph of another connected k-core, this means there does not exist $v \in V - S$ that has k or more neighbors in S. By PROPOSITION 11, S is invariant. $\diamond$

COROLLARY 13. Every k-core that is both maximal and minimal is a cell assembly.

Proof. This simply follows from the fact that a minimal k-core is tight, and a maximal k-core is its own closure. $\diamond$

7. WEAK SETS

Knowing that a set is weak requires iterating e an indefinite number of times until an empty set is returned; but this is not a very efficient algorithm for larger networks with greater connectivity, nor is it very interesting to us mathematically. Asking for a necessary and sufficient criteria for whether or not a set S is weak may be too much; after all, weakness requires that we know what happens to $e(S)$, $e^2(S)$, and so forth. For Palm's definition of a tight set, it may be just as instructive to know when a set is *not* weak.

PROPOSITION 14. Given $G = (V, E)$, let $S \subset V$. If $|S| \geq k$ and $|\bigcap_{v \in S} N(v)| \geq k$, then S is not weak.

Proof. Let $S \subset V$ such that $|S| \geq k$ and $|\bigcap_{v \in S} N(v)| \geq k$. Since each vertex in $\bigcap_{v \in S} N(v)$ is adjacent to all members in S, and $|S| \geq k$, $\bigcap_{v \in S} N(v) \subset e(S)$. Each vertex in S is adjacent to all members in $\bigcap_{v \in S} N(v)$, where $|\bigcap_{v \in S} N(v)| \geq k$; it follows then that $S \subset e(\bigcap_{v \in S} N(v)) \subset e^2(S)$. If we continue to apply e to the above sequence, we see that $S \subset e^{2n}(S)$ and $\bigcap_{v \in S} N(v) \subset e^{2n+1}(S)$, $\forall n \in \natural$. Thus $e^n(S) \neq \emptyset$, $\forall n$. S is not weak. $\diamond$

Certainly, the more iterations of e we apply to S, the better we can tell whether it is weak or not. Our goal is to be able to say something definitive about what is occurring at the tail of the sequence $S, e(S), e^2(S), \ldots$, simply by looking at the first few elements.

PROPOSITION 15. If $|e(S)| \geq |S| = k$, then S is not weak.

Proof. Let $|e(S)| \geq |S| = k$. Since each $v \in e(S)$ receives at least k edges from members in S and $|S| = k$, this implies each $v \in S$ is adjacent to each $w \in e(S)$. Thus, S itself is a set of k vertices such that $|\bigcap_{v \in S} N(v)| = |e(S)| \geq k$; by PROPOSITION 1, S is not weak. $\diamond$

Implicit in the above hypothesis is that $S \cap e(S) = \emptyset$. The next proposition tells us precisely when such is true.

PROPOSITION 16. $S \cap e(S) = \emptyset \Leftrightarrow \Delta(G[S]) < k$.

Proof. Let $S \cap e(S) = \emptyset$; this implies that there does not exist $v \in S$ that is adjacent to k or more other members of S, thus $\Delta(G[S]) < k$. Now let $\Delta(G[S]) < k$; this means that no $v \in S$ is adjacent to k or more other members in S; thus, $S \cap e(S) = \emptyset$. $\diamond$

PROPOSITION 17. Given $G = (V, E)$ and $S \subset V$, if $|N(v) \cap e(S)| \geq k$, $\forall v \in S$, then S is not weak.

Proof. If each $v \in S$ has k or more neighbors in $e(S)$, it follows that $S \subset e^2(S)$. Application of e over and over again yields

$$S \subset e^{2n}(S) \qquad e(S) \subset e^{2n+1}(S), \qquad \forall n \in \natural$$

Since $S \neq \emptyset$ and $e(S) \neq \emptyset$, it follows that $e^n(S) \neq \emptyset$. S is not weak. $\diamond$

We will use the following lemma (whose proof is left for the section on periods, so as not to disturb the flow of this one) to provide a necessary and sufficient condition for weakness of a certain class of vertex sets.

LEMMA. Given $G = (V, E)$, let $S \subset V$ ignite a period of length l, $\left(e^i(S)\right)_{i=0}^{l-1}$. Then $\bigcup_{i=0}^{l-1} e^i(S)$ induces a k-core in G.

PROPOSITION 18. Given $G = (V, E)$, and $S \subset V$ such that $e(S) \subset S$, S is weak if and only if no subset of it induces a k-core.

Proof. Let S not contain any subset which induces a k-core. We prove S is weak by showing that it cannot ignite a period or some invariant set. Assume that after l iterations of the e map, we arrive at an invariant set $e^l(S) = M = e(M)$; clearly M is a k-core. However, since $e(S) \subset S$, iteration of e yields $e^l(S) \subset ... \subset e(S) \subset S$, and so $M = e^l(S) \subset S$; this contradicts the fact that S does not contain a k-core. Thus, S does not ignite an invariant set M.

Now assume that S ignites a period $e^j(S), e^{j+1}(S), ..., e^m(S)$, where $e^{m+1}(S) = e^j(S)$. Again, since $e(S) \subset S$, then $e^j(S), e^{j+1}(S), ..., e^m(S) \subset S$; thus $S' = \bigcup_{i=j}^{m} e^i(S) \subset S$. By LEMMA, S' induces a k-core; this contradicts the fact that S does not contain any subset which induces a k-core. S does not ignite a period, thus by process of elimination S must be weak.

Let S be weak, so that $e^l(S) = \emptyset$. Assume that $S' \subset S$ induces a k-core; we know then that

$$S' \subset e(S') \subset ... \subset e^l(S') \subset e^l(S) = \emptyset$$

Thus, $S' = \emptyset$; there does not exist a subset of S that induces a k-core. $\diamond$

Thus, if we are given a set S such that $e(S) \subset S$, we merely need to implement the already mentioned greedy algorithm used to find the maximum k-core of a graph; if no set is returned we know then that S is weak, otherwise it is not.

8. PERIODS

As mentioned earlier, another way of understanding the notion of weakness is to see when a set is not weak; that is, when it ignites an invariant set or a period. It turns out that periods are extremely interesting, especially in the case of undirected graphs. Below are some results about periods.

PROPOSITION 19. If $e(S) \subset S$, then S does not ignite a period.

Proof. Let $e(S) \subset S$, and assume S does ignite a period $e^j(S), ..., e^m(S)$, where $e^{m+1}(S) = e^j(S)$. Iteration of e yields,

$$e^j(S) = e^{m+1}(S) \subset e^m(S) \subset ... \subset e^{j+1}(S) \subset e^j(S)$$

as a consequence $e^j(S) = e^{j+1}(S)$. S ignites an invariant set, not a period. $\diamond$

PROPOSITION 20. No k-regular graph contains a period of length $l > 2$.

Proof. Let $G = (V, E)$ be a k-regular graph, and assume $S \subset V$ ignites a period $S, e(S), \ldots, e^{l-1}(S)$. We know $\forall v \in e(S)$ have at least k neighbors in S. Since G is k-regular, we are guaranteed that $N(e(S)) \subset S$. However, $e^2(S) \subset N(e(S))$, implying $e^2(S) \subset S$. Repeated application of e yields $e^{2(n+1)}(S) \subset e^{2n}(S)$ and $e^{2n+3}(S) \subset e^{2n+1}(S)$, $\forall n$. We use this to show no period of length greater than 2 can be ignited by S.

CASE 1: Even length period. Let $e^j(S), e^{j+1}(S), \ldots, e^{j+2n-1}(S)$ be an even length period ignited by S, where $e^{j+2n}(S) = e^j(S)$. We know then that

$$e^j(S) = e^{j+2n}(S) \subset e^{j+2(n-1)}(S) \subset ... \subset e^{j+2}(S) \subset e^j(S)$$

Thus $e^{j+2}(S) = e^j(S)$. S ignites a period of length 2.

CASE 2: Odd length period. Let $e^j(S), e^{j+1}(S), \ldots, e^{j+2n}(S)$ be an odd length period ignited by S, where $e^{j+2n+1}(S) = e^j(S)$. We know then that

$$e^{j+1}(S) = e^{j+2n+2}(S) \subset e^{j+2n}(S) \subset \ldots \subset e^j(S)$$

$$e^j(S) = e^{j+2n+1}(S) \subset e^{j+2n-1}(S) \subset \ldots \subset e^{j+1}(S)$$

Thus $e^j(S) = e^{j+1}(S)$. S ignites an invariant set. $\diamond$

PROPOSITION 21. No complete graph contains a period of length $l > 2$.

Proof. Let $G = (V, E)$ be a complete graph, and $S \subset V$ initiate a period $S, e(S), ..., e^{l-1}(S)$. We know then that $|S| \geq k$.

CASE 1: $|S| > k$. In this case, $e(S) = V = e^2(S)$. S ignites an invariant set, not a period.

CASE 2: $|S| = k$ and $|V| = 2k$. In this case, $e(S) = V - S$, $|e(S)| = k$; $e^2(S) = V - (V - S) = S$. This is a period of length 2.

CASE 3: $|S| = k$ **and** $|V| < 2k$**.** Here $e(S) = V - S$, $|e(S)| < k$. Thus, $e^2(S) = \emptyset$, and so S is weak.

CASE 4: $|S| = k$ **and** $|V| > 2k$**.** Here, $e(S) = V - S$, $|e(S)| > k$. Thus, $e^2(S) = V$, and so S ignites an invariant set.

No period of length greater than 2 exists in a complete graph. $\diamond$

LEMMA 22. Given $G = (V, E)$, let $S \subset V$ such that $S \cap e(S) = \emptyset$. If $|e(S)| > \frac{k-1}{k}|S|$, then $S \cap e^2(S) \neq \emptyset$.

Proof. Let $|e(S)| > \frac{k-1}{k}|S|$ and $E_{S,e(S)}$ denote the set of edges between S and $e(S)$. Assume that $e^2(S) \cap S = \emptyset$; this implies that no $v \in S$ is adjacent to k or more members in $e(S)$, and so $|E_{S,e(S)}| \leq (k-1)|S|$. We know that each $w \in e(S)$ is adjacent to at least k members of S, so $|E_{S,e(S)}| \geq k|e(S)|$. It immediately follows that

$$(k-1)|S| \geq k|e(S)|$$

$$\frac{k-1}{k}|S| \geq |e(S)|$$

This is a contradiction, so our assumption that $S \cap e^2(S) = \emptyset$ is incorrect. $\diamond$

DEFINITION. A period $\left(e^i(S)\right)_{i=0}^{l-1}$ of length l in which $e^i(S) \cap e^j(S) = \emptyset$, $\forall i, j \in 0, 1, \ldots, l-1$, $i \neq j$, is called a disjoint period.

THEOREM 23. There does not exist a disjoint period of length $l > 2$.

Proof. Assume there exists a disjoint period $(S, e(S), e^2(S), \ldots, e^{l-1}(S))$ of length $l > 2$. Since $e^i(S) \cap e^{i+2}(S) = \emptyset$, $\forall\ i \in 0, 1, 2, \ldots, l-1$, the contrapositive of LEMMA 22 yields the following recurrence relation:

$$|e(S)| \leq \frac{k-1}{k}|S|$$

$$|e^2(S)| \leq \frac{k-1}{k}|e(S)| \leq \left(\frac{k-1}{k}\right)^2 |S|$$

$$\vdots$$

$$|e^{n+1}(S)| \leq \frac{k-1}{k}|e^n(S)| \leq \left(\frac{k-1}{k}\right)^{n+1} |S|$$

for any given disjoint period. Since $S = e^l(S)$, $|S| = |e^l(S)| \leq \left(\frac{k-1}{k}\right)^l |S| < |S|$; this is a contradiction. There does not exist a disjoint period of length $l > 2$. $\diamond$

We now want to find an inequality similar to the one in LEMMA 22 which will be valid if $S \cap e(S) \neq \emptyset$. If $S \cap e^2(S) = \emptyset$, then $|E_{S,e(S)}| \leq (k-1)|S|$. The goal is to force the minimum number of possible edges between S and $e(S)$ to be greater than $(k-1)|S|$; as a consequence, $S \cap e^2(S)$ will be forced to be non-empty. The minimum number of edges between sets S and $e(S)$ will be $k|e(S)| - \frac{|S\cap e(S)|(|S\cap e(S)|-1)}{2}$; the additional term corresponds to double counting of edges between vertices in $S \cap e(S)$. Thus, if $|E_{S,e(S)}| \geq k|e(S)| - \frac{(S\cap e(S)|)(|S\cap e(S)|-1)}{2} > (k-1)|S|$, then $S \cap e^2(S) \neq \emptyset$. Rearranging this inequality yields the following:

LEMMA 24. Given $G = (V, E)$ and $S, e(S) \subset V$, if

$$|e(S)| > \frac{k-1}{k}|S| + \frac{|S\cap e(S)|(|S\cap e(S)|-1)}{2k}$$

then $S \cap e^2(S) \neq \emptyset$.

Notice this inequality reduces to the one in LEMMA 22 when $S \cap e(S) = \emptyset$. Though the inequality looks a bit unwieldy, it does lead to some interesting results:

THEOREM 25. If $|S| > \frac{j(j-1)}{2}$, then S cannot ignite a period such that $|e^i(S) \cap e^{i+1}(S)| \leq j$ and $e^i(S) \cap e^{i+2}(S) = \emptyset$, $\forall i$.

Proof. Let $|S| > \frac{j(j-1)}{2}$, and assume there exists a period of length $l, \big(S, e(S), ..., e^{l-1}(S)\big)$, such that $|e^i(S) \cap e^{i+1}(S)| \leq j$ and $e^i(S) \cap e^{i+2}(S) = \emptyset$, $\forall i \in 0, 1, ..., l-1$. For convenience let $\alpha_i = \frac{|e^{i-1}(S)\cap e^i(S)|(|e^{i-1}(S)\cap e^i(S)|-1)}{2k} \geq 0$. The contrapositive of LEMMA 24 yields the following relation:

$$|e(S)| \leq \frac{k-1}{k}|S| + \alpha_1$$

$$|e^2(S)| \leq \frac{k-1}{k}|e(S)| + \alpha_2 \leq \left(\frac{k-1}{k}\right)^2 |S| + \left(\frac{k-1}{k}\right)\alpha_1 + \alpha_2$$

$$\vdots$$

$$|e^{n+1}(S)| \leq \left(\frac{k-1}{k}\right)^{n+1} |S| + \sum_{i=1}^{n+1} \alpha_i \left(\frac{k-1}{k}\right)^{n+1-i}$$

Since $S = e^l(S)$,

$$|S| = |e^l(S)| \leq \left(\frac{k-1}{k}\right)^{l} |S| + \sum_{i=1}^{l} \alpha_i \left(\frac{k-1}{k}\right)^{l-i}$$

$|e^i(S) \cap e^{i+1}(S)| \leq j$, thus $\alpha_i \leq \frac{j(j-1)}{2k}$, $\forall i$,

$$|S| \leq \left(\frac{k-1}{k}\right)^{l} |S| + \frac{j(j-1)}{2k} \sum_{i=1}^{l} \left(\frac{k-1}{k}\right)^{l-i}$$

$$|S|\left[1-\left(\frac{k-1}{k}\right)^l\right] \leq \frac{j(j-1)}{2k}\sum_{i=1}^{l}\left(\frac{k-1}{k}\right)^{l-i}$$

The sum on the right is geometric in $\frac{k-1}{k}$,

$$\sum_{i=1}^{l}\left(\frac{k-1}{k}\right)^{l-i} = \sum_{i=0}^{l-1}\left(\frac{k-1}{k}\right)^{i} = \frac{1-\left(\frac{k-1}{k}\right)^l}{1-\left(\frac{k-1}{k}\right)}$$

Therefore,

$$|S|\left[1-\left(\frac{k-1}{k}\right)^l\right] \leq \frac{j(j-1)}{2k}\left[\frac{1-\left(\frac{k-1}{k}\right)^l}{1-\left(\frac{k-1}{k}\right)}\right] = \frac{j(j-1)}{2}\left[1-\left(\frac{k-1}{k}\right)^l\right]$$

Cancellation of $1-\left(\frac{k-1}{k}\right)^l$ yields $|S| \leq \frac{j(j-1)}{2}$. This is a contradiction; no such period exists. $\diamond$

COROLLARY 26. There does not exist a period such that $|e^i(S) \cap e^{i+1}(S)| \leq 1$ and $e^i(S) \cap e^{i+2}(S) = \emptyset$, $\forall i$.

Proof. Setting $j = 1$ for the hypothesis of THEOREM 2, we see that if $|S| > 0$, then it cannot ignite such a period. $\diamond$

PROPOSITION 27. Given $G = (V, E)$, let $S \subset V$ ignite a period of length l, $\left(e^i(S)\right)_{i=0}^{l-1}$. Then $\bigcup_{i=0}^{l-1} e^i(S)$ induces a k-core in G.

Proof. Let $S' = \bigcup_{i=0}^{l-1} e^i(S) \subset V$. Choose $v \in e^j(S) \subset S'$; then v is adjacent to k or more vertices in $e^{j-1}(S) \subset S'$. Thus $deg_{G[S']}(v) \geq k$, $\forall v \in e^j(S)$. Since $e^j(S)$ was arbitrary, it follows that $deg_{G[S']}(v) \geq k$, $\forall v \in S'$. $\bigcup_{i=0}^{l-1} e^i(S)$ is a k-core. $\diamond$

COROLLARY 28. If a graph does not contain a k-core, then it does not contain a period.

Proof. Let $G = (V, E)$ not contain a k-core. By PROPOSITION 3, G cannot contain a period $S, e(S), ..., e^{l-1}(S)$, since $\bigcup_{i=0}^{l-1} e^i(S) \subset V$ induces a k-core. $\diamond$

PROPOSITION 3 shows that even periods can supply information about our main problem of finding all the k-cores, and ultimately the tight sets and cell assemblies, of a given network. It is interesting, as COROLLARY 28 shows, that periods have some sort of "dependence" on k-cores.

9. CONCLUSION

The most important consequence of this study has been the bridge made between Palm's notion of cell assembly and mainstream graph theory; the connection between k-persistent sets and k-cores allows for the use of precise tools in our analysis of networks. With these techniques, we have been able to elucidate a number of properties of persistent sets, weak sets, and periods. We hope to eventually consolidate all of this information into an effective algorithm for determining the tight sets, and eventually the cell assemblies, of a given network.

10. REFERENCES AND ACKNOWLEDGEMENTS

1. G. Palm: Towards a Theory of Cell Assemblies. *Biological Cybernetics*, 39:181-194, 1981.

2. S.B. Seidman. Network Structure and Minimum Degree. *Social Networks*, 5:269-287, 1983.

3. M. Garey and D. Johnson. *Computers and Intractibility: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Company, 1979.

The author would like to thank Dr. Steve Cox and Dr. Illya Hicks. This work was supported by NSF REU Grant DMS-0755294.