Binomial edge ideals and conditional independence statements

Jürgen Herzog\textsuperscript{a}, Takayuki Hibi\textsuperscript{b,}\textsuperscript{*}, Freyja Hreinsdóttir\textsuperscript{c}, Thomas Kahle\textsuperscript{d}, Johannes Rauh\textsuperscript{d}

\textsuperscript{a} Fachbereich Mathematik, Universität Duisburg–Essen, Campus Essen, 45117 Essen, Germany
\textsuperscript{b} Department of Pure and Applied Mathematics, Graduate School of Information Science and Technology, Osaka University, Toyonaka, Osaka 560-0043, Japan
\textsuperscript{c} School of Education, University of Iceland, Stakkahlid, 105 Reykjavik, Iceland
\textsuperscript{d} MPI for Mathematics in the Sciences, 04103, Leipzig, Germany

\textbf{A R T I C L E I N F O}

\textbf{Article history:}
Received 25 October 2009
Accepted 18 January 2010
Available online 19 February 2010

\textbf{MSC:}
primary 13P10, 60E05
secondary 13H10, 62H17, 60C05

\textbf{Keywords:}
Binomial ideals
Edge ideals
Cohen–Macaulay rings
Conditional independence ideals
Robustness

\textbf{A B S T R A C T}

We introduce binomial edge ideals attached to a simple graph $G$ and study their algebraic properties. We characterize those graphs for which the quadratic generators form a Gröbner basis in a lexicographic order induced by a vertex labeling. Such graphs are chordal and claw-free. We give a reduced squarefree Gröbner basis for general $G$. It follows that all binomial edge ideals are radical ideals. Their minimal primes can be characterized by particular subsets of the vertices of $G$. We provide sufficient conditions for Cohen–Macaulayness for closed and nonclosed graphs.

Binomial edge ideals arise naturally in the study of conditional independence ideals. Our results apply for the class of conditional independence ideals where a fixed binary variable is independent of a collection of other variables, given the remaining ones. In this case the primary decomposition has a natural statistical interpretation.

\textcopyright 2010 Elsevier Inc. All rights reserved.

\section{Introduction}

Let $G$ be a simple graph on the vertex set $[n] = \{1, \ldots, n\}$, that is to say, $G$ has no loops and no multiple edges. Furthermore let $K$ be a field and $S = K[x_1, \ldots, x_n, y_1, \ldots, y_n]$ be the polynomial ring...
in 2n variables. For \( i < j \) we set \( f_{ij} = x_i y_j - x_j y_i \). We define the \textit{binomial edge ideal} \( J_G \subset S \) of \( G \) as the ideal generated by the binomials \( f_{ij} = x_i y_j - x_j y_i \) such that \( i < j \) and \( \{i, j\} \) is an edge of \( G \). Note that if \( G \) has an isolated vertex \( i \), and \( G' \) is the restriction of \( G \) to the vertex set \( [n] \setminus \{i\} \), then \( J_G = J_{G'} \).

The class of binomial edge ideals is a natural generalization of the ideal of 2-minors of a \( 2 \times n \)-matrix of indeterminates. Indeed, the ideal of 2-minors of a \( 2 \times n \)-matrix may be interpreted as the binomial edge ideal of a complete graph on \( [n] \). Related to binomial edge ideals are the ideals of adjacent minors considered by Hoşten and Sullivant [9]. In the case of a line graph our binomial edge ideal may be interpreted as an ideal of adjacent minors. This particular class of binomial edge ideals has also been considered by Diaconis, Eisenbud and Sturmfels in [4] where they compute the primary decomposition of this ideal.

Binomial edge ideals, as they are defined in this paper, also arise in the study of conditional independence statements [5]. They generalize a class which has been studied by Fink [7].

Classically one studies edge ideals of a graph \( G \) which are generated by the monomials \( x_i x_j \) where \( \{i, j\} \) is an edge of \( G \). The edge ideal of a graph has been introduced by Villarreal [12] where he studied the Cohen–Macaulay property of such ideals. The purpose of this paper is to study the algebraic properties of binomial edge ideals in terms of properties of the underlying graph. In Section 1 we consider the Gröbner basis of \( J_G \) with respect to the lexicographic order induced by \( x_1 > x_2 > \cdots > x_n > y_1 > y_2 > \cdots > y_n \). We show in Theorem 1.1 that the property of \( J_G \) to have a quadratic Gröbner basis with respect to this monomial order can be characterized by a certain condition on the associated acyclic directed graph \( G^* \). By definition, \( (i, j) \) is an arrows of \( G^* \) if and only if \( (i, j, x) \) is an edge of \( G \) and \( i < j \). The condition in question is equivalent to saying that for any two distinct vertices \( i \) and \( j \) of \( G^* \), all shortest paths from \( i \) to \( j \) are directed, see Proposition 1.4. For easy reference in the further discussions we call a graph \( G \) closed with respect to the given labeling, if the associated directed graph satisfies this condition. In Proposition 1.6 we give a sufficient condition for a closed graph to have a Cohen–Macaulay binomial edge ideal. In Theorem 2.1 we compute explicitly the reduced Gröbner basis of \( J_G \) for any simple graph \( G \). This is one of the main results of this paper. As a consequence we see that the initial ideal of \( J_G \) is squarefree which in turn implies that \( J_G \) is a reduced ideal. Of course, Theorem 1.1 is a simple consequence of Theorem 2.1. But as the proof of Theorem 1.1 is quite simple and as it leads to the concept of closed graphs, we decided to present Theorem 1.1 independent from Theorem 2.1.

Section 3 is devoted to the study of the minimal prime ideals of \( J_G \). In Theorem 3.2 we write \( J = J(G) \) as a finite intersection of prime ideals which allows us to compute the dimension of \( S/J \). It turns out that if \( S/J \) is Cohen–Macaulay, then \( \dim S/J_G = |V(G)| + c \), where \( c \) is the number of connected components of \( G \). As a simple consequence of this, one sees that a circle of length \( n \) is unmixed or Cohen–Macaulay, if and only if \( n = 3 \). As a last result of Section 3 we identify in Corollary 3.9 the minimal prime ideals of \( J_G \). They are related to the cut-points of certain subgraphs of \( G \).

In the last section we discuss applications to the study of conditional independence ideals. For a class of conditional independence statements, suitable to model a notion of robustness, the results in the prior sections show that the corresponding ideal is a radical ideal. Furthermore, the primary decomposition can be computed, which yields a classification and parametrization of the set of probability distributions which satisfy these statements.

Terai informed the authors that M. Ohtani [10] independently obtained similar results for this class of ideals.

1. Edge ideals with quadratic Gröbner bases and closed graphs

We first study the question when \( J_G \) has a quadratic Gröbner basis.

**Theorem 1.1.** Let \( G \) be a simple graph on the vertex set \( [n] \), and let \( < \) be the lexicographic order on \( S = K[x_1, \ldots, x_n, y_1, \ldots, y_n] \) induced by \( x_1 > x_2 > \cdots > x_n > y_1 > y_2 > \cdots > y_n \). Then the following conditions are equivalent:

1. \( J_G \) has a quadratic Gröbner basis.
2. \( G \) is closed.
3. \( G \) is unmixed.
4. \( G \) is Cohen–Macaulay.
5. \( G \) is a complete graph.
6. \( G \) is a line graph.
7. \( G \) is a circuit.

We then consider the primary decomposition of \( J_G \) and we present an algorithm for computing it.
(a) The generators $f_{ij}$ of $J_G$ form a quadratic Gröbner basis;
(b) For all edges $[i, j]$ and $[k, l]$ with $i < j$ and $k < l$ one has $[j, l] \in E(G)$ if $i = k$, and $[i, k] \in E(G)$ if $j = l$.

**Proof.** (a) $\Rightarrow$ (b): Suppose (b) is violated, say, $[i, j]$ and $[i, k]$ are edges with $i < j < k$, but $[j, k]$ is not an edge. Then $S(f_{ik}, f_{ij}) = y_j f_{jk}$ belongs to $J_G$, but none of the initial monomials of the quadratic generators of $J_G$ divides $in_C(y_j f_{jk})$.

(b) $\Rightarrow$ (a): We apply Buchberger’s criterion and show that all $S$-pairs $S(f_{ij}, f_{kl})$ reduce to 0. If $i \neq k$ and $j \neq l$, then $in_C(f_{ij})$ and $in_C(f_{kl})$ have no common factor. It is well known that in this case $S(f_{ij}, f_{kl})$ reduces to zero. On the other hand, if $i = k$, we may assume that $l < j$. Then

$$S(f_{ij}, f_{kl}) = y_j f_{ik}$$

is the standard expression of $S(f_{ij}, f_{kl})$. Similarly, if $j = l$, we may assume that $i < k$. Then

$$S(f_{ij}, f_{kj}) = y_j f_{ik}$$

is the standard expression of $S(f_{ij}, f_{kj})$. In both cases the $S$-pair reduces to 0. □

Condition (b) of Theorem 1.1 does not only depend on the isomorphism type of the graph, but also on the labeling of its vertices. For example the graph $G$ with edges $\{1, 2\}$, $\{2, 3\}$, and the graph $G'$ with edges $\{1, 2\}$, $\{1, 3\}$ are isomorphic, but $G$ satisfies condition (b), while $G'$ does not.

In fact, condition (b) is a condition of the associated directed graph $G^*$ of $G$ which is defined as follows: the ordered pair $(i, j)$ is an arrow of $G^*$ if $[i, j]$ is an edge of $G$ with $i < j$. The directed graph $G^*$ is acyclic, that is, it has no directed cycles. Therefore we call $G^*$ also the associated acyclic directed graph of $G$.

An acyclic directed graph is also called an acyclic digraph or simply a DAG. Acyclic directed graphs constitute an important class of directed graphs and play an important role in the modeling of information flows in networks. Any acyclic directed graph arises in the same way as we obtained $G^*$ from $G$. Indeed, one of the fundamental results on acyclic directed graphs $G$ is that they admit an acyclic ordering of its vertices, that is, the vertices of $G$ can be ordered $v_1, \ldots, v_r$ such that for every arrow $(v_i, v_j)$ of $G$ we have $i < j$, see for example [2, Proposition 1.4.3]. An acyclic directed graph usually has many different acyclic orderings. In [11, Corollary 1.3] Stanley expressed the number of possible acyclic orderings in terms of the chromatic polynomial of $G$.

We say that a graph $G$ on $[n]$ is closed with respect to the given labeling of the vertices, if $G$ satisfies condition (b) of Theorem 1.1, and we say that a graph $G$ with vertex set $V(G) = \{v_1, \ldots, v_n\}$ is closed, if its vertices can be labeled by the integer $1, 2, \ldots, n$ such that for this labeling $G$ is closed.

**Proposition 1.2.** If $G$ is closed, then $G$ is chordal and has no induced subgraph consisting of three different edges $e_1, e_2, e_3$ with $e_1 \cap e_2 \cap e_3 \neq \emptyset$.

**Proof.** Suppose $G$ is not chordal, then $G$ contains a cycle $C$ of length $\geq 3$ with no chord. Let $i$ be the vertex of $C$ with $i < j$ for all $j \in V(C)$, and let $[i, j]$ and $[i, k]$ be the edges of $C$ containing $i$. Then $i < j$ and $i < k$, but $[j, k] \notin E(G)$.

Since $G$ is closed, any induced subgraph is closed as well. Suppose there exists an induced subgraph $H$ with three different edges $e_1, e_2, e_3$ such that three different edges $e_1, e_2, e_3$ with $e_1 \cap e_2 \cap e_3 \neq \emptyset$. Then there exists $i$ such that $e_1 \cap e_2 \cap e_3 = [i]$. Say, $e_1 = [i, j]$, $e_2 = [i, k]$ and $e_3 = [i, l]$. Then $i \neq \min\{i, j, k, l\}$, otherwise $H$ is not closed. If $j < i$, then $k > i$ and $l > i$, since $H$ is closed. But then $[k, j]$ must be an edge of $H$, a contradiction. □

A graph with three different edges $e_1, e_2, e_3$ such that $e_1 \cap e_2 \cap e_3 \neq \emptyset$ is called a claw. Hence Proposition 1.2 says that a closed graph is a claw-free chordal graph.
Corollary 1.3. A bipartite graph is closed if and only if it is a line.

Proof. A bipartite graph has no odd cycles. Since a closed graph is chordal, and since a chordal graph has no odd cycle, if it is a tree, a closed bipartite graph must be a tree. If the tree is not a line, then there exists an induced subgraph which is a claw. Thus a closed bipartite graph must be a line.

Conversely, if $G$ is a line of length $l$, then $G$ is closed for the labeling of the vertices such that $\{1, 2\}, \{2, 3\}, \ldots , \{l, l+1\}$ are the edges of $G$. 

The conditions for being a closed graph formulated in Proposition 1.2 are only sufficient. For example the graph with edges $\{a, b\}, \{b, c\}, \{a, c\}, \{a, x\}, \{b, y\}$ and $\{c, z\}$ is chordal without a claw, but is not closed.

In the following we give a characterization of graphs which are closed with respect to a given labeling. Let $G$ be a graph, and let $v$ and $w$ be vertices of $G$. A path $\pi$ from $v$ to $w$ is a sequence of vertices $v = v_0, v_1, \ldots, v_l = w$ such that each $\{v_i, v_{i+1}\}$ is an edge of the underlying graph. If $G$ is directed, then the path $\pi$ is called directed, if either $(v_i, v_{i+1})$ is an arrow for all $i$, or $(v_{i+1}, v_i)$ is an arrow for all $i$.

Proposition 1.4. A graph $G$ on $[n]$ is closed with respect to the given labeling, if and only if for any two vertices $i \neq j$ of the associated directed graph $G^*$, all paths of shortest length from $i$ to $j$ are directed.

Proof. Suppose all shortest paths from $i$ to $j$ in $G^*$ are directed. Let $(i, j)$ and $(i, k)$ be two arrow with $j < k$. Then $\{j, i\}, \{i, k\}$ is a path from $j$ to $k$ which is not directed. So it cannot be the shortest path. Hence there exists the arrow $(j, k)$. Similarly it follows that if $(i, k)$ and $(j, k)$ are arrows of $G^*$ with $i < j$, then there must exist the arrow $(i, j)$ in $G^*$. This shows that $G^*$ is closed.

Conversely, assume that $G$ is closed. Then there exists a labeling such that $G^*$ is closed. Let $i$ and $j$ be two distinct vertices and let $P$ be a path of shortest length from $i$ to $j$. Suppose $P$ is not directed. Then there exists a subpath $r, s, t$ of $P$ such that either $(r, s)$ and $(t, s)$, or $(s, r)$ and $(s, t)$ are arrows in $G^*$. In both cases we may assume that $r < t$. Then, since $G^*$ is closed, it follows that $(r, t)$ is an arrow in $G^*$. Replacing the subpath $r, s, t$ by $r, t$, we obtain a shorter path from $i$ to $j$, a contradiction.

In Proposition 1.4 it is important to require that all paths of shortest length from $i$ to $j$ are directed in order to conclude that $G^*$ is closed. Indeed, consider the graph $G$ with edges $\{1, 2\}, \{2, 3\}, \{3, 4\}$ and $\{1, 4\}$. Then the path 2, 3, 4 is directed, while 2, 1, 4 is not directed. But both paths are shortest paths between 2 and 4.

Proposition 1.5. Let $G$ be a simple graph on $[n]$. Then there exists a unique minimal (with respect to inclusion of edges) graph $\hat{G}$ on $[n]$ whose associated acyclic graph is closed with respect to the given labeling and such that $G$ is a subgraph of $\hat{G}$.

Proof. Consider the set $C$ of graphs on $[n]$ containing $G$ and whose associated acyclic graph is closed. This set is not empty, because the complete graph on $[n]$ belongs to this set. Since the intersection of any two graphs in $C$ belongs again to $C$, the assertion follows, as desired. 

The unique minimal closed graph $\hat{G}$ containing $G$ is called the closure of $G$.

One basic question is which of the binomial edge ideals are Cohen–Macaulay. For a graph $G$, this is the case if and only the binomial edge of each component is Cohen–Macaulay. Thus it is enough to consider connected graphs. A partial answer on the Cohen–Macaulayness of binomial edge ideals is given in

Proposition 1.6. Let $G$ be a connected graph on $[n]$ which is closed with respect to the given labeling. Suppose further that $G$ satisfies the condition that whenever $\{i, j + 1\}$ with $i < j$ and $\{j, k + 1\}$ with $j < k$ are edges of $G$, then $\{i, k + 1\}$ is an edge of $G$. Then $S \setminus J_G$ is Cohen–Macaulay.
Proof. We will show that $S/\text{in}_{\prec}(J_G)$ is Cohen–Macaulay. This will then imply that $S/J_G$ is Cohen–Macaulay as well.

Since the associated acyclic directed graph is closed, it follows from Theorem 1.1 that $\text{in}_{\prec}(J_G)$ is generated by the monomials $x_i y_j$ with $\{i, j\} \in E(G)$ and $i < j$. Applying the automorphism $\varphi : S \to S$ which maps each $x_i$ to $x_i$, and $y_j$ to $y_{j-1}$ for $j > 1$ and $y_1$ to $y_n$, $\text{in}_{\prec}(J_G)$ is mapped to the ideal generated by all monomials $x_i y_j$ with $\{i, j + 1\} \in E(G)$. This ideal has all its generators in $S' = K[x_1, \ldots, x_{n-1}, y_1, \ldots, y_{n-1}]$. Let $I \subset S'$ be the ideal generated by these monomials. Then $S/\text{in}_{\prec}(J_G)$ is Cohen–Macaulay if and only if $S'/I$ is Cohen–Macaulay. Note that $I$ is the edge ideal of the bipartite graph $\Gamma'$ on the vertex set $\{x_1, \ldots, x_{n-1}, y_1, \ldots, y_{n-1}\}$, and with $\{x_i, y_j\} \in E(\Gamma')$ if and only if $\{i, j + 1\} \in E(G)$. In [8] the Cohen–Macaulay bipartite graphs are characterized as follows: Suppose the edges of the bipartite graph can be labeled such that

- (i) $\{x_i, y_i\}$ are edges for $i = 1, \ldots, n$;
- (ii) if $\{x_i, y_j\}$ is an edge, then $i \leq j$;
- (iii) if $\{x_i, y_j\}$ and $\{x_j, y_k\}$ are edges, then $\{x_i, y_k\}$ is an edge.

Then the corresponding edge ideal is Cohen–Macaulay.

We are going to verify these conditions for our edge ideal. Condition (ii) is trivially satisfied, and condition (iii) is a consequence of our assumption that whenever $\{i, j + 1\}$ with $i < j$ and $\{j, k + 1\}$ with $j < k$ are edges of $G$, then $\{i, k + 1\}$ is an edge of $G$.

For condition (i) we have to show that $\{i, i + 1\} \in E(G)$ for all $i$. But this follows from Proposition 1.4 which says that all shortest paths from $i$ to $i + 1$ are oriented paths. If $i, i + 1$ would not be a path, then a shortest path from $i$ to $i + 1$ could not be oriented. Thus $i, i + 1$ is a path in $G$, and hence $\{i, i + 1\} \in E(G)$. □

Examples 1.7. (a) Any complete graph satisfies the conditions of Proposition 1.6, so that $S/J_G$ is Cohen–Macaulay. But of course this is well known because in this case $J_G$ is the ideal of 2-minors of a generic $2 \times n$-matrix.

(b) Any path with the natural order of the vertices satisfies conditions of Proposition 1.6. Actually $J_G$ is a complete intersection in this case.

(c) There are many more graphs satisfying the conditions of Proposition 1.6. For example the graph with edges $\{1, 2\}, \{2, 3\}, \{1, 3\}$ and $\{3, 4\}$.

(d) Not all closed graphs satisfy the conditions of Proposition 1.6. Such an example is the graph with edges $\{1, 2\}, \{1, 3\}, \{2, 3\}, \{2, 4\}$ and $\{3, 4\}$. For this graph we have that $\text{in}_{\prec}(J_G)$ and $J_G$ are not Cohen–Macaulay.

(e) A graph $G$ need not be closed for $S/J_G$ being Cohen–Macaulay. The graph given after Corollary 1.3 is such an example.

2. The reduced Gröbner basis of a binomial edge ideal

We now come to the main result of this paper. For this we need to introduce the following concept: let $G$ be a simple graph on $[n]$, and let $i$ and $j$ be two vertices of $G$ with $i < j$. A path $i = i_0, i_1, \ldots, i_r = j$ from $i$ to $j$ is called admissible, if

- (i) $i_k \neq i_\ell$ for $k \neq \ell$;
- (ii) for each $k = 1, \ldots, r - 1$ one has either $i_k < i$ or $i_k > j$;
- (iii) for any proper subset $(j_1, \ldots, j_s)$ of $(i_1, \ldots, i_{r-1})$, the sequence $i, j_1, \ldots, j_s, j$ is not a path.

Given an admissible path

$$\pi : i = i_0, i_1, \ldots, i_r = j$$
from $i$ to $j$, where $i < j$, we associate the monomial
\[ u_\pi = \left( \prod_{i_k > j} x_{i_k} \right) \left( \prod_{i_l < i} y_{i_l} \right). \]

**Theorem 2.1.** Let $G$ be a simple graph on $[n]$. Let $<$ be the monomial order introduced in Theorem 1.1. Then the set of binomials
\[ G = \bigcup_{i < j} \{ u_\pi f_{ij} : \pi \text{ is an admissible path from } i \text{ to } j \} \]
is a reduced Gröbner basis of $J_G$.

**Proof.** We organize this proof as follows: In First Step, we prove that $G \subseteq J_G$. Then, since $G$ is a system of generators, in Second Step, we show that $G$ is a Gröbner basis of $J_G$ by using Buchberger’s criterion. Finally, in Third Step, it is proved that $G$ is reduced.

**First Step.** We show that, for each admissible path $\pi$ from $i$ to $j$, where $i < j$, the binomial $u_\pi f_{ij}$ belongs $J_G$. Let $\pi$: $i = i_0, i_1, \ldots, i_{r-1}, i_r = j$ be an admissible path in $G$. We proceed with induction on $r$. Clearly the assertion is true if $r = 1$. Let $r > 1$ and $A = \{ i_k : i_k < i \}$ and $B = \{ i_\ell : i_\ell > j \}$. One has either $A \neq \emptyset$ or $B \neq \emptyset$. If $A \neq \emptyset$, then we set $i_{k_0} = \max A$. If $B \neq \emptyset$, then we set $i_{\ell_0} = \min B$.

Suppose $A \neq \emptyset$. It then follows that each of the paths $\pi_1$: $i_{k_0}, i_{k_0} - 1, \ldots, i_1, i_0 = i$ and $\pi_2$: $i_{\ell_0}, i_{\ell_0} + 1, \ldots, i_{r-1}, i_r = j$ in $G$ is admissible. Now, the induction hypothesis guarantees that each of $u_{\pi_1} f_{i_{k_0} i}$ and $u_{\pi_2} f_{i_{\ell_0} j}$ belongs to $J_G$. A routine computation says that the $S$-polynomial $S(u_{\pi_1} f_{i_{k_0} i}, u_{\pi_2} f_{i_{\ell_0} j})$ is equal to $u_\pi f_{ij}$. Hence $u_\pi f_{ij} \in J_G$, as desired.

When $B \neq \emptyset$, the same argument as in the case $A \neq \emptyset$ is valid.

**Second Step.** It will be proven that the set of those binomials $u_\pi f_{ij}$, where $\pi$ is an admissible path from $i$ to $j$, forms a Gröbner basis of $J_G$. In order to show this we apply Buchberger’s criterion, that is, we show that all $S$-pairs $S(u_\pi f_{ij}, u_{\sigma} f_{k\ell})$, where $i < j$ and $k < \ell$, reduce to zero. For this we will consider different cases.

In the case that $i = k$ and $j = \ell$, one has $S(u_\pi f_{ij}, u_{\sigma} f_{k\ell}) = 0$.

In the case that $(i, j) \cap (k, \ell) = \emptyset$, or $i = \ell$, or $k = j$, the initial monomials in $\prec (f_{ij})$ and in $\prec (f_{k\ell})$ form a regular sequence. Hence the $S$-pair $S(u_\pi f_{ij}, u_{\sigma} f_{k\ell})$ reduce to zero, because of the following more general fact: let $f, g \in S$ such that in $\prec (f)$ and in $\prec (g)$ form a regular sequence and let $u$ and $v$ be any monomials. Then $S(u f, v g)$ reduces to zero.

It remains to consider the cases that either $i = k$ and $j \neq \ell$ or $i \neq k$ and $j = \ell$. Suppose we are in the first case. (The second case can be proved similarly.) We must show that $S(u_\pi f_{ij}, u_{\sigma} f_{k\ell})$ reduces to zero. We may assume that $j < \ell$, and must find a standard expression for $S(u_\pi f_{ij}, u_{\sigma} f_{k\ell})$ whose remainder is equal to zero.

Let $\pi$: $i = i_0, i_1, \ldots, i_r = j$ and $\sigma$: $i = i_0', i_1', \ldots, i_\ell' = \ell$. Then there exist indices $a$ and $b$ such that
\[ i_a = i_b' \quad \text{and} \quad \{ i_{a+1}, \ldots, i_r \} \cap \{ i_{b+1}', \ldots, i_\ell' \} = \emptyset. \]

Consider the path
\[ \tau: j = i_r, i_{r-1}, \ldots, i_{a+1}, i_a = i_b', i_{b+1}', \ldots, i_{\ell-1}', i_\ell' = \ell \]
from $j$ to $\ell$. To simplify the notation we write this path as
\[ \tau: j = j_0, j_1, \ldots, j_\ell = \ell. \]
Let
\[ j_t(1) = \min\{ j_c : j_c > j, \ c = 1, \ldots, t \}, \]
and
\[ j_t(2) = \min\{ j_c : j_c > j, \ c = t(1) + 1, \ldots, t \}. \]
Continuing these procedures yield the integers
\[ 0 = t(0) < t(1) < \cdots < t(q) = t. \]
It then follows that
\[ j = j_t(0) < j_t(1) < \cdots < j_t(q) = \ell \]
and, for each \( 1 \leq c \leq t \), the path
\[ \tau_c : j_t(c-1), j_t(c-1)+1, \ldots, j_t(c)-1, j_t(c) \]
is admissible.

The highlight of the proof is to show that
\[ S(u_\pi f_{ij}, u_\sigma f_{\ell}) = \sum_{c=1}^{q} v_{\tau_c} u_{\tau_c} f_{j_t(c-1), j_t(c)} \]
is a standard expression of \( S(u_\pi f_{ij}, u_\sigma f_{\ell}) \) whose remainder is equal to 0, where each \( v_{\tau_c} \) is the monomial defined as follows: Let \( w = y_1 \text{lcm}(u_\pi, u_\sigma) \). Thus \( S(u_\pi f_{ij}, u_\sigma f_{\ell}) = -w f_{j_t}. \) Then

(i) if \( c = 1 \), then
\[ v_{\tau_1} = \frac{x_\ell w}{u_{\tau_1} x_{j_t(1)}}; \]
(ii) if \( 1 < c < q \), then
\[ v_{\tau_c} = \frac{x_j x_\ell w}{u_{\tau_c} x_{j_t(c-1)} x_{j_t(c)}}; \]
(iii) if \( c = q \), then
\[ v_{\tau_q} = \frac{x_j w}{u_{\tau_q} x_{j_t(q-1)}}. \]

Our work is to show that
\[ w f_{j_\ell} = \frac{wx_\ell}{x_{j_t(1)}} f_{j_t(1)} + \sum_{c=2}^{q-1} \frac{wx_j x_\ell}{x_{j_t(c-1)} x_{j_t(c)}} f_{j_t(c-1), j_t(c)} + \frac{wx_j}{x_{j_t(q-1)}} f_{j_t(q-1) \ell}. \]
is a standard expression of \( w(f_{i\ell}) \) with remainder 0. In other words, we must prove that

\[
(\sharp) \quad w(x_j y_\ell - x_\ell y_j) = \frac{wx_j}{x_{j(q-1)}} (x_j y_{j(1)} - x_{j(1)} y_j) + \sum_{c=2}^{q-1} \frac{wx_j x_c}{x_{j(c-1)} x_{j(c)}} (x_j y_{j(c-1)} - x_{j(c)} y_{j(c-1)}) + \frac{wx_j}{x_{j(q-1)}} (x_j y_{j(q-1)} - x_\ell y_{j(q-1)})
\]

is a standard expression of \( w(x_j y_\ell - x_\ell y_j) \) with remainder 0. Since

\[
w x_j y_\ell = \frac{wx_j}{x_{j(q-1)}} x_{j(q-1)} y_\ell > \frac{wx_j x_{j(q-2)}}{x_{j(q-1)}} x_{j(q-1)} y_{j(1)} > \cdots
\]

it follows that, if the equality \((\sharp)\) holds, then \((\sharp)\) turns out to be a standard expression of \( w(x_j y_\ell - x_\ell y_j) \) with remainder 0. If we rewrite \((\sharp)\) as

\[
w(x_j y_\ell - x_\ell y_j) = w \left( x_j x_{j(1)} y_{j(1)} - x_\ell y_j \right) + wx_j x_c \left( \frac{y_{j(c)}}{x_{j(c)}} - \frac{y_{j(c-1)}}{x_{j(c-1)}} \right) + w \left( x_j y_\ell - x_j x_{j(q-1)} \right)
\]

then clearly the equality holds.

**Third Step.** Finally, we show that the Gröbner basis \( G \) is reduced. Let \( u_\pi f_{ij} \text{ and } u_\sigma f_{k\ell} \), where \( i < j \) and \( k < \ell \), belong to \( G \) with \( u_\pi f_{ij} \neq u_\sigma f_{k\ell} \). Let \( \pi: i = i_0, i_1, \ldots, i_q = j \) and \( \sigma: k = k_0, k_1, \ldots, k_s = \ell \). Suppose that \( u_\pi x_i y_j \) divides either \( u_\sigma x_k y_\ell \) or \( u_\sigma x_\ell y_k \). Then \( \{i_0, i_1, \ldots, i_q\} \) is a proper subset of \( \{k_0, k_1, \ldots, k_s\} \).

Let \( i = k \) and \( j = \ell \). Then \( \{i_1, \ldots, i_{q-1}\} \) is a proper subset of \( \{k_0, k_1, \ldots, k_s\} \) and \( k_0, k_1, \ldots, k_s-1, \ell \) is an admissible path. This contradicts the fact that \( \sigma \) is an admissible path.

Let \( i = k \) and \( j \neq \ell \). Then \( y_j \) divide \( u_\sigma \). Hence \( j < k \). This contradicts \( i < j \).

Let \( \{i, j\} \cap \{k, \ell\} = \emptyset \). Then \( x_i y_j \) divide \( u_\sigma \). Hence \( i > \ell \) and \( j < k \). This contradicts \( i < j \).

**Corollary 2.2.** \( J_G \) is a radical ideal.

**Proof.** The assertion follows from Theorem 2.1 and the following general fact: let \( I \subset S \) be a graded ideal with the property that \( in_{<}(I) \) is squarefree for some monomial order \( < \). Then \( I \) is a radical ideal. Indeed, there exists an ideal \( \bar{I} \subset S[t] \) in the polynomial ring \( S[t] \) such that \( t \) is a nonzerodivisor on \( S[t]/\bar{I} \) with \( (S[t]/\bar{I})/(tS[t]/\bar{I}) \cong S/in_{<}(I) \) and such that \( \bar{I} S[t, t^{-1}] = IS[t, t^{-1}] \), and there are positive degrees on the variables of \( K[x_1, \ldots, x_n, t] \) such that \( I \) is a graded ideal with respect to this grading. Thus we may apply the graded version of Lemma 4.4.9 in [3] in order to conclude that \( \bar{I} \) is a radical ideal. From the equality \( IS[t, t^{-1}] = IS[t, t^{-1}] \), it follows that \( I \) is a radical ideal as well.

As a consequence of Theorem 2.1 we see that all admissible paths of a graph \( G \) can be determined by computing the reduced Gröbner basis of \( J_G \).
On the other hand, it is not the case that for each edge \( i, j \) in the closure of \( G \) there exists an admissible path from \( i \) to \( j \). For example, for the graph \( G \) with edges \( \{2, 3\}, \{1, 3\} \) and \( \{1, 4\} \), the edge \( \{2, 4\} \) belongs to the closure of \( G \), but the only path \( 2, 3, 1, 4 \) from 2 to 4 is not admissible. Thus the reduced Gröbner basis of \( J_G \) does not give the closure of \( G \).

3. The minimal prime ideals of a binomial edge ideal

Let \( G \) be a simple graph on \([n]\). For each subset \( S \subset [n] \) we define a prime ideal \( P_S(G) \). Let \( T = [n] \setminus S \), and let \( G_1, \ldots, G_{c(S)} \) be the connected component of \( G_T \). Here \( G_T \) is the induced subgraph of \( G \) whose edges are exactly those edges \( \{i, j\} \) of \( G \) for which \( i, j \in T \). For each \( G_i \) we denote by \( \tilde{G}_i \) the complete graph on the vertex set \( V(G_i) \). We set

\[
P_S(G) = \left( \bigcup_{i \in S} \{x_i, y_i\}, J_{\tilde{G}_1}, \ldots, J_{\tilde{G}_{c(S)}} \right).
\]

Obviously, \( P_S(G) \) is a prime ideal. In fact, each \( J_{\tilde{G}_i} \) is the ideal of 2-minors of a generic \( 2 \times n_j \)-matrix with \( n_j = |V(G_j)| \). Since all the prime ideals \( J_{\tilde{G}_i} \) as well as the prime ideal \((\bigcup_{i \in S} \{x_i, y_i\})\) are prime ideals in pairwise different sets of variables, \( P_S(G) \) is a prime ideal, too.

**Lemma 3.1.** With the notation introduced we have \( \text{height} \ P_S(G) = |S| + (n - c(S)). \)

**Proof.** We have

\[
\text{height} \ P_S(G) = \text{height} \left( \bigcup_{i \in S} \{x_i, y_i\} \right) + \sum_{j=1}^{c(S)} \text{height} \ J_{\tilde{G}_j}
\]

\[
= 2|S| + \sum_{j=1}^{c(S)} (n_j - 1)
\]

\[
= |S| + \left( |S| + \sum_{j=1}^{c(S)} n_j \right) - c(S)
\]

\[
= |S| + (n - c(S)),
\]

as required. \( \square \)

In [6] Eisenbud and Sturmfels showed that all associated prime ideals of a binomial ideal are binomial ideals. In our particular case we have

**Theorem 3.2.** Let \( G \) be a simple graph on the vertex set \([n]\). Then \( J_G = \bigcap_{S \subset [n]} P_S(G) \).

**Proof.** It is obvious that each of the prime ideals \( P_S(G) \) contains \( J_G \). We will show by induction on \( n \) that each minimal prime ideal containing \( J_G \) is of the form \( P_S(G) \) for some \( S \subset [n] \). Since by Corollary 2.2, \( J_G \) is a radical ideal, and since a radical ideal is the intersection of its minimal prime ideals, the assertion of the theorem will follow.

Let \( P \) be a minimal prime ideal of \( J_G \). We first show that \( x_i \in P \) if and only if \( y_i \in P \). For this part of the proof we may assume that \( G \) is connected. Indeed, if \( G_1, \ldots, G_r \) are the connected components of \( G \), then each minimal prime ideal \( P \) of \( J_G \) is of the form \( P_1 + \cdots + P_r \) where each \( P_i \) is a minimal prime ideal of \( J_{G_i} \). Thus if each \( P_i \) has the expected form, then so does \( P \). Let \( T = \{x_i: i \in [n], \ x_i \in P, \ y_i \not\in P\} \). We will show that \( T = \emptyset \). This will then imply that if \( x_i \in P \), then
Claim 3.1, and Corollary 3.2 yield the following

**Corollary 3.3.** Let $G$ be a simple graph on $[n]$. Then

$$\dim S/J_G = \max\{\dim S/J_G > n + c, S \subset \{1, \ldots, n\}\}. $$

In particular, $\dim S/J_G \geq n + c$, where $c$ is the number of connected components of $G$.

In general, this inequality is strict. For example, for our claw $G$ with edges $\{1, 2\}$, $\{1, 3\}$ and $\{1, 4\}$ we have $\dim S/J_G = 6$.

**Corollary 3.4.** Let $G$ be a simple graph on $[n]$ with $c$ connected components. If $S/J_G$ is Cohen–Macaulay, then $\dim S/J_G = n + c$.

**Proof.** Since $P_{\emptyset}(G)$ does not contain any monomials, it follows that $P_S(G) \subseteq P_{\emptyset}(G)$ for any nonempty subset $S \subset [n]$. Thus Theorem 3.2 implies that $P_{\emptyset}(G)$ is a minimal prime ideal of $J_G$. Since $\dim S/J_G = n + c$ and since $S/J_G$ is equidimensional, the assertion follows.

With the results obtained so far, we are able to identify the minimal prime ideals of the edge ideal of a path, and thereby recover a result of Diaconis, Eisenbud and Sturmfels [4, Theorem 4.3]. The conclusion obtained is also a simple consequence of Corollary 3.9.
and these ideals are known to be Cohen–Macaulay. Since it follows from Corollary 3.4 that \( \dim S/P = n + 1 \) for all minimal prime ideals of \( J_G \). Let \( S \) be any subset of \([n]\). Then Theorem 3.2 and Corollary 3.3 imply that the minimal prime ideals of \( J_G \) are exactly those prime ideals \( P_S(G) \) for which \( c(S) = |S| + 1 \). Let \( S \subseteq [n] \). Then there exists integers \( 0 \leq a_1 - 1 < b_1 < a_2 - 1 < b_2 < a_3 - 1 < b_3 < \cdots < a_r - 1 < b_r < n \) such that

\[
S = \bigcup_{i=1}^{r} [a_i, b_i]
\]

where for each \( i \), \([a_i, b_i] = \{ j \in \mathbb{Z} : a_i \leq j \leq b_i \} \). We see that \( |S| = \sum_{i=1}^{r} (b_i - a_i + 1) = \sum_{i=1}^{r} (b_i - a_i) + r \), and that

\[
c(S) = \begin{cases} 
    r - 1, & \text{if } a_1 = 1 \text{ and } b_r = n, \\
    r, & \text{if } a_1 \neq 1 \text{ and } b_r = n, \text{ or } a_1 = 1 \text{ and } b_r \neq n, \\
    r + 1, & \text{if } a_1 \neq 1 \text{ and } b_r \neq n.
\end{cases}
\]

Thus \( c(S) = |S| + 1 \) if and only if \( a_1 \neq 1, b_r \neq n \) and \( a_i = b_i \) for all \( i \). In other words, the minimal prime ideals of \( G \) are those \( P_S(G) \) for which \( S \) is a subset of \([n]\) of the form \( \{a_1, a_2, \ldots, a_r\} \) with \( 1 < a_1, a_r < n \) and \( a_i < a_{i+1} - 1 \) for all \( i \).

The question of when \( J_G \) is a prime ideal is easy to answer.

**Proposition 3.6.** Let \( G \) be a simple graph on \([n]\). Then \( J_G \) is a prime ideal if and only if each connected component of \( G \) is a complete graph.

**Proof.** Let \( G_1, \ldots, G_r \) be the connected components of \( G \), and suppose that \( J_G \) is a prime ideal. Since \( P_{\emptyset}(G) = (J_{G_1}, \ldots, J_{G_r}) \) is a minimal prime ideal of \( J_G \) and \( J_G \) is a prime ideal, it follows that \( J_G = (J_{G_1}, \ldots, J_{G_r}) \). On the other hand, \( J_G = (J_{G_1}, \ldots, J_{G_r}) \). Thus the desired conclusion is a consequence of the following observation. Suppose that \( G \) and \( G' \) are graphs on \([n]\) with \( V(G) \subset V(G') \). Then \( E(G) = E(G') \), if and only \( J_G = J_{G'} \). \( \square \)

**Corollary 3.7.** Let \( G \) be a cycle of length \( n \). Then the following conditions are equivalent:

(a) \( n = 3 \).
(b) \( J_G \) is a prime ideal.
(c) \( J_G \) is unmixed.
(d) \( S/J_G \) is Cohen–Macaulay.

**Proof.** Due to Proposition 3.6 the equivalence of (a) and (b) is clear, since a cycle of length \( n \) is a complete graph if and only if \( n = 3 \). It also follows from Proposition 3.6 that whenever \( J_G \) is a prime ideal, then \( J_G \) is Cohen–Macaulay, because if each of the components of \( G \) is a complete graph, then the binomial edge ideal of each component is the ideal of 2-minors of a \( 2 \times k \)-matrix for some \( k \), and these ideals are known to be Cohen–Macaulay. Since \( J_G \) is unmixed if \( S/J_G \) is Cohen–Macaulay, all implications follow once it is shown that (c) implies (b). One of the minimal prime ideals of \( G \) is \( P_{\emptyset}(G) \) and \( \dim S/P_{\emptyset}(G) = n + 1 \). Now let \( S \subset [n] \) with \( S \neq \emptyset \). We may assume that we have labeled the edges of the cycle counterclockwise, and that

\[
S = \bigcup_{i=1}^{r} [a_i, b_i] \quad \text{with } 0 = a_1 - 1 < b_1 < a_2 - 1 < b_2 < a_3 - 1 < b_3 < \cdots < a_r - 1 < b_r < n.
\]
Then \( c(S) = r \), and \( \dim S/P_S(G) = n - |S| + c(S) = n - \sum_{i=1}^{r} (b_i - a_i) - r + r \leq n \). Thus if \( J_G \) is unmixed, then \( P_S(G) \) is the only minimal prime ideal of \( J_G \), and hence since \( J_G \) is reduced it follows that \( J_G \) is a prime ideal, as required. \( \square \)

Now let \( G \) be an arbitrary simple graph. Which of the ideals \( P_S(G) \) are minimal prime ideals of \( J_G \)? The following result helps to find them.

**Proposition 3.8.** Let \( G \) be a simple graph on \([n]\), and let \( S \) and \( T \) be subsets of \([n]\). Let \( G_1, \ldots, G_s \) be the connected components of \( G_{[n] \setminus S} \), and \( H_1, \ldots, H_t \) the connected components of \( G_{[n] \setminus T} \). Then \( P_T(G) \subseteq P_S(G) \), if and only if \( T \subseteq S \) and for all \( i = 1, \ldots, t \) one has \( V(H_i) \setminus S \subseteq V(G_j) \) for some \( j \).

**Proof.** For a subset \( U \subseteq [n] \) we let \( L_U \) be the ideal generated by the variables \( \{x_i, y_i : i \in U\} \). With this notation introduced we have \( P_S(G) = (L_S, J_{G_1}, \ldots, J_{G_s}) \) and \( P_T(G) = (L_T, J_{H_1}, \ldots, J_{H_t}) \). Hence it follows that \( P_T(G) \subseteq P_S(G) \), if and only if \( T \subseteq S \) and \( (L_S, J_{G_1}, \ldots, J_{G_s}) \subseteq (L_S, J_{H_1}, \ldots, J_{H_t}) \).

Observe that \( (L_S, J_{H_1}, \ldots, J_{H_t}) = (L_S, J_{H_1}^t, \ldots, J_{H_t}^t) \) where \( H_i^t = (H_i)_{[n] \setminus S} \). It follows that \( P_T(G) \subseteq P_S(G) \) if and only if \( (L_S, J_{H_1}^t, \ldots, J_{H_t}^t) \subseteq (L_S, J_{G_1}, \ldots, J_{G_s}) \) which is the case if and only if \( (J_{H_1}^t, \ldots, J_{H_t}^t) \subseteq (J_{G_1}, \ldots, J_{G_s}) \), because the generators of the ideals \( (J_{H_1}^t, \ldots, J_{H_t}^t) \) and \( (J_{G_1}, \ldots, J_{G_s}) \) have no variables in common with the \( x_i \) and \( y_i \) for \( i \in S \).

Since \( V(H_i^t) = V(H_i) \setminus S \), the assertion will follow once we have shown the following claim: let \( A_1, \ldots, A_s \) and \( B_1, \ldots, B_t \) be pairwise disjoint subsets of \([n] \). Then

\[
(J_{A_1}, \ldots, J_{A_s}) \subseteq (J_{B_1}, \ldots, J_{B_t}).
\]

if and only if for each \( i = 1, \ldots, s \) there exists \( j \) such that \( A_i \subseteq B_j \).

It is obvious that if the conditions on the \( A_i \) and \( B_j \) are satisfied, then we have the desired inclusion of the corresponding ideals.

Conversely, suppose that \( (J_{A_1}, \ldots, J_{A_s}) \subseteq (J_{B_1}, \ldots, J_{B_t}) \). Without loss of generality we may assume that \( \bigcup_{j=1}^{t} B_j = [n] \). Consider the surjective \( K \)-algebra homomorphism

\[
\epsilon : S \rightarrow K[\{x_i, z_1\}_{i \in B_1}, \ldots, \{x_i, x_i z_t\}_{i \in B_t}] \subset K[x_1, \ldots, x_n, z_1, \ldots, z_t]
\]

with \( \epsilon(x_i) = x_i \) for all \( i \) and \( \epsilon(y_i) = x_i z_j \) for \( i \in B_j \) and \( j = 1, \ldots, t \). Then

\[
\operatorname{Ker}(\epsilon) = (J_{B_1}, \ldots, J_{B_t}).
\]

Now fix one of the sets \( A_i \) and let \( k \in A_i \). Then \( k \in B_j \) for some \( k \). We claim that \( A_i \subseteq B_j \). Indeed, let \( \ell \in A_i \) with \( \ell \neq k \) and suppose that \( \ell \in B_r \) with \( r \neq j \). Since \( x_k y_\ell - x_\ell y_k \in J_{A_i} \subseteq (J_{B_1}, \ldots, J_{B_t}) \), it follows that \( x_k y_\ell - x_\ell y_k \in \operatorname{Ker}(\epsilon) \), so that \( 0 = \epsilon(x_k y_\ell - x_\ell y_k) = x_k x_\ell z_j - x_k x_\ell z_r \), a contradiction. \( \square \)

Let \( G_1, \ldots, G_r \) be the connected components of \( G \). Once we know the minimal prime ideals of \( J_{G_i} \) for each \( i \) the minimal prime ideals of \( J_G \) are known. Indeed, since the ideals \( J_{G_i} \) are ideals in different sets of variables, it follows that the minimal prime ideals of \( J_G \) are exactly the ideals \( \sum_{i=1}^{r} P_i \) where each \( P_i \) is a minimal prime ideal of \( J_{G_i} \).

The next result detects the minimal prime ideals of \( J_G \) when \( G \) is connected.

**Corollary 3.9.** Let \( G \) be a connected simple graph on the vertex set \([n]\), and \( S \subseteq [n] \). Then \( P_S(G) \) is a minimal prime ideal of \( J_G \) if and only if \( S = \emptyset \) or \( S \neq \emptyset \) and for each \( i \in S \) one has \( c(S \setminus \{i\}) < c(S) \).

In the terminology of graph theory, the corollary says that if \( G \) is a connected graph, then \( P_S(G) \) is a minimal prime ideal of \( J_G \), if and only if each \( i \in S \) is a cut-point of the graph \( G_{([n] \setminus S) \cup \{i\}} \).
Proof of Corollary 3.9. Assume that $P_S(G)$ is a minimal prime ideal of $J_G$ and fix $i \in S$. Let $G_1, \ldots, G_r$ be the connected components of $G([n])_S$. We distinguish several cases.

Suppose that there is no edge $[i, j]$ of $G$ such that $j \in G_k$ for some $k$. Set $T = S \setminus \{i\}$. Then the connected components of $G([n])_T$ are $G_1, \ldots, G_r, \{i\}$. Thus $c(T) = c(S) + 1$. However this case cannot happen, since Proposition 3.8 would imply that $P_T(G) \subset P_S(G)$.

Next suppose that there exists exactly one $G_k$, say $G_1$, for which there exists $j \in G_1$ such that $[i, j]$ is an edge of $G$. Then the connected components of $G([n])_T$ are $G_1', G_2, \ldots, G_r$ where $V(G_1') = V(G_1) \cup \{i\}$. Thus $c(T) = c(S)$. Again, this case cannot happen since Proposition 3.8 would imply that $P_T(G) \subset P_S(G)$.

It remains the case that there are at least two components, say $G_1, \ldots, G_k$, $k \geq 2$, and $j_\ell \in G_\ell$ for $\ell = 1, \ldots, k$ such that $[i, j_\ell]$ is an edge of $G$. Then the connected components of $G([n])_T$ are $G_1', G_{k+1}, \ldots, G_r$, where $V(G_1') = \bigcup_{\ell=1}^k V(G_\ell) \cup \{i\}$. Hence in this case $c(T) < c(S)$.

Conversely, suppose that $c(S \setminus \{i\}) < c(S)$ for all $i \in S$. We want to show that $P_S(G)$ is a minimal prime ideal. Suppose this is not the case. Then there exists a proper subset $T \subset S$ with $P_T(G) \subset P_S(G)$. We choose $i \in S \setminus T$. By assumption, we have $c(S \setminus \{i\}) < c(S)$. The discussion of the three cases above show that we may assume that $G_1', G_{k+1}, \ldots, G_r$ are the components of $G([n]) \setminus \{i\}$ where $V(G_1') = \bigcup_{\ell=1}^k V(G_\ell) \cup \{i\}$ and where $k \geq 2$. It follows that $G([n])_T$ has one connected component $H$ which contains $G_1'$. Then $V(H) \setminus S$ contains the subsets $V(G_1')$ and $V(G_2)$. Hence $V(H) \setminus S$ is not contained in any $V(G_i)$. According to Proposition 3.8, this contradicts the assumption that $P_T(G) \subset P_S(G)$. □

As an example of Corollary 3.9 consider again the cycle $G$ of length $n$. Then, besides of the prime ideal $P_0(G)$ which is of height $n - 1$, the only other minimal prime ideals are the ideals $P_S(G)$ where $|S| > 1$ and no two elements $i, j \in S$ belong to the same edge of $G$. Each of these prime ideals has height $n$.

4. CI-ideals

Binomial equations and determinantal ideals are of fundamental importance in the theory of conditional independence. In this final section we will demonstrate the connection between binomial edge ideals and conditional independence (CI) statements.

We consider a random vector $X = (X_0, \ldots, X_N)$ of $N + 1$ discrete random variables, where the random variable $X_i$ takes values in the sets $[d_i]$ for some positive integers $d_i \in \mathbb{N}$. Then $X$ takes values in $\mathcal{X} := [d_0] \times \cdots \times [d_N]$. A joint probability distribution of $X$ is a nonnegative real valued function $p : \mathcal{X} \to \mathbb{R}_{\geq 0}$, such that $\sum_{x \in \mathcal{X}} p(x) = 1$. It can be represented by a real vector $p = (p_{x_0, \ldots, x_N})_{x_0, \ldots, x_N} \in \mathbb{R}^{\mathcal{X}}$, where $p_{x_0, \ldots, x_N}$ stands for the probability of the event $X_0 = x_0, X_1 = x_1, \ldots, X_N = x_N$. In the following we will consider polynomial equations in these $\prod_{i=0}^N d_i$ indeterminates, denoting $\mathbb{C}[p_x; x \in \mathcal{X}]$ the ambient polynomial ring.

For any subset $S \subseteq \{0, \ldots, N\}$ we write $X_S$ for the collection of random variables $X_i : i \in S$. Then $X_S$ is a random variable on the smaller state space $\mathcal{X}_S = \prod_{i \in S} [d_i]$. Given $x_T \in \mathcal{X}_T$, we denote $\{X_T = x_T\} := \{y \in \mathcal{X} : y_i = x_i, \forall i \in T\}$. The notation $p(X_T = x_T) := \sum_{x \in \{X_T = x_T\}} p_x$ is common and convenient and may be abbreviated by $p(x_T)$, if no confusion can arise.

Let $S$ and $S'$ be two disjoint subsets of $\{0, \ldots, N\}$, let $C \subseteq \mathcal{X}$, and fix a joint probability distribution $p$. We say that $X_S$ is conditionally independent of $X_{S'}$ given $C$ (under $p$) iff $p$ satisfies all equations of the form

$$p(x_S, x_{S'}; C) p(x'_S, x'_{S'}; C) = p(x_S, x'_{S'}; C) p(x'_S, x_{S'}; C) = 0.$$ 

(1)

where $x_S, x'_S \in \mathcal{X}_S$, $x_{S'}, x'_{S'} \in \mathcal{X}_{S'}$, and
Eqs. (1) are seen as equations among the elementary probabilities

Example 4.1. Consider for a simple example \( N = 2 \) and binary variables \( d_0 = d_1 = d_2 = 2 \). The polynomial ring is given as \( \mathbb{C}[p_{111}, p_{112}, p_{121}, p_{122}, p_{211}, p_{212}, p_{221}, p_{222}] \). The conditional independence \( X_0 \perp X_1 | X_2 \) describes the binomial ideal

\[
I_{X_0 \perp X_1 | X_2} = (p_{111}p_{221} - p_{121}p_{211}, p_{112}p_{222} - p_{122}p_{212}).
\]

In contrast to that, the independence \( X_0 \perp X_1 \) is given by the principal ideal

\[
I_{X_0 \perp X_1} = (p_{111} + p_{112})(p_{221} + p_{222}) - (p_{211} + p_{212})(p_{121} + p_{122})).
\]

Remark 4.2. A conditional independence \( X_5 \perp X_5' | C \) is usually defined differently: One requires

\[
p(X_5 = x_5, X_5' = x_5' \mid X \in C) = p(X_5 = x_5 \mid X \in C) p(X_5' = x_5' \mid X \in C)
\]

for all \( x_5 \in X_5 \) and \( x_5' \in X_5' \). Here,

\[
p(X_5 = x_5, X_5' = y_5' \mid X \in C) = \frac{p(X_5 = x_5, X_5 = y_5, X \in C)}{p(X \in C)},
\]

and so on. However, Eq. (3) is not well defined if \( p(X \in C) \) is zero, while Eq. (1) is defined for all joint distributions \( p \). It is an easy exercise to prove that Eqs. (1) and (3) are equivalent if \( p(X \in C) \) is nonzero.

We will now discuss a special case which makes it possible to apply the results of the first three sections. Namely, we assume \( d_0 = 2 \), i.e., \( X_0 \) is considered to be binary. In this case we can arrange the elementary probabilities \( p_x \) in a \( 2 \times d_1 \ldots d_N \)-matrix, where the columns are indexed by the state space \( X_{[N]} \) of \( X_{[N]} = (X_1, \ldots, X_N) \). The basic observation is that every 2-minor corresponds to one CI-statement; namely, the minor

\[
p_{1x}p_{2x'} - p_{2x}p_{1x'}
\]

of the two columns corresponding to \( x, x' \in X_{[N]} \) expresses exactly the CI-statement

\[
X_0 \perp X_{[N]} \mid \{X_{[N]} \in \{x, x'\}\}.
\]

In this way we can associate a collection of CI-statements to every graph on the vertex set \( X_{[N]} \).
Until now we did not use the fact that $X_{\left[N\right]}$ is a product of several random variables. Now let $S \cup T$ be a (disjoint) partition of $\left[N\right]$ and consider the CI-statement

$$X_0 \perp \perp X_S|X_T.$$  \hspace{1cm} (4)

For simplicity we assume that $S = \{1, \ldots, s\}$ for a moment. Then (4) is equivalent to the equations

$$p_{1xSxT}p_{2x'SxT} - p_{1x'SxT}p_{2xSxT} = 0$$

for all $x_S, x'_S \in X_S$ and $x_T \in X_T$. These equations come from all 2-minors with columns $x, x' \in X_{\left[N\right]}$ such that $x$ and $x'$ agree on their $T$-components. This means that we can associate with (4) the graph on $X_{\left[N\right]}$ with edges

$$E(G) = \{(x, x') : x, x' \in X_{\left[N\right]} \text{ agree on } T\}.$$  

More generally, when we have a collection $C = \{X_0 \perp \perp X_S|X_T\}$ of CI-statements corresponding to disjoint partitions $S_i \cup T_i$ of $\left[N\right]$, we can associate a graph $G_i$ with every single statement. If we define a graph $G$ on $X_{\left[N\right]}$ by $E(G) = \bigcup E(G_i)$, then the binomial edge ideal of $G$ equals the CI-ideal of $C$.

CI-statements of the form under consideration have the following natural interpretation in probabilistic modeling: We consider $X_0$ as the output node of a system which receives input from $X_1, \ldots, X_N$. Then we can ask how much information is lost when certain input nodes are not available. If $X_0 \perp \perp X_S'|X_T$, then all the relevant information can be reconstructed from $X_T$ alone: The system can dispense with the information from $X_S'$. In this way, a collection of CI-statements can be used to model a notion of robustness of probabilistic computation. In the language of [1] we study the probability distributions with vanishing exclusion dependence. Because of this interpretation we introduce the following notation:

**Definition 4.3.** A collection of CI-statements induced as above by a set of disjoint partitions $S_i \cup T_i = [N]$ will be called a robustness specification.

Theorems 2.2 and 3.2 imply two corollaries:

**Corollary 4.4.** The CI-ideal of a robustness specification with binary output is a radical ideal.

Now fix a robustness specification $C$. Owing to Theorem 3.2, each minimal prime is given by a subset $S \subseteq X_{\left[N\right]}$ which satisfies the conditions of Corollary 3.9. Such a subset $S$ defines events with zero probability: $p(X_{\left[N\right]} \in S) = 0$ if $p \in V(P_S(G))$, where $G = G_C$. In the language of statistical modeling, $S$ is a set of structural zeros.

**Corollary 4.5.** Let $I$ be the CI-ideal of a robustness specification. Each minimal prime $P$ of $I$ is characterized by a set $S$ of structural zeros in the distribution of $X_{\left[N\right]}$ which is common to all probability distributions lying in the component corresponding to $P$. The possible sets $S$ are characterized by Corollary 3.9.

The binomial generators $J_{G_1}, \ldots, J_{G_{c(S)}}$ in $P_S(G)$ also have a nice statistical interpretation: Namely $J_{G_i}$ expresses the CI-statement

$$X_0 \perp \perp X_{\left[N\right]}|X_{\left[N\right]} \in G_i).$$

This means: If we know $S$, then the knowledge in which component of $G_{X_{\left[N\right]}|S}$ the random vector $X_{\left[N\right]}$ lies contains all the relevant information about $X_0$. Once we know this component, the conditional probability distribution of $X_0$ is independent of any further information we may obtain. In other
words, if we know $G$ and $S$, then we can define a random variable $C$ which maps every outcome of $X$ with nonzero probability to the corresponding component in $[c(S)]$. We then have $X_0 \perp X_{[N]|C}$, a fact which can be depicted by the following Markov chain

$$X_{[N]} \rightarrow C \rightarrow X_0.$$ 

This corresponds to the classical result that each irreducible component of a binomial ideal is essentially a toric variety [6], and in particular each irreducible component has a rational parametrization. The most natural such parametrization in the statistical setting is the following: $p$ factors as a product of a distribution on the connected components $G_1, \ldots, G_{c(S)}$ and a distribution of $X_0$ for each of the connected components. This should be compared to the dimension $n - |S| + c(S)$ in Lemma 3.3.

Each binomial ideal $I \subset \mathbb{C}[p_x: x \in X]$ has the toric ideal $I$: $(\prod_{x \in X} p_x)^\infty$ as a minimal prime. It corresponds to $S = \emptyset$, and all distributions with full support $(p(x) > 0$ for all $x \in X$) satisfying the robustness specification are contained in the toric variety. We obtain the following

**Corollary 4.6.** Let $p$ be a probability distribution satisfying the robustness specification $C = \{X_0 \perp X_{S_i} | X_{T_i}: i = 1, \ldots, r\}$. If $p$ has full support (i.e., $p_x > 0$ for all $x \in X$), then

$$X_0 \perp X_{\bigcup_i S_i} \mid X_{\bigcap_i T_i}.$$ 

In particular, if $\bigcup_i S_i = [N]$ then $X_0 \perp X_{[N]}$ and $X_0$ is unconditionally independent of the input.

**Remark 4.7.** It is easy to prove this corollary directly using the intersection axiom [5].

This result is not surprising: If any combination of inputs in $X_{[N]}$ is possible, then we can’t deduce any missing information. Any distribution where $X_0$ is robust against perturbation of the inputs must make use of features of the input statistics.

**Examples 4.8.** Fix $k \in [N]$ and consider the collection of CI-statements

$$\left\{X_0 \perp X_S \mid [X_T]: S \in \binom{[N]}{k}\right\}$$

induced by all $k$-element subsets of $[N]$. Consider the graph $G_k$ with vertices $X_{[N]}$ and edges between any $x$ and $y$ which differ in at most $k$ components. In other words, $(x, y) \in E(G_k)$ if and only if the Hamming distance between $x$ and $y$ is at most $k$. The CI-ideal for the statements (5) is the binomial edge ideal of $G_k$.

(a) If $k = 1$ and $d_i = 2$, for all $i \in [N]$ we find the graph of the $N$-cube.

(b) If $k = 1$ and $N = 2$ we have just two CI-statements:

$$X_0 \perp X_1 \mid X_2 \quad \text{and} \quad X_0 \perp X_2 \mid X_1.$$ 

These statements have been studied by A. Fink [7]. In this case the minimal primes can be seen to correspond to bipartite graphs $\Gamma$ such that every connected component is a complete bipartite graph. The two groups of vertices in these graphs are $[d_1]$ and $[d_2]$. The corresponding prime is minimal if each vertex belongs to at least one edge. Such bipartite graphs are in bijection with pairs of partitions $[d_1] = I_1 \cup \cdots \cup I_c$ and $[d_2] = J_1 \cup \cdots \cup J_c$, where $c$ is the number of connected components of $\Gamma$, and $I_i$ resp. $J_i$ are the vertices in the $i$th component of $\Gamma$. Then $S = X_{[N]} \setminus \bigcup_{i=1}^c (I_i \times J_i)$ gives the link with our notation. In other words, the vertices of the connected components $G_1, \ldots, G_{c(S)}$ are given by $V(G_i) = I_i \times J_i$.

(c) The considerations of (b) generalize to the case $k = N - 1$: As above, the minimal primes correspond to partitions $[d_1] = I_{1,1} \cup \cdots \cup I_{1,c}$, where $S = X_{[N]} \setminus \bigcup_{j=1}^{k} (I_{1,j} \times \cdots \times I_{N,j})$, and the components
of $G_T$ satisfy $V(G_i) = I_{1,j} \times \cdots \times I_{N,j}$. We leave the verification of these results as an exercise to the reader. Unfortunately, the nice form of the connected components of $G_T$ does not generalize for $k < N - 1$.

Acknowledgments

We wish to thank Seth Sullivant, who established the contact between the two groups of authors, and Alex Fink, who carefully read the manuscript. The last two authors thank Nihat Ay for pointing them at conditional independence problems in robustness theory. Thomas Kahle had support by the Volkswagen Foundation.

References