EDGE-ERASURES AND CHORDAL GRAPHS

JARED CULBERTSON, DAN P. GURALNIK, AND PETER F. STILLER

Abstract. We give a new characterization of chordal graphs via edge deletion from a complete graph, where edges eligible for deletion are each properly contained in a single maximal clique. Interestingly, in this context such edges always arise in cycles, and we use this fact to give a modified version of Kruskal’s second algorithm for finding a minimum spanning tree in a finite metric space.

1. Introduction

In this short paper we prove several results about chordal graphs by focusing on edges which are each properly contained in a unique maximal clique; these we call exposed edges. Our first main result gives a new characterization of chordal graphs as those that can be produced through a sequence of exposed edge deletions starting from a complete graph. This is in contrast to the usual vertex-centric characterizations of chordal graphs in terms of elimination orderings or minimal separators (see [2] for a survey of these), and also distinct from edge-without-vertex elimination orderings of related graph classes (for example, see the characterization of strongly orderable graphs in [4]). Secondly, and perhaps most interestingly, we prove that exposed edges in a chordal graph always occur in a cycle of exposed edges. This, in turn, leads to a third main result, namely a variation of Kruskal’s second algorithm [7] for finding a minimum spanning tree in a weighted graph, with its attached relationship to ultrametrics and single-linkage clustering. This relationship is discussed in more detail in Section 3.

Our initial investigations were motivated by our theoretical work on data clustering (see [1]) and our search for an adequate notion of a minimal spanning complex for our $A^n$ clustering methods, analogous to the role played by minimal spanning trees for single-linkage clustering. Thus we were focused more in the realm of topology and the study of flag complexes obtainable by collapses from a simplex. The language in this case is interchangeable since the flag condition means the abstract simplicial complex is completely determined by its 1-skeleton. We have chosen the graph theoretical language for a more consistent presentation, but all the results can be restated topologically in terms of chordal complexes, which are flag complexes whose 1-skeleton is a chordal graph. For more on this topological perspective and the relationship to simplicial collapses, see the final section of the paper.
2. ERASURES

We begin by collecting some basic definitions, notation, and terminology.

**Definition 1.** Let $G = (V, E)$ be an undirected simple graph (with no loops or multiple edges) having finite vertex set $V$ and edge set $E$. The open $G$-neighborhood of a vertex $v \in V$ is

$$N_G(v) = \{ w \in V \mid vw \in E \}.$$ 

The closed neighborhood $N_G[v] = N_G(v) \cup \{v\}$. We will denote the induced subgraph on $A \subseteq V$ by $G[A]$. On occasion, we will simplify notation by understanding $N_G(v)$ or $N_G[v]$ to be the induced subgraph $G[N_G(v)]$ or $G[N_G[v]]$. Whether we are referring to the induced subgraph or just the vertex set will be clear from the context. Note $G[N_G(v)]$ is sometimes called the *link* of $v$, particularly in a more topological setting.

If $v_1, \ldots, v_k$ is an ordering on $V$, let $G_i = G[\{v_i, \ldots, v_k\}]$. A vertex $v$ is *simplicial* if the induced subgraph on $N_G[v]$ is complete. We say that a graph has a *perfect elimination ordering* if there is some ordering of $V$ such that $v_i$ is simplicial in $G_i$ for each $1 \leq i \leq k$. Recall also that a *bridge* is a cut-edge, *i.e.*, an edge whose removal increases the number of connected components of the graph. An equivalent characterization is that a bridge is an edge that does not lie in any cycle of the graph.

**Definition 2.** An undirected simple graph $G$ is *chordal* if every induced cycle has length three. Chordality is an induced-hereditary property.

There are many characterizations of chordal graphs available in the literature. We will not attempt here to give a full survey of the relevant results, but rather point the reader to [2], which provides an excellent guide to the related literature. However, there is one characterization that we will need in the sequel and one implication—we combine those as a theorem here.

**Theorem 3** ([3, 5]). A graph is chordal if and only if it has a perfect elimination ordering. Moreover, any chordal graph is either complete or has two non-adjacent simplicial vertices.

Borrowing from topology, and to simplify the exposition, we refer to an edge whose vertices induce a two-element maximal clique as a *facet edge*. The following lemma, however, shows that for chordal graphs, the notions of bridge and facet edge are equivalent; although this is not true for an arbitrary graph.

**Lemma 4.** Let $G$ be a graph. If an edge $xy \in G$ is a bridge, then it is a facet edge. Additionally, if $G$ is chordal, then the converse holds.

*Proof.* If $xy$ is a bridge, then it is not contained in any larger clique, since this would mean, for instance, that $xy$ is in a cycle through the vertices of that clique and removing $xy$ would not increase the number of connected components. Hence $xy$ is a facet edge.
If \( xy \) is not a bridge and \( G \) is chordal, then we can find a cycle \( yvxy \) in \( G \) (take any cycle containing \( xy \) and use chordality to shorten the cycle to this point). But this means that \( \{x, y, v\} \) forms a clique, and so \( xy \) is not a facet edge. \( \square \)

**Definition 5.** Let \( G \) be a graph. An edge \( xy \in G \) is said to be exposed, if \( xy \) is contained in a unique maximal clique and \( xy \) is not a facet edge.

**Definition 6.** Suppose \( G, H \) are graphs with vertex set \( V \). We say that \( H \) is obtained from \( G \) through an erasure, if \( G \) contains an exposed edge \( e \) such that \( H = G - e \).

The topological nature of an erasure, which can be described in terms of a strong deformation retraction, will be discussed in Section 4. We now provide a useful characterization of exposed edges.

**Lemma 7.** An edge \( vw \in G \) is exposed if and only if \( N_G(v) \cap N_G(w) \) is a nonempty clique in \( G \).

**Proof.** We remark that for any two vertices \( v, w \in G \), the intersection \( N_G(v) \cap N_G(w) \) is just the union of all maximal cliques which contain both \( v \) and \( w \), minus \( \{v, w\} \). The result follows in a straightforward way from this observation and the definitions. \( \square \)

The previous lemma highlights that our notion of an exposed edge is weaker than that of a simplicial edge [4], where the intersection is replaced by the union of the neighborhoods. Indeed, it follows from the lemma that an edge \( vw \) is exposed if and only if \( w \) is a non-isolated simplicial vertex of \( N_G(v) \), and vice versa.

**Theorem 8.** A graph \( H \) can be obtained from a complete graph through a sequence of erasures if and only if \( H \) is a connected chordal graph.

**Proof.** First, we can see that erasures from connected chordal graphs produce connected chordal graphs as follows. Suppose \( H = G - xy \), with \( G \) a connected chordal graph and \( xy \) an exposed edge in \( G \). If \( C \subseteq H \) is an induced cycle not containing \( \{x, y\} \), then \( C \) is also an induced cycle of \( G \) and so of length 3. Otherwise, suppose \( \{x, y\} \subseteq C \) and \( |C| > 3 \). Note that if \( |C| > 4 \), then the induced subgraph \( C' = C + xy \) of \( G \) has an induced cycle of length greater than 3, a contradiction. This leaves us with the case where \( C = xv_1yv_2x \) for some \( v_1, v_2 \). Since \( xy \) is exposed in \( G \), we must have \( v_1v_2 \in G \), otherwise \( xy \) would lie in two distinct maximal cliques and \( xy \) would not be exposed. However, \( v_1v_2 \in G \) (hence in \( H \)) means that \( C \) would not be an induced cycle in \( H \), a contradiction. As for connectedness, it is easy to see that an erasure does not disconnect a connected graph since by definition an exposed edge is not a facet edge.

Conversely, it suffices to show that for any non-complete connected chordal graph \( G \), we can add an edge \( e \) such that \( e \) is exposed in \( G + e \). Given such a \( G \), suppose \( v_1, \ldots, v_k \) is a perfect elimination ordering for \( G \). Let \( 1 \leq \ell \leq k \)
be the largest index such that \( G_i \) is complete for \( i > \ell \). Then there is some \( j > \ell \) with \( v_jv_j \notin G \), because \( G_{\ell} \) is not complete, but \( G_{\ell+1} \) is.

Let \( G' = G + v_jv_j \). We claim that \( v_1, \ldots, v_k \) is also a perfect elimination ordering for \( G' \). For \( i < \ell \), the neighbors of \( v_i \) in \( G' \) are just the same neighbors of \( v_i \) in \( G \), and \( N_{G'}(v_i) \) is a clique since \( v_i \) is simplicial in \( G_i \). In particular, \( \{v_j, v_{\ell}\} \notin N_{G'}(v_i) \) since \( v_jv_j \notin G \). Thus \( v_{\ell} \) is also simplicial in \( G'_{\ell} \).

On the other hand, for \( i > \ell \), \( G_i \) (and hence \( G'_{\ell} \)) is complete and so every vertex is simplicial. We still need to check that \( v_{\ell} \) is simplicial in \( G'_{\ell} \). This follows from the fact that \( v_{j}v_{n} \in G_{\ell}' \) for all \( n > \ell \) since \( G_{\ell+1}' \) is complete.

It remains to show that \( v_{j}v_{j} \) is exposed in \( G' \). Notice again that for \( i < \ell \), we must have that \( \{v_j, v_{\ell}\} \notin N_{G'}(v_i) \), since as noted above, \( \{v_j, v_{\ell}\} \notin N_{G'}(v_i) \). Hence

\[
N_{G'}(v_j) \cap N_{G'}(v_{\ell}) = N_{G'}(v_{\ell}) \setminus \{v_j\},
\]

which is a clique as shown above, because \( v_{\ell} \) is simplicial in \( G_{\ell}' \).

This result is similar in spirit to the result of Spinrad and Srinathanan showing that weakly chordal graphs can be recognized by the possibility of successively adding edges through their two-pair construction to arrive at a complete graph.

**Lemma 9.** If \( G \) has no facet edges and \( v \) is a simplicial vertex of \( G \), then every edge \( e \) incident on \( v \) is exposed.

**Proof.** Notice that every maximal clique (thought of as an induced subgraph) containing \( e \) necessarily contains \( v \). Because \( v \) is contained in a single maximal clique (specifically, \( N_{G}(v) \)), we must have that \( e \) is also contained only in \( N_{G}(v) \). Since \( G \) has no facet edges, \( e \) is not a facet edge, and so is exposed. \( \square \)

**Proposition 10.** Any connected chordal graph can be reduced through a sequence of erasures to a tree.

**Proof.** Let \( G \) be a connected chordal graph and \( G' \) be the resulting graph after removing all facet edges. Then the connected components of \( G' \) are chordal and hence each contains a simplicial vertex. Thus Lemma 9 shows that there are exposed edges in \( G' \) (which are then also exposed in \( G \)) unless all edges of \( G \) are facet edges. In this case, \( G \) is necessarily a tree. \( \square \)

**Lemma 11.** Let \( G \) be chordal and \( v \) a vertex in \( G \) that is in a maximal clique of size at least three. Then \( v \) is incident on at least two exposed edges.

**Proof.** Since \( v \) is in a maximal clique of size at least three, the induced subgraph \( N_{G}(v) \) is chordal and not edgeless. Thus there is some connected component of \( N_{G}(v) \) containing an edge, and so we can use Theorem 3 to find two simplicial vertices \( v_1, v_2 \) in \( N_{G}(v) \) in that component. But this implies that \( N_{G}(v) \cap N_{G}(v_i) \) is a (non-empty) clique for \( i = 1, 2 \), and so \( vv_1 \) and \( vv_2 \) are exposed edges in \( G \). Note that an isolated vertex \( u \) in \( N_{G}(v) \) would correspond to a facet edge \( uv \) in \( G \). \( \square \)
**Theorem 12.** If $G$ is a chordal graph, then every exposed edge of $G$ is contained in a cycle of exposed edges.

*Proof.* It is simple to check that the lemma holds when $G$ is either complete or has no more than four vertices. Now suppose that $G$ is a counterexample with a minimal number of vertices (so $|G| \geq 5$). Notice that $G$ is then bridgeless since if $e \in E(G)$ is a bridge, then $e$ is not exposed by definition and the connected components of $G - e$ have fewer vertices. However, every exposed edge in $G$ is an exposed edge in some component and so is in a cycle of exposed edges in that component by induction, which necessarily does not include $e$ since $e$ is a bridge.

Now since $G$ is chordal, Theorem 3 allows us to find non-adjacent simplicial vertices $u, v \in G$. Then $G - u$ has fewer vertices and so every exposed edge of $G - u$ is contained in a cycle of exposed edges in $G - u$. Notice that for vertices $x, y \in G - u$, we have that

$$N_{G - u}(x) \cap N_{G - u}(y) = [N_G(x) \cap N_G(y)] \setminus \{u\}.$$

and so using Lemma 7 we see that if $\{x, y\} \not\subseteq N_G[u]$, then $xy$ is exposed in $G - u$ if and only if $xy$ is exposed in $G$. Thus in this case, if $xy$ is exposed in $G$ we can find a cycle $C = xyv_1 \cdots v_kx$ in $G - u$ of exposed edges in $G - u$. If none of the edges in $C$ are in $N_G(u)$, then $C$ is also a cycle of exposed edges in $G$. However, if some edges of $C$ are contained in $N_G(u)$, then let $i$ be the smallest index with $v_i v_{i+1}$ in $N_G(u)$. Similarly, let $j$ be the largest index with $v_j v_{j+1}$ in $N_G(u)$. Notice that $j > i$, but we could have $j = i + 1$ if there is a single edge of $C$ in $N_G(u)$ (in order to make the notation consistent, we are treating $v_0$ as $y$ and $v_{k+1}$ as $x$). We can use Lemma 8 to see that the cycle $C' = xyv_1 \cdots v_{i-1}v_i uv_j v_{j+1} \cdots v_kx$ is a cycle of exposed edges in $G$ containing $xy$.

The remaining case to be considered is when $\{x, y\} \subseteq N_G[u]$. Here notice that $\{x, y\} \not\subseteq N_G(v)$, since $xy$ is exposed and $u$ and $v$ are not adjacent. Hence, by the same reasoning as before, we can find a cycle of exposed edges in $G$ containing $xy$ by modifying a cycle of exposed edges in $G - v$ containing $xy$. \qed

3. $d$-erasures and Weighted Chordal Graphs

We turn now to an application of these results in the setting of finite metric spaces, and show a connection with single-linkage clustering through minimum spanning trees.

**Definition 13.** Let $d$ be a metric on a finite set $X$ and suppose $G, H$ are weighted graphs with vertex set $X$ and edge weighting given by $d$. We say that $H$ is obtained from $G$ through a $d$-erasure, if $H$ is obtained from $G$ through an erasure of an exposed edge $e$ such that $d(e) \geq d(e')$ for any exposed edge $e'$ of $G$.

Observe that given any sequence of erasures $G_0, G_1, \ldots, G_m$, with $G_0$ a complete graph and deleted edges $e_0, e_1, \ldots, e_{m-1}$, we can define a metric
(V(G), d) such that G_0, G_1, ..., G_m is also a sequence of d-erasures. For example, for \( \varepsilon < \frac{1}{m-1} \), we could define

\[
d_{xy} = \begin{cases} 
1 & \text{if } xy \in G_m \\
2 - j \varepsilon & \text{if } xy = e_j
\end{cases}
\]

and obtain G_0, G_1, ..., G_m as a sequence of d-erasures.

Recall that a minimum spanning tree for a weighted graph G is a spanning tree which minimizes the sum of the weights over the edges of the tree. The following theorem shows that d-erasure preserves the existence of a minimum spanning tree.

**Theorem 14.** Let (X, d) be a finite metric space and let \( m \geq 0 \) be an integer. If G_0, G_1, ..., G_m is a sequence of graphs with vertex set X, where G_0 is the complete graph and each G_{i+1} is obtained from G_i through a d-erasure, then G_m contains a minimum spanning tree of (X, d).

**Proof.** First, recall as stated above that all of the G_i are connected since exposed edges are by definition not bridges, and these are the only edges removed at each stage.

G_0 contains a minimum spanning tree of d, providing the base for an induction on m. It is also clear that the lemma holds true whenever \( |X| \leq 3 \). Let m \geq 1 and suppose that G_i contains a minimum spanning tree T of (X, d) for 0 \leq i < m. Suppose G_m = G_{m-1} - xy, where xy is an exposed edge of G.

Let T_x, T_y denote the connected components of x and y, respectively, in T - xy. Let E denote the set of all G_{m-1} edges zw with z \in T_x and w \in T_y, excluding the edge xy. Since G_m is connected, E intersects G_m. For any zw \in E, the graph T' := T - xy + zw is a spanning tree of X, implying d_{zw} \geq d_{xy}.

On the other hand, if zw is exposed in G_{m-1}, then d_{zw} \leq d_{xy} and so any exposed edge in E has equal weight with xy. Now we can appeal to Theorem 12 to see that xy is contained in a cycle of exposed edges of G_{m-1} which necessarily intersects E, say at zw. Thus T' = T - xy + zw is another minimum spanning tree contained in G_m.

In his seminal paper on minimum spanning trees [7], Kruskal proposed two algorithms for computing such a tree. The second of which (the less efficient one) proceeds as follows: starting with the complete graph G_0 on X endowed with the weight w, for each i \geq 0 remove from G_i a heaviest edge (that is, one whose w-value is maximal) among those not separating the current graph to obtain G_{i+1}. The process terminates after stage \( t = \binom{|X| - 1}{2} \) with G_{t+1} a tree. Using the cut property of minimum spanning trees, it is easy to argue that every minimum spanning tree of (X, w) may be obtained in this way. The previous theorem then allows us to show that, surprisingly, when restricting this algorithm to only exposed edges, we are nonetheless able to recover all minimum spanning trees, having replaced a global eligibility criterion for
an edge to be removed with a local one. Some differences between the two algorithms are illustrated in Figure 1.

Corollary 15. Let \((X, d)\) be a finite metric space. Then a maximal sequence of \(d\)-erasures produces a minimum spanning tree for \((X, d)\). Any minimum spanning tree for \((X, d)\) can be obtained in this way.

Proof. The first statement is a direct consequence of Proposition 10 and Theorem 14. For the second, let us start with a given minimal spanning tree \(T\) for \((X, d)\), and a sequence \(G_0, \ldots, G_k\) of graphs obtained by erasure, with \(G_0\) the complete graph on \(X\) and \(G_i\) containing \(T\) for each \(0 \leq i \leq k\). If \(G_k \neq T\), then for any exposed edge \(xy\) in \(T\) of maximal weight (among the exposed edges of \(G_k\)), the same reasoning as in the proof of Theorem 14 (and using the same notation), shows that there must be another exposed edge \(zw\) in \(G_k\) with equal weight as \(xy\) and \(z \in T_x, w \in T_y\) (in particular, \(zw \notin T\)). Thus we can extend the sequence by setting \(G_{k+1} = G_k - zw\).

4. Connections to the Topological Viewpoint

Topologically, we can view the characterization of chordality given in Theorem 3 in terms of perfect elimination orderings as providing the basis for realizing chordal graphs as the 1-dimensional skeleta of simplicial flag complexes assembled through successive “coning-off” of existing simplices; or (by reversing the perspective) of simplicial flag complexes which admit an exceedingly tame kind of strong-deformation retraction to a vertex through a sequence of “vertex-collapses.” Put in the language of simple homotopy theory (see, e.g. [6], Definition 6.13 and the ensuing discussion), erasing a simplicial vertex \(w\) of a chordal graph \(G\) is realized in the polyhedron \(|K|\) of the subtended complex \(K\) as the straight-line homotopy from the identity mapping of \(|K|\) to the (realization of the) simplicial map \(K \to \text{sd}(K)\). This homotopy fixes all vertices of \(K - w\) and maps \(w\) to the barycenter of its
opposing face in $K$, which is the face subtended by the collection of the
neighbors of $w$ in $G$, see Figure 2(left).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{Collapsing an ‘exposed’ vertex $w$ in a 3-facet (left); an exposed edge $uv$ in a 2-facet (center); and an exposed edge in a 3-facet (right).}
\end{figure}

Using the same language, the erasure process described in this paper can
be understood as another restricted type of strong deformation retraction
characterized, at the level of one-dimensional skeleta, by the removal of
exposed edges. Indeed, at the level of the complexes $K_i = K_{G_i}$, it quickly
becomes evident that erasing an arbitrary edge of $G_i$ to obtain $G_{i+1}$ (as
required by Kruskal’s algorithm) does not guarantee a strong deformation
retraction of $K_i$ onto $K_{i+1}$, unless the edge being removed is exposed—\textit{i.e.}, it
is properly contained in a unique maximal simplex of $K_i$. Then it is possible
to eliminate the edge by “pressing in” in the form of an \textit{edge-collapse}, see
Definition 6.13 in [6] and Figure 2(center,right). Homotopy equivalences
of this kind have been studied by combinatorial algebraic topologists since
the introduction of the notions of collapsibility and simple homotopy types
by Whitehead [10, 9] (also see [6], Chapter 6, for an overview and more
modern treatment). Our results, then, provide an understanding of chordal
graphs as 1-skeleta of connected flag complexes arising as strong deformation
retractions of a simplex, providing an interpretation of chordality from a
standpoint of extendibility.

We close by briefly noting that this approach could be generalized by
considering simplicial complexes other than the simplex as starting points,
or \textit{ambient complexes}, for the erasure process. For example, an interesting
replacement would be the standard triangulation of the $n$-cube induced by
its isomorphism with the Hasse diagram of the inclusion order in a power
set. The corresponding question, then, is to identify which families of com-
plexes/graphs might be characterized as emerging from some ambient com-
plex $S$ by excavating them out of $S$ via repeated application of a restricted
family of collapses, subject to a suitable stopping condition.

\textbf{References}

[1] J. Culbertson, D. P. Guralnik, and P. F. Stiller, \textit{Functorial hierarchical clus-
tering with overlaps}, arXiv:1609.02513, (2017).

[2] P. De Carla, \textit{A joint study of chordal and dually chordal graphs}, PhD thesis, Uni-
versidad Nacional de La Plata, 2012.

[3] G. A. Dirac, \textit{On rigid circuit graphs}, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg, 25 (1961), pp. 71–76.
[4] F. F. Dragan, *Strongly orderable graphs: A common generalization of strongly chordal and chordal bipartite graphs*, Discrete Appl. Math., 99 (2000), pp. 427–442.

[5] D. R. Fulkerson and O. A. Gross, *Incidence matrices and interval graphs*, Pacific J. Math., 15 (1965), pp. 835–855.

[6] D. Kozlov, *Combinatorial algebraic topology*, vol. 21, Springer Science & Business Media, 2007.

[7] J. B. Kruskal, *On the shortest spanning subtree of a graph and the traveling salesman problem*, Proceeding of the American Mathematical Society, 7 (1956), pp. 48–50.

[8] J. Spinrad and R. Sritharan, *Algorithms for weakly triangulated graphs*, Discrete Appl. Math., 59 (1995), pp. 181–191.

[9] J. Whitehead, *Simple homotopy types*, American Journal of Mathematics, 72 (1950), pp. 1–57.

[10] J. H. C. Whitehead, *Simplicial spaces, nuclei and m-groups*, Proceedings of the London mathematical society, 2 (1939), pp. 243–327.

Sensors Directorate, Air Force Research Laboratory, 2241 Avionics Circle, Building 620, Wright-Patterson Air Force Base, Ohio 45433-7302, USA. Email: jared.culbertson@us.af.mil

Electrical & Systems Engineering Dept., University of Pennsylvania, 200 S. 33rd st., 203 Moore Bldg., Philadelphia, Pennsylvania 19104-6314, USA. Email: guraldan@seas.upenn.edu

Department of Mathematics, MS3368, Texas A&M University, College Station, Texas 77843-3368, USA. Email: stiller@math.tamu.edu