Detecting a long odd hole

Maria Chudnovsky[1]
Princeton University, Princeton, NJ 08544

Alex Scott[2]
Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK

Paul Seymour[3]
Princeton University, Princeton, NJ 08544.

January 31, 2019; revised April 28, 2019

[1]Supported by NSF grant DMS 1763817 and US Army Research Office Grant W911NF-16-1-0404.
[2]Supported by a Leverhulme Trust Research Fellowship.
[3]Supported by NSF grant DMS-1800053, and partially supported by the Simons Foundation and by the Mathematisches Forschungsinstitut Oberwolfach.

Abstract

For each integer $\ell \geq 5$, we give a polynomial-time algorithm to test whether a graph contains an induced cycle with length at least ℓ and odd.

1 Introduction

All graphs in this paper are finite and have no loops or parallel edges. A *hole* of G is an induced subgraph of G that is a cycle of length at least four, and an *antihole* is an induced subgraph whose complement is a cycle of length at least four. In 2005, two of us, with Cornuéjols, Liu and Vušković [4], gave an algorithm to test whether an input graph G has an odd hole or odd antihole, and thereby to test whether G is perfect, with running time at most polynomial in $|G|$. ($|G|$ denotes the number of vertices of G.) At that time we were unable to separate the test for odd holes from the test for odd antiholes, and testing for odd holes in poly-time has remained open until very recently. Indeed, it seemed quite likely that testing for an odd hole was NP-complete; for instance, D. Bienstock [1, 2] showed that testing if a graph has an odd hole containing a given vertex is NP-complete. So it was something of a surprise when recently we found a poly-time algorithm to test for odd holes [3].

In this paper we extend that result: for each integer $\ell \geq 5$ we give a poly-time algorithm to test whether G has an odd hole of length at least ℓ. More exactly:

1.1 *For each integer $\ell \geq 5$, there is an algorithm with the following specifications:*

Input: *A graph G.*

Output: *Decides whether G has an odd hole of length at least ℓ.*

Running time: $O(|G|^{20\ell+43})$.

We have not tried very hard to optimize the exponent in the running time (although getting the exponent to be linear in ℓ took some effort).

The new algorithm once again uses "cleaning", as does the algorithm of [3] and several other algorithms to detect special induced subgraphs. Indeed it was modelled on the algorithm of [3], but it is considerably more complicated.

Here is an outline of the method. Throughout the paper, $\ell \geq 5$ is a fixed number, and throughout, a *long* path or hole means a path or hole of length at least ℓ. If C is a hole in G, a vertex v of $V(G) \setminus V(C)$ is *C-major* if there is no three-vertex path of C containing all the neighbours of v in $V(C)$. A hole C is *clean* if it has no C-major vertex.

- First we test for the presence in the input graph G of certain kinds of induced subgraphs ("short" long odd holes, "long pyramids" and "long jewels") that we can test for in polynomial time, and whose presence would imply that G contains a long odd hole. We call these three kinds of subgraphs "easily-detected configurations". We may assume these tests are unsuccessful.

- Second, we generate a "cleaning list", a list of polynomially-many subsets of $V(G)$, such that if G has a long odd hole, and C is a long odd hole of minimum length (a *shortest long odd hole*) then some set X in the list contains all the C-major vertices and contains no vertex of C itself. This relies on the fact that G contains none of the easily-detected configurations.

- Third, for each X in the cleaning list, we test whether $G \setminus X$ has a clean shortest long odd hole. (More exactly, we give an algorithm that either decides that $G \setminus X$ has a long odd hole, or decides that $G \setminus X$ has no clean shortest long odd hole.) This again relies on the absence of the easily-detected configurations.

The reader familiar with the method of [3] will see the similarity of the two algorithms.

But part of the approach is significantly different. To generate the cleaning list in [3], we used a theorem that if C is a shortest odd hole, and M is a set of C-major vertices such that one of them is nonadjacent to all the others, then there is a "heavy edge" in C, an edge uv of C such that every vertex in M is adjacent to one of u, v. We tried to extend this to the long odd hole situation, but failed. For our purposes, a "heavy path" of C of bounded length (that is, such that every vertex in M has a neighbour in the path) would be just as good as a heavy edge; and this extension might be true, but we were unable to prove it. In its place we had to use a considerably more complicated method, proving that there is a bounded set of paths of C, each of bounded length, such that every vertex in M has a neighbour in one of the paths; and we could only prove this when the exceptional vertex of M was carefully chosen.

The paper is organized as follows. First, we explain how to test for the easily-detected configurations; this is a straightforward adaptation of the algorithms in [4] to test for pyramids and jewels. Then we give the algorithm for the third step above; and finally we show how to generate the cleaning list.

Let us remark, finally, that if we want to test for a long hole, rather than a long odd hole, then this is easy: enumerate all induced paths of G of length $\ell - 1$, and for each one, test directly if it can be extended to a hole. This has running time $O(|G|^{\ell+1})$, and so both this and our algorithm for 1.1 have running time $|G|^{O(\ell)}$. We do not know if either can be substantially improved.

2 The easily-detected configurations

We begin with the test for what we earlier called "short" long odd holes:

2.1 *There is an algorithm with the following specifications:*

Input: *A graph G, and an integer $k \geq 0$.*

Output: *Decides whether there is a long odd hole in G of length at most k.*

Running time: $O(|G|^k)$.

Proof. We enumerate all sets of at most k vertices, and for each one, check whether it induces a long odd hole. ∎

If $X \subseteq V(G)$, we denote the subgraph of G induced on X by $G[X]$. If X is a vertex or edge of G, or a set of vertices or a set of edges of G, we denote by $G \setminus X$ the graph obtained from G by deleting X. Thus, for instance, if b_1b_2 is an edge of a hole C, then $C \setminus \{b_1, b_2\}$ and $C \setminus b_1b_2$ are both paths, but one contains b_1, b_2 and the other does not. If P is a path, we denote by P^* the *interior* of P, the set of vertices of the path P that are not ends of P. If P is a path and $x, y \in V(P)$, we denote the subpath with ends x, y by x-P-y. The *length* of a path or cycle is the number of edges in it.

Let $u, v \in V(G)$, and let Q_1, Q_2 be induced paths between u, v, of different parity. Let P be an induced path between u, v of length at least ℓ, such that no vertex of P^* equals or is adjacent to any vertex of Q_1^*, Q_2^*. We say the subgraph induced on $V(P \cup Q_1 \cup Q_2)$ is a *long jewel* of *order* $\max(|V(Q_1)|, |V(Q_2)|)$, *formed by* Q_1, Q_2, P. Any graph containing a long jewel has a long odd hole, since the holes $P \cup Q_1, P \cup Q_2$ are both long and one of them is odd. The next result extends theorem 3.1 of [4]:

2.2 *There is an algorithm with the following specifications:*

Input: *A graph G, and an integer $k \geq 0$.*

Output: *Decides whether there is a long jewel in G of order at most k.*

Running time: $O(|G|^{2k+\ell})$.

Proof. We enumerate all triples of induced paths Q_1, Q_2, R of G, such that:

- Q_1, Q_2 join the same pair of vertices, say u, v;
- one of Q_1, Q_2 is odd and the other is even, and each has at most k vertices;
- R has length $\ell - 2$, and has one end u and the other some vertex w say;
- no vertex of $V(R) \setminus \{u\}$ equals or has a neighbour in $V(Q_1 \cup Q_2) \setminus \{u\}$.

For each such triple of paths, let X be the set of vertices of G that are different from and nonadjacent to each vertex of $V(Q_1 \cup Q_2 \cup R) \setminus \{v, w\}$. We test whether there is a path in $G[X \cup \{w, v\}]$ between w, v. If so we output that G contains a long jewel of order at most k. If no triple yields this outcome, we output that G has no such long jewel.

To see the correctness of the algorithm, certainly the output is correct if G contains no long jewel of order at most k. Suppose then it does, say formed by Q_1, Q_2, P. Let u, v be the ends of P, and let R be the subpath of P of length $\ell - 2$ with one end u. When the algorithm tests the triple Q_1, Q_2, R, it will discover there is a path in $G[X \cup \{w, v\}]$ between w, v, because the remainder of P is such a path. Consequently the output is correct.

The running time is $O(|G|^2)$ for each triple of paths, and there are at most $|G|^{2k+\ell-2}$ such triples, so the running time is as claimed. This proves 2.2. ∎

Many of the algorithms in this paper follow the same outline; we enumerate all subgraphs, or sequences of vertices, of some prescribed type, and for each one, perform some test on it. (Critically, there must be only polynomially many such subgraphs to test.) If the test is successful, we have found a subgraph of the desired type, and if it is never successful we will apply a theorem that says that then there is no subgraph of the desired type. For brevity we call the process of enumerating all these subgraphs and testing them one-by-one "guessing"; thus we would describe the long jewel algorithm above as "guessing the two paths Q_1, Q_2 and an initial subpath of P".

Let $v_0 \in V(G)$, and for $i = 1, 2, 3$ let P_i be an induced path of G between v_0 and v_i, such that

- P_1, P_2, P_3 are pairwise vertex-disjoint except for v_0;
- $v_1, v_2, v_3 \neq v_0$, and at least two of P_1, P_2, P_3 have length at least ℓ;
- v_1, v_2, v_3 are pairwise adjacent; and
- for $1 \leq i < j \leq 3$, the only edge between $V(P_i) \setminus \{v_0\}$ and $V(P_j) \setminus \{v_0\}$ is the edge $v_i v_j$.

We call the subgraph induced on $V(P_1 \cup P_2 \cup P_3)$ a *long pyramid*, with *apex* v_0 and *base* $\{v_1, v_2, v_3\}$, *formed by* P_1, P_2, P_3.

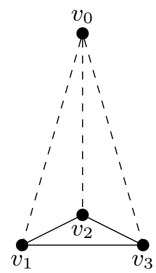

Figure 1: A long pyramid. The dashed lines represent paths, of indeterminate length, but two of them must have length at least ℓ.

If G has a long pyramid then G has a long odd hole (because two of the paths P_1, P_2, P_3 have the same length modulo two, and they induce a long odd hole). The next result extends theorem 2.2 of [4], and is proved similarly:

2.3 *There is an algorithm with the following specifications:*

Input: *A graph G.*

Output: *Decides whether there is a long pyramid in G.*

Running time: $O(|G|^{6\ell+8})$.

Proof. Suppose that G contains a long pyramid; then it contains a "smallest" one, one with the fewest vertices, say with apex v_0 and base $\{v_1, v_2, v_3\}$, formed by the path P_1, P_2, P_3. For $i = 1, 2, 3$, let m_i be a vertex of P_i that divides it into two paths with lengths differing by at most one. For $i = 1, 2, 3$,

- if P_i has length at least ℓ, let A_i be a subpath of P_i of length ℓ with one end v_0, and let B_i be a subpath of P_i of length ℓ with one end v_i;
- if P_i has length less than ℓ, let $A_i = B_i = P_i$.

Let A_i have ends v_0, a_i, and let B_i have ends v_i, b_i for $i = 1, 2, 3$.

The algorithm proceeds as follows. If there is a pyramid as above, we guess the vertex v_0 and for $i = 1, 2, 3$ we guess the vertices v_i, m_i, a_i, b_i and the paths A_i, B_i. Let X be the set of all these vertices (including the vertices of the paths A_i, B_i for $1 \leq i \leq 3$). For each i such that m_i does not belong to $V(A_i)$ we choose a shortest path A'_i between m_i, a_i such that its interior is disjoint from $X \setminus \{m_i, a_i\}$ and contains no vertex with a neighbour in $X \setminus \{m_i, a_i\}$, and let $Q'_i = A_i \cup A'_i$. If $m_i \in V(A_i)$ let $Q'_i = A_i$. Similarly, if $m_i \notin V(B_i)$ we choose a shortest path B'_i between m_i, b_i such that its interior is disjoint from $X \setminus \{m_i, b_i\}$ and contains no vertex with a neighbour in $X \setminus \{m_i, b_i\}$; and let $R'_i = B_i \cup B'_i$. If $m_i \in V(B_i)$ let $R'_i = B_i$.

Now for $1 \leq i \leq 3$ we test whether $Q'_i \cup R'_i$ is an induced path between v_0, v_1, and if this is true for each i, and the three paths form a pyramid, we return that G contains a long pyramid. To prove the correctness, we must now prove a theorem that starting from a smallest pyramid as described, $Q_i' \cup R_i'$ is indeed a path between v_0, v_i for $1 \le i \le 3$, and these three paths form a (possibly different) smallest pyramid.

Let Π be a smallest pyramid in G, formed by paths P_1, P_2, P_3 as above. For $i = 1, 2, 3$, let Q_i, R_i be the subpaths of P_i between m_i, v_0 and between m_i, v_i respectively.

(1) *With Q_1' chosen as in the algorithm, no vertex of Q_1' belongs to or has a neighbour in $V(P_2 \cup P_3)$.*

Suppose that this is false. Consequently $Q_1' \neq A_1$, and so $m_1 \notin V(A_1)$ and m_1 is an end of Q_1'. Let S be a minimal subpath of Q_1' with one end m_1 such that its other end has a neighbour in $V(P_2 \cup P_3)$. Let S have ends m_1, s. There are three cases.

First, suppose that s has a unique neighbour t in $P_2 \cup P_3$. We may assume that $t \in V(P_2)$ from the symmetry; let Π' be the pyramid with apex t and base $\{v_1, v_2, v_3\}$, formed by the paths t-P_2-v_2, t-p_2-v_0-P_3-v_3 and an induced path between t, v_1 with interior in $V(S \cup R_1)$. This is indeed a pyramid, from the minimality of S; it is long, since all three of the paths have length at least ℓ (because they include the paths B_2, B_3, B_1 respectively); and it has fewer vertices than Π, since $|V(S)| \le |V(Q_1)| - \ell$. This is impossible from the choice of Π.

Second, suppose that s has two nonadjacent neighbours in $V(P_2 \cup P_3)$. Let Π' be the pyramid with apex s and base $\{v_1, v_2, v_3\}$ formed by the induced paths between s, v_2 and between s, v_3, both with interior in $V(P_2 \cup P_3\}$ (these are unique), and a path between s, v_1 with interior in $V(S \cup R_1)$. Again, this is a long pyramid with fewer vertices than Π, a contradiction.

Third, suppose that s has exactly two neighbours t_1, t_2 in $V(P_2 \cup P_3)$ and they are adjacent. We may assume that $t_1, t_2 \in V(P_2)$ and t_2 is closer to v_2 in P_2. Let Π' be the pyramid with apex v_0 and base $\{s, t_1, t_2\}$ formed by the paths v_0-P_2-t_1, v_0-P_3-v_3-v_2-P_2-t_2 and a path between v_0, s with interior in $V(S \cup Q_1)$. Again, this is a long pyramid (because the three paths include A_2, A_3, A_1 respectively), and it has fewer vertices than Π, because $|V(S)| \le |V(Q_1)| - \ell < |V(R_1)|$, a contradiction. This proves (1).

(2) *With R_1' chosen as in the algorithm, no vertex of R_1' belongs to or has a neighbour in $V(P_2 \cup P_3)$.*

The proof is similar and we omit it.

From (1) and (2), it follows that if P_1' is an induced path between v_0, v_1 with interior in $V(A_1 \cup Q_1' \cup R_1' \cup B_1)$, then with P_2, P_3, it forms a long pyramid with at most as many vertices as Π. Consequently equality holds, and so $A_1 \cup Q_1' \cup R_1' \cup B_1$ is an induced path between v_0, v_1, say P_1'; and P_1', P_2, P_3 form a smallest long pyramid. Similarly $A_2 \cup Q_2' \cup R_2' \cup B_2$ is an induced path between v_0, v_2, say P_2'; and P_1', P_2', P_3 form a smallest long pyramid. And similarly for Q_3', R_3'. This proves the correctness of the algorithm.

For its running time, we are guessing a sequence of at most $6(\ell + 1)$ vertices, so the running time is as claimed. This proves 2.3. ∎

Let us say G is a *candidate* if G contains no long pyramid, no long jewel of order at most $\ell + 2$, and no long odd hole of length at most $2\ell + 2$. In view of 2.1, 2.2, and 2.3, we have:

2.4 *There is an algorithm with the following specifications:*

Input: *A graph G.*

Output: *Decides whether G is a candidate.*

Running time: $O(|G|^{6\ell+8})$.

Any graph that is not a candidate has a long odd hole, so now we just need to find a poly-time algorithm to test whether candidates have long odd holes.

3 Detecting a clean shortest long odd hole

The following was proved in [4]:

3.1 *Let G be a graph containing no jewel or pyramid, and let C be a clean shortest odd hole in G. Let $u, v \in V(C)$ be distinct and nonadjacent, and let L_1, L_2 be the two subpaths of C joining u, v, where $|E(L_1)| < |E(L_2)|$. Then:*

- *L_1 is a shortest path in G between u, v, and*
- *for every shortest path P in G between u, v, $P \cup L_2$ is a shortest odd hole in G.*

This was crucial in that paper. Happily, the exact analogue holds for clean shortest long odd holes:

3.2 *Let G be a graph containing no long jewel of order at most $\ell + 2$, and no long pyramid, and with no long odd hole of length at most $2\ell + 2$. Let C be a clean shortest long odd hole in G. Let $u, v \in V(C)$ be distinct and nonadjacent, and let L_1, L_2 be the two subpaths of C joining u, v, where $|E(L_1)| < |E(L_2)|$. Then:*

- *L_1 is a shortest path in G between u, v, and*
- *for every shortest path P in G between u, v, $P \cup L_2$ is a clean shortest long odd hole in G.*

First we prove the first assertion of 3.2. If u, v are vertices of a graph G, $d_G(u, v)$ denotes the length of the shortest path of G joining u, v ($d_G(u, v) = \infty$ if there is no such path).

3.3 *Let G be a graph containing no long pyramid, no long jewel of order at most $\ell + 2$, and no long odd hole of length at most 2ℓ. Let C be a clean shortest long odd hole in G. Then $d_G(u, v) = d_C(u, v)$ for all $u, v \in V(C)$.*

Proof. Suppose the result is false; then there is an induced path Q with vertices $q_1\text{-}\cdots\text{-}q_k$ in order, such that some vertex $u \in V(C)$ is adjacent to q_1, some $v \in V(C)$ is adjacent to q_k, and $d_C(u, v) > k + 1$. Choose such a path Q with k minimum. Since $d_C(u, v) > k + 1 \geq 2$, and q_1 is not C-major, it follows that $q_1 \neq q_k$, and so $k \geq 2$. If some vertex of Q belongs to $V(C)$, say $q_i \in V(C)$, then from the choice of k, $d_C(u, q_i) \leq i$, and $d_C(q_i, v) \leq k - i + 1$, and so

$$d_C(u, v) \leq d_C(u, q_i) + d_C(q_i, v) \leq k + 1,$$

a contradiction. Thus $Q \cap C$ is null. There are two paths of C that join u, v; one, L_1 say, of length $d_C(u, v)$, and the other, L_2 say, longer and of opposite parity.

Since q_1 is not C-major, there is a path P_1 of C of length at most two such that all neighbours of q_1 in $V(C)$ lie in $V(P_1)$; choose P_1 minimal, and consequently both its ends are adjacent to q_1. (Possibly P_1 has only one vertex.) Define P_2 similarly for q_k.

(1) *P_1, P_2 are vertex-disjoint.*

Suppose that $P_1 \cap P_2$ is non-null. Thus $d_C(u,v) \leq 4$, and since $d_C(u,v) > k+1 \geq 3$ it follows that $k=2$ and $d_C(u,v) = 4$, and P_1, P_2 both have length exactly two. Hence $P_1 \cup P_2 = L_1$, and the three paths L_1, u-q_1-q_2-v, and L_2, form a long jewel of order four, a contradiction. Thus $P_1 \cap P_2$ is null. This proves (1).

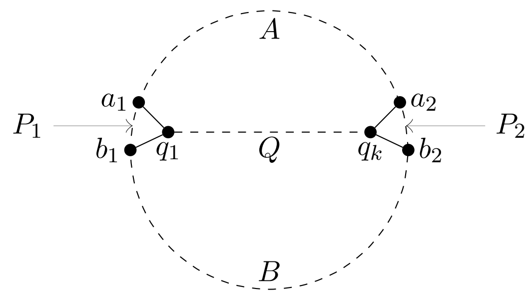

Figure 2: The paths P_1, P_2, A, B of C.

Let P_1 have ends a_1, b_1, and let P_2 have ends a_2, b_2, where a_1, b_1, b_2, a_2 are in order in C (possibly $a_1 = b_1$ or $a_2 = b_2$). Let A, B be two paths of C, where A has ends a_1, a_2, and B has ends b_1, b_2, and the four paths A, B, P_1, P_2 are pairwise edge-disjoint and have union C.

(2) *One of $q_2, \ldots, q_{k-1}$ has a neighbour in $V(C)$.*

Suppose not. The hole formed by adding the path a_1-q_1-$\cdots$-q_k-a_2 to A is shorter than C, so either it has length less than ℓ or it has even length; so either $|A| + k < \ell$, or $|A| + k$ is even. (We recall that $|A|$ denotes the number of vertices of A, not its length.) Similarly either $|B| + k < \ell$, or $|B| + k$ is even. Since $|A| + |B| + 2 \geq |C| > 2\ell$, we may assume that $|B| \geq \ell$. In particular, $|B| + k \geq \ell$, and so $|B| + k$ is even. Now the paths b_1-q_1-$\cdots$-q_k-b_k, b_1-P_1-a_1-A-a_2-P_2-b_2 and B form a long jewel, of order the maximum of the lengths of the first two paths. But the first path has length at least $d_C(u,v) \geq k+2$, so the first path is longer; and so the long jewel has order $|A| + |P_1| + |P_2| - 3$. Consequently $|A| + |P_1| + |P_2| - 3 > \ell + 2$, and so $|A| \geq \ell$. Hence $|A| + k \geq \ell$, and so $|A| + k$ is even. Since C is odd, it follows that one of P_1, P_2 is odd and the other is even, and we may assume that P_2 has length one; and so q_k has exactly two neighbours in $V(C)$, a_2 and b_2. If $a_1 = b_1$ then since $|A| + |P_1| + |P_2| \geq \ell + 6$, it follows that $|A| \geq \ell + 3$, and so the three paths A, B and a_1-q_1-Q-q_k form a long pyramid, a contradiction. Thus P_2 has exactly three vertices. Since $|A| + |P_1| + |P_2| \geq \ell + 6$, it follows that $|A| > \ell + 1$, and so the three paths q_1-a_1-A-a_2, q_1-b_1-B-b_2 and Q form a long pyramid, a contradiction. This proves (2).

(3) *None of $q_2, \ldots, q_{k-1}$ has a neighbour in $V(L_2)$.*

For suppose that q_i has a neighbour $w \in V(L_2)$ say. Thus $w \neq u, v$. Let R_2, S_2 be the subpaths of L_2 between w, u and between w, v respectively. From the minimality of k, $d_C(u,w) < d_C(u,v)$, and so the path of C between u, w of length $d_C(u,v)$ is a subpath of L_2, that is, R_2 has length $d_C(u,w)$. Similarly S_2 has length $d_C(v,w)$, and so $d_C(u,w) + d_C(v,w) = |E(L_2)|$. From the minimality of k, $d_C(u,w) \leq i+1$ (because otherwise $d_G(u,w) \leq i+1 < d_C(u,w)$ and $d_G(u,w) < k+1$, contrary to the minimality of k), and similarly and $d_C(v,w) \leq k-i+2$, and so $|E(L_2)| \leq k+3$. But $|E(L_2)| > |E(L_1)| = d_C(u,v) \geq k+2$, and so equality holds throughout. In particular, $|E(L_2)| = |E(L_1)| + 1 = k+3$. Moreover, R_2 has length $d_C(u,w) = i+1$ and S_2 has length $k-i+2$.

The union of L_1 and the path u-q_1-Q-q_k-v therefore is a cycle of length $|C|-2$. This is odd, less than C, and at least ℓ, so this cycle is not induced. Hence there exists $j \in \{1, \ldots, k\}$ such that q_j has a neighbour in the interior of L_1, say x. From the symmetry we may assume that $j \geq i$. Let R_1, S_1 be the subpaths of L_1 between x, u and between x, v respectively. The path $S_1 \cup S_2$ has length more than the length of S_2 and hence at least $k-i+3$; and the path w-q_i-$\cdots$-q_j-x has length $j-i+2 \leq k-i+2$. Thus $S_1 \cup S_2$ is longer than w-q_i-$\cdots$-q_j-x. But from the minimality of k, it follows that the length of w-q_i-$\cdots$-q_j-x is at least $d_C(w,x)$, and so at least the length of $R_1 \cup R_2$. Thus $j-i+1 \geq |A_1|+|A_2|+2$. Also, the minimality of k implies that the path x-q_j-$\cdots$-q_k-v has length at least the length of S_1, and so $k-j+2 \geq |S_1|-1$. Adding, we deduce that $k-i+3 \geq |R_1|+|R_2|+|S_1|+1$. But S_2 has length $k-i+2$, so $|S_2| \geq |R_1|+|R_2|+|S_1|+1$. In particular, S_2 is longer than the path $R_1 \cup R_2 \cup S_1$, contradicting that S_2 has length $d_C(v,w)$. This proves (3).

We may assume that u, v are chosen, adjacent to q_1, q_k respectively, with $d_C(u,v)$ maximum. Then we have

(4) *q_1, q_k have no neighbours in L_2^*.*

Suppose this is false; then we may assume that $a_2 \in L_2^*$. From the choice of u, v, $d_C(v, a_2) \leq d_C(u,v)$, and since q_1 is not C-major, it follows that $d_C(u,v) = (|C|-1)/2$. By (2) and (3), one of $q_2, \ldots, q_{k-1}$ (say q_i) has a neighbour in L_1^*, say w. From the minimality of k, $w \notin V(P_1 \cup P_2)$, and so $w \in A^*$. Let R, S be the subpaths of A between w and a_1, b_2 respectively. From the minimality of k, the path a_1-q_1-$\cdots$-q_i-w has length at least the length of $P_1 \cup R$; and w-q_i-$\cdots$-q_k-v has length at least that of S. Adding, we deduce that $k+3$ is at least the length of $A \cup P_1$. But $d_C(u,v) \geq k+2$, and $d_C(u,v)$ is strictly less than the length of $A \cup P_1$; so $d_C(u,v) = k+2$, and u, a_1 are adjacent. Thus $d_C(a_1, v) = d_C(u,v)$, so the pair a_1, v satisfies the hypotheses defining u, v; and by exchanging u, a_1, we see that the edge $q_i w$ violates (3). This proves (4).

(5) *L_1 has length $k+2$.*

From (2) and (3), there exists $i \in \{2, \ldots, k-1\}$ such that q_i has a neighbour $w \in L_1^*$. From the minimality of k, the path u-q_1-$\cdots$-q_i-w has length at least the length of the subpath of L_1 between u, w; and the path w-q_i-$\cdots$-q_k-v has length at least the length of the subpath of L_1 between w, v. Adding, we deduce that $k+3$ is at least the length of L_1. But adding the path u-q_1-$\cdots$-q_k-v to L_2 makes a hole, of length at least ℓ since L_2 has length at least $|C|/2$; and this hole is shorter than C, and so has even length. Consequently $k+1$ has the same parity as the length of L_2, and hence k has the same parity as the length of L_1. Since the length of L_1 is at most $k+3$, it equals $k+2$. This proves (5).

Let the vertices of L_1 in order be $u = w_0$-w_1-w_{k+1}-$w_{k+2} = v$.

(6) *For* $1 \le i \le k$ *and* $1 \le j \le k+1$, *if* q_i, w_j *are adjacent then* $j \in \{i, i+1\}$.

If $j < i$ then the path w_j-q_i-$\cdots$-q_k-v has length less than the length of w_j-w_{j+1}-$\cdots$-w_{k+1}-v, since the first path has length $k-i+2$ and the second has length $k+2-j = d_C(w_j, v)$. This is contrary to the minimality of k. Similarly, if $j > i+1$, the path u-q_1-$\cdots$-q_i-w_j is shorter than u-w_1-$\cdots$-w_j, again a contradiction. This proves (6).

From (2), we may assume (exchanging u, v if necessary) that there exists $i \in \{2, \ldots, k-1\}$ such that q_i is adjacent to w_i. Let C' be the union of L_2 and the path

$$u\text{-}w_1\text{-}\cdots\text{-}w_i\text{-}q_i\text{-}\cdots\text{-}q_k\text{-}v.$$

From (3), (4) and (6), C' is a hole of length $|C|$, and so is a shortest long odd hole. Let G' be the subgraph of G induced on $V(C) \cup V(Q)$. Then C' is a hole in G', and moreover, in G' C' is a clean hole, because of (6). But the path u-q_1-$\cdots$-q_i is shorter then the path u-w_1-$\cdots$-w_i-q_i, and the latter has length $d_{C'}(u, q_i)$; and this contradicts the minimality of k. This proves 3.3. ∎

We need the following lemma.

3.4 *Let G be a graph containing no long jewel of order at most k, and no long odd hole of length less than $k+\ell$. Let C be a shortest long odd hole, and let $v \in V(G)$ be C-major. Then every path of C that contains all the neighbours of v in $V(C)$ has length more than k.*

Proof. Suppose that P is a path of C, with ends a, b say, containing all the neighbours of v in $V(C)$, and P has length at most k. Since v is C-major, it follows that P has length at least three. Let Q be the other path of C with ends a, b. Since by hypothesis, C has length at least $k+\ell$, it follows that Q has length at least ℓ. Adding the path a-v-b to Q therefore gives a long hole, and it is shorter than C since P has length at least three. Consequently this hole is not odd; so Q is even and so P is odd. But then the three paths P, a-v-b and Q form a long jewel of order at most k, a contradiction. This proves 3.4. ∎

Now we prove the second statement of 3.2, in the following.

3.5 *Let G be a candidate, and let C be a clean shortest long odd hole in G. Let $u, v \in V(C)$ be nonadjacent, and let Q be a shortest path in G joining u, v. Let L_1, L_2 be the paths of C that join u, v, where L_1 is shorter than L_2. Then $L_2 \cup Q$ is a clean shortest long odd hole.*

Proof. Let Q have vertices u-q_1-$\cdots$-q_k-v in order. We proceed by induction on k. By 3.3, L_1 and Q have the same length. If some vertex q_i of Q^* belongs to L_2^*, then by two applications of 3.3, $d_C(u, q_i) \le i$ and $d_C(q_i, v) \le k-i+1$, so L_2 has length at most $k+1$, which is impossible since L_1 has length $k+1$ and L_2 is longer. Thus $Q^* \cap L_2^* = \emptyset$, and so $L_2 \cup Q$ is a cycle, with the same length as C.

(1) $L_2 \cup Q$ *is induced.*

Suppose it is not induced. Then there exist $i \in \{1, \dots, k\}$ and $w \in L_2^*$ adjacent. Let R, S be the subpaths of L_2 between w and u, v respectively. From 3.3, $d_C(u, w) \le i+1$, and $d_C(w, v) \le k-i+2$, so L_2 has length at most $k + 3$. Since L_1 has length $k + 1$ and L_1, L_2 have opposite parity, it follows that L_2 has length $k + 2$, and either $d_C(u, w) = i + 1$ or $d_C(w, v) = k - i + 2$; and from the symmetry we may assume that $d_C(w, v) = k - i + 2$ and $d_C(u, w) \le i + 2$. Also, R has length $d_C(u, w)$ and S has length $d_C(w, v)$. Since the path w-q_i-$\cdots$-q_k-v has the same length as S, and the latter is a shortest path between v, w by 3.3, it follows that w-q_i-$\cdots$-q_k-v is also a shortest path between v, w. Suppose $i > 1$; then from the inductive hypothesis, the union of the path $L_1 \cup R$ and w-q_i-$\cdots$-q_k-v is a clean shortest long odd hole C' say. The two subpaths of C' between u, q_i have lengths $|E(L_1)| + k - i = 2k + 1 - i$ and $|E(R)| + 1$. By 3.3, one of these paths has length at most the length of u-q_1-$\cdots$-q_i, that is, at most i. But $2k + 1 - i > i$ since $i \le k$; and $|E(R)| + 1 > i$ since $|E(R)| = |E(L_2)| - |E(S)| = k + 2 - (k - i + 2) = i$, a contradiction. This proves (1).

Let $C' = L_2 \cup Q$; then C' is a shortest long odd hole. It only remains to check that C' is clean. Suppose not. Then there is a C'-major vertex x. Since x is not C-major, x has a neighbour in the interior of Q. Since Q is a shortest path, there is a subpath P_1 of Q of length at most two containing all neighbours of x in $V(Q)$; choose P_1 minimal. Since x is C'-major, it has a neighbour in the interior of L_2. Since x is not C-major, there is a path P_2 of L_2, of length at most two, containing all neighbours of x in $V(C)$; choose P_2 minimal. Let P_1 have ends a_1, b_1, where u, a_1, b_1, v are in order in Q, and let P_2 have ends a_2, b_2, where u, a_2, b_2, v are in order in L_2. Let A be the path of C' between a_1, a_2 that contains u, and let B be the path of C' between b_1, b_2 that contains v. By 3.4, the path $P_1 \cup A \cup P_2$ has length at least $\ell + 3$, and so does the path $P_1 \cup B \cup P_2$. In particular, P_1, P_2 are vertex-disjoint and do not contain u or v. Since C' has length at least $2\ell + 3$, one of A, B has length at least ℓ, say A. Hence the hole obtained by adding a_1-x-a_2 to A is long, and shorter than C, so even; and hence A has even length. The path $P_1 \cup B \cup P_2$ therefore has odd length, and so this path, a_1-x-a_2 and A form a long jewel, of order the length of $P_1 \cup B \cup P_2$; and so this path has length at least $\ell + 3$. If P_1 has length two, let it have vertices q_{i-1}-q_i-q_{i+1}; then the path

$$u\text{-}q_1\text{-}\cdots\text{-}q_{i-1}\text{-}v\text{-}q_{i+1}\text{-}\cdots\text{-}q_k\text{-}v$$

is a path between u, v with the same length as Q, and violates (1). Thus P_1 has length at most one. Consequently B has length at least ℓ; and so B is even. Since C' is odd, and A, B are both even, it follows that exactly one of P_1, P_2 is odd. If P_2 is odd, that is, if a_2, b_2 are adjacent, then since P_2 is even and has length at most one, it follows that $a_1 = b_1$; and the three paths A, B and a_1-x form a long pyramid. Thus P_2 is even and P_1 is odd. If $a_2 = b_2$ then similarly the three paths A, B and a_2-x form a long pyramid; so P_2 has length two. Since x is not C-major, it has no neighbours in L_1. Since P_1 is odd, it follows that $k \ge 2$, so from the inductive hypothesis, the hole obtained from C by replacing the middle vertex of P_2 by x is a clean shortest long odd hole, violating (1). This proves that C' is clean, and so completes the inductive proof of 3.5. ∎

Now we can give the main result of this section. (In the following, we could bring the running time down to $O(|G|^4)$, but there is no need.)

3.6 *There is an algorithm with the following specifications:*

Input: *A candidate G.*

Output: *Decides either that G has a long hole, or that there is no clean shortest long odd hole in G.*

Running time: $O(|G|^5)$.

Proof. If C is a shortest long odd hole in G, let $u, v, w \in V(C)$ be chosen such that each of $d_C(u,v), d_C(v,w), d_C(w,u)$ equals either $\lfloor |C|/3 \rfloor$ or $\lceil |C|/3 \rceil$. Here is the algorithm: guess u, v, w, find a shortest path between each pair of them, and test whether these three paths make a long odd hole. If so, output that G has an odd hole. After checking all triples, if none has produced an odd hole, output that G has no clean shortest long odd hole. It follows immediately from 3.2 that the output is correct. ∎

4 Covering by a short path

Now we begin the third step of the main algorithm: cleaning a shortest long odd hole. In this section we prove some preliminary lemmas. Let C be a shortest long odd hole in G, and let x be C-major. An *x-gap* means a path of C of length at least two, with both ends adjacent to x and with no interior vertices adjacent to x; so if P is an x-gap then adding x gives a hole. If x, y are distinct nonadjacent C-major vertices in G, an *(x, y)-gap* means a path P of C such that $V(P)$ is the interior of an induced path of G between x and y. If x, y are adjacent C-major vertices, an *(x, y)-gap* means a path P of C such that $V(P)$ is the interior of an induced path of $G \setminus e$ between x and y, where e is the edge xy. Two sets of vertices X, Y are *anticomplete* if they are disjoint and there are no edges between them. We use the same term for two subgraphs P, Q; thus we say P is *anticomplete* to Q if $V(P)$ is anticomplete to $V(Q)$.

4.1 *Let C be a shortest long odd hole in a candidate G, and let u, v be nonadjacent C-major vertices. Then there is a (u, v)-gap of length less than $\ell/2 - 1$.*

Proof. Let A, B be the sets of neighbours of u, v in $V(C)$ respectively. We may assume that u, v have no common neighbour in $V(C)$, because that would make an (x, y)-gap of length zero; and so there are an even number of (u, v)-gaps. Let us number them $D_1, \ldots, D_{2k}$ say, in order in C. We may assume that each D_i has length at least $\ell/2 - 1$, and in particular, have length at least two, since $\ell \geq 5$. For $1 \leq i \leq 2k$, let D_i have ends $a_i \in A$ and $b_i \in B$. We say $i, j \in \{1, \ldots, 2k\}$ are *consecutive* if either $|j - i| = 1$ or $|j - i| = 2k - 1$.

(1) *For all $i, j \in \{1, \ldots, 2k\}$, if D_i, D_j have different parity then i, j are consecutive.*

Suppose not; then we may assume that D_1, D_i have opposite parity where $3 \leq i \leq 2k - 1$. Let a_1' be the vertex of C adjacent to a_1 that is not in D_1, and define b_1' similarly. Since $i \neq 2, 2k$ it follows that $a_1, b_1 \notin V(D_i)$. If $a_1' \in V(D_i)$ then since $i, 1$ are not consecutive, it follows that $\{a_1, a_1'\}$ includes the vertex set of a (u, v)-gap; but this is impossible since all (u, v)-gaps have length at least two. So $a_1', b_1' \notin V(D_i)$, and adding u, v to $D_1 \cup D_i$ gives a long odd hole, shorter than C, a contradiction. This proves (1).

(2) $D_1, \dots, D_{2k}$ *all have the same parity.*

Suppose not; then $k \le 2$, and we may assume that D_i has the same parity as i for $1 \le i \le 2k$. If say $a_1 = a_2$, then adding v to $D_1 \cup D_2$ gives a long odd hole shorter than C (note that $b_1 \neq b_2$ and they are nonadjacent, since u has at least two neighbours); so $D_1, \dots, D_{2k}$ are pairwise vertex-disjoint. Since D_1, D_2 have opposite parity, they are not anticomplete, and we may assume that a_1 is adjacent to one of a_2, b_2; and not to b_2 since all (u,v)-gaps have length at least two. So a_1, a_2 are adjacent. Since u is C-major, it follows that $k \ge 2$, and so $k = 2$. Similarly either a_2, a_3 are adjacent, or b_2, b_3 are adjacent, and the first is impossible since a_2 is adjacent to a_1 and has a neighbour in the interior of D_2. So b_2b_3, a_3a_4, b_4b_1 are edges; but then C has even length, a contradiction. This proves (2).

From (2), an even number of edges of C belong to (u,v)-gaps. Let F be the graph obtained from C by deleting the edges and internal vertices of every (u,v)-gap. Since C is odd, it follows that $|E(F)|$ is odd, and so some component P of F is odd. Now the ends of P belong to two consecutive (u,v)-gaps, D_1, D_2 say; and so we may assume that P has ends a_1, a_2. But then $D_1 \cup P \cup D_2$ is odd, and adding v makes a long odd hole of length less than C, a contradiction. This proves 4.1. ∎

A similar proof yields:

4.2 *Let C be a shortest long odd hole in a candidate G, and let u, v be nonadjacent C-major vertices. Suppose that all (u,v)-gaps are odd; then for every long (u,v)-gap Q, there is a (u,v)-gap of length at most $\ell - 5$, anticomplete to Q.*

Proof. Since an even number of edges belong to (u,v)-gaps, it follows as in the proof of 4.1 (and using the notation of that proof) that some component P of F is odd, and we may assume that P has ends a_1, a_2. Then $D_1 \cup P \cup D_2$ is odd, and adding v makes an odd hole of length less than that of C, which therefore has length less than ℓ. Since P is odd, it follows that the sum of the lengths of D_1, D_2 is at most $\ell - 4$, and so each has length at most $\ell - 5$. Suppose that neither of them is anticomplete to Q. Since Q exists and $Q \neq D_1, D_2$, there are at least four (u,v)-gaps; let $Q = D_i$ say where $3 \le i \le 2k$. Since D_i, D_2 are not anticomplete, it follows that $i \le 4$; and similarly $i \ge 2k - 1$, and so $k = 2$, and we may assume that $i = 3$ from the symmetry. Since D_1, D_3 are not anticomplete, it follows that $b_1 = b_4$, $a_3 = a_4$, and a_4, b_4 are adjacent. Also since D_2, D_3 are not anticomplete, the vertices b_2, b_3 are either equal or adjacent; and since C is odd and D_1, D_2, D_3, D_4 and P are odd, it follows that $b_2 = b_3$. Now the three paths $D_1 \cup P \cup D_2$, b_4-v-b_2 and D_3 form a long jewel of order $|V(D_1 \cup P \cup D_2)| < \ell$, a contradiction. This proves 4.2. ∎

The next result is crucial, and provides the machinery behind all the proof of 1.1 that is novel.

4.3 *Let P be a path of odd length, with ends p, p'. Let $A_1, \dots, A_k \subseteq V(P)$ be nonempty, and let us say a subpath Q of P is "covering" if $V(Q) \cap A_s \neq \emptyset$ for $1 \le s \le k$. Suppose that the minimal covering subpath with one end p, and the minimal covering subpath with one end p', have the same parity. Then there is an odd subpath Q of P such that Q is covering, and for some (possibly equal) $s, t \in \{1, \dots, k\}$, one end of Q belongs to A_s, the other end belongs to A_t, and A_s, A_t contain no other vertex of Q.*

Proof. Let P have vertices p_0-$\cdots$-p_n in order. Thus n is odd. Choose $d \in \{0, \dots, n\}$ minimum such that p_0-$\cdots$-p_d is covering. For $d \le j \le n$, choose $i \le j$ maximum such that p_i-$\cdots$-p_j is covering and define $m(j) = j - i$. If $m(d)$ is odd, then the subpath of P between $p_{d-m(d)}$ and p_d satisfies the theorem; so we assume that $m(d)$ is even. Also, from the hypothesis, $m(n)$ and d have the same parity, so $m(n) + n$ and $m(d) + d$ have different parity. Consequently there exists i with $d+1 \le i \le n$ such that $m(i)+i$ has different parity from $m(i-1)+i-1$. It follows that $m(i), m(i-1)$ have the same parity. Choose $h \le i-1$ maximum such that p_h-$\cdots$-p_{i-1} is covering. Consequently $i - 1 - h = m(i-1)$. From the maximality of h, there exists $t \in \{1, \dots, k\}$ such that $p_h \in A_t$, and $p_{h+1}, \dots, p_{i-1} \notin A_t$. Since $m(i), m(i-1)$ have the same parity, it follows that one of $p_{h+1}, \dots, p_{i-1}, p_i$ belongs to A_t, and so $p_i \in A_t$. If i, h have opposite parity, the path p_h-$\cdots$-p_i satisfies the theorem, so we assume they have the same parity. Choose $g \le i$ maximum such that p_g-$\cdots$-p_i is covering. Thus $i - g = m(i)$. From the maximality of g, there exists $s \in \{1, \dots, k\}$ such that $p_g \in A_s$ and $p_{g+1}, \dots, p_i \notin A_s$. Since $m(h-1), m(h)$ have the same parity, it follows that $i-1-h$ and $i-g$ have the same parity, that is, $g - h$ is odd. But i, h have the same parity, so $i - g$ is odd, and the path p_g-$\cdots$-p_i satisfies the theorem. This proves 4.3. ∎

5 Bases

In this section we prepare to apply 4.3 to generate a cleaning list.

5.1 *Let C be a shortest long odd hole in a candidate G, and let u, v be nonadjacent C-major vertices. There is not both a long odd (u, v)-gap and a long even (u, v)-gap.*

Proof. Let P be a long odd (u, v)-gap, and let Q be a long even (u, v)-gap. They are not anticomplete, since otherwise adding u, v to their union gives a long odd hole, shorter than C, a contradiction. If they share a vertex, then their union is either a long odd u-gap or a long odd v-gap, a contradiction. So they are vertex-disjoint. Let P have ends p_1, p_2, and let Q have ends q_1, q_2, where u is adjacent to p_1, q_1 and v to p_2, q_2. We may assume that one of q_1, q_2 is adjacent to p_2. If q_1 is adjacent to p_2 then v-p_2-q_1-Q-q_2-v is a long odd hole shorter than C. So q_2 is adjacent to p_2. Let R be the path of C joining p_1, q_1 that does not contain p_2, q_2. Thus R has odd length, and so the three paths R, p_1-u-q_1 and p_1-P-p_2-q_2-Q-q_1 form a long jewel; and therefore this jewel has order at least $\ell + 3$. Consequently R has at least $\ell + 4$ vertices. Let r_1, r_2 be the neighbours of p_1, p_2 respectively in R. Suppose that v has no neighbours in $V(R) \setminus \{p_1, r_1, r_2, p_2\}$. Since v is C-major, it is adjacent to at least one of r_1, r_2; not to exactly one, since it would make a long pyramid with C, and not with both since then v-r_1-R-r_2-v is a long odd hole shorter than C; in each case a contradiction. So v has a neighbour in $V(R) \setminus \{p_1, r_1, r_2, p_2\}$. Now suppose that u has no neighbour in this set. If u is adjacent to neither of r_1, r_2 then u-p_1-R-p_2-u is a long odd hole shorter than C; if u is adjacent to both r_1, r_2 then u-r_1-R-r_2-u is a long odd hole shorter than C; and if u is adjacent to exactly one of r_1, r_2, it makes a long pyramid with C, in each case a contradiction. So u also has a neighbour in $V(R) \setminus \{p_1, r_1, r_2, p_2\}$. But then there is a path joining u, v with interior in $V(R) \setminus \{p_1, r_1, r_2, p_2\}$; and this path, with u-p_1-P-p_2 and u-q_1-Q-q_2, forms a long pyramid, a contradiction. This proves 5.1. ∎

Let C be a shortest long odd hole in a candidate G. If C is not clean, there is a maximal path D of C such that some C-major vertex (x say) is adjacent to its ends and not to any of its internal vertices. We call (x, D) a *base* (for C in G). A base (x, D) is *remote* if D has length at least 2ℓ, and no C-major vertex different from x has has a neighbour $w \in V(C)$ such that $d_C(w, d_i) < \ell$ for some $i \in \{1, 2\}$, where d_1, d_2 are the ends of D. (If there is a C-major vertex different from x, the second condition just given implies the first.) It is easy to "arrange" algorithmically that a base is remote: we just guess the two paths of C of length 2ℓ with middle vertex d_1, d_2 respectively, and delete all vertices not in these paths with a neighbour in the interior of one of them, except x. (This is safe, because no vertex of C will be deleted.) If any C-major vertex different from x remains, then the base has become remote (and if they all have been deleted then we have won). So the theorems to come will often assume that (x, D) is a remote base. If (x, D) is a remote base then D is even, since adding v to D gives a long hole shorter than C.

In what follows, when we have a base (x, D), we will always denote the ends of D by d_1, d_2. If v is a C-major vertex nonadjacent to x, then v has a neighbour in $V(D)$ from the maximality of D, and since (x, D) is remote, there are two (x, v)-gaps included in D, $D_1(v)$ and $D_2(v)$ say, where for $i = 1, 2$, $D_i(v)$ has ends d_i and $d_i(v)$ say. Both of them are long, since (x, D) is remote, so they have the same parity by 5.1.

We will use this notation throughout the paper without defining it again.

5.2 *Let C be a shortest long odd hole C in a candidate G, and let (x, D) be a remote base. Then for every C-major vertex v nonadjacent to x, all (x, v)-gaps have the same parity.*

Proof. Let D' be the path of C different from D with ends d_1, d_2. Since D is even (because adding x gives a long hole shorter than C) it follows that D' is odd. Thus the three paths D', d_1-x-d_2 and D form a long jewel, and so D' has length at least $\ell + 3$. Consequently x has a neighbour in C that is different from and nonadjacent to d_1, d_2 (to see this, suppose not; then x has at most four neighbours in C, and if three then it makes a long pyramid, and if two or four then it makes a long odd hole shorter than C). Hence every (x, v)-gap different from $D_1(v), D_2(v)$ is anticomplete to one of $D_1(v), D_2(v)$, and therefore has the same parity as $D_1(v), D_2(v)$. This proves 5.2. ∎

Let (x, D) be a remote base for C. If v is a C-major vertex nonadjacent to x, we say the *x-parity* of v is the common parity of all (x, v)-gaps.

5.3 *Let C be a shortest long odd hole in a candidate G, and let (x, D) be a remote base. Let v_1, v_2 be nonadjacent C-major vertices, nonadjacent to x, and with the same x-parity, and let Q be a long odd (v_1, v_2)-gap edge-disjoint from D. Then x has no neighbour in Q.*

Proof. Let Q have ends q_1, q_2, where v_1q_1 and v_2q_2 are edges. Since q_i is adjacent to v_i, it follows that $d_C(q_i, d_j) \geq \ell$ for $j = 1, 2$; and so Q is anticomplete to D.

Suppose x has a neighbour in $V(Q)$. If x has no neighbour in the interior of Q, then x is adjacent to one of q_1, q_2, so all (x, q_1)- and (x, q_2)-gaps are even, and consequently Q is not an (x, q_1) or (x, q_2)-gap; and therefore x is adjacent to both q_1, q_2, and adding x to Q would give a long odd hole shorter than C, a contradiction. So x has a neighbour in the interior of Q.

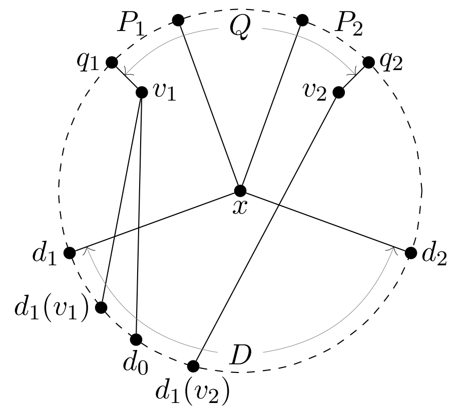

Figure 3: For 5.3.

Consequently there is an (x, v_1)-gap P_1 that is anticomplete to D and v_2 has no neighbour in $V(P_1)$; and also there is P_2 similarly. For $i = 1, 2$, let P_i^+ be the path between v_i, x with interior $V(P_i)$. Now every (v_1, v_2)-gap included in D is anticomplete to Q, and since Q is odd and has length at least ℓ, it follows that every (v_1, v_2)-gap included in D has odd length. In particular, $d_1(v_1) \neq d_1(v_2)$. From the symmetry we may assume that $D_1(v_2)$ is a proper subpath of $D_1(v_1)$. Choose a vertex d_0 of the path $d_1(v_1)$-D-$d_1(v_2)$, adjacent to v_2, such that the subpath $d_1(v_1)$-D-d_0 is minimal. The latter is a (v_1, v_2)-gap, and so odd. On the other hand, since v_1, v_2 have same x-parity, the paths $D_1(v_1), D_1(v_2)$ have the same parity, and so $d_1(v_1)$-D-$d_1(v_2)$ is even; and hence $d_1(v_2)$-D-d_0 is odd. If $d_1(v_2), d_0$ are nonadjacent, then the path d_0-D-$d_1(v_1)$-v_1-P_1^+-x-d_1-D-$d_1(v_1)$ can be extended to a hole by adding either the path $d_1(v_2)$-D-d_0 or $d_1(v_2)$-v_2-d_0, and since these two paths have different parity, one of these holes is odd. It is long since d_1-D-$d_1(v_2)$ has length at least ℓ; and it is shorter than C, a contradiction. Thus $d_1(v_2), d_0$ are adjacent. Choose a minimal subpath R of D with one end d_2 such that the other end, r say, is adjacent to one of v_1, v_2. Thus R has length at least ℓ, and only one of v_1, v_2 has a neighbour in $V(R)$. If r is adjacent to v_1, the three paths x-d_2-R-r-v_1-$d_1(v_1)$-D-d_0, x-d_1-D-$d_1(v_2)$ and x-P_2^+-v_2 form a long pyramid (omitting v_1 from the first if $d_1(v_1), d_0$ are equal or adjacent). If r is adjacent to v_2, the three paths x-d_2-R-r-v_2, x-d_1-D-$d_1(v_1)$ and x-P_1^+-v_1-$d_1(v_1)$-D-d_0 form a long pyramid, a contradiction. This proves 5.3. ∎

6 Catch and clean

Let C be a shortest long odd hole, and let P be a path of C. If v is a vertex, we say that P *catches* v if $v \notin V(P)$ and v is adjacent to an internal vertex of P. The point is, if P is a path of C then we are sure that no vertices it catches belong to $V(C)$; they might not all be C-major, but it does no harm to delete them. This is a quite effective way to clean C of C-major vertices: we guess a path of bounded length (or a bounded number of such paths) and delete all the vertices each one catches. For instance, if we could prove that there is such a set of paths of C (of bounded size, and each of bounded length) that together catch all the C-major vertices, we could clean C and be done. We have not been able to prove or disprove this. Nonetheless, catching C-major vertices by paths is a useful technique, and we will develop it in this section. If $\mathcal{F}$ is a set of paths of C, its *cost* is the number of vertices in the union of the paths; and it *catches* v if one of its paths catches v.

6.1 *Let C be a shortest long odd hole in a candidate G, and let (x, D) be a remote base. Let M be a set of C-major vertices, all nonadjacent to x and with the same x-parity. Let $v_1, v_2 \in M$ be adjacent, and let Q be an odd (v_1, v_2)-gap, with ends q_1, q_2, edge-disjoint from D, where v_i, q_i are adjacent for $i = 1, 2$. Suppose that x has a neighbour in $V(Q)$ and every vertex in M has a neighbour in $V(Q)$. Then there is a set $\mathcal{F}$ of paths of C with cost at most $5\ell + 15$ that catches all the vertices in M.*

Proof.

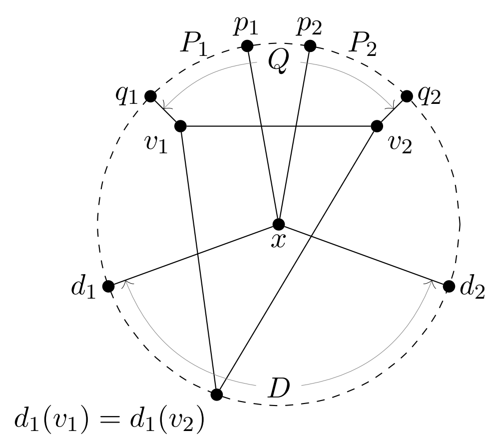

Figure 4: For 6.1.

For $i = 1, 2$, let P_i be the (x, v_i)-gap included in Q, and let its ends be q_i, p_i.

(1) *$d_1(v_1) = d_1(v_2)$.*

Suppose not; then we can assume that $D_1(v_1)$ is a proper subpath of $D_1(v_2)$. But then the holes v_2-$d_1(v_2)$-$D_1(v_2)$-d_1-x-p_2-P_2-q_2-v_2 and v_2-v_1-$d_1(v_1)$-$D_1(v_1)$-d_1-x-p_2-P_2-q_2-v_2 have opposite parity, and are both long and shorter than C, a contradiction. This proves (1).

(2) *p_1, p_2 are distinct and adjacent.*

Since P_1, P_2 have the same parity, it follows that the path p_1-Q-p_2 is odd, and in particular $p_1 \neq p_2$. If p_1, p_2 are nonadjacent, there is a long pyramid formed by the paths x-p_1-P_1-q_1-v_1, x-p_2-P_2-q_2-v_2 and x-d_1-$D_1(v_1)$-$d_1(v_1)$, a contradiction. This proves (2).

For each $v \in M$ and for $i = 1, 2$, let $Q_i(v)$ be the (v, v_i)-gap in Q, and let $q_i(v)$ be the neighbour of v in $Q_i(v)$. Let M_0 be the set of $v \in M$ such that one of $Q_1(v), Q_2(v)$ has length less than ℓ, or one of the (x, v)-gaps in Q has length less than ℓ. We can catch M_0 with a set of three paths, two of length $\ell + 1$ and one of length $2\ell + 1$, so with cost at most $4\ell + 6$.

(3) *Every vertex in $M \setminus M_0$ is adjacent to one of v_1, v_2.*

Suppose that $v \in M \setminus M_0$ and v is nonadjacent to both v_1, v_2. Suppose first that v has a neighbour in P_1 and a neighbour in P_2. Then there is a long pyramid formed by the two (x, v)-gaps in Q each extended by v, and the path v-$d_1(v)$-$D_1(v)$-d_1-x, a contradiction. Thus we may assume that all neighbours of v in $V(Q)$ belong to $V(P_1)$. Since x has neighbours in $Q_2(v)$, and $Q_2(v)$ is long, it is even by 5.3; but then by (2), the (v, x)-gap and the (x, v_2)-gap in $Q_2(v)$ have different parities, a contradiction. This proves (3).

For $i = 1, 2$, let M_i be the set of $v \in M \setminus M_0$ that are adjacent to v_i. By 4.1, there is a (v_1, x)-gap T of length less than $\ell/2 - 1$; let T^+ be the path between v_1, x with interior $V(T)$.

(4) *Every vertex in M_1 is adjacent to $d_1(v_1)$ or to a vertex in $V(T)$.*

Let $v \in M_1$, and suppose it is not adjacent to $d_1(v_1)$ and has no neighbour in $V(T)$. In particular, $d_1 \neq d_1(v_1)$, so one of $D_1(v), D_1(v_1)$ is a proper subpath of the other. If $D_1(v_1)$ is a proper subpath of $D_1(v)$, choose a path S joining v, x with interior in the interior of Q; then since S and $D_1(v_1)$ have the same parity, it follows that the hole v-v_1-$d_1(v_1)$-$D_1(v_1)$-d_1-S-v is a long odd hole shorter than C, a contradiction. If $D_1(v)$ is a proper subpath of $D_1(v_1)$, then since T and $D_1(v)$ have the same parity, the hole v_1-v-$d_1(v)$-$D_1(v)$-d_1-x-T^+-v_1 is a long odd hole shorter than C, a contradiction. This proves (4).

From (4), we can catch M_1 with a set of paths of C with cost at most $\ell/2 + 9/2$. The same holds for M_2. Since we can catch M_0 with cost $4\ell + 6$, and $M_0 \cup M_1 \cup M_2 = M$, the result follows. This proves 6.1. ∎

As a consequence we have:

6.2 *Let C be a shortest long odd hole in a candidate G, and let (x, D) be a remote base. Then there is a set of paths of C with cost at most $5\ell + 15$ that catches all C-major vertices that both are nonadjacent to x and have even x-parity.*

Proof. Let M be the set of all C-major vertices nonadjacent to x with even x-parity. Let P be the path of C obtained by deleting the interior of D. Every vertex in $M \cup \{x\}$ has at least two neighbours in $V(P)$. Moreover, for each $v \in M \cup \{x\}$, the shortest subpath of P with one end d_1 and the other adjacent to v has the same parity as $D_1(v)$ and so is even; and the same holds for d_2. Thus $M \cup \{x\}$ and P satisfy the hypotheses of 4.3, and so there is an odd path Q of C with ends q_1, q_2, edge-disjoint from D, and there are vertices $v_1, v_2 \in M \cup \{x\}$, such that v_1q_1 and v_2q_2 are edges and there are no other edges between $\{v_1, v_2\}$ and $V(Q)$; and every vertex in $M \cup \{x\}$ has a neighbour in $V(Q)$. If Q has length less than ℓ the result follows, so we assume Q is long. Hence $v_1 \neq v_2$. Now x is nonadjacent to all vertices in M, and they all have even x-parity; so $v_1, v_2 \neq x$. By 5.3, v_1, v_2 are adjacent, and the result follows from 6.1. This proves 6.2. ∎

6.3 *Let C be a shortest long odd hole in a candidate G, and let (x, D) be a remote base. Then there is a set of paths of C with cost at most $16\ell + 31$ that catches all C-major vertices nonadjacent to x.*

Proof. By 6.2 there is a set $\mathcal{F}_1$ of paths of C with cost at most $5\ell + 15$ that catches all C-major vertices nonadjacent to x with even x-parity. Let M_1 be the set of all C-major nonneighbours of x not caught by $\mathcal{F}_1$. Now let us apply 4.3 to the path P of C obtained by deleting the interior of D, and the set $M_1 \cup \{x\}$. We may assume that $M_1 \neq \emptyset$. Again, the hypotheses of 4.3 are satisfied, since for each $v \in M_1$, the shortest subpath of P with one end d_1 and the other adjacent to v has the same parity as $D_1(v)$ and so is odd; and the same holds for d_2; and so for $i = 1, 2$ the shortest subpath of P with one end d_i containing a neighbour of each vertex in $M_1 \cup \{x\}$ is odd. (Note that although x is adjacent to d_i, the members of M_1 are not adjacent to d_i, and $M_1 \neq \emptyset$, so this shortest subpath does not have length zero.)

It follows that there is an odd path Q of C with ends q_1, q_2, edge-disjoint from D, and there are vertices $v_1, v_2 \in M \cup \{x\}$, such that v_1q_1 and v_2q_2 are edges and there are no other edges between $\{v_1, v_2\}$ and $V(Q)$; and every vertex in $M_1 \cup \{x\}$ has a neighbour in $V(Q)$. If Q has length less than ℓ the result follows, so we assume Q is long. Hence $v_1 \neq v_2$.

First, suppose that $v_1, v_2 \neq x$. By 5.3, v_1, v_2 are adjacent; and by 6.1 the result holds. So we may assume that $v_2 = x$ say. Let B be the v_1-gap that includes Q. Then (v_1, B) is a base for C in $G \setminus N_1$, where N is the set of C-major neighbours of x, and N_1 is the union of N with the set of C-major vertices caught by $\mathcal{F}_1$. Let $\mathcal{F}_2$ be the set of the two paths of C of length 2ℓ with middle vertices the ends of F, and let N_2 be the union of N_1 with the set of C-major vertices caught by $\mathcal{F}_2$; then (v_1, B) is a remote base for C in $G \setminus N_2$. By 6.2 there is a set $\mathcal{F}_3$ of paths of C with cost at most $5\ell + 15$ that catches all C-major vertices v of $G \setminus N_2$ nonadjacent to v_1 and with even v-parity. Let N_3 be the union of N_2 with the set of C-major vertices caught by $\mathcal{F}_3$.

By 4.2 there is an (x, v_1)-gap T of length at most $\ell - 5$ anticomplete to Q; and it follows that T is vertex-disjoint from and anticomplete to D. let T^+ be the path between x, v_1 with interior $V(T)$. Let T have ends t_1, t_2, where t_1v and t_2x are edges. For each C-major vertex v of $G \setminus N_3$, let $Q_1(v)$ be the (v_1, v)-gap in Q, with ends $q_1(v)$ and q_1, and let $Q_2(v)$ be the (x, v)-gap in Q, with ends $q_2(v)$ and q_2. Note that $Q_1(v)$ is long, since (v_1, B) is a remote base.

(1) *For every C-major vertex v of $G \setminus N_3$, either v has a neighbour in $V(T)$, or $Q_2(v)$ has length less than ℓ.*

Let v be a C-major vertex of $G \setminus N_3$ adjacent to v_1, and suppose v has no neighbour in $V(T)$, and $Q_2(v)$ is long. Suppose first that v, v_1 are adjacent. Now T is odd, since v_1 has odd x-parity; and also $Q_2(v)$ is odd, since v has odd x-parity. But v_1 has no neighbour in $Q_2(v)$, since q_1 is the only neighbour of v_1 in Q, and $q_1 \notin V(Q_2(v))$ since v, q_1 are not adjacent. Since T is anticomplete to Q and hence to $Q_2(v)$, the hole v_1-v-$q_2(v)$-$Q_2(v)$-q_2-x-T^+-v_1 is a long odd hole shorter than C, a contradiction.

Now suppose that v, v_1 are nonadjacent. Since Q, $Q_1(v)$ and $Q_2(v)$ are odd, and $Q_1(v)$, $Q_2(v)$ are edge-disjoint, it follows that the path $q_1(v)$-Q-$q_2(v)$ is odd. There is an induced path R between q_1, q_2 with interior in $V(T) \cup \{x, v_1\}$, and there are no edges between the interior of R and the interior of Q, or between the interior of R and v. Consequently if $q_1(v), q_2(v)$ are nonadjacent, the paths $q_1(v)$-v-$q_2(v)$, $q_1(v)$-Q-q_2 and $q_1(v)$-$Q_1(v)$-$q)1$-R-q_2-$Q_2(v)$-$q_2(v)$ form a long jewel, and so there is a long odd hole of length less than C, a contradiction. Hence $q_1(v), q_2(v)$ are adjacent. Since the paths $Q_1(v)$ and T are disjoint and anticomplete, there is a long pyramid formed by v_1-q_1-$Q_1(v)$-$q_1(v)$, a path between v_2, q_2 with interior in $V(T) \cup \{x\} \cup V(Q_2(v))$ and a path between v_1, v with interior in the interior of D, a contradiction. This proves (1).

It follows from (1) that there is a set $\mathcal{F}_4$ of paths of C with cost at most $2\ell - 1$ that catches every C-major vertex of $G \setminus N_3$. Then $\mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{F}_4$ catches every C-major vertex of G that is nonadjacent to x. This set has cost at most

$$(5\ell + 15) + (4\ell + 2) + (5\ell + 15) + (2\ell - 1) = 16\ell + 31.$$

This proves 6.3. ∎

7 The algorithm

In this section we will use the results of the previous sections to give our main theorem. We need

7.1 *There is an algorithm with the following specifications:*

Input: *A candidate G.*

Output: *Either decides that G has a long odd hole, or decides that G does not have a shortest long odd hole C and a remote base (x, D) such that all C-major vertices are equal or adjacent to x.*

Running time: $O(|G|^9)$.

Proof. Suppose that C is a shortest long odd hole with a remote base (x, D), and every other C-major vertex is adjacent to x. Here is an algorithm: we guess x and the ends d_1, d_2 of the path D, and its middle vertex d_0 (D has even length). Let Z be the set consisting of x and all its neighbours. For $i = 1, 2$, compute the set D_1 of internal vertices of all shortest paths with interior in $G \setminus Z$ between d_i, d_0. Let Y be the set of vertices of G not in $D_1 \cup D_2 \cup \{d_0, d_1, d_2\}$ and with a neighbour in $D_1 \cup D_2 \cup \{d_0\}$. Apply 3.6 to $G \setminus (Y \cup \{x\})$. If it finds a long odd hole, output that G has a long odd hole. If (after checking all possible guesses) we did not find a long odd hole, output that G has no shortest long odd hole C and base (x, D) such that every other C-major vertex is adjacent to x.

To see correctness, we only need to check correctness when G has a shortest long odd hole C and a remote base (x, D), and every other C-major vertex is adjacent to x. In this case, when we guess correctly, the interior of D contains no vertices in Z; and since the subpath of D between d_1, d_0 has length less than $|C|/2$, it is a shortest path of G between d_1, d_0 containing no C-major vertices, by 3.2. Any shortest path with interior in $G \setminus Z$ between d_1, d_0 will contain no C-major vertices, because they all belong to Z. Hence the interior of D is a subset of the set $D_1 \cup D_2 \cup \{d_0\}$ computed by the algorithm. Also by 3.2, $V(C) \setminus V(D)$ is anticomplete to $D_1 \cup D_2 \cup \{d_0\}$; and so the set Y computed by the algorithm contains no vertices of C. But it does contain all the C-major vertices except x; because they all have neighbours in the interior of D, and do not belong to $D_1 \cup D_2$ since $D_1 \cup D$ is disjoint from Z. Hence after deleting $Y \cup \{x\}$, all C-major vertices have been deleted and C has become a clean shortest long odd hole, and the algorithm of 3.6 will detect a long odd hole; and so the output is correct.

For running time, there are $|G|^4$ guesses to check, and each one takes time $O(|G|^5)$ (because we are applying 3.6). Thus the total running time is $O(|G|^9)$. This proves 7.1. ∎

Now our main result 1.1, which we restate:

7.2 *There is an algorithm with the following specifications:*

Input: *A graph G.*

Output: *Decides whether G has a long odd hole.*

Running time: $O(|G|^{20\ell+43})$.

Proof. First we apply 2.4, and we may assume we determine that G is a candidate. Next we apply 3.6 to G, and we assume we find that there is no clean shortest long odd hole. So either G has no long odd hole, or it has a shortest long odd hole C with a base (x, D). We assume the latter. Let R_1, R_2 be the paths of C of length 2ℓ with middle vertices the ends d_1, d_2 of D. We guess x and R_1, R_2, and delete all vertices caught by $\{R_1, R_2\}$ different from x, producing H say, and run 3.6 on $H \setminus \{x\}$; and assuming it still does not find a long odd hole, and G has a long odd hole, then when we guess correctly, H has a shortest long odd hole C with a remote base (x, D).

Now by 6.3 there is a set $\mathcal{F}$ of paths of C, with cost at most $16\ell + 31$, that catches every C-major vertex of H nonadjacent to x. We guess the paths in $\mathcal{F}$, and delete all the vertices they catch. Then we apply 7.1 to the resulting graph. If it finds a long odd hole we are done; and it will do so in the case when we guess correctly. If after all guesses we never find a long odd hole, we return that there is none.

The total cost of guesses for the paths is $(4\ell + 2) + (16\ell + 31) = 20\ell + 33$, and we also have to guess x; and checking each guess takes time $O(|G|^9)$, since we are applying 7.1. Thus the total running time is $O(|G|^{20\ell+43})$. This proves 7.2. ∎

References

[1] D. Bienstock, "On the complexity of testing for odd holes and induced odd paths", *Discrete Math.* 90 (1991), 85–92.

[2] D. Bienstock, "Corrigendum: On the complexity of testing for odd holes and induced odd paths". *Discrete Math.* 102 (1992), 109.

[3] M. Chudnovsky, A. Scott, P. Seymour and S. Spirkl, "Detecting an odd hole", submitted for publication (manuscript December 2018), `arXiv:1903.00208`.

[4] M. Chudnovsky, G. Cornuéjols, X. Liu, P. Seymour and K. Vušković, "Recognizing Berge graphs", *Combinatorica*, 25 (2005), 143–186.