EVEN KERNELS

Aviezri Fraenkel[1]
Curtin University
School of Mathematics and Statistics
GPO Box U 1987, Perth, WA 6001, Australia

ABSTRACT. Given a graph $G = (V, E)$, an *even kernel* is a nonempty independent subset $V' \subseteq V$, such that every vertex of G is adjacent to an even number (possibly 0) of vertices in V'. It is proved that the question of whether a graph has an even kernel is NP-complete. The motivation stems from combinatorial game theory. It is known that this question is polynomial if G is bipartite. We also prove that the question of whether there is an even kernel whose size is between two given bounds, in a given bipartite graph, is NP-complete. This result has applications in coding and set theory.

1 Introduction

EVEN KERNEL (EVEK). Given an undirected graph $G = (V, E)$. Is there a nonempty independent subset $V' \subseteq V$ such that every $u \in V$ has even degree with respect to V', i.e., u has an even number (possibly 0) of neighbors in V'?

EXAMPLE. In the graph depicted in Fig. 1, the subset $\{u_1, u_3, u_7, u_9\}$ is an even kernel. So is its subset $\{u_1, u_3\}$; and also $\{u_2, u_4, u_5, u_6, u_8\}$ is an even kernel. Thus an even kernel may exist nonuniquely. A triangle has no even kernel.

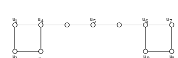


FIGURE 1. Even kernels in a graph $G = (V, E)$.

The notion of an even kernel was defined in Fraenkel, Scheinerman and Ullman [1993],

[1]Permanent address: Dept. of Applied Mathematics and Computer Science, The Weizmann Institute of Science, Rehovot 76100, Israel. The main part of this work was done at the University of Pennsylvania and later at the University of Calgary, during 1993.

with the motivation that the vertices of an even kernel are P-positions (second player win positions) in a game called "Edge Geography". It was shown there that EK is polynomially decidable if G is bipartite. In §2 we prove

Theorem 1. EVEK *is* NP*-complete even for graphs with maximum degree 3.*

This result is best possible, in the sense that for a graph with maximum degree ≤ 2, the question can be decided in linear time, since a simple path or a simple circuit each have an even kernel if and only if the path or circuit has even length (an even number of edges).

The notion of an even kernel is not all that new, though our terminology for it might be. Berlekamp, McEliece and van Tilborg [1978] showed that the problem of whether a binary matrix A contains exactly L rows such that each column of these L rows has an even number of 1-bits (i.e., whether there is a binary vector X such that $XA \equiv 0 \pmod 2$) is NP-complete, and asked about the status of the problem when "exactly L" is replaced by "$\leq L$" ($\leq$ DECOD). See also Garey and Johnson [1979, DECODING OF LINEAR CODES]. They also asked in "OPEN5" about the following EVEN SET (EVES) problem: "Given a collection C of subsets of a finite set X and $L \in \mathcal{Z}^+$, is there a nonempty subcollection $C' \subseteq C$ with $|C'| \leq L$, such that each $x \in X$ belongs to an even number (possibly 0) of sets in C'?". They stated that EVES is equivalent to $\leq$ DECOD. It is easy to see that both EVES and $\leq$ DECOD are equivalent to asking whether a given bipartite graph $G = (V_1, V_2; E)$ with disjoint and independent parts V_1 and V_2 has an even kernel $K \subseteq V_1$ with $|K| \leq L$.

Define the problem

Even Single Bipartite Kernel (ESBIK). Given $A, C \in \mathcal{Z}^+$ with $A \leq C$ and a bipartite graph $G = (V_1, V_2; E)$, where V_1, V_2 are disjoint independent subsets of vertices. Is there a subset $K \subseteq V_1$, with $A \leq |K| \leq C$ such that every vertex has an even number of neighbors (possibly 0) in K?

In §3 we prove

Theorem 2. ESBIK *is* NP*-complete even for graphs with maximum degree 3.*

A related problem is

Even Double Bipartite Kernel (EDBIK). Given $A, C \in \mathcal{Z}^+$ with $A \leq B$ and a bipartite graph $G = (V_1, V_2; E)$, where V_1, V_2 are disjoint indpendent subsets of vertices. Is there a subset $K \subseteq V_1 \cup V_2$ with $A \leq |K| \leq C$, such that every vertex has an even number of neighbors (possibly 0) in K?

In §4 we prove

Theorem 3. EDBIK *is* NP*-complete even for graphs with maximum degree 3.*

All our reductions are made from 1-3SAT, defined below. A Boolean formula is *positive* if it contains no negated variables. A Boolean formula is in 3CNF if it is a conjunction of clauses, where each clause is a disjunction of three literals.

ONE-IN-THREE 3SAT (1-3SAT). Given a positive Boolean 3CNF formula B. Is B *1-satisfiable*, i.e., is there a truth assignment for B such that each clause in B has precisely one true variable?

Schaefer [1978] proved that 1-3SAT is NP-complete. See also Garey and Johnson [1979].

For all the three proofs, we associate with any positive 3CNF-formula $B = c_1 \wedge \cdots \wedge\ c_m$ with clauses $c_1, \ldots, c_m$ and variables $x_1, \ldots, x_n$, a graph $G(B)$ whose vertex set is $\{c_1 \ldots c_m, x_1, \ldots, x_n\}$ and there is an edge (x_j, c_i) if and only if $x_j \in c_i$, i.e., x_j is in c_i (Fig. 2). We shall make a standard modification on $G(B)$, so as to preserve the degree-at-most-3 requirement throughout the construction.

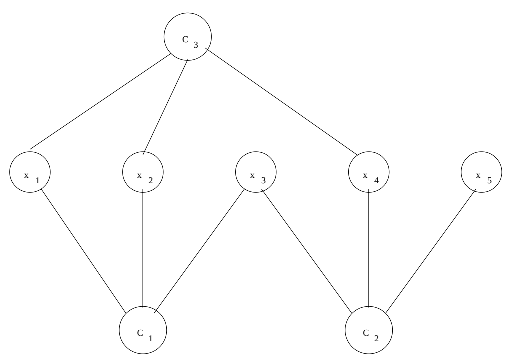


FIGURE 2. The graph $G(B)$ for $B = (x_1 \vee x_2 \vee x_3) \wedge (x_3 \vee x_4 \vee x_5) \wedge (x_1 \vee x_2 \vee x_4)$.

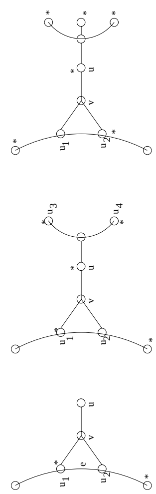
It is clear that each of the problems EVEK, ESBIK and EDBIK is in *NP*.

NOTATION. A vertex belonging to a given fixed even kernel will be termed *marked*. Otherwise it is *unmarked*. In the figures below, we use asterisks to indicate marked vertices.

The *length* of a simple path in a graph is the number of its edges, so it is 1 less than the number of its vertices. An *n-path* is a path of length n.

2 Proof of Theorem 1

1. Injectors. An *injector* injects a mark onto a vertex. Typically, one end of an injector is a two-pronged "or"-gate consisting of two adjacent vertices u_1, u_2, constituting one edge $e = (u_1, u_2)$ of a circuit in which alternate vertices are marked. Thus precisely one of u_1, u_2 is marked. The two vertices u_1, u_2 are both adjacent to the *focus* v of the injector. The focus is adjacent to the other end of the injector, which is a single vertex u (Fig. 3(i)), possibly joined to an or-gate of several vertices, an odd number of them being marked (Fig. 3 (ii),(iii)). The latter vertices may be adjacent to each other (implying further mark and adjacency restrictions) or not. Note that an injector injects a mark in either direction, and in fact may be completely symmetric relative to its mid-vertex (Fig. 3(ii), if u_3 and u_4 are adjacent).



(i) Simple injector | (ii) Injector with 2-pronged or-gate | (iii) Injector with 3-pronged or-gate

Figure 3. Various manifestations of injectors.

We wish to make sure that the even kernel induced by our construction is a "full" even kernel, rather than only some subset of an even kernel, such as pointed out in Fig. 1. The injectors see to this.

2. Variable Circuits. Let $m(j)$ be the total number of occurrences of the variable x_j in B. Construct a simple circuit of $2(2m(j)+2)$ vertices for x_j $(1 \le j \le \ n)$, where

alternate vertices are *labeled* $x_{1j}, x_{2j}, \ldots, x_{2m(j)+2,j}$; the vertices in-between are *unlabeled.* The circuits for x_{j-1} and x_j are joined by an injector with a 2-pronged or-gate on both of its sides (Fig. 4), where, here and below, x_{ij} is indicated by ij. The locations of the ends of the injector on any variable circuit are such that if the vertices of a variable circuit are traversed in clockwise direction, then the first vertex of the injector which is encountered in this traversal is labeled. (Note the distinction between marked and labeled vertices.)

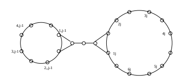


FIGURE 4. Two adjacent variable circuits joined via an injector.

3. CLAUSE CIRCUITS. A clause circuit c_i $(1 \leq i \leq m)$ is a network consisting of 12 vertices interconnected as shown in Fig. 5. (It would be more compact, if the degree constraint would be relaxed to $d \leq 4$.) It has four *terminals*, a, b, d and g.

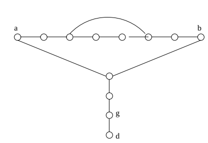


FIGURE 5. A clause circuit c_i.

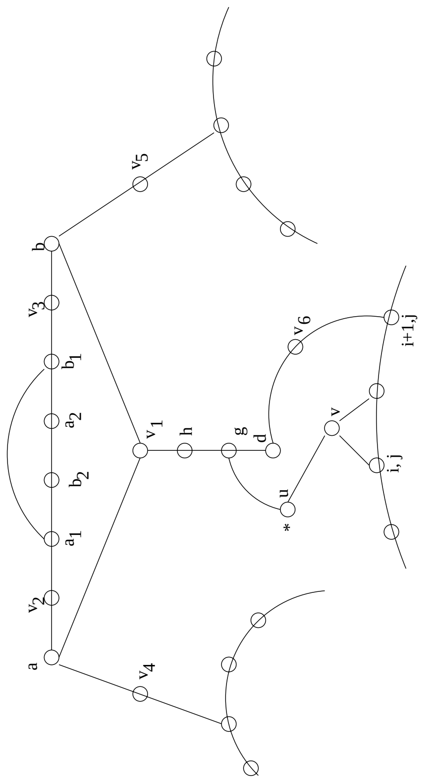


FIGURE 6. Joining variable circuits with a clause circuit.

In the global construction, x_{ij} is joined via a 2-path to precisely one of the terminals a, b or d of c_i, if and only if x_j is in c_i. In addition, one injector, based on an x_j-variable circuit with x_j in c_i is joined, via a 2-path, to terminal g of c_i (Fig. 6). Also here the location of the ends of the injector is such that if the vertices of a variable circuit are traversed in clockwise direction, then the first vertex of the injector encountered in this traversal is labeled. Since there are $m(j)$ 2-paths between the variable circuit of x_j and the c_i, and at most $m(j) + 2$ injectors on it, the $2(2m(j) + 2)$ vertices on the variable circuit suffice to insure that the degrees on any variable circuit are at most 3. The global construction (Fig. 7) is clearly polynomial.

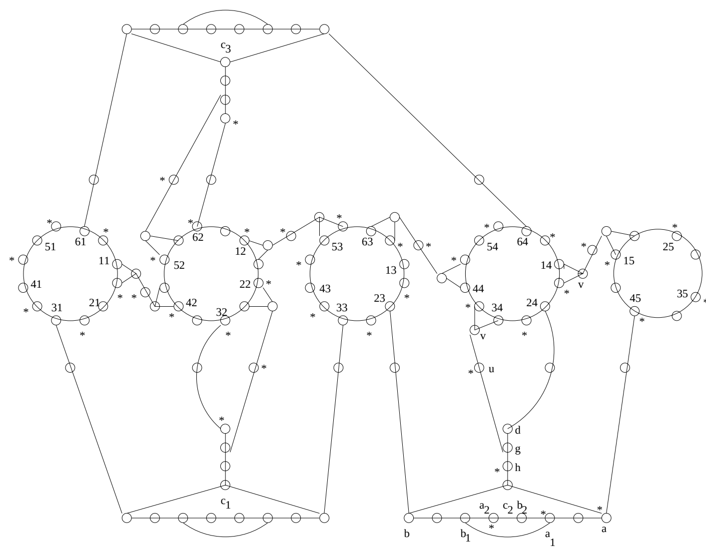


FIGURE 7. The global construction for $B = (x_1 \vee x_2 \vee x_3) \wedge (x_3 \vee x_4 \vee x_5) \wedge (x_1 \vee x_2 \vee x_4)$.

Now assume that 1-3SAT is 1-satisfiable, i.e., each clause in B has precisely one true variable. In the variable circuits, mark the x_{ij} $(1 \leq i \leq 2m(j)+2)$ if $x_j = 1$ for some truth assignment for which B is 1-satisfiable, and the unlabeled vertices between the x_{ij} $(1 \leq i \leq 2m(j)+2)$ if $x_j = 0$. Then precisely one of the x_{ij} for x_j in c_i is marked for each i, and precisely one of a, b, d is marked, namely, the one at the end of a 2-path whose other end is a marked x_{ij}. Also the vertices u (Figs. 6 and 7) and the mid-vertices on the injectors connecting adjacent variable circuits get injected marks. If d has been marked, no further vertex of c_i is marked. But if a or b has been marked, then h is marked; and so is a_1 and a_2 (if a is marked) or b_1 and b_2 (if b is marked). It is easily verified that the marked vertices constitute an even kernel of the constructed graph G.

Conversely, assume that G has an even kernel. We begin by collecting a few properties of G and its even kernel.

Proposition 1. *If a labeled vertex of a variable circuit is marked, then all labeled vertices of that variable circuit are marked.*

Proof. Proceed in clockwise direction from a marked vertex x_{ij}, and note that the mark necessarily "propagates" along the circuit at the labeled vertices, including those which impinge on 2-paths joined to the c_i, and at the beginnings of injectors, all of which are labeled. □

Proposition 2. *No focus v of any injector can be marked.*

Proof. Suppose v is marked. Then both x_{ij} in clockwise direction and the unlabeled vertex w in counterclockwise direction of a variable circuit (Fig. 8) are marked. By Proposition 1, all labeled vertices are marked, in particular $x_{i-1,j}$. This is a contradiction, since $x_{i-1,j}$ is adjacent to both v and w. □

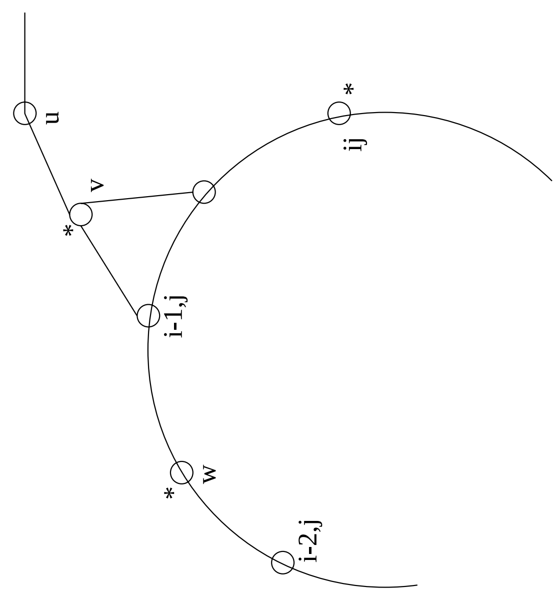


FIGURE 8. An impossible situation.

Proposition 3. *For every clause circuit c_i, none of the vertices g and v_j $(1 \le j \le 6)$ can be marked (Fig. 6).*

Proof. For g, v_1 and v_6 this follows directly from Proposition 2. If v_2 were marked, then both b_1 and b_2 would be marked, so a_1 would have an odd number of marked neighbors, a contradiction. By symmetry, also v_3 cannot be marked. But then also neither v_4 nor v_5 can be marked. □

Proposition 4. *If an unlabeled vertex w of a variable circuit is marked, then all unlabeled vertices of that variable circuit are marked.*

Proof. The only neighbors of the vertices of a variable circuit which lie outside that circuit, are vertices of the type v, v_4, v_5 and v_6 (Fig. 6). By Propositions 2 and 3, none of these is marked. Thus the mark at w necessarily propagates along the variable circuit itself. □

Propositions 1 and 4 imply that if any vertex on any variable circuit is marked, then because of the injectors between adjacent variable circuits, all variable circuits are marked: either all $2m(j)+2$ labeled or all $2m(j)+2$ unlabeled vertices are marked in every variable circuit. Moreover, these two possible markings are independent of each other in the n variable circuits.

We now show that any even kernel of G necessarily intersects a variable circuit. A mark on any of a, b or d injects a mark into a variable circuit via v_4, v_5 or v_6 respectively, a mark on u does so via v, and a mark on h via d or u. Also a mark on b_1 or b_2 injects a mark into a variable circuit via b, and a mark on a_1 or a_2 does so via a. By Propositions 2 and 3, no other vertex outside the variable circuits and their interconnecting injectors can be marked. Since an even kernel is nonempty, some vertex of a variable circuit must be marked. Hence all the variable circuits are marked; in fact, each one has all its labeled or else all its unlabeled vertices marked.

Each clause circuit c_i receives a mark that is injected via some u. Then precisely one of d and h is marked, otherwise g would have 3 marked neighbors. Assume first that d is marked. Since h is then unmarked, a is marked if and only if b is marked. But if both a and b are marked, then so are the adjacent vertices a_1 and b_1, a contradiction. Thus a and b are both unmarked. Secondly, assume that h is marked. Then precisely one of a and b is marked (a_1 and a_2 are marked if a is marked; b_1 and b_2 are marked if b is marked).

It follows that for every c_i, precisely one of the three terminals a, b or d is marked. Hence precisely one of the three variable circuits connecting to the terminals via 2-paths has all its labeled vertices marked, and the other two have all their unlabeled vertices marked. Putting $x_j = 1$ if and only if the j^{th} variable circuit has all its labeled vertices marked, thus constitutes a consistent truth assignment which 1-satisfies the given instance of 1-3SAT. □

3 Proof of Theorem 2

We make again a reduction from 1-3SAT. Since the construction is similar to that used for proving Theorem 1, we give less detail and refer the reader to Fig. 9, where the global construction is illustrated. We also use the same notation as in the proof of Theorem 1. The main difference between the two constructions is that in the present case the injector cannot be used, as it is not bipartite. Its function is emulated, in part, by long chains (paths) whose length may cause certain subgraphs to be marked or unmarked.

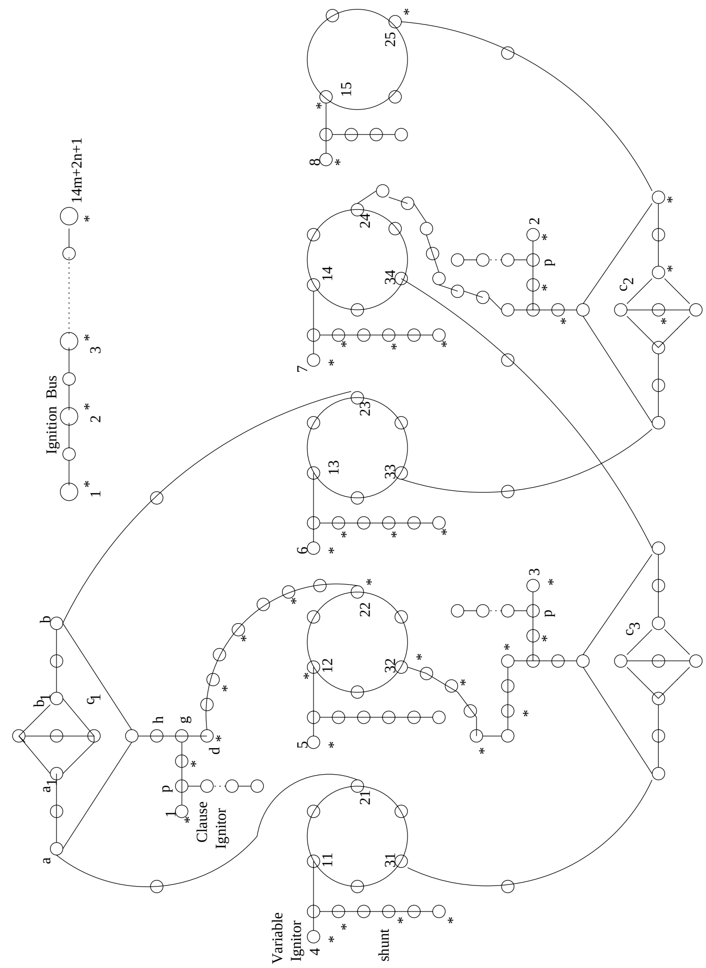


FIGURE 9. The global construction for $B = (x_1 \vee x_2 \vee x_3) \wedge (x_3 \vee x_4 \vee x_5) \wedge (x_1 \vee x_2 \vee x_4)$.

1. IGNITION BUS. This is a path of length $4(7m+n)$, where alternate vertices, including the two end vertices, are numbered $1, \ldots, 14m+2n+1$. In a proper labeling, the numbered vertices are "ignited", i.e., marked. The vertices numbered $1, \ldots, m$ feed into the m clause circuits $c_1, \ldots, c_m$, and the vertices numbered $m+1, \ldots, m+n$ feed into the n variable circuits via *ignitors* (see below). Each of the vertices $i\,(1 \le i \le m+n)$ appears twice in Fig. 9, but it is one and the same vertex. Its split into two was precipitated only by the desire to avoid the many intersecting edges which would otherwise clutter the drawing.

2. IGNITORS. The variable-ignitors are 2-paths feeding into the variable circuits, and the clause-ignitors are 3-paths feeding into the clause circuits. The vertices numbered i are marked in proper operation ($i \in \{1, \ldots, m+n\}$). From each vertex labeled p on each clause-ignitor, there emanates a simple path L of length $2(22m+3n+1)$. In proper

operation, all vertices of the paths L remain unmarked.

3. Variable Circuits. The variable circuit for x_j contains $2(m(j)+1)$ vertices ($1 \leq j \leq n$), and again alternate vertices are labeled. There is a *shunt* of length $2(m(j)+1)-1$ connected to the jth variable circuit. If the ignition bus is marked, then either the variable circuit or its shunt are marked, but not both.

4. Clause Circuits. There are terminals a, b, d, g as in the previous construction, but the clause circuits are now bipartite.

The construction is complete by putting $A = 14m + 2n + 1$ and $C = 22m + 3n + 1$. It is clearly polynomial and produces a bipartite graph G with degrees at most 3.

Suppose B is 1-satisfiable. Mark the numbered vertices on the ignition bus. Mark the labeled vertices in the variable circuit of x_j if and only if $x_j = 1$ in a truth assignment which makes B 1-satisfiable. In all other variable circuits, alternate vertices on the shunts are marked. Since B is 1-satisfiable, exactly one of a, b, d gets an induced mark from a variable circuit. Three additional vertices on c_i or on the path leading from d to a variable circuit are then marked. The resulting set of marked vertices forms an even kernel K of size

$$\begin{array}{rcl} |K| = & 14m + 2n + 1 & \text{(ignition bus)} \\ + & 4m + m & \text{(clause circuits and their ignitors)} \\ + & \sum_{j=1}^{n}(m(j)+1) & \text{(variable circuits/shunts)} \\ & & = 22m + 3n + 1 = C. \end{array}$$

Conversely, assume that G has an even kernel K of size $A \leq |K| \leq C$. First note that none of the vertices labeled p can be marked: if any were marked, we would already have an entire path marked, contributing $22m + 3n + 2 > C$ marks. Secondly, suppose that the ignition bus is unmarked. The largest K could then be is when each c_i contributes 8 labels (d is marked; so is precisely one of a and b), and the labeled vertices on all the variable circuits and shunts are marked. Then

$$|K| \leq 8m + \sum_{j=1}^{n} 2(m(j)+1) = 14m + 2n < A.$$

We could have unlabeled vertices in all the variable circuits and the two neighbors of v in all the c_i marked. But this clearly produces an even kernel of size $< A$.

Thus alternate vertices on the ignition bus have to be marked. It is easy to see that then in each c_i either one or all three terminals a, b, d are marked. If precisely one is marked, then $|K| = B$ as we saw in the first part of the proof. The marked terminal induces marks on the labeld vertices in the variable circuit it is connected to via a two-path. Putting $x_j = 1$ if and only if the labeled vertices are marked in the j-th variable circuit, thus constitutes a 1-satisfable solution to B. If even one c_i has 3 marked terminals, then

$|K| > B$, so this is not possible. So the existence of an even kernel K of size $A \le |K| \le |C|$ implies that B is 1-satisfiable. □

4 Proof of Theorem 3

The construction is similar to the constructions used for Theorems 1 and 2, especially to the latter. See Fig. 10 for the global picture. On the variable circuits we now have two shunts. The one connected to a labeled vertex is termed shunt s, and the other, connected to an unlabeled vertex, is shunt s'. From each of the vertices labeled p on the clause-ignitors, there emanates a path L of length $2(39m+6n+2)$. There are two ignition buses, numbered 1 and 2, each of length $4(7m+n)$. We put $A = 31m + 5n + 2$ and $C = 39m+6n+2$, and note that the construction is polynomial and produces a bipartite graph.

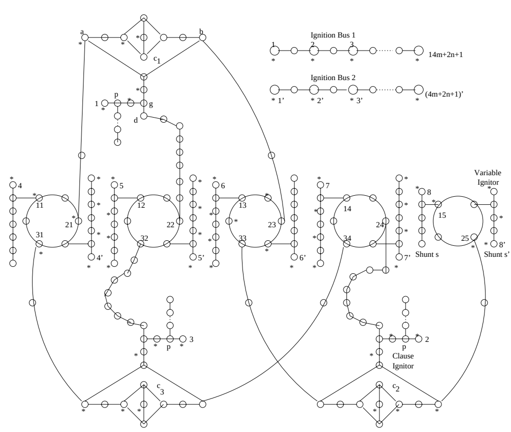


FIGURE 10. The global construction for $B = (x_1 \vee x_2 \vee x_3) \wedge (x_3 \vee x_4 \vee x_5) \wedge (x_1 \vee x_2 \vee x_4)$.

Suppose B is 1-satisfiable. Mark the numbered vertices on the ignition buses. Mark the labeled vertices in a variable circuit of x_j and alternate vertices on shunt s' if $x_j = 1$; the unlabeled vertices in the variable circuit or alternate vertices on s' (but not both), and alternate vertices on shunt s if $x_j = 0$ for a given truth assignment that renders B 1-satisfiable. Then exactly one of a, b, d in each c_i is marked, leading to a total of 4 marked vertices in each c_i. The result is an even kernel K of size

$$|K| = 28m + 4n + 2 + 5m + 2\sum_{j=1}^{n}(m(j) + 1) = 39m + 6n + 2 = C.$$

Conversely, assume that G has an even kernel K of size $A \le |K| \le C$. None of the vertices labeled p can be labeled, for otherwise we would already have a kernel of size $\ge 39m + 6n + 3 > C$. Now suppose that at most one of the ignition buses is marked, say ignition bus 2. Then each c_i can contribute at most 8 to K and each variable circuit at most $3(m(j) + 1)$, so

$$|K| \le 14m + 2n + 1 + 8m + 3\sum_{j=1}^{n}(m(j) + 1) = 31m + 5n + 1 < A.$$

If ignition bus 1 rather than 2 is marked, we get a smaller even kernel. It follows that the numbered vertices of both ignition buses have to be marked. Then each variable circuit has the labeled vertices and alternate vertices on shunt s' marked; or else alternate vertices on s, and either the unlabeled vertices or alternate vertices on s' (but not both). Each c_i has either precisely one of a, b, d labeled or else all three of them. In the first case we have then a kernel K of size

$$|K| = 28m + 4n + 2 + 5m + 2\sum_{j=1}^{n}(m(j) + 1) = 39m + 6n + 2 = C.$$

If even a single c_i has all of the a, b, d marked, then the kernel would obviously be larger than C. □

REFERENCES

[1] E.R. Berlekamp, R.J. McEliece and H.C.A. van Tilborg [1978], On the inherent intractability of certain coding problems, *IEEE Trans. Inform. Theory* **IT-24**, 384-386.

[2] A.S. Fraenkel, E.R. Scheinerman and D. Ullman [1993], Undirected edge geography, *Theoret. Comput. Sci. (Math. Games)* **112**, 371-381.

[3] M.R. Garey and A.S. Johnson [1979], *Computers and Intractability: A guide to the Theory of* NP-*Completeness*, Freeman, New York.

[4] T.J. Schaefer [1978], The complexity of satisfiability problems, *Proc. 10th Annual ACM Symp. on Theory of Computing*, Assoc. Comput. Mach., New York, 216-226.