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 only one-hop information in a communication network, making it faster than clique-finding algorithms that increase the number of computations necessary for hole localization. This makes it an effective pre-processing step to finding network coverage holes.**

***Index Terms*—Hole Localization, Simplicial Complex, Simplicial Collapse, Sensor Network Coverage, 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 discovering holes in a coverage region is that it doesn't take into account specific sensor locations. In practice, acquiring precise sensor locations is resource-intensive. Many applications, ranging from cell phone coverage areas to tracking in security and defense involve battery-powered remote devices, so energy conservation is paramount. 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 define the *relevance* of a node, and use this notion to 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. Furthermore, the algorithms will only use one-hop information to execute the collapse.

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 (vertices) 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 equivalence classes of connected components, loops, 3-D voids, and their generalizations, the higher dimensional *cycles* that don't bound a subcomplex 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 as β_i. We shall not go into the details of computing homology, which can be found in most introductory algebraic topology texts, such as [10].

B. Constructing the Flag Complex of a Graph

Corresponding to a graph $G=(V,E)$ is a natural simplicial complex structure 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 [2], [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 in (X^T, V, L) 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 necessarily 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

$$\check{q}(\Delta_l) := \max_{j \in L}\{\dim(\Delta_l \cap \Delta_j)\}.$$

That is, $\check{q}(\Delta_l)$ is the dimension of a maximal face of Δ_l shared with any other labelled simplex $\Delta_j \in X$. In the event that Δ_l intersects no labelled simplices Δ_j, we define $\check{q}(\Delta_l) = -1$, so that $\mathrm{ecc}(\Delta_l) \in [0, 1]$). It immediately follows that a simplex Δ_l has eccentricity 0 if and only if $\hat{q}(\Delta_l) = \check{q}(\Delta_l)$. In other words, $\Delta_l \subset \Delta_j$. Since Δ_l is not locally maximal or is a repeated label 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 $(\tilde{X}, \tilde{L}, \tilde{V})$.

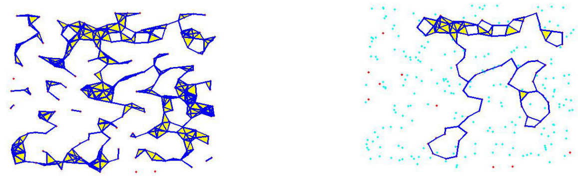
The *strong collapse* of a labelled simplicial complex (X, L, V) is as follows: (X^T, V, L) is constructed and has all of its 0-eccentricity simplex labels removed, giving $\left(\widetilde{X^T}, \tilde{L}, \tilde{V}\right)$. Then, the conjugate of the resulting complex is constructed again. Finally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

$$\begin{tikzcd}
(X,L,V) \arrow[r, "Conjugate"] & (X^T,V,L) \arrow[r, "0-ecc", "Removal"'] & (\widetilde{X^T},\tilde{V},\tilde{L}) \arrow[ddl, "Conjugate"] \\
\cup & & \\
\left(\widetilde{\left(\widetilde{X^T}\right)^T}, \tilde{\tilde{L}}, \tilde{\tilde{V}}\right) & \left(\left(\widetilde{X^T}\right)^T, \tilde{L}, \tilde{V}\right) \arrow[l, "0-ecc"', "Removal"] &
\end{tikzcd}$$

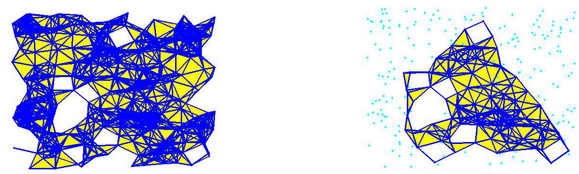
Two theorems proven in [15] show the value of this collapse:

Theorem 1. *The strong collapse produces a subcomplex of X with isomorphic homology to X.*

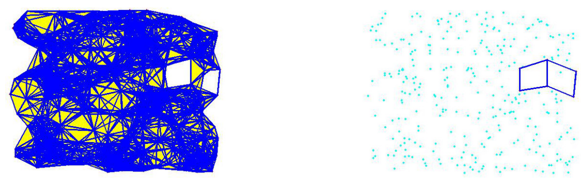
Theorem 2. *The strong collapse preserves at least one of the shortest paths around each hole and void in X.*



(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

Combining the first theorem with the facts that nodes may only be deleted by an iteration of the collapse, and that every bounded monotonic sequence converges, yields the result that the collapse must converge in a finite number of iterations to a *core complex*. The second theorem 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 $\mathcal{F}(G)$ (and will assume maximal cliques to be the only labelled simplices) 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 needs to be run. The value of 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:

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

It is useful to find a direct geometric interpretation of

$\hat{q}(\Delta_v^T)$ and $\check{q}(\Delta_v^T)$: $\hat{q}(\Delta_v^T)$ is the number of locally maximal simplices incident to v, while $\check{q}(\Delta_v^T)$ is the maximal number of locally maximal simplices shared by v with some other vertex w. Therefore, $\text{rel}(v) = 0$ 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 knows the identity of each neighbor in its neighbor set N_v, and that $v \in N_v$.

Theorem 3. *Given adjacent vertices v and w, $N_v \subset N_w$ 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 $x \in N_w$.

($\Rightarrow$) Given a maximal n-simplex Δ incident to v, without loss of generality, let $\Delta = \langle x_1, x_2, \cdots, x_n, v\rangle$. 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 graph stabilizes. Given that the sensors have unique IDs, denoted $v_1, \ldots, v_M$, and that sensor v_i has m_i neighbors, each sensor v_i executes the following steps at each iteration:

Broadcast $N_{v_i} = \{v_{i_j}\}_{j=1}^{m_i}$ to neighbors.
for $j = 1 \to m_i$ **do**
  v_i receives $N_{v_{i_j}}$
  Compare $N_{v_{i_j}}$ with N_{v_i}
  if $N_{v_{i_j}} \subset N_{v_i}$ **then**
    Broadcast TURN OFF command to v_{i_j}
    if TURN OFF command received from v_{i_j} **then**
      Handshake to determine which sensor turns off
    end if
  end if
end for
if TURN OFF received OR Handshake decides v_i turns OFF **then**
  v_i stops broadcasting
else
  Update neighbor set N_{v_i}, 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 maximum distance at which each 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 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 $\cup_{v_i \in S} D_i$, the sensor coverage region. Moreover, the generators of $H_1(\check{C}(S, r_s))$ 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 requires specific geometric information, including the coordinates of the nodes, 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 below which 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, $R(S, r_c) := \mathcal{F}(G(S))$. 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 $H_1(R(S, r_c))$ are geometrically small. They can be contained in triangles making up a very small percentage of the total area covered by the network.

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

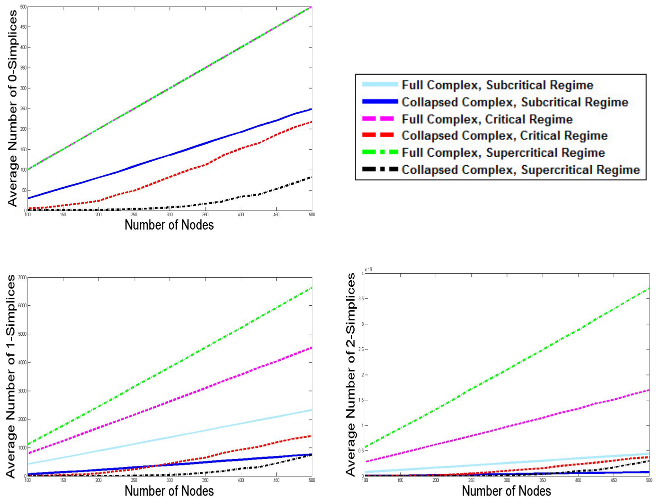


Fig. 2. Average number of 0-, 1-, and 2-simplices before and after the collapse in each regime

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.

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 distance. It is known that geometric random graphs can be categorized by node density into 3 regimes [14]. When the points are not very dense, the network is in the *subcritical* regime. In the *supercritical* regime, the points are very densely packed into the coverage space, which often results in complete coverage. Between these two extremes is the *critical* regime, in which the radius of the coverage balls centered at the nodes is explicitly related to the density of the nodes in the region. 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. We generated 500 examples of each network, examples of which may be found in Figure 1, and collapsed them; 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. Since 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 $O\big(\sum_v |N_v|^2\big)$ 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, only nodes whose neighbor sets have changed execute this step. Finally, the number of iterations before the complex converges to its core is bounded by the diameter of the communication graph.

V. Summary and Conclusion

We presented a simple distributed algorithm for reducing the number of sensors needed to accurately detect the topology of the coverage region of a sensor network. We showed through the new concept of node relevance 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, while converging in finite time. 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 to 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*, 4, 1972, pp. 139–167.

[2] J. A. Barmak and E. G. Minian, “Strong homotopy types, nerves, and collapses,” *Discrete & Computational Geometry*, 47, 2012, pp. 301–328.

[3] K. Borsuk, “On the imbedding of systems of compact in simplicial complexes,” *Fundamenta Mathematicae*, 35, issue 1, 1948, pp. 217-234.

[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*, Helsinki, Finland, Springer, 2012, pp. 189–200.

[5] G. Carlsson, “Topology and data,” *Bulletin of the American Mathematical Society*, 46, 2009, pp. 255–308.

[6] H. Chintakunta and H. Krim, “Divide and conquer: Localizing coverage holes in sensor networks,” in *Proceedings 7th Annual IEEE SECON*, Boston, MA, 2010, pp. 1–8.

[7] C. H. Dowker, “Homology groups of relations,” *The Annals of Mathematics*, 56, 1952, pp. 84–95.

[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)*, 2012, pp. 1–5.

[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*, 8, 1981, pp. 77–86.

[12] R. M. Karp, “Reducibility among combinatorial problems,” in *Complexity of Computer Computations*, R.E. Miller and J. W. Thatcher, eds., Plenum Press, 1972, pp. 85-103.

[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, 5102, Springer: Heidelberg, 2008, pp. 568–575.

[14] M. Penrose, *Random Geometric Graphs, Volume 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 IEEE ICASSP*, Vancouver, Canada, May 2013, pp. 341–362.