A Distributed Collapse of a Network's Dimensionality

Adam C. Wilkerson, Harish Chintakunta, Hamid Krim

Department of Electrical and Computer Engineering
North Carolina State University
Raleigh, North Carolina, USA

Terrence J. Moore, Ananthram Swami

Army Research Lab
Adelphi, Maryland, USA

***Abstract*—Algebraic topology has been successfully applied to detect and localize sensor network coverage holes with minimal assumptions on sensor locations. These methods all use a computation of topological invariants called homology spaces. We develop a distributed algorithm for collapsing a sensor network, hence simplifying its analysis. We prove that the collapse is equivalent to a previously developed strong collapse in that it preserves coverage hole locations. In this way, the collapse simplifies the network without losing crucial information about the coverage region. We show that the algorithm requires one-hop information in a communication network, making it faster than the clique-finding algorithms that slow topological computations for hole localization. This makes it an effective pre-processing step to finding network coverage holes.**

***Index Terms*—Applied Topology, Simplicial Complex, Simplicial Collapse, Flag Complex, Homology**

I. Introduction

The tools of simplicial complexes and homology have seen increased applications in recent years in the modeling and analysis of sensor networks [9], [8]. The advantage of using algebraic topology for network coverage hole discovery is that it doesn't take into account specific sensor locations. In practice, acquiring precise sensor locations is power-intensive. Power conservation is paramount, as the sensors in question are often battery-powered and remotely located. Unfortunately, the topological tools used for hole-localization can be expensive, so minimizing the number of simplices in the complex is desirable [2], [15]. We propose a distributed algorithm that reduces the sensor field to the minimal set needed to accurately count and find the coverage holes in a network, which can also be applied to general flag complexes of graphs. Distributed algorithms are particularly desirable in sensor network applications, as they allow the nodes to aggregate local data into global information without having to spend time and energy broadcasting that information to a central hub for computation.

The paper is organized as follows: in Section II, we introduce the algebraic topology needed for network analysis. In Section III, we review the strong collapse [15], introduce the distributed version, and prove their equivalence. In Section IV, we precisely define the application of the collapse to sensor networks, and include simulation results and complexity analysis. Finally, we provide concluding remarks in Section V.

II. Flag Complexes and Homology

There is a well-developed toolbox from algebraic topology that is useful for analyzing a network using simplicial complexes and homology. Homology reveals high-dimensional structure in a network, and provides a precise definition of coverage holes.

A. Simplicial Complexes and Homology

A *simplicial complex* is a mathematical structure that can be seen as a generalization of a graph: it contains vertices and edges, and in addition may contain higher-dimensional structures like triangles, tetraheda, etc. More formally, a simplicial complex is any collection of sets which is closed under the subset operation. A set with $k+1$ elements in this collection is referred to as a k-*simplex*. Geometrically, a k-simplex is the convex hull of $k+1$ points in an ambient space, and is said to have *dimension* k. Any subset of a k-simplex is called a *face* of that simplex. It is easy to see that 0- and 1-simplices of any complex form a graph.

Simplicial homology, often simply referred to as homology, is an algebraic tool for studying simplicial complexes: given a simplicial complex X, its *homology spaces* are a sequence of real vector spaces $\{H_0(X), H_1(X), H_2(X), \dots\}$, whose ranks respectively count the connected components, loops, 3-D voids, and their generalizations, the higher dimensional *cycles* in the complex. We define the i^{th} *betti number*, $\beta_i(X)$ as the rank of $H_i(X)$, and when no confusion may arise, we denote it β_i. The computation of the homology of a complex is involved and requires a great deal of linear algebra, and a good introduction can be found in any introductory algebraic topology text, such as [10].

B. Constructing the Flag Complex of a Graph

The vertex set V of any graph $G=(V,E)$ yields a natural simplicial complex structure relying on data from the edges E called the *flag complex* of the graph, denoted $\mathcal{F}(G)$. $\mathcal{F}(G)$ has 0-simplices V and 1-simplices E. Then, the 2-simplices are the 3-cliques in G, and the k-simplices are the $(k+1)$-cliques in G.

III. Strong Collapses

The strong collapse for general simplicial complexes was developed [15] using the notions of eccentricity from Q-analysis [1], [11] and a duality construction called the conjugate complex [7].

We define a *labelled simplicial complex* (X, L, V) as a simplicial complex X with a vertex set V, equipped with labels L on some of the simplices in X. The only caveat on L is that every *locally maximal* simplex (one which is not the face of another simplex) must be labelled. For any label $l \in L$, we denote the simplex bearing it Δ_l. Given a labelled simplicial complex (X, L, V) we can construct the *conjugate complex*, denoted (X^T, V, L): X^T is a simplicial complex with vertices corresponding to the elements of L, and a labelled simplex corresponding to each $v \in V$, denoted Δ_v^T. The vertices $l_i \in L$ of Δ_v^T correspond to the faces Δ_{l_i} that v belongs to in (X, L, V). It should be noted that not every labelled simplex in (X^T, V, L) is locally maximal.

Given a simplex $\Delta \in X$, we define its *eccentricity* [13] as

$$\mathrm{ecc}(\Delta) := \frac{\hat{q}(\Delta) - \check{q}(\Delta)}{\hat{q}(\Delta) + 1},$$

where $\hat{q}(\Delta)$ is the dimension of Δ, and $\check{q}(\Delta)$ is given by

$$\check{q}(\Delta) := \max_{l \in L}\{\dim(\Delta \cap \Gamma_l)\}.$$

That is, $\check{q}$ is the dimension of a maximal face of Δ shared with any other labelled simplex $\Delta_l \in X$. In the event that Δ intersects no labelled simplices Δ_l, we define $\check{q}(\Delta) = -1$, so that $\mathrm{ecc}(\Delta) \in [0, 1]$). It immediately follows that a simplex Δ has eccentricity 0 if and only if $\hat{q} = \check{q}$. In other words, $\Delta \subset \Delta_l$. Since Δ is not locally maximal in this case, removing its label from L changes nothing about the underlying complex X, including the homology of X. The reduced labelled complex obtained from removing all eccentricity 0 labels is denoted $(\tilde{X}, \tilde{L}, \tilde{V})$.

The *strong collapse* of a labelled simplicial complex (X, L, V) is as follows: (X^T, V, L) is constructed and has all of its 0-eccentricity simplex labels removed, giving $(\widetilde{X^T}, \tilde{L}, \tilde{V})$. Then, the conjugate of the resulting complex is constructed again. Finally all the eccentricity 0 simplices are removed there, resulting in $\left(\widetilde{(\widetilde{X^T})^T}, \tilde{\tilde{L}}, \tilde{\tilde{V}}\right) \subset (X, L, V)$. That is,

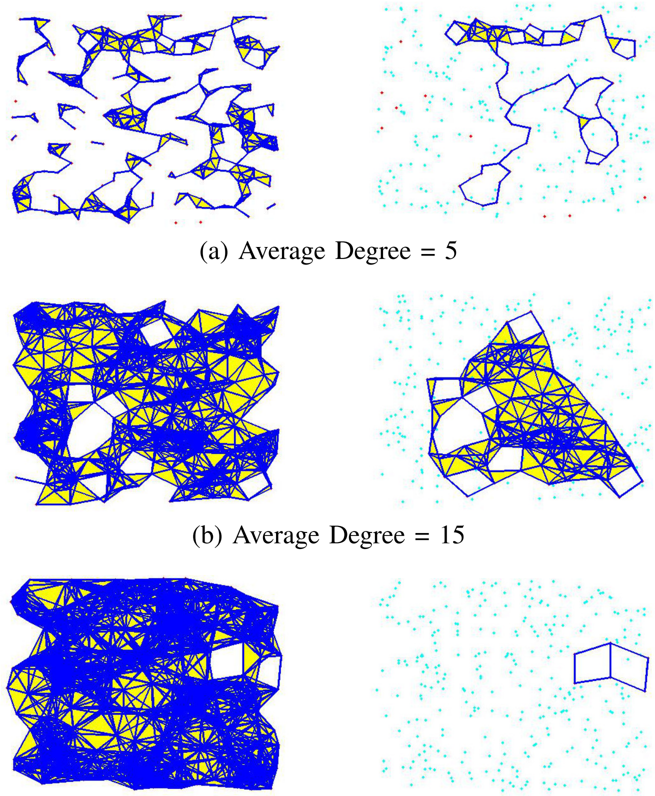
$$\begin{array}{ccccc}
(X, L, V) & \longrightarrow & (X^T, V, L) & \longrightarrow & (\widetilde{X^T}, \tilde{V}, \tilde{L}) \\
\cup & & & \swarrow & \\
\left(\widetilde{(\widetilde{X^T})^T}, \tilde{\tilde{L}}, \tilde{\tilde{V}}\right) & \longleftarrow & \left((\widetilde{X^T})^T, \tilde{L}, \tilde{V}\right) & &
\end{array}$$

This process is iterated until the complex stabilizes. Two theorems introduced in [15] show the value of this collapse:

Theorem 1. *The strong collapse leaves the homology of X invariant.*

Theorem 2. *The strong collapse preserves at least one of the shortest paths around each hole and void in X.*

These facts highlight that the strong collapse not only preserves the homology of the complex X, but that it also maintains the tightest bounding path around any "holes" in X.



(a) Average Degree = 5

(b) Average Degree = 15

(c) Average Degree = 25

Fig. 1. Examples of the collapse of the Rips complex of sensor networks at various average degrees

This second property allows us to collapse a sensor network without fear of losing track of coverage hole locations.

A. The Distributed Algorithm

The general strong collapse requires full *a priori* knowledge of the entire simplicial complex. In the sensor network case, this means that a preprocessing step is needed to find all the cliques in the network, which causes an expensive [12] bottleneck in computing homology. We exploit the fact that every clique in the graph G yields a simplex in the flag complex $\mathcal{F}(G)$ to create an algorithm to execute a collapse that is not only equivalent to the strong collapse, but is also implemented distributively and only requires one-hop information at each node. More importantly, the collapse takes place before any clique-finding algorithm need be run. The value in this property is that the remaining network will be sparser than what we started with, thus tremendously simplifying clique-finding.

Before continuing with the construction of the distributed algorithm, we need one more important definition: the *relevance* of a node v in a simplicial complex X:

$$\mathrm{rel}(v) := \mathrm{ecc}\big(\Delta_v^T\big) = \frac{\hat{q}\big(\Delta_v^T\big) - \check{q}\big(\Delta_v^T\big)}{\hat{q}\big(\Delta_v^T\big) + 1}.$$

It is useful to find a direct geometric interpretation of $\hat{q}\big(\Delta_v^T\big)$ and $\check{q}\big(\Delta_v^T\big)$: $\hat{q}\big(\Delta_v^T\big)$ is the number of locally maximal simplices incident to v, while $\check{q}\big(\Delta_v^T\big)$ is the maximal number of locally maximal simplices shared by v with some other vertex w. Therefore, $\mathrm{rel}(v) = 0$ only when every maximal simplex incident to v is also incident to some other vertex w. This property is equivalent to the notion of v being *dominated* by w, as described in [2]. While the original algorithm works through conjugate complexes to eliminate all those vertices with relevance 0, the distributed algorithm will exploit this updated definition to avoid such intricate, expensive calculations. It follows that any vertex w sharing any faces with v must be adjacent to v in the underlying graph. We assume that each sensor v contains complete knowledge of its neighbor set N_v.

Theorem 3. *For v and w adjacent vertices, $N_v \subset N_w$ if and only if every maximal simplex incident to v is also incident to w, that is, rel$(v) = 0$.*

Proof: ($\Leftarrow$) If every maximal simplex incident to v is also incident to w, then the edge spanning v and w must be in the complex, meaning that $w \in N_v$. Thus, the 1-simplex spanning w and v, denoted $\langle w, v\rangle$, is in the complex. For a vertex $x \in N_v$, let Δ be a maximal simplex with $\langle x, v\rangle \subset \Delta$. Δ is incident to v, and so it's incident to w by assumption. Hence, by the subset closure property of simplicial complexes, $\langle x, w\rangle$ is a 1-simplex in the complex, and so $x \in N_w$.

($\Rightarrow$) Furthermore, given a maximal n-simplex Δ incident to v, without loss of generality, $\Delta = \langle x_1, x_2, \cdots, x_n, v\rangle$. Hence, $x_i \in N_v$ for every $i \in \{1, \cdots, n\}$. Thus, $x_i \in N_w$ for every $i \in \{1, \cdots, n\}$ by assumption. Therefore, $\Delta \cup \{w\}$ is a simplex in the complex. Thus, by the maximality of Δ, $Delta \cup \{w\} \subset \Delta$. Hence, there is some $j \in \{1, \cdots, n\}$ for which $w = x_j$. Δ is therefore incident to w, thus concluding the proof. ■

We exploit this fact to construct the following algorithm, which is iterated until the communication graph stabilizes. Given that the sensors in the network are labelled $v_1, \ldots, v_M$, each sensor v executes the following steps each iteration:

```
Broadcast N_v = {v_{i_j}}_{j=1}^m to immediate neighbors.
for j = 1 → m do
    Receive N_{v_{i_j}}
    Compare N_{v_{i_j}} with N_v
    if N_{v_{i_j}} ⊂ N_v then
        Broadcast OFF signal to v_{i_j}
        if OFF signal received from v_{i_j} then
            Handshake to determine which sensor turns off
        end if
    end if
end for
if OFF received OR Handshake determined v turns OFF
then
    v stops broadcasting
else
    Update neighbor set N_v, omitting OFF neighbors
end if
```

IV. Applications

Here, we provide simulation results and a precise mathematical definition of the sensor network application of the collapse, along with complexity analysis of the algorithm.

A. Sensor Network Coverage and Rips Complexes

Given a distribution of sensors S in some compact region of $\mathbb{R}^2$, we can define the sensing radius r_s as the distance about each sensor in which the sensor can detect targets. That is, for a sensor $v_i \in S$, there is a *coverage disc* D_i centered at v_i with radius r_s within which v_i can detect targets. Then, the *coverage region* spanned by S is well defined as $\cup_{v_i \in S} D_i$. Given this information, we can construct the *Čech complex* of the coverage region, $\check{C}(S, r_s)$: this simplicial complex is constructed iteratively from the 0-simplices, defined to be the sensors S. Following that, we include an n-simplex in the complex spanning any set of $(n+1)$ sensors $\{v_{i_j}\}_{j=0}^n$ for which the coverage discs $\{D_{i_j}\}_{j=0}^n$ share a common intersection point. A classical result called the *nerve theorem* [3] states that $\check{C}(S, r_s)$ has the same homology as $\cup_{v_i \in S} B_i$, the sensor coverage region. Moreover, the generators of $H_1\big(\check{C}(S, r_s)\big)$ bound the coverage holes in the coverage region, thus giving us a convenient, computable definition of a "hole" in a sensor network. The problem is that computing the Čech complex is expensive and requires specific hole location, and so a more computable approximation is required, motivating the Rips complex construction.

Given the same network S, we now define the communication radius r_c as the distance so that any two sensors $s_i, s_j \in S$ with $d(s_i, s_j) < r_c$ can communicate. A natural construction called the *communication graph* $G(S)$ follows: we construct the graph (S, E) with vertices S and an edge e_{ij} between every pair of vertices s_i, s_j with $d(s_i, s_j) < r_c$. We then define the *Rips complex* of S, $R(S, r_c)$, as the flag complex of the communication graph of S, that is, $R(S, r_c) := \mathcal{F}\big(G(S)\big)$. This complex is distributively computable. In addition, even though it doesn't perfectly model the coverage region of S in general, it can be shown that for $r_s = \frac{r_c}{2}$, $\check{C}(S, r_s) \subset R(S, r_c) \subset \check{C}(S, r_c)$ [5]. Furthermore, for $r_s \geq \frac{r_c}{2}$, the coverage holes undetected by the first homology space $H_1\big(R(S, r_c)\big)$ are geometrically small.

Even though the homology of the Rips complex $R(S)$ is distributively computable, doing so is still expensive, as are distributed hole localization methods [6]. The major advantage of the distributed strong collapse is that it can be executed before computing any cliques in the communication graph. It simply takes one-hop information within the network and turns off the irrelevant nodes before finding any cliques.

B. Complexity Analysis

Because homology computations are essentially nullity calculations of a matrix, the complexity of computing the homology of a simplicial complex with n simplices is on the same order as computing the rank of a matrix, $O(n^{2.37})$ [4]. The benefit of the collapse to hole-localization is therefore reflected by the degree to which the number of simplices in the complex is reduced. Furthermore, we are only interested in finding the coverage holes in the coverage region C, so we only need information regarding $H_1(C)$. From the construction of homology [10], the only simplices that contribute to the construction of $H_1(C)$ are the 0-, 1-, and 2-simplices.

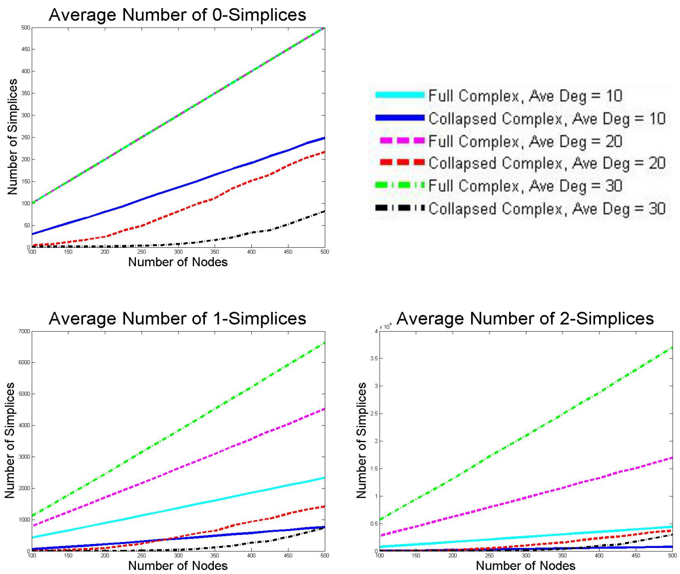


Fig. 2. Number of 0-, 1-, and 2-simplices before and after the collapse in each regime

We studied the effect of the collapse on simplex counts by generating geometric random graphs with average degrees ranging from 5 to 35, and the number of nodes ranging from 100 to 500. Geometric random graphs are effective models of sensor networks, constructed by randomly scattering nodes in a unit square, and building the communication graph by connecting pairs of nodes which are within a certain radius. The average degree of the nodes in the network is closely related to the radius chosen above [14]. We generated 500 examples of each network in Figure 1. We discovered 3 regimes among these networks. The cause of this partitioning into regimes is a complex question, and will be the subject of future work. There is a subcritical regime, in which the nodes don't form a large connected component, and so the coverage area has few holes, but isn't covering a great deal of the area, either. Next, there is the critical regime, in which the area is mostly covered except for a few holes, which homology detects. Finally, there is the supercritical regime, in which the entire region is covered due to the sheer density of the node distribution in the region. The average reduction in 0-, 1-, and 2-simplices in each regime is displayed in Figure 2.

This algorithm runs with a message-passing complexity of $O(|N_v|^2)$ for the sensor v, where $|N_v|$ is its number of neighbors. This is because each node must pass a signal of size $|N_v|$ to each of its neighbors in N_v. Because each node must update its neighbor list to delete all nodes which turned off in the current iteration, the overall message-passing complexity of the algorithm is $O(\sum\limits_{v} |N_v|^2 + \sum\limits_{w \text{ Collapsed}} |N_w|)$ for the first iteration. Because nodes can only be turned off in the algorithm, the per-iteration message-passing complexity is bounded by the complexity of this first iteration. Furthermore, since each node is quick-sorting its neighbor list and comparing it to another such list, the per-node computational complexity of the algorithm is $O(|N_v|^2 \log |N_v|)$ per iteration. It should be noted that after the first iteration, the only nodes executing this step are those whose neighbor sets have changed. Finally, the number of iterations needed for the algorithm to stabilize is bounded by the diameter of the communication graph.

V. Summary and Conclusion

We presented here a distributed algorithm for reducing the number of sensors needed to accurately detect the topology of the coverage region of a sensor network. We showed that it is equivalent to the previously developed strong collapse, and that it therefore inherits the properties of preserving the topology and the precise locations of holes in a network. These properties guarantee that the resulting collapsed complex can be used to locate holes in the original network by way of locating them in the collapsed network. The algorithm was derived solely from the properties of a flag complex, and therefore, it can be used to collapse the flag complex of any graph. We justified the collapse with simulations demonstrating the degree to which the network is collapsed in various density regimes, and showed that with one-hop information, the network can be minimized prior the computational bottleneck of finding cliques in the network.

References

[1] R. H. Atkin, "From cohomology in physics to q-connectivity in social sciences," *International Journal of Man-Machine Studies*, vol. 4, pp. 341–362, 1972.
[2] J. A. Barmak and E. G. Minian, "Strong homotopy types, nerves, and collapses," *Discrete & Computational Geometry*, vol. 47, pp. 301–328, 2012.
[3] K. Borsuk, "On the imbedding of systems of compact in simplicial complexes," *Fundamenta Mathematicae*, vol. 35, issue 1, pp. 217-234, 1948.
[4] O. Busaryev, S. Cabello, C. Chen, T. K. Dey, and Y. Wang, "Annotating simplices with a homology basis and its applications," *Algorithm Theory–SWAT 2012*, pp. 189–200, Springer, 2012.
[5] G. Carlsson, "Topology and data," *Bulletin of the American Mathematical Society*, vol. 46, pp. 255–308, 2009.
[6] H. Chintakunta and H. Krim, "Divide and conquer: Localizing coverage holes in sensor networks," in *Proceedings of the 2010 7th Annual IEEE Communications Society Conference Sensor Mesh and Ad Hoc Communications and Networks (SECON)*, Boston, MA, 2010, pp. 1–8.
[7] C. H. Dowker, "Homology groups of relations," *The Annals of Mathematics*, vol. 56, pp. 84–95, 1952.
[8] J. Gamble, H. Chintakunta, and H. Krim, "Applied topology in static and dynamic sensor networks," *Proceedings of the 2012 International Conference on Signal Processing and Communications (SPCOM)*, pp. 1–5, 2012.
[9] R. Ghrist and A. Muhammad, "Coverage and hole-detection in sensor networks via homology," in *Proceedings of the 4th international symposium on information processing in sensor networks*, Piscataway, NJ, 2005, Article 34.
[10] A. Hatcher, *Algebraic Topology*. Cambridge University Press: Cambridge, 2001.
[11] J. J. Johnson, "Some structures and notation in q-analysis," *Environment and Planning B*, vol. 8, pp. 77–86, 1981.
[12] R. M. Karp, "Reducibility among combinatorial problems," http://cgi.di.uoa.gr/ sgk/teaching/grad/handouts/karp.pdf
[13] S. Maletić, M. Rajković, and D. Vasiljević, "Simplicial complexes of networks and their statistical properties," In: M. Bubak, G. D. van Albada, J. Dongarra, P. M. A. Sloot (eds.) ICCS 2008, Part II. LNCS, vol. 5102, pp. 568–575. Springer: Heidelberg, 2008.
[14] M. Penrose, *Random Geometric Graphs*, Vol. 5, Oxford: Oxford University Press, 2003.
[15] A. C. Wilkerson, T. J. Moore, A. Swami, A. H. Krim, "Simplifying the homology of networks via strong collapses," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013