Recognizing Balanceable Matrices

Michele Conforti[*], Giacomo Zambelli[†]

August, 2004
Revised November 2004

Abstract

A $0/\pm 1$ matrix is balanced if it does not contain a square submatrix with exactly two nonzero entries per row and per column in which the sum of all entries is 2 modulo 4. A 0/1 matrix is balanceable if its nonzero entries can be signed ± 1 so that the resulting matrix is balanced. A signing algorithm due to Camion shows that the problems of recognizing balanced $0/\pm 1$ matrices and balanceable 0/1 matrices are equivalent. Conforti, Cornuéjols, Kapoor and Vušković gave an algorithm to test if a $0/\pm 1$ matrix is balanced. Truemper has characterized balanceable 0/1 matrices in terms of forbidden submatrices. In this paper we give an algorithm that explicitly finds one of these forbidden submatrices or shows that none exists.

1 Introduction

A $0/\pm 1$ matrix A is *balanced* if A does not contain a square submatrix with exactly two nonzero entries per row and column in which the sum of all entries is 2 modulo 4. This notion was introduced by Berge [1] for 0/1 matrices and extended to $0/\pm 1$ matrices by Truemper [14]. Balanced matrices have rich connections with integer programming, in particular with the set packing and set covering models (see, for example, Berge [2], Conforti and Cornuéjols [5], Fulkerson, Hoffman and Oppenheim [13]). The first known polynomial time recognition algorithm for 0/1 balanced matrices is due to Conforti, Cornuéjols and Rao [8], followed by a similar algorithm for the general $0/\pm 1$ case due to Conforti, Cornuéjols, Kapoor, Vušković [7].

A 0/1 matrix A is *balanceable* if the nonzero entries of A can be signed $+1$ or -1 so that the resulting $0/\pm 1$ matrix A' is balanced. Truemper [14] gave a co-NP characterization for this class of matrices by showing that, if a matrix is not balanceable, then it must contain

[*]Dipartimento di Matematica, Universitá degli Studi di Padova, Via Belzoni 7, 35131, Padova, Italy
[†]Department of Combinatorics and Optimization, University of Waterloo, 200 University Ave., Waterloo, Ontario, Canada, N2J 3G1

some well described submatrix. This indicates that one could recognize balanceable matrices by looking for these forbidden submatrices. An algorithm that does precisely this will be described in the next sections. A similar algorithm for recognizing balanced matrices was given in [16] by one of the authors, and it was inspired by techniques introduced by Chudnovsky and Seymour [4].

1.1 Camion's signing algorithm

In this paper all graphs are simple. We refer the reader to the book of West [15] for standard graph theory terminology. In the remainder of the paper, we work with the bipartite representation of a matrix. Given a 0/1 matrix A, the *bipartite representation of* A is the simple bipartite graph $G(A)$ where the two sides of the bipartition are the sets R and C of rows and columns of A, respectively, and $i \in R$ and $j \in C$ are adjacent if and only if $a_{ij} = 1$. Conversely, any bipartite graph G is the bipartite representation of a 0/1 matrix, the matrix being unique, up to row and column permutation and transposition.

A *signing* σ of the edges of a bipartite graph G is a function from $E(G)$ to $\{1, -1\}$. A *hole* H in G is a chordless cycle. Given a chordless path (resp. hole) Q in G, we denote $\sigma(Q) = \sum_{e \in Q} \sigma(e)$.

Since there is a one-to-one correspondence between minimal square submatrices of A with exactly two ones per row and per column and holes in $G(A)$, and each signing of the nonzero entries of A corresponds to a signing of the edges of $G(A)$, then a matrix A is balanceable if and only if its bipartite representation $G(A)$ admits a signing σ such that, for each hole H in $G(A)$, $\sigma(H) \equiv 0 \bmod 4$.

Therefore we say that a signed bipartite graph is *balanced* if $\sigma(H) \equiv 0 \bmod 4$ every hole H. We also say that a bipartite graph is *balanceable* if it can be signed to be balanced.

Camion [3] gave a polynomial time algorithm to sign the edges of a bipartite graph G so that the resulting signed graph (G, σ) is balanced whenever G is balanceable. Given a set S of nodes of G, we denote by $(S, \bar{S})$ the set of edges with one endnode in S and the other in $\bar{S} = V \setminus S$. Camion's algorithm is based on the following observation:

Given a set S of nodes in a signed graph (G, σ), multiplying the signs of the edges in $(S, \bar{S})$ by -1 gives a signing σ' so that $\sigma(H) \equiv \sigma'(H) \bmod 4$, for every hole H. Therefore (G, σ) is balanced if and only if (G, σ') is balanced. Since, given a maximal forest F of G and an edge e of F, there exists a cut $(S, \bar{S})$ of G such that $E(F) \cap (S, \bar{S}) = \{e\}$, this implies that any signing of F can be extended to a signing σ of G so that (G, σ) is balanced whenever G is balanceable (in fact, given a signing σ such that (G, σ) is balanced, and an edge $e \in F$ where $\sigma(e)$ differs from the prescribed signing of F, we can choose S such that $(S, \bar{S}) \cap F = \{e\}$, and change sign to all edges in $(S, \bar{S})$ so that the resulting signing σ' coincide with σ on all edges of F, except e).

Camion's Algorithm:

- **Input** A bipartite graph G, a maximal forest F of G, and a signing of F.

- **Output** A signing σ of G, extending the signing of F.

For each edge e of F, let $\sigma(e)$ be the given signing of e. Let $G_0 = F$, $n = |E(G)|$.
For $i = 0, \ldots, n - |E(F)| - 1$, do the following:

1. Choose an edge $e_i \in E(G) \setminus E(G_i)$ and a path P_i in G_i between its two endnodes so that $|P_i|$ is minimum over all possible choices of e_i and P_i;

2. Define $\sigma(e_i) \equiv -\sigma(P_i) \bmod 4$, and $G_{i+1} = (V(G), E(G_i) \cup \{e_i\})$.

Lemma 1.1 *The signed graph (G, σ) produced by the algorithm is balanced whenever G is balanceable and the signing σ, extending the signing of F, is unique.*

Proof: At each iteration, the edge e_i and the path P_i form a hole H_i of G_{i+1} which, by the choice of e_i and P_i, is also a hole in G. The only way to extend the signing constructed so far so that $\sigma(H_i) \equiv 0 \bmod 4$ is to assign $\sigma(e_i) \equiv -\sigma(P_i) \bmod 4$. Since we know that there exists a balanced signing of G which extends the signing of F, then the signing produced by the algorithm is the only possible. $\square$

Theorem 1.2 *The problems of testing if a $0/1$ matrix is balanceable and of testing if a $0/\pm 1$ matrix is balanced are polynomially reducible one to the other.*

Proof: Suppose we have a polynomial time algorithm to test if a matrix is balanceable, and we wish to test if a given $0, \pm 1$ matrix A is balanced. Let B be the support matrix of A. Test if B is balanceable. If it is not, then output that A is not balanced. Else, let F be a maximal forest in the bipartite representation of B and let a_{ij} be the signing of $ij \in E(F)$. Apply Camion's algorithm to B, F, and the signing of F, to obtain a balanced matrix B'. Since B' is unique, then A is balanced if and only if $A = B'$.

To test if a $0/1$ matrix A is balanceable, one can apply Camion's algorithm to produce a signed copy A' and then test if A' is a balanced $0/\pm 1$ matrix. A is balanceable if and only if A' is balanced. $\square$

1.2 Truemper's theorem

We say that a graph G *contains* a graph F, whenever G contains an induced subgraph isomorphic to F. Given a set X of nodes of G, we denote by $G[X]$ the subgraph of G induced by X. Given a subgraph F of G and a node x of G, we denote the set of neighbors of x in F by $N_F(x)$.

Given a path or a hole Q, we will denote by $|Q|$ the length of Q. Given a graph F and two nodes x, y of F, $d_F(x, y)$ denotes the length of the shortest path between x and y contained in F. Also, if P is a chordless path and x, y are two nodes of P, we will denote by $P(x, y)$ the unique subpath of P between x and y. The *interior* of P is the set of all nodes of P except the endnodes of P.

The following two graphs will play an important role in the remainder of the paper. Given two nonadjacent nodes a and b in distinct sides of the bipartition, a *3-path configuration* between a and b is a graph consisting of three chordless paths P_1, P_2, P_3 between a and b, all of length greater than one, such that, for every $1 \leq i < j \leq 3$, no node in the interior of P_i belongs to or has a neighbor in the interior of P_j. We say that P_1, P_2, P_3 form a 3-path configuration. A *wheel* consists of a hole H and a node v with at least 3 distinct neighbors in H, and is denoted by (H, v). The node v is called the *center* of the wheel. A wheel (H, v) for which v has k neighbors in H is said a k-wheel. A k-wheel is an *odd wheel* if k is odd.

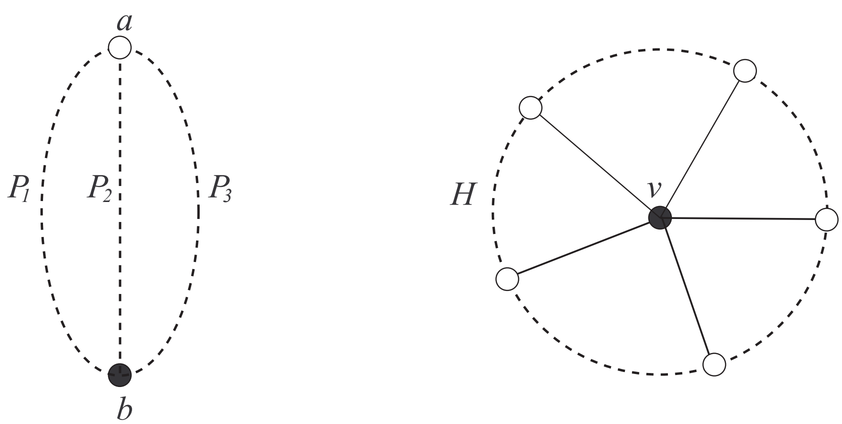

Figure 1: A 3-path configuration and a wheel.

It is easy to see that if G contains a 3-path configuration or an odd wheel, then G is not balanceable. In fact, if F is a 3-path configuration or a wheel contained in G, then F contains an odd number of edges, and each edge is contained in exactly 2 holes. Denote by $\mathcal{H}$ the family of all holes in F. For any signing σ of F, $\sum_{H\in\mathcal{H}} \sum_{e\in E(H)} \sigma(H) = 2\sum_{e\in E(F)} \sigma(e) \equiv 2 \bmod 4$, therefore there exists an odd number of holes H such that $\sum_{e\in E(H)} \sigma(e) \equiv 2 \bmod 4$.

Truemper showed that the converse is also true.

Theorem 1.3 (Truemper [14]) *A bipartite graph is balanceable if and only if it does not contain a 3-path configuration or an odd wheel.*

A nice proof of Theorem 1.3 can be found in [9].

In the remainder of the paper, we will describe an $O|V(G)|^9$ algorithm to test whether a bipartite graph G contains a 3-path configuration or an odd wheel. By Theorem 1.3, this is equivalent to test it G is balanceable.

In sections 2 and 3 we describe two algorithms, both given in [16], to recognize whether G has a 3-path configuration or a detectable 3-wheel (which is a special type of odd wheel), respectively. In section 4 we show how to produce, from a graph G with no 3-path configurations, a family of polynomially many sets of nodes of G, such that for some odd wheel (H, x) of G (if any exists), one of these sets is disjoint from H and contains all nodes y that are centers of odd wheels (H, y). In section 5 we show how to detect some "special" odd wheel (H, x) in G, provided that G does not contain any 3-path configuration or any detectable 3-wheel, and that we are given a set X disjoint from H containing all centers y of odd wheels (H, y). In section 6 we finally provide the $O(|V(G)|^9)$ algorithm to test if a bipartite graph is balanceable.

The results of sections 2 and 3 where proven in [16], and we report them here for the sake of completeness.

2 Detecting a 3-path configuration

We say that a 3-path configuration is *smallest* in G if it contains the minimum number of nodes among all 3-path configurations in G. We denote by C and R the two sides of the bipartition of G.

Lemma 2.1 *Let Π be a smallest 3-path configuration in G. Assume Π consists of the paths $P_i = a, a_i, \dots, b_i, b$, $i \in \{1, 2, 3\}$, where $a \in R$, $b \in C$. For every $i \in \{1, 2, 3\}$, let m_i be a node of P_i such that $|d_{P_i}(a_i, m_i) - d_{P_i}(b_i, m_i)| \leq 1$. Let X be the set of nodes of G with no neighbors in $\{a, b, a_2, a_3, b_2, b_3\}$, and P be a shortest path between a_1 and m_1 in $G[X \cup \{a_1, m_1\}]$. Then $P'_1 = a, a_1, P, m_1, P_1(m_1, b_1), b_1, b$ is a chordless path and P'_1, P_2, P_3 form a smallest 3-path configuration.*

Symmetrically, analogous statements hold for every P_i, $i \in \{1, 2, 3\}$, and all possible pairs a_i, m_i and m_i, b_i

Proof: Let $P = p_1, \dots, p_k$ where $a_1 = p_1$ and $m_1 = p_k$. If $a_1 = m_1$ or a_1 is adjacent to m_1, then the statement holds trivially, hence we may assume $|P_1| \geq 5$ and $m_1 \neq b_1$, therefore m_1 has no neighbors in P_2 or P_3.

If no node in the interior of P belongs to or has a neighbor in P_2 or P_3 then, given P'_1 the shortest path between a and b with interior in $V(P \cup P_1(m_1, b_1))$, P'_1, P_2, P_3 form a 3-path configuration between a and b which, by the minimality of Π and the choice of P, must have the same cardinality as Π, hence $P'_1 = a, a_1, P, m_1, P_1(m_1, b_1), b_1, b$ and we are done.

Assume, then, that there exists h, $2 \leq h \leq k-1$, such that p_h belongs to or has a neighbor in P_2 or P_3, and let h be maximum with this property. Note that, by definition, p_h does not belong to P_2 or P_3.
Suppose p_h has at least two distinct neighbors in $P_2 \cup P_3$. If $p_h \in R$, let P'_1 be the shortest path between p_h and b in $P(p_h, p_k) \cup P_1(m_1, b)$, let P'_2 be the (unique) shortest path between p_h and b in $(p_h \cup P_2 \cup P_3) \setminus b_3$ and P'_3 be the (unique) shortest path between p_h and b in $(p_h \cup P_2 \cup P_3) \setminus b_2$. Then P'_1, P'_2, P'_3 form a 3-path configuration between p_h and b which is strictly shorter than Π since $|P'_1| < |P_1|$ and $|P'_2| + |P'_3| \leq |P_2| + |P_3|$. Similarly,

if $p_h \in C$, let P'_1 be the shortest path between a and p_h in $P(p_h, p_k) \cup P_1(a, m_1)$, let P'_2 be the (unique) shortest path between a and p_h in $(p_h \cup P_2 \cup P_3) \setminus a_3$ and P'_3 be the (unique) shortest path between a and p_h in $(p_h \cup P_2 \cup P_3) \setminus a_2$. Then P'_1, P'_2, P'_3 form a 3-path configuration Π' between a and p_h. Since $|P_1(a_1, m_1)| \leq |P_1(b_1, m_1)| + 1$ and $h \geq 2$, then

$$\begin{aligned} |P'_1| &\leq |P| - 1 + |P_1(a, m_1)| \leq |P_1(a, m_1)| + |P_1(a_1, m_1)| - 1 \\ &\leq |P_1(a, m_1)| + |P_1(m_1, b_1)| < |P_1|. \end{aligned} \tag{1}$$

Furthermore, $|P'_2| + |P'_3| \leq |P_2| + |P_3|$, hence Π' has cardinality strictly smaller than Π, a contradiction.
Therefore, we may assume that p_h has a unique neighbor x in $P_2 \cup P_3$, say $x \in V(P_2)$. If $x \in R$, then let P'_1 be the shortest path between x and b in $x \cup P(p_h, m_1) \cup P_1(m_1, b)$, let $P'_2 = x, P_2(x, b), b$ and $P'_3 = x, P_2(x, a), a, P_3, b$. Then P'_1, P'_2, P'_3 form a 3-path configuration between x and b which has cardinality strictly smaller than Π since $|P'_2| + |P'_3| = |P_2| + |P_3|$ and

$$|P'_1| \leq |P| - 1 + |P_1(m_1, b)| + 1 \leq |P_1(a_1, m_1)| + |P_1(m_1, b)| < |P_1|.$$

If $x \in C$, then let P'_1 be the shortest path between x and a in $x \cup P(p_h, m_1) \cup P_1(a, m_1)$, let $P'_2 = a, P_2(a, x), x$ and $P'_3 = a, P_3, b, P_2(b, x), x$. Then P'_1, P'_2, P'_3 form a 3-path configuration Π' between x and a. If $h = 2$, then $|P'_1| = 3 < |P_1|$, otherwise $h \geq 3$ and $|P'_1| \leq |P| + |P_1(a, m_1)| - 1 < |P_1|$. Since $|P'_2| + |P'_3| = |P_2| + |P_3|$, then Π' has cardinality strictly smaller than Π, a contradiction. $\square$

Lemma 2.2 *There exists a $O(|V(G)|^9)$ algorithm with the following specifications:*

- **Input** *A bipartite graph G.*
- **Output** *Either:*
  1. *a 3-path configuration Π, or*
  2. *it determines that G does not contain any 3-path configurations.*

Algorithm:
For every 6 tuple $a_1, a_2, a_3, b_1, b_2, b_3$ such that:

- $a_i \in R$, $b_i \in C$ for every $i \in \{1, 2, 3\}$,
- a_i is nonadjacent to b_j for every $i \neq j$,
- there exist nonadjacent nodes x and y such that x is adjacent to a_1, a_2, a_3 and y is adjacent to b_1, b_2, b_3;

do the following:

1. For $i = 1, 2, 3$, compute the set $X(i)$ of nodes that are not adjacent to any of x, y, a_j or b_j for $j \neq i$.

2. For $i = 1, 2, 3$, for every node $m \in X(i)$, compute the paths $P_i'(m)$ and $P_i''(m)$ (if they exist), where $P_i'(m)$ is the shortest path between a_i and m in $G[X(i) \cup a_i]$ and $P_i''(m)$ is the shortest path between b_i and m in $G[X(i) \cup b_i]$.

3. For $i = 1, 2, 3$, for every node $m \in X(i) \cup a_i$, define $P_i(m)$ as follows: if a_i is adjacent to b_i, then $P_i(a_i) = a_i, b_i$ and $P_i(m)$ is undefined for every $m \in X(i)$; else $P_i(a_i)$ is undefined and for every $m \in X(i)$ satisfying the following

   (i) $P_i'(m)$ and $P_i''(m)$ both exist

   (ii) No node in $P_i'(m)$, except m, belongs to or has a neighbor in $P_i''(m)$

   let $P_i(m) = x, a_i, P_i'(m), m, P_i''(m), b_i, y$, else, if $P_i'(m)$ and $P_i''(m)$ do not satisfy (i) and (ii), $P_i(m)$ is undefined.

4. For every $m \in X(i) \cup a_i$ such that $P_i(m)$ is defined, compute the set $Y_i(m)$ of nodes that do not belong or have a neighbor in the interior of $P_i(m)$.

5. For every $1 \leq i < j \leq 3$, and for every $m_i \in X(i) \cup a_i$ and every $m_j \in X(j) \cup a_j$, verify that the interior of $P_j(m_j)$ is contained in $Y_i(m_i)$. If this is the case, say that the pair m_i, m_j is (i, j)-good.

6. Verify if there exists a triple m_1, m_2, m_3 such that $m_i \in X(i) \cup a_i$ for $i \in \{1, 2, 3\}$ and such that m_i, m_j is (i, j)-good for every $1 \leq i < j \leq 3$. If such a triple exist, output the graph Π induced by $P_1(m_1), P_2(m_2), P_3(m_3)$ and stop.

Otherwise output the fact that G contains no 3-path configuration.

Correctness: It takes time $O(|V(G)|^8)$ to compute all possible 6-tuples a_1, a_2, a_3, b_1, b_2, b_3 as above, and there are $O(|V(G)|^6)$ of them. For each 6-tuple, each step from 1 through 6 takes time $O(|V(G)|^3)$, therefore the total running time is $O(|V(G)|^9)$.
If for some 6-tuple, in step 6 the algorithm outputs a graph Π induced by $P_1(m_1)$, $P_2(m_2)$, $P_3(m_3)$, then Π is a 3-path configuration between x and y, since step 3 ensures that $P_i(m_i)$ is a chordless path between x and y for every $i \in 3$, while steps 5 and 6 guarantee that no node in the interior of $P_i(m_i)$ belongs to or has a neighbor in the interior of $P_j(m_j)$ for every $1 \leq i < j \leq 3$.
We only need to verify that, if G contains some 3-path configuration, then the algorithm will detect one. Let $\tilde{\Pi}$ be a smallest 3-path configuration in G. Let $\tilde{P}_1, \tilde{P}_2, \tilde{P}_3$ be the 3-paths inducing $\tilde{\Pi}$, where $\tilde{P}_i = a, a_i, \dots, b_i, b$. Then there exist nonadjacent nodes x and y such that x is adjacent to a_i and y is adjacent to b_i for every $i \in \{1, 2, 3\}$ (since $x = a$ and $y = b$ would satisfy such condition). For every $i \in \{1, 2, 3\}$, let P_i be the shortest path between x and y with interior contained in the interior of $\tilde{P}_i$. Then P_1, P_2, P_3 form a 3-path configuration Π with at most as many nodes as $\tilde{\Pi}$, hence Π and $\tilde{\Pi}$ must have the same cardinality and $P_i = x, a_i, \dots, b_i, y$. For every $i \in \{1, 2, 3\}$, let m_i be a node of P_i such that $|d_{P_i}(a_i, m_i) - d_{P_i}(b_i, m_i)| \leq 1$, in particular we may assume that, when a_i and b_i are adjacent, $m_i = a_i$. Then, by Lemma 2.1, given $P_1' = x, a_1, P_1'(m_1), m_1, P_1(m_1, b_1), b_1, x$,

where $P_1'(m_1)$ is the path computed in step 2 of the algorithm, P_1', P_2, P_3 forms a 3-path configuration between x and y. By repeating the argument, we conclude that the paths $P_1(m_1), P_2(m_2), P_3(m_3)$ computed by the algorithm form a 3-path configuration between x and y, hence the algorithm would have output the correct answer. □

3 Detectable 3-wheels

A 3-wheel (H, v) is *detectable* if two of the neighbors of v in H have distance two in H. If (H, v) has the minimum number of nodes among all detectable 3-wheels in G, we say that (H, v) is a *smallest* detectable 3-wheel.

Lemma 3.1 *Let G be a bipartite graph containing no 3-path configurations. Let (H, v) be a smallest detectable 3-wheel in G. Let u, v_1 and v_2 be the neighbors of v in H, where v_1 and v_2 are both adjacent to a node w in H. Let u_1 and u_2 be the two neighbors of u in H such that the two maximal paths P_1 and P_2 in $H \setminus \{u, w\}$ have endpoints u_1, v_1 and u_2, v_2, respectively. Let s be the neighbor of u_1 in P_1. Let X be the set of nodes of G with no neighbors in $\{u, v, w, u_2, v_2\}$. Let P be a shortest path between v_1 and s in $G[X \cup \{v_1, s\}]$. Then $H' = v_1, P, s, u_1, u, u_2, P_2, v_2, w, v_1$ is a hole and (H', v) is a smallest detectable 3-wheel.*

Proof: Let $P = p_1, \ldots, p_k$, where $p_1 = v_1$ and $p_k = s$. Let C and R be the sides of the bipartition of G. W.l.o.g., $v \in R$ and $u \in C$. If no node in the interior of P belongs to or has a neighbor in P_2, then $H' = v_1, P, s, u_1, u, u_2, P_2, v_2, w, v_1$ is a hole, hence by construction (H', v) is a detectable 3-wheel which is smallest since $|P| \leq d_{P_1}(s, v)$. We may therefore assume that there exists h, $2 \leq h \leq k-1$, such that p_h belongs to or has a neighbor in P_2. Assume h is the highest such index. Then p_h does not belong to P_2. Suppose p_h has exactly one neighbor in P_2, say x. If $x \in R$, then let Q_1 be the shortest path between x and u in $P(p_h, p_k) \cup x, u_1, u$, let $Q_2 = x, P_2(x, v_2), v_2, v, u$ and $Q_3 = x, P_2(x, u_2), u_2, u$. Then $|Q_i| \geq 3$ and Q_1, Q_2, Q_3 form a 3-path configuration between x and u, a contradiction. If $x \in C$, then let Q_1 be the shortest path between x and v in $P(p_h, p_k) \cup P_1(s, v_1) \cup \{v, x\}$, $Q_2 = x, P_2(x, v_2), v_2, v$ and $Q_3 = x, P_2(x, u_2), u_2, u, v$. Q_1, Q_2, Q_3 form a 3-path configuration between x and v. Hence we may assume that p_h has at least 2 neighbors in P_2. Let x and y be the neighbors of p_h in P_2 that are closest, respectively, to v_2 and u_2. If $p_h \in R$, let Q_1 be the shortest path between p_h and u in $P(p_h, p_k) \cup u_1, u$, let $Q_2 = p_h, x, P_2(x, v_2), v_2, v, u$ and $Q_3 = p_h, y, P_2(y, u_2), u_2, u$. Then Q_1, Q_2, Q_3 form a 3-path configuration between p_h and u. If $p_h \in C$, then let Q_1 be the shortest path between p_h and v in $P(p_h, p_k) \cup P_1(s, v_1) \cup v$, $Q_2 = p_h, x, P_2(x, v_2), v_2, v$ and $Q_3 = p_h, y, P_2(y, u_2), u_2, u, v$. Q_1, Q_2, Q_3 form a 3-path configuration between x and v, a contradiction. □

Lemma 3.2 *There exists a $O(|V(G)|^9)$ algorithm with the following specifications:*

- **Input** *A bipartite graph G containing no 3-path configuration.*

- **Output** *Either:*
  1. *a detectable 3-wheel, or*
  2. *it determines that G does not contain any detectable 3-wheel.*

Algorithm:
For every 7-tuple $u_1, u_2, v, v_1, v_2, w, s$ such that:

- v and w are both adjacent to v_1 and v_2
- there exists a node x such that x is adjacent to v, u_1, u_2 but not to w or s
- s is adjacent to u_1
- either $s = v_1$ or no node in $\{u_2, v, v_2, x, w\}$ is coincident with or adjacent to s.

do the following:

1. Compute the set X of nodes of G that do not belong to or have a neighbor in $\{u_2, v, v_2, x, w\}$.
2. Compute the shortest path P, if one exists, between v_1 and s in $G[X \cup \{v_1\}]$.
3. Verify that the only neighbor of u_1 in P is s, if this is the case let $P_1 = v_1, P, s, u_1$, otherwise P_1 is undefined.
4. If P_1 is not undefined, compute the set Y of all nodes that do not belong to or have a neighbor in $P_1 \cup \{v, w, x\}$.
5. Compute, if one exists, a chordless path P_2 between u_2 and v_2 with interior contained in Y. If P_2 exists, then let $H = w, v_1, P_1, u_1, x, u_2, P_2, v_2, w$, output (H, v) and stop.

Otherwise output the fact that G does not contain any detectable 3-wheel.

Correctness: It takes time $O(|V(G)|^8)$ to compute all possible 7-tuples u_1, u_2, v, v_1, v_2, w, s as above, and there are $O(|V(G)|^7)$ of them. For every 7-tuple, step 4 takes time $O(|V(G)|^2)$, while all other steps take linear time, thus the overall running time is $O(|V(G)|^9)$.
Obviously, when the algorithm outputs a graph (H, v), such graph is a detectable 3-wheel. Suppose that G contains some detectable 3-wheel. We want to show that the algorithm will output one. Let $(\tilde{H}, v)$ be a smallest detectable 3-wheel in G. Let u, v_1 and v_2 be the neighbors of v in $\tilde{H}$, where v_1 and v_2 are both adjacent to a node w in $\tilde{H}$. Let u_1 and u_2 be the two neighbors of u in $\tilde{H}$ such that the two maximal paths $\tilde{P}_1$ and $\tilde{P}_2$ in $\tilde{H} \setminus \{u, w\}$ have endpoints u_1, v_1 and u_2, v_2, respectively. Let s be the neighbor of u_1 in $\tilde{P}_1$. Then the 7-tuple $u_1, u_2, v, v_1, v_2, w, s$ satisfies the properties described in the algorithm, hence at some stage the algorithm will examine it. Let x be a node adjacent to v, u_1, u_2 but

not to w or s (such a node exists since $x = u$ satisfies such condition). Let u'_1 and u'_2 be neighbors of x in $\tilde{H}$, such that u'_i is closest possible to v_i in $\tilde{P}_i$, $i = 1, 2$, and let P'_i be the path between v_i and u'_i in $\tilde{P}_i$. Then $H' = w, v_1, P'_1, u'_1, x, u'_2, P'_2, v_2, w$ is a hole and (H', v) is a detectable 3-wheel with at most as many nodes as $(\tilde{H}, v)$, therefore $P'_i = \tilde{P}_i$, for $i = 1, 2$. Let P be the shortest path between v_1 and s in $G[X \cup v_1]$ computed by the algorithm in step 2. Then, by Lemma 3.1, $P_1 = v_1, P, s, u_1$ is a path and the algorithm will verify this in step 3. Finally, there exists a chordless path P_2 between u_2 an v_2 with interior in the set Y computed at step 5 of the algorithm, since $\tilde{P}_2$ is such a path, therefore $H = w, v_1, P_1, u_1, x, u_2, P_2, v_2, w$ is a hole and (H, v) is detectable 3-wheel. $\square$

4 Cleaning

Throughout this section we assume that G is a bipartite graph that does not contain a 3-path configuration, and we will denote with R, C the two sides of the bipartition of G.

We say that (H, x) is a *smallest odd wheel* in G, if (H, x) is an odd wheel in G with the minimum number of nodes. We say that a vertex $v \in V(G) \setminus V(H)$ is *major* for H if $N_H(v)$ is not contained in a subpath of H of length 2, and denote by $M(H)$ the set of major nodes for H. A set $X \subseteq V(G) \setminus V(H)$ is a *cleaner for* H if $M(H) \subseteq X$. A set X is said to be a *cleaner for* G if either G is balanceable or G contains a smallest odd wheel (H, x) such that X is a cleaner for H.

We will give an algorithm, running in time $O(|V(G)|^7)$, that, given a graph G containing no 3-path configuration, constructs a family $\mathcal{C}$ of subsets of $V(G)$ containing $O(|V(G)|^6)$ members such that, if (H, x) is a smallest odd wheel in G, then $\mathcal{C}$ contains a cleaner for H.

First, we need to prove five lemmas. Given a chordless path or a hole Q and a set $X \subseteq V(G)$ with at least two distinct elements in Q, an *X-sector of Q* is a maximal subpath of Q whose interior does not contain a node in X.

Lemma 4.1 *Let H be a hole in G. Let X, Y be subsets of $V(H)$ such that $|X|$ is odd, $|Y|$ is even, and $|X|, |Y| \geq 2$. Then one of the following holds:*

- (i) *there exists a Y-sector of H, containing an odd number of elements of X, that contains an element of X in its interior.*
- (ii) *there exists an X-sector of H, containing an odd number of elements of Y, that contains an element of Y in its interior.*

Proof: If $X \cap Y = \emptyset$, then, since X has an odd number of elements, there exist an odd number of Y-sectors of H containing an odd number of elements of X, thus (i) holds.

Since $(X \cup Y) \setminus (X \cap Y)$ has an odd number of elements in H, there exist $q_1, q_2 \in X \cap Y$ such that either q_1, q_2 are distinct endnodes of an $X \cap Y$-sector Q of H containing an odd number of elements of $(X \cup Y) \setminus (X \cap Y)$, or $|X \cap Y| = 1$ and $q_1 = q_2$ (in which case we denote $Q = H$). In particular Q contains an even number of elements of $X \cup Y$ if $q_1 = q_2$, odd otherwise.

Let $S_1, \ldots S_k$ be the $X \cup Y$-sectors of Q with one endnode in $X \setminus Y$ and the other in $Y \setminus X$ (if any), in the order they appear traversing Q from q_1 to q_2, and let $S_i = s_i \ldots, t_i$ for $1 \leq i \leq k$ (notice that if $q_1 = q_2$ one such sector exists). For ease of notation, define $S_0 = s_0, \ldots t_0$, where $s_0 = t_0 = q_1$ and $S_{k+1} = s_{k+1}, \ldots t_{k+1}$, where $s_{k+1} = t_{k+1} = q_2$.

For every $0 \leq i \leq k$ there exist distinct $U, V \in \{X, Y\}$ such that $t_i, s_{i+1} \in U$, and $Q(s_i, t_{i+1})$ is a V-sector containing $Q(t_i, s_{i+1})$. Therefore we may assume that $Q(t_i, s_{i+1})$ contains a positive even number of elements of U (and thus an even number of elements of $X \cup Y$), otherwise $Q(s_i, t_{i+1})$ is a V-sector containing an odd number of elements of U, that contains an element of U in the interior, and the statement holds.

Since $|V(Q) \cap (X \cup Y)| = \sum_{i=0}^{k} |V(Q(t_i, s_{i+1})) \cap (X \cup Y)| - |\{q_1\} \cap \{q_2\}|$, then Q contains an odd number of elements of $X \cup Y$ if $q_1 = q_2$, even otherwise, a contradiction. $\square$

Lemma 4.2 *Let H be a hole in G. Let X, Y, Z be subsets of $V(H)$ of odd cardinality, such that $|X|, |Y|, |Z| \geq 3$ and $X \cap Y \cap Z = \emptyset$. Then there exist distinct $U, V, W \in \{X, Y, Z\}$ such that one of the following holds:*

- (i) *there exists a U-sector S of H, containing an odd number of elements of W, that contains an element of W in its interior.*
- (ii) *there exists a $U \cup V$-sector S of H containing an odd number of elements of W, such that one endnode of S is in $U \setminus V$ and the other in $V \setminus U$.*

Proof: We may assume $X \cap Y \neq \emptyset$, otherwise there exist an odd number of X-sectors of H containing an odd number of elements of Y, and (i) holds.

Since Z has an odd number of elements in H, there exist $q_1, q_2 \in X \cap Y$ such that either q_1, q_2 are distinct endnodes of an $X \cap Y$-sector Q of H containing an odd number of elements of Z, or $|X \cap Y| = 1$ and $q_1 = q_2$ (in which case we denote $Q = H$). By assumption, $q_1, q_2 \notin Z$.

Let $S_1, \ldots S_k$ be the $X \cup Y$-sectors of Q with one endnode in $X \setminus Y$ and the other in $Y \setminus X$ (if any), in the order they appear traversing Q from q_1 to q_2, and let $S_i = s_i \ldots, t_i$ for $1 \leq i \leq k$ (notice that if $q_1 = q_2$ one such sector exists). For ease of notation, define $S_0 = s_0, \ldots t_0$, where $s_0 = t_0 = q_1$ and $S_{k+1} = s_{k+1}, \ldots t_{k+1}$, where $s_{k+1} = t_{k+1} = q_2$. We may assume that Z has an even number of elements in S_i, $0 \leq i \leq k+1$, otherwise (ii) holds if $1 \leq i \leq k$, while it is trivially true if $i = 0$ or $i = k+1$.

If $Q(s_i, t_{i+1})$ contains an odd number of elements of Z, $0 \leq i \leq k$, we may assume it has exactly one element z in Z, and $z = s_i$ or $z = t_{i+1}$, otherwise (i) holds (since $Q(s_i, t_{i+1})$ is either an X-sector or a Y-sector of Q). By symmetry, we may assume $z = s_i$, thus z is the unique node of Z in S_i, a contradiction. Thus, for $0 \leq i \leq k$, $Q(s_i, t_{i+1})$ has an even number of elements in Z. By the inclusion-exclusion principle

$$|V(Q) \cap Z| = |\bigcup_{0}^{k} (V(Q(s_i, t_{i+1})) \cap Z)| = \sum_{i=0}^{k} |V(Q(s_i, t_{i+1})) \cap Z| - \sum_{i=1}^{k} |V(S_i) \cap Z|$$

therefore Z has an even number of elements in Q, a contradiction. $\square$

Lemma 4.3 *Let (H, x) be a smallest odd wheel in G and y be a major node for H nonadjacent to x. If for some choice of $u, v \in \{x, y\}$, v has an odd number of neighbors in some $N(u)$-sector S of H, then v has exactly one neighbor in S. Furthermore, if u, v are in the same side of the bipartition, then v is adjacent to one endnode of S.*

Proof: If v has at least 3 neighbors in $S = s_1, \ldots, s_h$, then $V(S) \cup \{u, v\}$ induces an odd wheel (H', v), and (H', v) has less nodes than (H, x) since x and y are both major nodes. Therefore v has exactly one neighbor s_i, $1 \leq i \leq h$, in S. Assume that u and v are in the same side of the bipartition. If $i = 1$ or $i = h$ we are done, hence we may assume $3 \leq i \leq h-2$. Suppose u and v both have neighbors in $V(H) \setminus V(S)$. Then there exists a path P between u and v with interior in $V(H) \setminus V(S)$. But then $P_1 = s_i, v, P, u$, $P_2 = s_i, S(s_i, s_h), s_h, u$, $P_3 = s_i, S(s_1, s_i), s_1, u$ induce a 3-path configuration between s_i and x. Since v is major, v has at least one neighbor in $V(H) \setminus V(S)$, therefore u has exactly two neighbors in H, so $u = y$ and $v = x$. Let x', x'' be the neighbors of x closest to s_1 and s_h, respectively, in the path Q induced by $V(H) \setminus \{s_2, \ldots, s_h\}$. Let P' and P'' be the unique paths in Q between s_1 and x', and s_h and x'', respectively. Then $P_1 = x, x', P', s_1$, $P_2 = x, s_i, Q(s_i, s_1), s_1$ and $P_3 = x, x'', P'', s_h, y, s_1$ induce a 3-path configuration between x and s_1, a contradiction. □

Lemma 4.4 *Let (H, x) be a smallest odd wheel in G. If y is a major node for H, then y has an odd number of neighbors in H.*

Proof: Suppose, by contradiction, that y has an even number of neighbors in H. Let $X = N_H(x)$ and $Y = N_H(y)$.

Case 1: x and y are in the same side of the bipartition.

Clearly, X and Y satisfy the hypothesis of Lemma 4.1, thus there exists $u, v \in \{x, y\}$ such that v has an odd number of neighbors in some $N(u)$-sector S of H, and v has a neighbor in the interior of S, contradicting Lemma 4.3.

Case 2: x and y are in distinct sides of the bipartition.

Assume x and y are adjacent. One can easily verify that there exists $u, v \in \{x, y\}$, $u \neq v$, such that u has a positive even number of neighbors in some $N_H(v)$-sector $S = s_1, \ldots, s_k$ of H. Thus, given $H' = v, s_1, S, s_k, v$, (H', u) is an odd wheel, and $|H'| < |H|$ since x and y are major, a contradiction.

Henceforth we may assume that x and y are nonadjacent. Since x has an odd number of neighbors in H, then there exists a Y-sector $S = s_1, ..., s_k$ of H containing an odd number of neighbors of x. By Lemma 4.3, x has exactly one neighbor, say s_i, in S. Let $z', z'' \in V(H) \setminus V(S)$ be the nodes in $X \cup Y$ that are closer to s_1 and s_k, respectively, in the path Q induced by $V(H) \setminus \{s_2, \ldots, s_{k-1}\}$. Let P' and P'' be the unique paths in Q between z' and s_1, and z'' and s_k, respectively.

(4.4.1) *At least one of z' and z'' is adjacent to y.*

Suppose not. Then z' and z'' are adjacent to x. If $i \geq 3$ or $i \leq k-2$, say $i \leq k-2$, then there is a 3-path configuration induced by the paths $P_1 = s_i, S(s_i, s_k), s_k$,

$P_2 = s_i, x, z'', P'', s_k$, $P_3 = s_i, S(s_i, s_1), s_1, y, s_k$. So $S = s_1, s_2, s_3$ and $i = 2$. Let $H' = x, z', P', s_1, y, s_3, P'', z'', x$; (H', s_i) is an odd wheel with at most as many nodes as (H, x). Thus $|H'| = |H|$, since (H, x) is a smallest odd wheel, and z', z'' have a common neighbor in Q. Since y has an even number of neighbors in H, then s_1, s_3 are the only neighbors of y in H, a contradiction since y is major for H. This concludes the proof of (4.4.1).

Thus we may assume, w.l.o.g., that z' is adjacent to y. Let S' be the X-sector containing s_1 and z', and let x' be the endnode of S' distinct from s_i. Since y has at least two neighbors in S', then by Lemma 4.3 y must have an even number of neighbors in S'. Therefore, since x has an odd number of neighbors in H and y as an even number of neighbors in H, both x and y have neighbors in $V(H) \setminus (V(S') \cup V(S))$, so there exists a path P between x and y with interior in $V(H) \setminus (V(S) \cup V(S'))$. Let y' be the neighbor of y closest to x' in S'. Consider the paths $P_1 = x, P, y$, $P_2 = x, s_i, S(s_i, s_1), s_1, y$, $P_3 = x, x', S'(x', y'), y', y$. P_1, P_2 and P_3 induce a 3-path configuration unless the neighbor y'' of y in P is adjacent to x'. Therefore y'' is the only neighbor of y in $V(H) \setminus (V(S) \cup V(S'))$, so s_k and y'' are the endnodes of a Y-sector S'' of H containing an odd number of neighbors of x. Thus S'' contains exactly one neighbor of x. Now s_i and x' are the nodes of $(V(H) \setminus V(S'')) \cap (X \cup Y)$ closest to s_k and y'', respectively, in the subpath induced by $V(H)$ minus the interior of S''; but s_i and x' are both adjacent to x, contradicting (4.4.1). □

By Lemma 4.4, if (H, X) is a smallest odd wheel, then (H, y) is a smallest odd wheel for every major node y for H.

Lemma 4.5 *Let (H, x) be a smallest odd wheel in G. There exist $a \in V(H) \cap R$ and $b \in V(H) \cap C$ such that $N(a) \supset M(H) \cap C$ and $N(b) \supset M(H) \cap R$.*

Proof: The statement is obvious if $|H| = 6$, hence we may assume $|H| \geq 8$. By symmetry, we only need to prove the statement for $M(H) \cap C$. We will proceed by induction on $|M(H) \cap C|$.

(4.5.1) *Lemma 4.5 holds if $|M(H) \cap C| = 2$.*

Let $\{x, y\} = M(H) \cap C$. By Lemma 4.4, x has an odd number number of elements in H, thus there exists an $N_H(y)$-sector of H where x has an odd number of neighbors, so by Lemma 4.3 this sector contains a common neighbor of x and y. This concludes the proof of (4.5.1).

Assume $|M(H) \cap C| = 3$ and let $\{x, y, z\} = M(H) \cap C$. By contradiction, suppose that there is no node in $N_H(x) \cap N_H(y) \cap N_H(z)$. By Lemma 4.2, there exist $u, v, w \in \{x, y, z\}$ such that either there exists an $N(u)$-sector of H containing an odd number of neighbors of w that contains one neighbor of w in the interior, contradicting Lemma 4.3, or there exists an $N(u) \cup N(v)$-sector $S = s_1, \ldots, s_h$ of H such that s_1 is adjacent to u and not v, s_h is adjacent to v and not u, and S contains an odd number of neighbors of w. By (4.5.1), there exists a node $t \in V(H)$ adjacent to both u and v, therefore $t \notin V(S)$ and $H' = t, u, s_1, S, s_h, v, t$ is a hole of length smaller than H. Since w is not adjacent to

t, then w has an odd number of neighbors in H', so w must have exactly one neighbor in H', say s_i, $1 \leq i \leq h$. We may assume, w.l.o.g., that $i > 1$. By (4.5.1), there exists a node $r \in V(H)$ adjacent to both u and w, therefore $r \notin V(S)$ and $r \neq t$. The paths $P_1 = u, s_1, S(s_1, s_i), s_i$, $P_2 = u, r, w, s_i$, $P_3 = u, t, v, s_h, S(s_h, s_i), s_i$ induce a 3-path configuration, a contradiction.

Henceforth we may assume $|M(H) \cap C| \geq 4$. Let $x_1, x_2, x_3, x_4 \in M(H) \cap C$. By induction, there exist nodes $s_1, s_2, s_3 \in V(H)$ such that s_i is adjacent to every node in $M(H) \cap C$ except x_i, $i = 1, 2, 3$. Thus $H' = x_1, s_2, x_3, s_1, x_2, s_3, x_1$ is a hole of length 6 and (H', x_4) is an odd wheel (since x_4 is adjacent to s_1, s_2, s_3), a contradiction. □

Lemma 4.6 *There exists a $O(|V(G)|^7)$ algorithm with the following specifications:*

- **Input** *A bipartite graph G containing no 3-path configurations.*
- **Output** *A family $\mathcal{C}$ of $O(|V(G)|^6)$ subsets of $V(G)$ such that, if (H, x) is a smallest odd wheel in G, then there exists a member of $\mathcal{C}$ that is a cleaner for (H, x).*

Algorithm:

For every 6-tuple of nodes $u_1, \ldots, u_6$, such that u_1, u_2, u_3 and u_4, u_5, u_6 induce paths, and $u_2 \in C$, $u_5 \in R$, compute

$$X(u_1, \ldots, u_6) = (N(u_2) \cup N(u_5)) \setminus \{u_1, u_2, u_3, u_4, u_5, u_6\}.$$

Let $\mathcal{C}$ be the family containing $X(u_1, \ldots, u_6)$ for every possible choice of $u_1, \ldots, u_6$.

Correctness: The running time of the algorithm is obviously $O(|V(G)|^7)$ and $\mathcal{C}$ has $O(|V(G)|^6)$ members. We only need to show that, if G contains a smallest odd wheel (H, x), then $\mathcal{C}$ contains a cleaner for H. By Lemma 4.5, there exists two nodes $u_2 \in V(H) \cap C$, $u_5 \in V(H) \cap R$, such that every node in $M(H) \cap R$ is adjacent to u_2 and every node in $M(H) \cap C$ is adjacent to u_5. Let u_1, u_3 be the neighbors of u_2 in H and u_4, u_6 be the neighbors of u_5 in H. Then the algorithm will examine the 6-uple $u_1, \ldots, u_6$, and clearly $X(u_1, \ldots, u_6)$ is a cleaner for (H, x). □

5 Detecting a smallest odd wheel

Given two smallest odd wheels (H, x) and (H', y), we say that (H, x) *dominates* (H', y) (or (H', y) *is dominated by* (H, x)) if $M(H') \subseteq M(H)$.

Lemma 5.1 *Let G be a bipartite graph containing no 3-path configuration and no detectable 3-wheel. Let (H, x) be a smallest odd wheel of G, u and v be two nonadjacent nodes of H and P_1, P_2 be the two internally node-disjoint subpaths of H between u and v, where $|P_1| \leq |P_2|$. Let P be a shortest path between u and v in $G' = G \setminus M(H)$. Then the following hold:*

(i) $|P| = |P_1|$ *(i.e. $d_{G'}(u, v) = d_H(u, v)$)*

(*ii*) *Either* $H' = u, P, v, P_2, u$ *is a hole, and* (H', x) *is a smallest odd wheel dominated by* (H, x)*; or* $|P_1| = |P_2|$, $H'' = u, P, v, P_1, u$ *is a hole, and* (H'', x) *is a smallest odd wheel dominated by* (H, x).

Proof: We will prove Lemma 5.1 by induction on $d_{G'}(u, v)$. Let $H = h_1, \ldots, h_{2s}, (h_{2s+1} = h_1)$ where $h_1 = u$, $s \geq 3$. Let $\overrightarrow{H}$ be the directed cycle obtained by orienting the edges of H from h_i to h_{i+1} for every $1 \leq i \leq 2s$. For any two distinct nodes a and b in H, let $H(a, b)$ be the underlying graph of the directed path from a to b in $\overrightarrow{H}$. W.l.o.g., $P_1 = H(u, v)$ and $P_2 = H(v, u)$, and $v = h_m$ for some $3 \leq m \leq s+1$. Let $P = p_0, \ldots, p_{k+1}$, where $p_0 = u$ and $p_{k+1} = v$. Clearly $d_{G'}(u, v) = k + 1$.

If $d_{G'}(u, v) = 2$, then $k = 1$ and p_1 is adjacent to u and v. Since p_1 is not major for H, then u, v are the only neighbors of p_1 in H and they are contained in a subpath of H of length 2, say u, w, v. Hence $H' = u, p_1, v, P_2, u$ is a hole of the same length as H. Suppose (H', x) is not an odd wheel, then x is adjacent to exactly one node in $\{p_1, w\}$. Let $y, z \in \{p_1, w\}$, such that x is adjacent to y and nonadjacent to z, and let u', v' be the neighbors of x closest to u and v in P_2, respectively. Then $C = u, z, v, H(v, v'), v', x, u', H(u', u), u$ is a hole and (C, y) is a detectable 3-wheel, a contradiction. Finally we need to prove that (H', x) is dominated by (H, x). Assume not and let $y \neq x$ be a major node for H' that is not major for H. Since y is not major for H, then y is adjacent to p_1 but not to w, y has exactly 2 neighbors in P_2 and they are contained in a subpath of H of length 2. But then (H', y) is a detectable 3-wheel, a contradiction.

Hence we may assume $d_{G'}(u, v) \geq 3$.

(5.1.1) *Either:*

(*i*) $H' = u, P, v, P_2, u$ *is a hole; or*

(*ii*) $|P_1| = |P_2|$ *and* $H'' = u, P, v, P_1, u$ *is a hole.*

If (*i*) does not hold, there exists a node of $P(p_1, p_k)$ that belongs to or has a neighbor in $H(h_{m+1}, h_{2s})$, thus there exists j, $m+1 \leq j \leq 2s$, such that there are chordless paths Q_1 and Q_2 between h_j and u and h_j and v, respectively, with interior contained in the interior of P.

Therefore

$$\begin{aligned} |Q_1| + |Q_2| &\leq k + 3 \leq m + 1 \leq 2s + 3 - m = (2s + 2 - j) + (j - m + 1) \\ &\leq (|H(h_j, u)| + 1) + (|H(v, h_j)| + 1). \end{aligned} \tag{2}$$

First we show that either $|Q_1| \leq |H(h_j, u)|$, $|Q_1| < |P|$, and $j < 2s$, or $|Q_2| \leq |H(v, h_j)|$, $|Q_2| < |P|$, and $j > m + 1$. Since $|Q_1|$ has the same parity as $|H(h_j, u)|$ and $|Q_2|$ has the same parity as $|H(v, h_j)|$, then, by (2), either $|Q_1| \leq |H(h_j, u)|$ or $|Q_2| \leq |H(v, h_j)|$. Clearly $|Q_1|, |Q_2| \leq |P|$. Suppose $|Q_1| \leq |H(h_j, u)|$ and $|Q_1| = |P|$, then $Q_1 = u, P(u, p_k), p_k, h_j$ and $Q_2 = h_j, p_k, v$, hence $|Q_2| \leq |H(v, h_j)|$ and $|Q_2| < |P|$. Analogously, if $|Q_2| \leq |H(v, h_j)|$ and $|Q_2| = |P|$, then $|Q_1| \leq |H(h_j, u)|$ and $|Q_1| < |P|$.

Thus either $|Q_1| \leq |H(h_j,u)|$ and $|Q_1| < |P|$, or $|Q_2| \leq |H(v,h_j)|$ and $|Q_2| < |P|$. If $|Q_1| \leq |H(h_j,u)|$, $|Q_1| < |P|$, and $j = 2s$, then $|Q_2| < |P| \leq |H(v,h_j)| + 1$ (since p_1 cannot be adjacent to both h_1 and h_{2s}), and $j > m+1$. Analogously, if $|Q_2| \leq |H(v,h_j)|$, $|Q_2| < |P|$, and $j = m+1$, then $|Q_1| \leq |H(h_j,u)|$, $|Q_1| < |P|$, and $j < 2s$.

By symmetry, we may assume $|Q_1| \leq |H(h_j,u)|$, $|Q_1| < |P|$ and $j < 2s$.

Since $V(Q_1) \subseteq G'$, then $d_{G'}(h_j,u) \leq |Q_1| < |P| = d_{G'}(u,v)$, thus, by inductive hypothesis,

$$d_{G'}(h_j,u) = d_H(h_j,u) = \min(2s+1-j, j-1).$$

Since $d_{G'}(h_j,u) \leq |Q_1| < |P| \leq m-1 < j-1$, then $d_{G'}(h_j,u) = 2s+1-j = |Q_1| < s$. Therefore Q_1 is a shortest path between u and h_j in G', thus, by induction, $C = u, H(u,h_j), h_j, Q_1, u$ is a hole and (C,x) is a smallest odd wheel dominated by (H,x). We obtain a directed cycle $\overrightarrow{C}$ by orienting the edges of C to agree with the orientation of the edges in $H(u,h_j)$, and define $C(a,b)$ for every a,b in C as before.
Let u' be the neighbor of h_j in Q_1. Then $P' = P(u',v)$ is a path between u' and v of length $k+2-|Q_1| = k+j+1-2s < k+1$, and P' does not contain a node that is major for C, since it does not contain a node that is major for H and (H,x) dominates (C,x). Since $|P'| < |P|$ then, by induction, either $|C(u',v)| \leq |P'|$ or $|C(v,u')| \leq |P'|$. But $|C(u',v)| > |H(u,v)| \geq k+1 > |P'|$, hence $|C(v,u')| \leq k+j+1-2s$. This implies $j-m+1 \leq k+j+1-2s$, so $2s \leq k+m$, but $m \leq s+1$ and $k \leq s-1$, hence $m = s+1$, $k = s-1$, $d_{G'}(u',v) = d_C(u',v) = |P'|$. By induction, $C' = u', C(u',v), v, P', u'$ is a hole, and (C',x) is a smallest odd wheel. Clearly, $C' = H''$, where $H'' = u, P, v, P_1, u$. This concludes the proof of (5.1.1).

By (5.1.1) and symmetry, we may assume that no node of $P(p_1,p_k)$ belongs to or has a neighbor in $H(h_{m+1},h_{2s})$.

(5.1.2) *Either*

- *(i)* $H' = u, P, v, P_2, u$ *is a hole, and* (H',x) *is a smallest odd wheel; or*
- *(ii)* $|P_1| = |P_2|$, $H'' = u, P, v, P_1, u$ *is a hole, and* (H'',x) *is a smallest odd wheel.*

Case 1: no node in $P(p_1,p_k)$ belongs to or has a neighbor in $H(h_2,h_{m-1})$.

Then u and v must be on the same side of the bipartition, else $H(u,v)$, $H(v,u)$, and P would induce a 3-path configuration between u and v. If (H'',x) is an odd wheel, then it is a smallest one, since $|H''| \leq |H|$, and case *(ii)* occurs. Thus, for $i = 1,2$, $V(P_i) \cup V(P)$ contains either exactly one neighbor of x, or an even number of neighbors of x. We may assume that x is not adjacent to both u and v, otherwise the number of neighbors of x in $H(u,v)$ and in $H(v,u)$ have distinct parities, so either (H',x) or (H'',x) is an odd wheel. Thus we may assume x and v are nonadjacent. If x is adjacent to u, we may assume that either $H(u,v) \setminus u$ and $H(v,u) \setminus u$ both contain neighbors of x, or $P \setminus u$ contains a neighbor of x, otherwise all neighbors of x in H are contained in $H(u,v)$ or $H(v,u)$, and no neighbor of x is contained in $P \setminus u$, but then either (H',x)

or (H'', x) is a smallest odd wheel. Thus, if u_1, u_2, u_3 are the neighbors of x closest to v in $H(u, v)$, $H(v, u)$ and P, respectively, then u_1, u_2, u_3 are pairwise distinct, and the paths $Q_1 = x, u_1, H(u_1, v), v$, $Q_2 = x, u_2, H(v, u_2), v$, $Q_3 = x, u_3, P(u_3, v), v$ induce a 3-path configuration (since x and v are in distinct sides of the bipartition, because x and u are adjacent). Thus we may assume that both u and v are nonadjacent to x. This implies that the number of neighbors of x in $H(u, v)$ and in $H(v, u)$ have distinct parities, so x has an odd number of neighbors on the hole C, where $C = H'$ or $C = H''$. This implies that x has exactly one neighbor, say x' in C, while x has at least 2 neighbors in the chordless path P' between u and v contained in H whose interior is disjoint from C. Let u' and v' be the neighbors of x in P' closest to u and v, respectively. Clearly u, v, u', v' are pairwise distinct. Let Q and Q' be the two distinct subpaths of C between x' and u such that v is in Q'. If u and x are in distinct sides of the bipartition, then the paths $Q_1 = u, Q, x', x$, $Q_2 = u, P(u, u'), u', x$ and $Q_3 = u, Q'(u, v), v, P(v, v'), v', x$ form a 3-path configuration, a contradiction. Thus x' and u are in distinct sides of the bipartition. Since $|H(u, v)|, |H(v, u)|, |P| \geq 3$, then x' is not adjacent to both u and v, say, w.l.o.g., x' is nonadjacent to u. The paths Q, Q', $Q'' = u, P(u, u'), x, x'$ induce a 3-path configuration.

Case 2: Some node in $P(p_1, p_k)$ belongs to or has a neighbor in $H(h_2, h_{m-1})$.

Then there exists j, $2 \leq j \leq m - 1$, such that there are chordless paths Q_1 and Q_2 between h_j and u and h_j and v, respectively, with interior contained in the interior of P. We have

$$|Q_1| + |Q_2| \leq k + 3 \leq m + 1 = j + (m + 1 - j) = (|H(u, h_j)| + 1) + (|H(h_j, v)| + 1)$$

and, by an argument similar to the one used in the proof of (5.1.1), we may assume $|Q_1| \leq j - 1$, $|Q_1| < |P|$ and $j > 2$. By induction, $|Q_1| = d_{G'}(u, h_j) = d_H(u, h_j) = j - 1$, $C = u, Q_1, h_j, H(h_j, u), u$ is a hole and (C, x) is a smallest odd wheel dominated by (H, x). We obtain a directed cycle $\overrightarrow{C}$ by orienting the edges of C to agree with the orientation of the edges in $H(u, h_j)$, and define $C(a, b)$ for every a, b in C as usual.
Let u' be the neighbor of h_j in Q_1 and let P' be the path between u' and v in P. Then

$$|P'| = k + 2 - |Q_1| = k - j + 3 \leq |C(u', v)|$$

thus, by induction, $C' = u', P', v, C(v, u'), u'$ is a hole and (C', x) is a smallest odd wheel. Clearly, $C' = H'$. This concludes the proof of (5.1.2).

By (5.1.2) and by symmetry, we may assume that (H', x) is a smallest odd wheel. To conclude the proof of Lemma 5.1 we only need to show that (H', x) is dominated by (H, x). Suppose there exists a major node y for H' that is not major for H. Then the neighbors of y in H are contained in a subpath of H of length 2. Also, the neighbors of y in P are contained in a subpath of P of length 2, otherwise let i, j, $0 \leq i < j \leq k + 1$ be the minimum and maximum index, respectively, such that p_i and p_j are adjacent to y; then $P' = u, P(u, p_i), p_i, y, p_j, P(p_j, v), v$ is a path in G' strictly shorter then P, a contradiction. Therefore y has at most 3 neighbors in H', and two of them are contained in a subpath of H' of length 2. Thus (H', y) is a detectable 3-wheel, a contradiction. $\square$

Lemma 5.2 *There exists a $O(|V(G)|^5)$ algorithm with the following specifications:*

- **Input** *A bipartite graph G containing no 3-path configuration and no detectable 3-wheel, and a set X of vertices of G.*
- **Output** *Either*
  - *(i) An odd wheel (H, x),*
  - *(ii) Determines that either G is balanceable or X is not a cleaner for G.*

Algorithm:
Let $G' = G \setminus X$.
For every possible triple of nodes of G', u_1, u_2, u_3, do the following

1. Compute the shortest path $P(u_i, u_j)$ between u_i and u_j in G' for every $1 \leq i < j \leq 3$.
2. If, for each $1 \leq i < j \leq 3$, $1 \leq h < k \leq 3$, $(i, j) \neq (h, k)$, the interior of $P(u_i, u_j)$ and the interior of $P(u_h, u_k)$ are disjoint and have no edges between them, then let $H(u_1, u_2, u_3) = u_1, P(u_1, u_2), u_2, P(u_2, u_3), u_3, P(u_1, u_3), u_1$.
3. For each $x \in X$, check if $(H(u_1, u_2, u_3), x)$ is an odd wheel. If it is, output $(H(u_1, u_2, u_3), x)$ and stop.

Otherwise output that either G is balanceable or X is not a cleaner for G.

Correctness: for each triple u_1, u_2, u_3 the algorithm performs steps 1-3, so these steps are performed $O(|V(G)|^3)$ times. Step 1 consists of 3 shortest path computations, thus it is performed in time $O(|V(G)|^2)$, and steps 2 and 3 are both performed in time $O(|V(G)|^2)$. Therefore the total running time is $O(|V(G)|^5)$ as claimed.

Obviously, when the algorithm outputs an odd wheel, it is correct. We need to verify that the algorithm is always correct when it outputs that either G is balanceable or X is not a cleaner for G. Assume G is not balanceable and X is a cleaner for some smallest odd wheel (H, x). Let u_1, u_2, u_3 be three nodes in H with the property that the three subpaths Q_{ij} of H between u_i and u_j, $1 \leq i < j \leq 3$, such that Q_{ij} does not contain u_k, $k \neq i, j$, have the property that $|Q_{ij}| < |H|/2$ for each $1 \leq i < j \leq 3$. Eventually, the algorithm will check the triple u_1, u_2, u_3, and compute the paths $P(u_i, u_j)$, $1 \leq i < j \leq 3$. Since G does not contain a 3-path configuration or a detectable 3-wheel, then, by Lemma 5.1, $H' = u_1, P(u_1, u_2), u_2, Q_{23}, u_3, Q_{13}, u_1$ is a hole, (H', x) is a smallest odd wheel, and X is a cleaner for (H', x). Repeating the argument, one argues that $H(u_1, u_2, u_3) = u_1, P(u_1, u_2), u_2, P(u_2, u_3), u_3, P(u_1, u_3), u_1$ is a hole and $(H(u_1, u_2, u_3), x)$ is a smallest odd wheel. Since $x \in X$, the algorithm will detect it. $\square$

6 The recognition algorithm

At this point we are ready to describe an algorithm for checking if a bipartite graph is balanceable. One can apply algorithm of Lemma 2.2 to G. If G contains a 3-path configuration, then output the fact that G is not balanceable. At this point, apply algorithm of Lemma 3.2. If G contains a detectable 3-wheel, then output the fact that G is not balanceable. Now, since G does not contain any 3-path configuration, we can apply algorithm of Lemma 4.6 to determine a family $\mathcal{C}$ of subsets of $V(G)$. For each $X \in \mathcal{C}$ apply algorithm of Lemma 5.2 to G and X. If G contains an odd wheel, then for some choice of X, X will be a cleaner for some smallest odd wheel, thus the algorithm of Lemma 5.2 will output an odd wheel, so we will conclude that G is not balanceable. If G is balanceable, then for each $X \in \mathcal{C}$ algorithm of Lemma 5.2 will output correctly that G is balanceable (since every $X \in \mathcal{C}$ is a cleaner by definition), thus we will conclude that G is balanceable. The running time of this algorithm is $O(|V(G)|^{11})$, but one can reduce the time complexity to $O(|V(G)|^9)$ by "mixing" the algorithms of Lemma 4.6 and Lemma 5.2 as described in the next statement.

Theorem 6.1 *There exists a $O(|V(G)|^9)$ algorithm with the following specifications:*

- **Input** *A bipartite graph G.*
- **Output** *Determines whether G is balanceable or not.*

Algorithm:

1. Apply the algorithm of Lemma 2.2. If G contains a 3-path configuration, then output the fact that G is not balanceable and stop.

2. Apply the algorithm of Lemma 3.2. If G contains a detectable 3-wheel, then output the fact that G is not balanceable and stop.

3. For every 7-tuple of nodes $u_1, \ldots, u_7$, such that u_1, u_2, u_3 and u_4, u_5, u_6 induce a path, $u_2 \in C$, $u_5 \in R$, u_7 is nonadjacent to u_2, do the following

   (a) Compute $X(u_1, \ldots, u_6) = (N(u_2) \cup N(u_5)) \setminus \{u_1, u_2, u_3, u_4, u_5, u_6\}$.

   (b) Compute the shortest paths $P_1(u_1, \ldots, u_7)$ and $P_2(u_1, \ldots, u_7)$ between u_1 and u_7 and u_3 and u_7 in $G \setminus X(u_1, \ldots, u_6)$, respectively.

   (c) If no node in the interior of $P_1(u_1, \ldots, u_7)$ belongs to or has a neighbor in $P_2(u_1, \ldots, u_7)$, define:

   $$H(u_1, \ldots, u_7) = u_1, P_1(u_1, \ldots, u_7), u_7, P_2(u_1, \ldots, u_7), u_3, u_2, u_1;$$

   (d) For each $x \in X(u_1, \ldots, u_6)$ check if $(H(u_1, \ldots, u_7), x)$ is an odd wheel. If it is, output that G is not balanceable and stop.

   Otherwise output that G is balanceable.

Correctness: Both step 1 and step 2 take time $O(|V(G)|^9)$. Step 3 performs computations (a)-(d) at most $|V(G)|^7$ times. Steps (a) can be performed in time $O(|V(G)|)$, while steps (b)-(d) can be performed in time $O(|V(G)|^2)$, thus the running time is $O(|V(G)|^9)$ as claimed.

We need to show that the algorithm is correct. If G contains a 3-path configuration or a detectable 3-wheel, then by Lemmas 2.2 and 3.2 the algorithm will output correctly that G is not balanceable. We only need to prove that, if G does not contain a 3-path configuration or a detectable 3-wheel, but G contains an odd wheel, then step 3 will output that G is not balanceable. Let (H, x) be a smallest odd wheel in G. Then by Lemma 4.5 there exist two subpaths u_1, u_2, u_3 and u_4, u_5, u_6 of H such that every major node for H is adjacent to u_2 or u_5. The set $X(u_1, ..., u_6)$ computed in step (a) is a cleaner for H, as shown in the proof of Lemma 4.6. Let u_7 be the node at distance $|H|/2$ from u_2 in H. Clearly, the paths Q_1 and Q_2 between u_1 and u_7 and between u_3 and u_7 in H, respectively, have length strictly less then $|H|/2$, thus, by an argument similar to the one in the proof of Lemma 5.2, $H(u_1, \ldots, u_7) = u_1, P_1(u_1, \ldots, u_7), u_7, P_2(u_1, \ldots, u_7), u_3, u_2, u_1$ is a hole and $(H(u_1, \ldots, u_7), x)$ is a smallest odd wheel, where $P_1(u_1, \ldots, u_7)$ and $P_2(u_1, \ldots, u_7)$ are the paths computed in step (b). Since $x \in X(u_1, \ldots, u_6)$, then step (d) will output that G is not balanceable. $\Box$

References

[1] C. Berge, Sur certains hypergraphes généralisant les graphes bipartites, in *Combinatorial Theory and its Applications I* (P. Erdös, A. Rényi and V. Sós, eds.), *Colloquia Mathematica Societatis János Bolyai*, Vol. 4, North Holland, Amsterdam (1970), 119-133.

[2] C. Berge, Balanced Matrices, *Mathematical Programming* **2** (1972), 19-31.

[3] P. Camion, Characterization of totally unimodular matrices, *Proceedings of the American Mathematical Society* **16** (1965), 1068-1073.

[4] M. Chudnovsky, P. Seymour, Recognizing Berge Graphs, manuscript, 2002.

[5] M. Conforti, G. Cornuéjols, Balanced $0, \pm 1$ matrices, Bicoloring and Total Dual Integrality, *Mathematical Programming* **71** (1995), 249-258.

[6] M. Conforti, G. Cornuéjols, A. Kapoor, K. Vušković, Balanced $0, \pm 1$ Matrices I. Decomposition, *Journal of Combinatorial Theory B* **81** (2001), 243-274.

[7] M. Conforti, G. Cornuéjols, A. Kapoor, K. Vušković, Balanced $0, \pm 1$ Matrices II. Recognition Algorithm, *Journal of Combinatorial Theory B* **81** (2001), 275-306.

[8] M. Conforti, G. Cornuéjols, M. R. Rao, Decomposition of Balanced Matrices, *Journal of Combinatorial Theory B* **77** (1999), 292-406.

[9] M. Conforti, A.M.H. Gerards, A. Kapoor, A Theorem of Truemper, *Combinatorica* **20** (2000), 15-26.

[10] M. Conforti, M. R. Rao, Properties of Balanced and Perfect Matrices, *Mathematical Programming* **55** (1992), 35-49.

[11] M. Conforti, M. R. Rao, Structural Properties and Decomposition of Linear Balanced Matrices, *Mathematical Programming* **55** (1992), 129-169.

[12] M. Conforti, M. R. Rao, Testing Balancedness and Perfection of Linear Matrices, *Mathematical Programming* **61** (1993), 1-18.

[13] D. R. Fulkerson, A. Hoffman, R. Oppenheim, On Balanced Matrices, *Mathematical Programming Study* **1** (1974), 120-132.

[14] K. Truemper, Alpha-balanced Graphs and Matrices and $GF(3)$-representability of Matroids, *Journal of Combiantorial Theory B* **32** (1982), 112-139.

[15] D. West, *Introduction to Graph Theory*, Prentice Hall (1996).

[16] G. Zambelli, A Polynomial Recognition Algorithm for Balanced Matrices, preprint, May 2003.