A Polynomial Recognition Algorithm for Balanced Matrices

Giacomo Zambelli*

May, 2003
Revised February 2005

Abstract

A $0, \pm 1$ matrix is balanced if it does not contain a square submatrix with two nonzero elements per row and column in which the sum of all entries is 2 modulo 4. Conforti, Cornuéjols and Rao [9], and Conforti, Cornuéjols, Kapoor and Vušković [6], provided a polynomial algorithm to test balancedness of a matrix. In this paper we present a simpler polynomial algorithm, based on techniques introduced by Chudnovsky and Seymour in [3] for Berge graphs.

1 Introduction

A $0, \pm 1$ matrix A is *balanced* if it does not contain a square submatrix with two nonzero elements 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 [16]. The first known polynomial time recognition algorithm for $0, 1$ balanced matrices is due to Conforti, Cornuéjols and Rao [9], followed by a similar algorithm for the general $0, \pm 1$ case due to Conforti, Cornuéjols, Kapoor, Vušković [5, 6]. Balanced matrices are of special interest in integer programming, as several polytopes arising in classical optimization problems have only integral vertices when the constraint matrix is balanced. Particularly important examples are the *generalized set packing, set covering* and *set partitioning* polytopes, which are defined, respectively, by the systems $\{x \in \mathbb{R}^n \,|\, Ax \leq \mathbf{1} - n(A), \mathbf{0} \leq x \leq \mathbf{1}\}$, $\{x \in \mathbb{R}^n \,|\, Ax \geq \mathbf{1} - n(A), \mathbf{0} \leq x \leq \mathbf{1}\}$ and $\{x \in \mathbb{R}^n \,|\, Ax = \mathbf{1} - n(A), \mathbf{0} \leq x \leq \mathbf{1}\}$, where A is an $m \times n$ $0, \pm 1$ matrix, and $n(A)$ is the m-dimensional vector whose ith entry is the number of -1s in the ith row of A. All these polytopes are integral when A is balanced, as proven by Berge [2] for the $0, 1$ case, and by Conforti and Cornuéjols [4] for $0, \pm 1$ matrices. More generally, Fulkerson,

*Department of Combinatorics and Optimization, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada, N2L 3G1.

Hoffman and Oppenheim [14], showed that the inequalities defining such polytopes form totally dual integral systems if A is a $0,1$ balanced matrix, and an analogous statement was proven by Conforti and Cornuéjols [4] for the $0,\pm 1$ case.

The recognition algorithms in [6] and [9] are both based on decomposition theorems characterizing the class of balanced matrices. This decomposition approach was used in the literature for other classes of matrices related to balanced matrices, such as, for example, Seymour's decomposition of regular matroids [15]. A novelty element in the recognition algorithm for balanced matrices, was the use of "weak" decompositions, such as multiple star-cutsets, which do not preserve balancedness. To overcome this problem, Conforti and Rao [12, 13] devised a pre-processing technique, called cleaning, which consists of generating polynomially many submatrices of the input matrix A, with the property that A is balanced if and only if all the matrices generated are balanced, and, if A is not balanced, then there exists an unbalanced matrix in this family for which the "natural" decomposition algorithm would give the correct answer. The same approach has been successfully applied to other problems such as deciding whether a graph contains a chordless cycle of even length [8] or recognizing perfect graphs [3]. Interestingly, Chudnovsky and Seymour observed in [3], in their version of the algorithm for recognizing perfect graphs, that there was no need to apply any decomposition, but cleaning by itself would allow to test perfection in a more direct way.

A natural question is whether techniques used by Chudnovsky and Seymour in [3] could also be applied to the problem of recognizing balanced matrices, to design algorithms that do not make use of any structural knowledge about such class. In this paper we answer to this question positively, by providing a self-contained exposition of a polynomial time recognition algorithm for balancedness. This is desirable for several reasons. The main reason is that proving the validity of the decomposition based algorithm requires to prove a structural theorem for the class of balanced matrices, which is extremely difficult, technical, and lengthy. From this point of view, the new algorithm is simpler and more transparent. Also, avoiding to use decomposition is beneficial for the running time, since the algorithm which will be presented in the remainder runs in time $O(n^9)$ (n being the number of rows and columns of the input matrix) for the $0,1$ case, which, although not implementable by any means, is significantly lower than the running time of the previously known algorithm.

1.1 Notations and definitions

It will be convenient, in the remainder of the paper, to work with the bipartite representation of a matrix. Given a $0,1$ matrix A, the *bipartite representation of* A is the bipartite graph G where the two sides of the bipartition are the sets R and C of rows and columns of A, respectively, and there is an edge between $i \in R$ and $j \in C$ if and only if $a_{ij} = 1$. Clearly, A is balanced if and only if its bipartite representation does not contain a *hole* of length 2 modulo 4 as an induced subgraph (a hole is a chordless cycle). A bipartite graph is *balanced* if it does not contain any hole of length 2 modulo 4. For general $0,\pm 1$, matrices, the most convenient setting to work with, is their signed bipartite representation. A *signed bipartite graph* is a pair (G, σ) where G is a bipartite graph and σ is a *signing* of the edges, that is a function from $E(G)$ to $\{1, -1\}$. Given a $0, \pm 1$ matrix A, the *signed bipartite representation of* A is the signed bipartite graph (G, σ) where G is the bipartite representation of the $0, 1$ matrix underlying A and σ is defined, for each edge ij, by $\sigma(ij) = a_{ij}$. For any subgraph F of G, we define

$$\sigma(F) = \sum_{e \in E(F)} \sigma(e).$$

It is immediate to verify that a $0, \pm 1$ matrix A is balanced if and only if its signed bipartite representation does not contain a hole H such that $\sigma(H) \equiv 2 \bmod 4$ as an induced subgraph. We will say that such a hole is *unbalanced*, and a signed bipartite graph is *balanced* if it contains no unbalanced hole. Observe that, given a cut $(S, \bar{S})$ of G (where S is a subset of the nodes of G), if we define a signing σ' by

$$\sigma'(ij) = \begin{cases} \sigma(ij) & \text{if } ij \notin (S, \bar{S}) \\ -\sigma(ij) & \text{if } ij \in (S, \bar{S}) \end{cases},$$

it is easy to see that (G, σ) is balanced if and only if (G, σ') is balanced (since, for any hole H, $\sigma(H) \equiv \sigma'(H) \bmod 4$). We call this operation *scaling* along the cut $(S, \bar{S})$.

In the remainder, G will always be a bipartite graph. Every time we say that a graph G *contains* a graph F, we will always mean that G contains a graph isomorphic to F as an induced subgraph. 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, that is the number of its edges. Given a graph F and two nodes x and 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 and 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 endpoints 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 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 outside H with at least 3 distinct neighbors in H, and is denoted by (H, v). A wheel (H, v) for which v has k neighbors in H is said to be a k-wheel. A *sector* of (H, v) is a maximal subpath of H with no neighbors of v in its interior. The *spokes* of (H, v) are the edges of G of the form uv where u is a neighbor of v in H. (H, v) is an *odd wheel* if it is a wheel and v has an odd number of neighbors in H.

The importance of 3-path configurations and odd wheels is explained by the next remark.

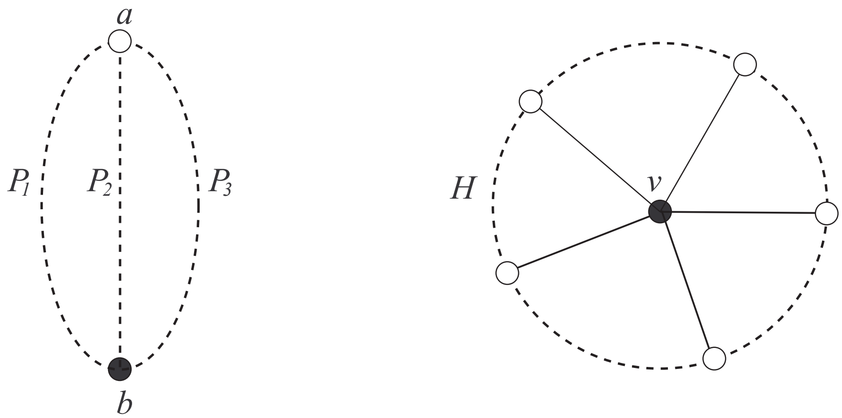

Figure 1: A 3-path configuration and a wheel.

1.1 *If G is a bipartite graph containing a 3-path configuration or an odd wheel, then (G,σ) is not balanced for any signing σ.*

Proof: Let σ be a signing of G. If P_1, P_2, P_3 form a 3-path configuration in G between nodes a and b of G, then, w.l.o.g., $\sigma(P_1) = \sigma(P_2)$. Since a and b are in distinct sides of the bipartition, then P_i has odd length for every $i \in [3]$, hence $\sigma(P_1) \equiv k \bmod 4$ for some $k \in \{1,3\}$, but then $H = a, P_1, b, P_2, a$ is an unbalanced hole of (G,σ).
Let(H,v) be a wheel. We may assume, by scaling along $(S,\bar{S})$ where $S \supseteq \{u \in N_H(v) \,|\, \sigma(uv) = -1\}$ and $\bar{S} \supseteq \{v\} \cup \{u \in N_H(v) \,|\, \sigma(uv) = 1\}$, that all spokes of (H,v) have sign 1. If $\sigma(H) \equiv 2 \bmod 4$ then we are done, so assume $\sigma(H) \equiv 0 \bmod 4$. Since (H,v) has an odd number of sectors and there is an even number of sectors with total sign 2 modulo 4, then there exists a sector Q, say with endpoints u and u', such that $\sigma(Q) \equiv 0 \bmod 4$, but then $H' = v, u, Q, u', v$ is an unbalanced hole. □

1.2 Outline of the paper

We will provide two algorithms: one to recognize balanced graphs, and one to solve the more general problem of recognizing balanced signed graphs. In section 2 we will provide an algorithm to recognize whether a bipartite graph has a 3-path configuration, while in section 3 an algorithm is presented, to recognize if a bipartite graph not containing any 3-path configuration, contains a detectable 3-wheel (which is a special type of odd wheel). In these two sections we will not be concerned with signing, since 1.1 guarantees that, whenever we find any of these configurations, the graph will not be balanced with respect to any signing. In section 4 we show how to produce, from a graph G, a family of polynomially many induced subgraphs of G, in which the recognition problem is "easy" (we will give a rigorous definition of cleaning in section 4, while in 5 it will be explained what "easy" means) with the property that G is balanced if and only if every member of the family is balanced. In particular, section 4 deals separately with cleaning in bipartite graphs and in signed bipartite graphs, since in the unsigned case we will be able to give a more efficient algorithm (i.e. the family of graphs it produces in the unsigned case is smaller, thus improving the overall running time in this case). For cleaning in the signed case, we will use an algorithm of Conforti, Cornuéjols, Kapoor and Vušković [6]. Finally, section 5 shows how to solve directly the recognition problem in a clean graph which does not contain any 3-path configuration or detectable 3-wheel. In section 6 we put all the pieces together by giving a formal description of the two algorithms for the signed and the unsigned case.

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.

2.1 *Let* Π *be a smallest 3-path configuration in* $G = (R, C; E)$. *Assume* Π *is formed by the paths* $P_i = a, a_i, \ldots, b_i, b$, $i \in [3]$, *where* $a \in R$, $b \in C$. *For every* $i \in [3]$, *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, $i \in [3]$, *and all possible pairs* a_i, m_i *and* m_i, b_i.

Proof: Let $P = p_1, \ldots, p_k$ 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 the shortest path P'_1 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 of h, 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. $\square$

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 $i \in [3]$,
- 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 $i \in [3]$ and such that m_i, m_j is (i, j)-good for every $1 \leq i < j \leq 3$. If such a triple exists, 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 $\tilde{P}_i = a, a_i, \ldots, b_i, b$. 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 $i \in [3]$ (since $x = a$ and $y = b$ would satisfy such condition). For $i = 1, 2, 3$, 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 $P_i = x, a_i, \ldots, b_i, y$. For every $i \in [3]$, 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 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. $\square$

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 possible detectable 3-wheels, we say that (H, v) is a *smallest* detectable 3-wheel.

3.1 *Let $G = (R, C; E)$ 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 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 $p_k = s$. 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 $|P| \leq |P_1| - 1$. 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 $|Q_i| \geq 3$ for $i \in [3]$ (since $x \neq u_2$, because $p_h \in X$ is nonadjacent to u_2), 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. $\square$

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 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 $G[X \cup \{v_1\}]$. If no such path exists, select a different 7-tuple.
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 select a different 7-tuple.
4. Compute the set Y of all nodes that do not belong to or have a neighbor in $P_1 \cup \{w, x\}$.
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, steps 2 and 4 take 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 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. □

4 Cleaning a smallest unbalanced hole

Given a signed graph (G, σ), an unbalanced hole is *smallest* if it has minimum length among all unbalanced holes. Let H be an unbalanced hole. We say that a vertex $x \in V(G) \setminus V(H)$ is major for H if $N_H(x)$ is not contained in a subpath of H of length 2. Let $M(H)$ be the set of major vertices for H. We say that H is *clean* if $M(H) = \emptyset$. A set $X \subseteq V(G) \setminus V(H)$ is a *cleaner* for H if $M(H) \subseteq X$ (i.e. if H is clean in $G \setminus X$). A signed bipartite graph (G, σ) is *clean* if G is either balanced or it contains a clean smallest unbalanced hole.

It should be noted here that this definition of a clean graph is slightly different from the one proposed by Conforti, Cornuéjols and Rao [9]. Their definition is more demanding, in that it requires for the property we described above to hold not just for one smallest unbalanced hole H, but for a family of smallest unbalanced holes that can be derived from H by some simple operation. However, this stronger property will hold in clean graphs (according to our definition) that do not contain 3-path configurations or detectable 3-wheels. Therefore, in light of sections 2 and 3, this more restrictive concept of cleaning will be sufficient for our purposes.

4.1 *Let H be a smallest unbalanced hole in (G, σ). If $x \in V(G) \setminus V(H)$ has a positive even number of neighbors in H, then there exists a subpath u, v, w of H of length 2 such that u and w are the only neighbors of x in H and $H \cup x \setminus v$ is an unbalanced hole. In particular, all major nodes for H have an odd number of neighbors in H.*

Proof: Possibly by scaling, we may assume that $\sigma(xu) = 1$ for every $u \in N_H(x)$. Since x has an even number of neighbors and H is unbalanced, then there exists a subpath Q of H such that $\sigma(Q) \equiv 0 \bmod 4$, containing no neighbor of x except the endnodes u and v that are adjacent to x. But then $H' = x, u, Q, v, x$ is an unbalanced hole strictly smaller than H unless Q has length $|H| - 2$, hence u and v have a common neighbor w in H and $Q = H \setminus w$. □

4.1 Cleaning unsigned graphs

In this section we will deal with unsigned graphs (or, if you prefer, signed graphs where all the edges have sign 1). We will provide an algorithm running in time $O(n^5)$ that will construct a family $\mathcal{C}$ of subsets of $V(G)$ containing $O(n^4)$ elements such that if H is a smallest unbalanced hole in G, then $\mathcal{C}$ contains a cleaner for H. The following is due to Conforti and Rao [11].

4.2 *Let H be a smallest unbalanced hole 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 $|H| \geq 10$. By symmetry, we only need to prove the statement for $M(H) \cap C$. The proof is by induction on $|M(H) \cap C|$. If $|M(H) \cap C| = 2$, let x, y be the two nodes in $M(H) \cap C$ and suppose, by contradiction, that x and y have no neighbor in common in H. Let $\mathcal{I}$ be the family of all maximal subpaths of H with no neighbor of x or y in the interior. By definition the endnodes of every element I in $\mathcal{I}$ are neighbors of x or y and every path in $\mathcal{I}$ has even length. Since every major node for H has an odd number of neighbors in H, then $\mathcal{I}$ has an even number of elements and, since H has length 2 modulo 4, then there is an odd number of paths of $\mathcal{I}$ of length 2 modulo 4. Also, one can readily verify that there is an even number of paths I in $\mathcal{I}$ such that both endnodes a, a' of I are adjacent either to x or to y, and every such path I must have length 2 modulo 4, otherwise x, a, I, a', x or y, a, I, a', y would be an unbalanced hole of length strictly smaller than H. Therefore there is an odd number of paths of $\mathcal{I}$ of length 2 modulo 4 having one endnode adjacent to x and the other adjacent to y. If there are exactly 2 paths I_1 and I_2 with one endnode adjacent to x, say a_1 and a_2, respectively, and the other adjacent to y, say b_1 and b_2 respectively, then one of the two has length 2 modulo 4 and the other has length 0 modulo 4, and one can easily verify that I_1 and I_2 have no nodes in common, therefore $H' = x, a_1, I_1, b_1, y, b_2, I_2, a_2, x$ is an unbalanced hole strictly smaller than H, a contradiction. Therefore there are at least 4 elements in $\mathcal{I}$, say I_1, I_2, I_3, I_4, such that I_i has an endpoint, say a_i, adjacent to x and the other, say b_i, adjacent to y, for every $i \in [4]$, and such that I_1 has length k modulo 4, for some $k \in \{0, 2\}$ while I_2, I_3, I_4 have length $k + 2$ modulo 4. Without loss of generality, I_1 and I_2 do not have any node in common, hence $H' = x, a_1, I_1, b_1, y, b_2, I_2, a_2, x$ is an unbalanced hole strictly smaller than H, a contradiction.
Thus $|M(H) \cap C| \geq 3$. Let x, y, z be 3 distinct nodes in $M(H) \cap C$. By induction, there are nodes x', y' and z' in H that are adjacent to all nodes in $M(H) \cap C \setminus x$, $M(H) \cap C \setminus y$ and $M(H) \cap C \setminus z$ respectively. We may assume that x', y' and z' are all distinct, otherwise, if $x' = y'$, then x' is adjacent to every node in $M(H) \cap C$ and we are done. But then $H' = x, y', z, x', y, z', x$ is a 6 hole, a contradiction. $\square$

For the sake of the running time, we will also need the following.

4.3 *Let H be a smallest unbalanced hole in G. Then one of the following holds:*

(*i*) *Every node in $M(H) \cap R$ is adjacent to every node in $M(H) \cap C$.*

(*ii*) *There exist two adjacent nodes a and b in H such that every major node for H is adjacent to a or b.*

Proof: If $|H| = 6$, case (ii) must always occur, hence we may assume $|H| \geq 10$. Assume that there exist $a \in M(H) \cap R$ and $b \in M(H) \cap C$ such that a and b are not adjacent. Fix an orientation on H and let $a_1, \ldots, a_k$ be the neighbors of a in H in the order they appear according to such orientation starting from a_1. For $1 \leq i \leq k$ let A_k be the subpath of H between a_i and a_{i+1} (where $a_{k+1} = a_1$) containing no neighbors of a in its interior.

(4.3.1) *Up to symmetry, b is adjacent to the neighbor of a_1 in A_1, say b_1, all neighbors of b distinct from b_1 are contained in A_k and the neighbor of b closest to a_k in A_k has distance 3 modulo 4 from a_k.*

Let $\mathcal{I}$ be the family of all maximal subpaths of H with no neighbor of a or b in the interior. By definition the endnodes of every element I in $\mathcal{I}$ are neighbors of a or b. Since every major node for H has an odd number of neighbors in H, then $\mathcal{I}$ has an even number of elements. Every path in $\mathcal{I}$ of even length has both endnodes adjacent either to a or b and it must have length 2 modulo 4 (else either $I \cup a$ or $I \cup b$ induce an unbalanced hole strictly smaller than H). Since H has length 2 modulo 4, then there must be an even number of paths in $\mathcal{I}$ with odd length, and the sum of all lengths of such paths must be 2 modulo 4. Obviously, every odd path must have one endnode adjacent to a and the other adjacent to b. Suppose there are exactly 2 odd paths in $\mathcal{I}$, say I' and I'' with endnodes $a' \in N(a)$, $b' \in N(b)$ and $a'' \in N(a)$, $b'' \in N(b)$ respectively. Then a', b', a'', b'' are all distinct otherwise, w.l.o.g., $a' = a''$ and the subpath of H between b' and b'' not containing a' contains at least one neighbor of a, therefore there exists another path in $\mathcal{I}$ of odd length. Also, a' and b'' are not adjacent, otherwise a', b'' would be an odd path in $\mathcal{I}$. Analogously a'' and b' are not adjacent, hence $H' = a, a', I', b', b, b'', I'', a'', a$ is an unbalanced hole smaller than H. Thus there are at least 4 paths of odd length in $\mathcal{I}$. Furthermore there exist 3 paths I_1, I_2, I_3 in $\mathcal{I}$ each of length q modulo 4 for some $q \in \{1, 3\}$ (since either $\mathcal{I}$ contains at least 6 odd paths, or $\mathcal{I}$ contains exactly 4 odd paths whose total length must be 2 modulo 4). Since $\mathcal{I}$ contains at least 6 elements, we may assume that I_2 and I_3 have no node in common. If no node in I_2 is adjacent to a node in I_3, then $I_2 \cup I_3 \cup \{a, b\}$ induces an unbalanced hole strictly smaller than H. Thus an endnode of I_2, say a_1 w.l.o.g., is adjacent to an endnode of I_3, say b_1. If $I_1 \neq a_1, b_1$, then by the previous argument one endnode of I_1 must be either coincident or adjacent to the endnode of I_2 distinct from a_1, and the other endnode of I_1 must be either coincident or adjacent to the endnode of I_3 distinct from b_1. Since $\mathcal{I}$ has at least 6 elements, then it must be the case that $\mathcal{I}$ has exactly 6 elements, namely I_1, I_2, I_3 and 3 sectors of length 1, say I_4, I_5, I_6. Thus I_1, I_2, I_3 also have length congruent to 1 modulo 4, and there exist $1 \leq i < j \leq 6$ such that no node of I_i belongs to or has a neighbor in I_j, a contradiction. Thus $I_1 = a_1, b_1$, I_2 and I_3 have length 1 modulo 4, and there are no other paths in $\mathcal{I}$ of length congruent to 1 modulo 3. This argument also shows that there exists a unique path I_4 in $\mathcal{I}$ of length 3 modulo 4, therefore we may assume that b_1 is in A_1, all neighbors of b distinct from b_1 are in A_k, and I_4 is the shortest path in A_k between a_k and a neighbor of b. This proves 4.3.1.

If every major node for H is adjacent to a_1 or b_1, then we are done. Hence we may assume that there exists $x \in M(H) \cap C$ nonadjacent to b_1.

(4.3.2) *x and b are both adjacent to the neighbor of a_1 in A_k, say b_2, and $|A_1| > 2$.*

By 4.2 there exists a node b' in H adjacent to both b and x. Assume first that x is adjacent to a, then $b' = b_2$, else a, x, b', b, b_1, a_1, a is a 6-hole. If x has no neighbors in A_1, then $H' = a, a_2, A_1(a_2, b_1), b_1, b, b_2, x, a$ is an unbalanced hole strictly smaller than H, as one can readily verify. Hence x has a neighbor in A_1, distinct from b_1 by assumption, therefore $|A_1| > 2$. Hence we may assume that x is not adjacent to a. Since $b' \neq b_1$ and all neighbors of b in H distinct from b_1 are contained in A_k, then $b' \in A_k$. Since a and x are not adjacent, then by 4.3.1 we have two cases.
Case (1): every neighbor of x in H except one, say b'', is contained in A_k, and b'' is either the neighbor of a_1 in A_1 or the neighbor of a_k in A_{k-1}. Since $b'' \neq b_1$ by assumption, then b'' is the neighbor of a_k in A_{k-1}, but then, given the path I between b_1 and b'' in $H \setminus A_k$, b, b_1, I, b'', x, b', b is an unbalanced hole strictly smaller than H, a contradiction.
Case (2): b' is the only neighbor of x in A_k. In this case, either $b' = b_2$ and all neighbors of x in H distinct from b_2 are contained in A_1, hence $|A_1| > 2$ and we are done, or b' is the neighbor of a_k in A_k, contradicting 4.3.1, since b' is also adjacent to b and the neighbor of b closest to a_k in A_k has distance 3 modulo 4 from a_k itself.
This conclude the proof of 4.3.2.

By 4.3.2, every node in $M(H) \cap C$ is adjacent to b_1 or b_2. If there exists $y \in M(H) \cap C$ such that y is not adjacent to b_2, then by 4.2, x and y have a common neighbor b' in H and $x, b', y, b_1, a_1, b_2, x$ is a 6 hole, a contradiction. Thus b_2 is adjacent to every node of $M(H) \cap C$. If a_1 is adjacent to every node of $M(H) \cap R$ then we are done, therefore, by 4.3.2 and by symmetry, every node in $M(H) \cap R$ is adjacent to the neighbor of b_1 in H distinct from a_1, say a'. In particular $a' = a_2$ and $|A_1| = 2$, contradicting 4.3.2. □

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

- **Input** *A bipartite graph G.*
- **Output** *A family $\mathcal{C}$ of $O(|V(G)|^4)$ subsets of $V(G)$ such that, if H is a smallest unbalanced hole in G, then there exists an element of $\mathcal{C}$ that is a cleaner for H.*

Algorithm:

1. For every chordless path P of length 3, $P = u_1, u_2, u_3, u_4$ define
   $X(P) = (N(u_2) \cup N(u_3)) \setminus V(P)$ and
   $Y(P) = (N(u_1) \cap N(u_3)) \cup (N(u_2) \cap N(u_4))$.

2. Let $\mathcal{C}$ be the family containing $X(P)$ and $Y(P)$ for every chordless path P of length 3.

Correctness: The running time of the algorithm is obviously $O(|V(G)|^5)$ and $\mathcal{C}$ has $O(|V(G)|^4)$ elements. We only need to show that, if G contains a smallest unbalanced hole H, then $\mathcal{C}$ contains a cleaner for H. If H contains two adjacent nodes u_2 and u_3 such that every major node for H is adjacent to u_2 or u_3, then let u_1 be the neighbor of u_2 in H distinct from u_3, and u_4 be the neighbor of u_3 in H distinct from u_2. $X(u_1, u_2, u_3, u_4)$ is obviously a cleaner for H.
Otherwise, by 4.3, every node in $M(H) \cap R$ is adjacent to every node in $M(H) \cap C$. By 4.2, there exist nodes u_1 and u_4 in H such that u_1 is adjacent to every node in $M(H) \cap R$ and u_4 is adjacent to every node in $M(H) \cap C$. Let a', a'' be the neighbors of u_1 in H and b', b'' the neighbors of u_4 in H. Then there exist $x', x'' \in M(H) \cap C$ such that x' is not adjacent to a' and x'' is not adjacent to a''. If $x' \neq x''$, then $u_4, x', a'', u_1, a', x'', u_4$ is a 6-hole, a contradiction. Let $u_3 = x' = x''$. Analogously, there exists a node $u_2 \in M(H) \cap R$ that is nonadjacent to both b' and b''. It is immediate to verify that $Y(u_1, u_2, u_3, u_4)$ is a cleaner for H. □

4.2 Cleaning signed graphs

Let (G, σ) be a signed bipartite graph. We state, without proving it, the following lemma due to Conforti, Cornuéjols, Kapoor and Vušković [6].

4.5 *Let (G, σ) be a signed graph. Let G be the smallest unbalanced hole of (G, σ). Then there exist two edges u_1u_2 and v_1v_2 of H such that every major node for H is adjacent to one of u_1, u_2, v_1, v_2.*

This provides the following cleaning algorithm.

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

- **Input** *A signed bipartite graph (G, σ).*
- **Output** *A family $\mathcal{C}$ of $O(|V(G)|^8)$ subsets of $V(G)$ such that, if H is a smallest unbalanced hole in G, then there exists an element of $\mathcal{C}$ that is a cleaner for H.*

Algorithm:
For every pair (P_1, P_2) of chordless paths of length 3, where $P_1 = u_0, u_1, u_2, u_3$, $P_2 = v_0, v_1, v_2, v_3$, define

$$X(P_1, P_2) = (N(u_1) \cup N(u_2) \cup N(v_1) \cup N(v_2)) \setminus (V(P_1) \cup V(P_2)).$$

Let $\mathcal{C}$ be the family containing $X(P_1, P_2)$ for every pair (P_1, P_2) of chordless paths of length 3.

Correctness: follows immediately from 4.5. □

5 Detecting a clean smallest unbalanced hole

5.1 *Let (G, σ) be a signed bipartite graph containing no 3-path configuration and no detectable 3-wheel. Let H be a clean smallest unbalanced hole of (G, σ), 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. Then the following hold:*

(*i*) $|P| = |P_1|$

(*ii*) *Either $\sigma(P) = \sigma(P_1)$ and $H' = u, P, v, P_2, u$ is a clean smallest unbalanced hole, or $|P_1| = |P_2|$, $\sigma(P) = \sigma(P_2)$ and $H'' = u, P, v, P_1, u$ is a clean smallest unbalanced hole.*

Proof: The statement is obvious if H is an unbalanced hole of length 4, hence we may assume $|H| \geq 6$. 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 $1 \leq i \leq h_{2s}$. For any two distinct nodes x and y in H, let $H(x, y)$ be the underlying graph of the directed path from x to y 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 $p_{k+1} = v$. We will prove 5.1 by induction on k.

If $k = 1$, then p_1 has exactly two neighbors in H, namely u and v (since H is clean), and they are contained in a subpath of H of length 2, say u, w, v. Hence, by 4.1, $H' = u, p_1, v, P_2, u$ is a an unbalanced hole of the same length as H. We only need to prove that H' is clean. Assume not and let x be a major node for H'. Since x is not major for H, then x is adjacent to p_1 but not to w, x has exactly 2 neighbors in P_2 and they are contained in a path of length 2. But then (H', x) is a detectable 3-wheel, a contradiction. Hence we may assume $k \geq 2$.

(5.1.1) *Either:*

(*i*) *no node of $P(p_1, p_k)$ belongs to or has a neighbor in $H(h_{m+1}, h_{2s})$, or*

(*ii*) *$|P| = |H(u, v)| = |H(v, u)| = s$, $\sigma(P) = \sigma(H(v, u))$ and no node of $P(p_1, p_k)$ belongs to or has a neighbor in $H(h_2, h_{m-1})$.*

Assume that there is a node of $P(p_1, p_k)$ that belongs to or has a neighbor in $H(h_{m+1}, h_{2s})$, then there exists a 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

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

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 symmetry, we may assume $|Q_1| \leq |H(h_j, u)|$. We can also argue that either

$|Q_1| < |P|$ and $j < 2s$, or $|Q_2| \leq |H(v, h_j)|$, $|Q_2| < |P|$ and $j > m+1$. In fact, if $|Q_1| = |P|$, then $Q_1 = u, p_1, \ldots, p_k, h_j$ and $Q_2 = h_j, p_k, v$, hence $|Q_2| \leq |H(v, h_j)|$ and $|Q_2| < |P|$. Furthermore, if $j = 2s$, then $|Q_2| < |P| \leq |H(v, h_j)| + 1$ (since h_{2s} cannot be adjacent to both h_1 and p_1) and $j > m+1$. Thus, by symmetry, we may assume $|Q_1| \leq |H(h_j, u)|$, $|Q_1| < |P|$ and $j < 2s$. By inductive hypothesis,

$$d_G(h_j, u) = d_H(h_j, u) = \min(2s + 1 - j, j - 1)$$

and $d_G(h_j, u) \leq |Q_1| < |P| \leq m - 1 < j - 1$, hence $d_G(h_j, u) = 2s + 1 - j = |Q_1| < s$. By induction, $\sigma(Q_1) = \sigma(H(h_j, u))$ and $H' = u, H(u, h_j), h_j, Q_1, u$ is a clean smallest unbalanced hole. We obtain a directed cycle $\overrightarrow{H'}$ by orienting the edges of H' to agree with the orientation of the edges in $H(u, h_j)$, and define $H'(x, y)$ for every x, y in H' as before.
Let u' be the neighbor of h_j in Q_1. Then there exists a subpath P' of P between u' and v of length $k+2-|Q_1| = k+j+1-2s < k+1$. By induction, $|H'(u', v)| \leq k+j+1-2s$ or $|H'(v, u')| \leq k + j + 1 - 2s$. But $|H'(u', v)| > |H(u, v)| \geq k + 1 > |P'|$, hence $|H'(v, u')| \leq k + j + 1 - 2s$. 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 $m = s + 1$, $k = s - 1$, $d_G(u', v) = d_{H'}(u', v) = |P'|$ and $\sigma(P') = \sigma(H'(v, u'))$. By induction, $H'' = u', H'(u', v), v, P', u'$ is a clean smallest unbalanced hole. Since $H(u, v)$ is contained in $H'(u', v)$, then no node in the interior of P' belongs to or has a neighbor in $H(h_2, h_{m-1})$. Since every node in $P(p_1, p_k)$ is either a node of Q_1 or a node of P', then no node of $P(p_1, p_k)$ belongs to or has a neighbor in $H(h_2, h_{m-1})$. Finally

$$\begin{aligned} \sigma(P) &= \sigma(P(u, u')) + \sigma(P(u', v)) = \sigma(Q_1) - \sigma(u'h_j) + \sigma(P') \\ &= \sigma(H(h_j, u)) - \sigma(u'h_j) + \sigma(H'(v, u')) \\ &= \sigma(H(h_j, u)) - \sigma(u'h_j) + \sigma(H(v, h_j)) + \sigma(u'h_j) \\ &= \sigma(H(v, u)) \end{aligned} \tag{2}$$

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*) $|H(u, v)| = |P|$ *and* $\sigma(H(u, v)) = \sigma(P)$, *or*

(*ii*) $|P| = |H(u, v)| = |H(v, u)| = s$, $\sigma(P) = \sigma(H(v, u))$ *and no node of* $P(p_1, p_k)$ *belongs to or has a neighbor in* $H(h_2, h_{m-1})$.

Clearly, if $\sigma(H(u, v)) = \sigma(P)$, then $H' = u, P, v, H(v, u,), u$ is an unbalanced hole of length at most $|H|$, hence $|P| = |H(u, v)|$ and 5.1.2 holds. Suppose then that $\sigma(H(u, v)) \neq \sigma(P)$. If 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 $H(u, v), H(v, u), P$

would induce a 3-path configuration between u and v. Thus P has even length and $H'' = u, H(u,v), v, P, u$ is an unbalanced hole strictly smaller than H unless $|P| = |H(u,v)| = |H(v,u)| = s$, thus case (ii) holds. Therefore 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. By 5.1.1 and symmetry, we may assume $m \leq s$. 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 the same argument as in 5.1.1, we may assume $|Q_1| \leq j-1$, $|Q_1| < |P|$ and $j > 2$. By induction, $|Q_1| = d_G(u,h_j) = d_H(u,h_j) = j-1$, $\sigma(Q_1) = \sigma(H(u,h_j)$ and $H' = u, Q_1, h_j, H(h_j,u), u$ is a clean smallest unbalanced hole.
Let u' be the neighbor of h_j in Q_1 and let P' be the path between u' and v in P. Then

$$|P'| = k+2-|Q_1| = k-j+3$$

thus, by induction, $\sigma(P') = \sigma(H'(u',v))$. Finally, with a calculation very similar to the one in (2),

$$\begin{aligned} \sigma(P) &= \sigma(P(u,u')) + \sigma(P(u',v)) \\ &= \sigma(H(u,h_j)) - \sigma(u'h_j) + \sigma(H(h_j,v)) + \sigma(u'h_j) \\ &= \sigma(H(u,v)). \end{aligned}$$

This completes the proof of 5.1.2.

By 5.1.1, 5.1.2 and by symmetry, we may assume that $H' = u, P, v, H(v,u), u$ is a smallest unbalanced hole. To conclude the proof of 5.1 we only need to show that H' is clean. Suppose, by contradiction, that H' is not clean, and let x be a major vertex for H'. If x has at least two neighbors in $H(v,u)$, then such neighbors are contained in a subpath of H of length 2, thus x is adjacent to h_i and h_{i+2} for some $m \leq i \leq 4s+1$ and has no other neighbors in H. Thus $H'' = h_i, x, h_{i+2}, H(h_{i+2}, h_i), h_i$ is a clean smallest unbalanced hole and the interior of P contains a neighbor of x, whence, by 5.1.1 applied to H'' and P, it must be the case that $|P| = s$, $\sigma(P) = \sigma(H'(v,u)) = \sigma(H(v,u))$ and no node of $P(p_1,p_k)$ belongs to or has a neighbor in $H'(h_2,h_{m-1}) = H(h_2,h_{m-1})$. Since $\sigma(H) \equiv 2 \bmod 4$, and $\sigma(H(u,v)) = \sigma(P) = \sigma(H(v,u))$, then u and v are in distinct sides of the bipartition and $H(u,v), H(v,u), P$ induce a 3-path configuration between u and v. Thus x has at most one neighbor in $H(v,u)$ and at least 2 neighbors in the interior of P. Let p_i and p_j be the neighbors of x in P of lowest and highest index, respectively. Then $j = i+2$, else $u, P(u,p_i), p_i, x, p_j, P(p_j,v), v$ is a path between u and v strictly shorter than P. But then x has exactly 3 neighbors in H', and two of these neighbors have distance 2 in H', so (H',x) is a detectable 3 wheel, a contradiction. □

5.2 *There exists a $O(|V(G)|^4)$ algorithm with the following specifications:*

- **Input** *A clean signed bipartite graph (G,σ) containing no 3-path configuration and no detectable 3-wheel.*

- **Output** *Either*
  - (i) *An unbalanced hole H, or*
  - (ii) *determines that (G, σ) is balanced.*

Algorithm:
For every quadruple of nodes, check if they induce an unbalanced hole H. If this is the case, then output H and stop.

For every possible pair of nodes u_1, u_2, do the following:

1. compute the shortest path P between u_1 and u_2.
2. compute the set X of nodes that do not belong to or have a neighbor in the interior of P.
3. for every node u_3 in X at distance 2 from u_1 in $G[X \cup \{u_1\}]$, compute the shortest paths $P_1(u_3)$ and $P_2(u_3)$ between u_1 and u_3 in $G[X \cup \{u_1\}]$ and between u_2 and u_3 in $G[X \cup \{u_2\}]$ (if one exists), respectively.
4. for every such $u_3 \in X$, verify that no node in $P_1(u_3) \backslash u_3$ belongs to or has a neighbor in $P_2(u_3) \backslash u_3$. If this is the case, define $H(u_1, u_2, u_3) = u_1, P, u_2, P_2(u_3), u_3, P_1(u_3), u_1$, otherwise let $H(u_1, u_2, u_3)$ be undefined.
5. If $\sigma(H(u_1, u_2, u_3)) \equiv 2 \bmod 4$, then output the unbalanced hole $H = H(u_1, u_2, u_3)$ and stop.

Otherwise output that G is balanced.

Correctness: checking if (G, σ) has an unbalanced hole of length 4 takes time $O(|V(G)|^4)$. For every possible pair u_1 and u_2, the running time of steps 1 through 4 is $O(|V(G)|^2)$ (in fact, step 4 takes linear time for every choice of u_3 since $P_1(u_3)$, by definition, has constant length 2). Hence the overall running time is $O(|V(G)|^4)$. Obviously, when the algorithm outputs an unbalanced hole it is correct. We need to verify that the algorithm is always correct when it outputs that G is balanced. Assume G is not balanced. Since G is clean, there exists a clean smallest unbalanced hole H, with $|H| = 2s$, and $s \geq 3$ since $|H| \geq 6$. Let u_1, u_2, u_3 be three nodes in H such that $d_H(u_1, u_3) = 2$, while $d_H(u_1, u_2) = d_H(u_2, u_3) = s - 1$. Let P, $P_1(u_3)$, $P_2(u_3)$ be the paths computed by the algorithm for the triple u_1, u_2, u_3. Let Q_1 and Q_2 be the subpaths of H between u_1 and u_2 such that Q_1 does not contain u_3 and Q_2 contains u_3. Then, by our choice of u_1, u_2, u_3, $|Q_1| < |Q_2|$, hence, by 5.1, $H' = u_1, P, u_2, Q_2, u_1$ is a clean smallest unbalanced hole. By repeating the same argument for $P_1(u_3)$ and $P_2(u_3)$, we argue that $H(u_1, u_2, u_3) = u_1, P, u_2, P_2(u_3), u_3, P_1(u_3), u_1$ is a clean smallest unbalanced hole, hence the algorithm would have output it correctly. □

6 The recognition algorithm

6.1 Unsigned graphs

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

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

Algorithm:

1. Apply the algorithm in 2.2. If G contains a 3-path configuration, then output the fact that G is not balanced and stop.
2. Apply the algorithm in 3.2. If G contains a detectable 3-wheel, then output the fact that G is not balanced and stop.
3. Apply the algorithm in 4.4 to obtain a family $\mathcal{C}$ of subsets of $V(G)$.
4. For every element X of $\mathcal{C}$ apply the algorithm in 5.2 to $G \setminus X$. If, for some X, the algorithm outputs an unbalanced hole, then output the fact that G is not balanced; else output the fact that G is balanced.

Correctness: The running time of the algorithm is $O(|V(G)|^9)$. In fact both step 1 and step 2 takes time $O(|V(G)|^9)$, step 3 takes time $O(|V(G)|^5)$, while step 4 takes time $O(|V(G)|^4)$ for each of the $O(|V(G)|^4)$ elements of $\mathcal{C}$, hence it takes time $O(|V(G)|^8)$ overall.
The correctness follows immediately by 1.1 and by the correctness of the algorithms in 2.2, 3.2, 4.4 and 5.2. □

6.2 Signed graphs

The algorithm in 6.1 can be generalized to an algorithm to test balancedness for signed bipartite graphs simply by using the cleaning step 4.6 instead of 4.4. Since the family of sets produced this way is of order $O(n^8)$, then the overall running time is $O(n^{12})$. One can reduce the running time by combining algorithms 4.6 and 5.2 to obtain an $O(n^{11})$ algorithm to test for balancedness for signed bipartite graphs with no 3-path configurations or detectable 3-wheels as follows.

6.2 *There exists a $O(|V(G)|^{11})$ algorithm with the following specifications:*

- **Input** *A signed bipartite graph (G, σ) containing no 3-path configuration and no detectable 3-wheel.*
- **Output** *Determines whether G is balanced or not.*

Algorithm: For every pair (P_1, P_2) of chordless paths of length 3, where $P_1 = u_0, u_1, u_2, u_3$, $P_2 = v_0, v_1, v_2, v_3$, compute $X(P_1, P_2) = (N(u_1) \cup N(u_2) \cup N(v_1) \cup N(v_2)) \setminus (V(P_1) \cup V(P_2))$. For every $x \in V(G) \setminus X(P_1, P_2)$, do the following:

1. In $G \setminus X(P_1, P_2)$, compute the shortest paths Q and Q' between x and v_0 and x and v_2, respectively.

2. Verify that v_1 does not belong to or has a neighbor in $Q \setminus v_0$, and $Q' \setminus v_2$, and that no node of $Q \setminus v_0$ belongs to or has a neighbor in $Q' \setminus v_2$. In this case let $H = v_0, v_1, v_2, Q', x, Q, v_0$.

3. If $\sigma(H) \equiv 2 \bmod 4$, then output the unbalanced hole H.

Otherwise output the fact that G is balanced.

Correctness: The running time is $O(|V(G)|^{11})$, since there are $O(|V(G)|^9)$ choices for P_1, P_2 and x, and each of the steps 1-3 takes at most $O(|V(G)|^2)$. If the algorithm outputs an unbalanced hole, then clearly it answers correctly, hence we only need to verify that the algorithm is always correct when it outputs that G is balanced. Suppose G is not balanced and let H' be a smallest unbalanced hole. By Lemma 4.5, there exist two edges v_1v_2 and u_1u_2 of H', such that every major node for H' is adjacent to a node in $\{u_1, u_2, v_1, v_2\}$. If $P_1 = u_0, u_1, u_2, u_3$ and $P_2 = v_0, v_1, v_2, v_3$ are subpaths of H', then H' is clean in $G \setminus X(P_1, P_2)$. Let x be the node of H' such that $d_{H'}(v_0, x) = d_{H'}(v_2, x)$. Then, by 5.1, given the paths Q and Q' computed by the algorithm, $H = v_0, v_1, v_2, Q', x, Q, v_0$ is an unbalanced hole, thus the algorithm would have output it in the iteration relative to P_1, P_2, x. □

The following is an immediate consequence of the previous claim.

6.3 *There exists a $O(|V(G)|^{11})$ algorithm with the following specifications:*

- **Input** *A signed bipartite graph (G, σ).*

- **Output** *Determines whether G is balanced or not.*

Algorithm:

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

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

3. Apply the algorithm in 6.2. If (G, σ) contains an unbalanced hole, then output the fact that (G, σ) is not balanced; else output the fact that (G, σ) is balanced.

Acknowledgments: I would like to thank Gerard Cornuéjols for several helpful discussions.

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] M. Chudnovsky, G. Cornuéjols, X. Liu, P. Seymour, K. Vušković, Recognizing Berge Graphs, *Combinatorica*, to appear.

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

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

[6] 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.

[7] M. Conforti, G. Cornuéjols, A. Kapoor, K. Vušković, Even-Hole-Free Graphs Part I: Decomposition Theorem, *Journal of Graph Theory* **39** (2002), 6-49.

[8] M. Conforti, G. Cornuéjols, A. Kapoor, K. Vušković, Even-Hole-Free Graphs Part II: Recognition Algorithm, *Journal of Graph Theory* **20** (2002), 238-266.

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

[10] M. Conforti, G. Cornuéjols, K. Vušković, Decomposition of Odd-Hole-Free Graphs by Double Star Cutsets and 2-Joins, preprint (2002), to appear in a special issue of *Discrete Applied Mathematics* dedicated to the Brazilian Symposium on Graphs, Algorithm and Combinatorics, Fortaleze, March 17-19, 2001.

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

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

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

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

[15] P. Seymour, Decomposition of Regular Matroids, *Journal of Combinatorial Theory B* **28** (1980), 305-359.

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