Local improvement algorithms for a path packing problem: A performance analysis based on linear programming


Abstract

Given a graph, we wish to find a maximum number of vertex-disjoint paths of length 2. We propose a series of local improvement algorithms for this problem, and present a linear-programming-based method for analyzing their performance.

1. Introduction

Given a graph $G = (V, E)$, we wish to find a maximum number of vertex-disjoint paths of length 2. For this NP-hard problem [7], we propose a series of local improvement algorithms. The basic heuristic, denoted $H_0$, applies the greedy algorithm to obtain a maximal path packing. For $k \geq 1$, the $k$th heuristic in the series, denoted $H_k$, starts from a maximal path packing and attempts to improve it by replacing any $k$ paths of length 2 by $k + 1$ paths of length 2; when no further improvements are found, $H_k$ terminates. For fixed $k$, $H_k$ runs in polynomial time, but $H_0$ is unlikely to run in time polynomial in $k$, unless P = NP [2].

In this paper we are interested in the worst-case performance ratios $\rho_k$ of the heuristics $H_k$ ($k \geq 0$) to the optimum. We will show that $\rho_0 = 1/3$, $\rho_1 = 1/2$, $\rho_2 = 5/9$, $\rho_3 = 7/11$, and $\rho_4 = 2/3$, and that these ratios are tight.

In Section 2 we review related work. In Section 3 we present classes of instances for which $H_0, \ldots, H_4$ meet their claimed performance ratios. In Sections 4-6 we establish lower bounds on $\rho_0, \ldots, \rho_4$ that match the upper bounds provided by these examples. The lower bounds are obtained by solving linear programs, where the inverse of the performance ratio is the objective and an analysis of certain configurations that can or cannot be locally improved yields the constraints.

Our research was motivated by the relation between the path packing problem and the test cover problem with tests of size at most 2; see [2].

2. Earlier work

Hurkens and Schrijver [6] consider a series of analogous local improvement algorithms for the more general problem of packing vertex-disjoint subgraphs on $t$ vertices in a given graph. They derive a lower bound $\phi_k$ on the performance ratio of their $k$th heuristic, and prove that it is tight if the subgraph is a clique. In particular, for $t = 3$,

$$\phi_k = \begin{cases} \frac{2k+1}{2k+2} & \text{if } k \text{ is even}, \\ \frac{2k+2}{2k+3} & \text{if } k \text{ is odd}. \end{cases}$$

Since a path of length 2 is a subgraph on three vertices, we know that $\rho_k \geq \phi_k$. Table 1 lists the values of $\phi_k$ and $\rho_k$ for $k = 0, \ldots, 4$. Note that $\lim_{k \to \infty} \phi_k = \rho_k$. The limiting value of $\rho_k$ is unknown, but it is

Table 1: Values of $\phi_k$ and $\rho_k$ for $k = 0, \ldots, 4$.
is likely to be strictly smaller than 1, since the problem of packing
paths of length 2 is APX-hard [2].

For $k = 0, 1, 2$, the lower bound $\phi$ on $\rho_{k}$ matches the upper bound provided by the examples given in Section 3, which proves part of our result. For $k = 3, 4$, the proof is more involved.

Hassin and Rubinstein [4] study packing 3-edge paths of maximum
total weight and give a $\frac{3}{4}$-approximation algorithm. They present a $(\frac{80}{109} - \epsilon)$-approximation for the weighted triangle pack-
ing problem in [5], and show that this algorithm also gives a
performance ratio of $(\frac{80}{109} - \epsilon)$ for the maximum weight packing of
2-edge paths. Bafna et al. [1] analyze the k-local search algorithm for computing maximum weighted independent sets in (k + 1)-
claw free graphs to show a performance ratio of $(k - 1 + \frac{1}{k})$.

Fernau and Raible [3] give a parameterized algorithm for deciding
whether a 2-edge path packing of size $k$ exists in time $O^*(2.448^{3k})$
where the $O^*$ hides polynomial bounds on the size of the instance
like $n^r$ for some constant $c$.

From now on, any path consists of three vertices and two
edges.

3. Upper bounds: worst-case examples

Fig. 1 shows instances for which the heuristics $H_0, \ldots, H_4$ may
perform badly and meet the performance ratios claimed above.
For each $k$, the solid edges form an optimal solution, and the
dashed edges form a locally optimal solution, produced by $H_k$. The latter solution can be improved by sacrificing $k + 1$ paths in the
top-left, in return for $k + 2$ or more alternative paths, indicated by solid lines. Sacrificing only $k$ paths does not give room for
improvement. We leave it to the reader to verify this assertion.

Note that we obtain infinite classes of such examples by creating
multiple copies of the graphs.

4. Lower bounds: outline of approach

Before getting into details, we will outline our approach of
proving lower bounds on the performance of the heuristics $H_0, \ldots, H_4$.

For a given value of $k$, $0 \leq k \leq 4$, let $P$ denote the path packing
found by $H_k$, and let $Q$ denote an arbitrary maximal packing. To investigate the relation between $P$ and $Q$, we will give each
vertex in $P$ a label that quantifies the interaction between the
path in $Q$ to which it belongs and the paths in $P$. When the
middle vertex of a path in $P$ has label $\beta$ and its end vertices have
labels $\alpha$ and $\gamma$ with $\alpha \leq \gamma$, then the path is said to be of type $\alpha\beta\gamma$. A combinatorial analysis enables us to reduce the number
of different path types to 49.

Next, for each path type $\alpha\beta\gamma$, we define a variable $x(\alpha\beta\gamma)$,
whose value is equal to the fraction of paths in $P$ of that type.
We then formulate a linear program in these variables. The objective
will be given by $|Q|/|P|$. The constraints capture conditions that are implied by the local optimality of the solutions. These
conditions are rather straightforward for $H_0, H_1$ and $H_2$ but
ask for some case analysis for $H_3$ and $H_4$. It will be seen that the
optimum solution values match the inverses of the performance
ratios $\rho_0, \ldots, \rho_4$ given in Table 1.

5. Lower bounds: vertex labels

For a given $k \geq 0$, let $P$ denote the path packing found by
$H_k$, and let $Q$ denote an arbitrary maximal packing. A $P$-path is a
path in $P$, a $Q$-path is a path in $Q$. For a vertex $v$, $p(v)$ denotes the
$P$-path containing $v$, if it exists, and $q(v)$ denotes the $Q$-path
containing $v$, if it exists.

Without loss of generality, we make the following assumptions.

(1) $G$ contains no other edges than those in $P$ and $Q$. Otherwise,
delete the edges that are neither in $P$ nor in $Q$ from $G$. In the new
graph, $P$ is still locally optimal and $Q$ remains maximal.

(2) $|P| < |Q|$. Otherwise, $|P|/|Q| \geq 1$, and we are done.

(3) Each $P$-path intersects a $Q$-path. Otherwise, $Q$ is not maxi-
mal.

(4) Each $Q$-path intersects a $P$-path. Otherwise, $P$ is not locally
optimal.

(5) No $P$-path and $Q$-path cover the same three vertices. Other-
wise, delete those three vertices. In the new graph, $P$ is locally
optimal, $Q$ is maximal, and $|P|/|Q|$ is smaller.

(6) Each node of degree $\geq 3$ is in both $P$ and $Q$. For a vertex
$v$ that is an endpoint of $p(v)$ or $q(v)$, the statement is evidently
true, using (1). The case remains where $v = b$ is the midpoint of
a path $a\rightarrow b\rightarrow c$, that is either in $P$ or in $Q$. Then $a$ or $c$ or both
must be covered twice, as otherwise $P$ or $Q$ is not maximal. Assume
that $c$ is covered twice. By (1), $b$ has no neighbors besides $a$ and $c$.
Modify the graph by changing the path $a\rightarrow b\rightarrow c$ to $a\rightarrow c\rightarrow b$ (removing
the edge $ab$ and adding if needed the edge $ac$). In the new setting
$b$ has only one neighbor, $c$. Let $P'$ and $Q'$ denote the new packings
and suppose $P'$ can be improved to $P''$. $P''$ must involve the edge
$a\rightarrow c$, as it would otherwise also exist in the original graph. If $P''$
contained the path $a\rightarrow c\rightarrow d$, we can change it to $a\rightarrow c\rightarrow b$, since then
$b$ is not covered in $P''$. But then, by changing $a\rightarrow c\rightarrow b$ to $a\rightarrow b\rightarrow c$,
the improvement is also possible in the original graph.

We give each vertex $v$ in $P$ a label, which expresses the inter-
action of its $Q$-path $q(v)$ with the $P$-paths. Vertices not covered
by $P$ do not get a label. Vertex $v$ in $P$ receives label

- 0 if $v$ is not on any $Q$-path;
- 1 if $q(v)$ intersects exactly one $P$-path $p(v)$, in one vertex $v$;
- 2 if $q(v)$ intersects exactly two $P$-paths $p(v)$ and $p(w)$ in
  the two vertices $v$ and $w$ only, with $v$ being the middle vertex of
  $q(v)$; ($w$ receives label 3) note that, by assumption (6),
  either $v$ or $w$ is the middle vertex of $q(v)$;
- 3 if $q(v)$ intersects exactly two $P$-paths $p(v)$ and $p(w)$ in
  vertices $v$ and $w$, where $w$ receives label 2 or 5;
- 4 if $q(v)$ intersects three $P$-paths;
- 5 if $q(v)$ and $p(v)$ intersect in two vertices $v$ and $w$, and the
  third vertex $z$ on $q(v)$ is on another $P$-path $p(z)$; $w$ receives
  label 5 as well; $z$ receives label 3;
- 6 if $q(v)$ and $p(v)$ intersect in two vertices $v$ and $w$, and the
  third vertex on $q(v)$ is not on another $P$-path; $w$ receives
  label 6 as well.

The case that $q(v)$ and $p(v)$ intersect in three vertices is excluded
by assumption (5). Note that the labeling depends on the triples
of vertices that form a $P$-path. It is irrelevant which edges are
used in a $P$-path. Further note that label 5 comes in pairs, and so
does label 6. Fig. 2 illustrates some typical labelings.
We make some further assumptions.

(7) Vertex pairs labeled 5 or 6 have a Q-path and a P-path with coinciding middle vertices. Consider a P-path on a, b, c and a Q-path on a, c, d, with a and c labeled 5 or 6. Change $G$, replacing the paths by $b \rightarrow a \rightarrow c$ and $d \rightarrow a \rightarrow c$, respectively. Any improvement in the new graph must sacrifice $b \rightarrow a \rightarrow c$ and use, w.l.o.g., $d \rightarrow a \rightarrow c$. This improvement applies to the old graph too, replacing the P-path on a, b, c by the Q-path on a, c, d.

(8) No vertex has label 6. Consider a P-path $a \rightarrow b \rightarrow c$ with labels 5, 5, intersecting a Q-path $d \rightarrow b \rightarrow c$, with d labeled 3. Change $G$, introducing a new vertex $c'$ and replacing the P-path by $a \rightarrow b \rightarrow c'$, with $c'$ labeled 0 and b labeled 1. Any improvement of P in the new graph implies an improvement in the old graph.

(9) No P-path is labeled 155. Consider a P-path $a \rightarrow b \rightarrow c$ with labels 1, 5, 5, intersecting a Q-path $d \rightarrow b \rightarrow c$, with d labeled 3. Change $G$, introducing a new vertex $a'$ and replacing the P-path by $a' \rightarrow a \rightarrow b$, which is labeled 012. Vertex c becomes unlabeled. If the new graph allows for an improvement of P, the new solution uses w.l.o.g. the paths $q(a)$ and $q(b)$, and hence exists in the old graph as well.
(10) If a $P$-path has a vertex labeled 1, its middle vertex has label 1. Otherwise, consider a $P$-path $a \cdots c$ with $c$ labeled 1. Change $G$ as under assumption (6), and replace $a \cdots c$ by $P$. Any improvement of $P$ in the new graph must sacrifice this path and use, w.l.o.g., the path $q(c)$, and hence not the edges $(a,c)$ and $(c,b)$. This improvement could have been made in the old graph as well. Note that by assumption (9), vertices $a$ and $b$ are not labeled 5. Hence, assumption (7) is maintained.

A $P$-path is said to be of type $\alpha \beta \gamma$ if its middle vertex has label $\beta$ and its end vertices have labels $\alpha$ and $\gamma$, with $\alpha \leq \gamma$. A $P$-path is said to be of form $[\alpha \beta \gamma]$ if its type is a permutation of $\alpha \beta \gamma$. A wildcard $v$ may be used to match any label. As a consequence of the assumptions made, there exist 49 different labelings.

A vertex $v$ that is labeled 1, 2, 3, or 4 is called black if $p(v)$ contains a vertex with label 1, or if $v$ is an endpoint of $p(v)$ and $p(v) \setminus v$ contains a vertex labeled 2. In terms of local improvement, a black vertex $v$ can be ‘set free’ by replacing $p(v)$ by a path not covering $v$. The replacing path uses one or two edges from a $Q$-path, the middle vertex of which lies on $p(v) \setminus v$, and one or two endpoints not on any $P$-path. We will mark a black vertex by an overscore.

It follows from the labeling of $p(v)$ whether vertex $v$ is black or not. We may, for instance, have paths labeled $111$, $214$, $233$, $424$, $434$ and $255$. From assumptions (6) and (10) we have that a vertex labeled 0, 3, 3, or 4 is never the middle vertex of a $P$-path.

6. Lower bounds: LP formulations

We define $x(\alpha \beta \gamma)$ as the fraction of $P$-paths of type $\alpha \beta \gamma$, $n(\sigma)$ as the number of occurrences of label $\sigma$, and $|\alpha \beta \gamma|$ as the number of paths of type $\alpha \beta \gamma$. Then

$$n(\sigma) = \frac{|P|}{\sum_{(\alpha \beta \gamma)} x(\alpha \beta \gamma) [\chi(\sigma, \alpha) + \chi(\sigma, \beta) + \chi(\sigma, \gamma)]} \quad (*)$$

where $\chi(\sigma, \tau) = 1$ if $\sigma = \tau$, and 0 otherwise. If $\sigma$ has an overscore, we restrict the count to occurrences of a black $\sigma$.

A short path is a $Q$-path that intersects exactly two $P$-paths (and hence is labeled -23 or 355, where - denotes an unlabeled node). A black short path is a short path with at least one black vertex (2 or 3).

**Heuristic $H_0$**

$$|Q| = n(1) + \frac{1}{2} n(2) + \frac{1}{2} n(3) + \frac{1}{4} n(4) + \frac{1}{4} n(5) \quad (0a)$$

$$n(3) = n(2) + \frac{1}{2} n(5) \quad (0b)$$

$$\sum_{(\alpha \beta \gamma)} x(\alpha \beta \gamma) = 1 \quad (0c)$$

Dividing Eqs. (0a) and (0b) by $|P|$ and coupling the number of occurrences $n(\sigma)$ to the fractions $x$ yields an LP-model. Its optimum solution value is $|Q|/|P| = 3$, for $x(111) = 1$. (Note that a 111 $P$-path is colored black: $111$.)

The full version of the LP is given below. It has 49 variables $x(\alpha \beta \gamma)$, a variable $q = |Q|/|P|$ (which is the objective), and variables $v(k) = n(k)/|P|$. The first three constraints are equivalent to (0a), (0b), (0c); the six constraints defining $v(0), \ldots, v(5)$ follow from $(*)$.

**Heuristic $H_1$**

No black node is labeled 1:

$$n(1) \leq 0 \quad (1)$$

The LP extended by this additional constraint has an optimum solution value $|Q|/|P| = 2$ for $x(11\overline{2}) = 1$ or $x(2\overline{1}2) = x(\overline{3}3) = 0.5$.

**Heuristic $H_2$**

No two black nodes are on a short path:

$$n(\overline{2}) + n(3) \leq n(3) \quad (2)$$

To justify this constraint, consider black vertices $u$, $v$ on a black short path $q(u) = q(v)$. Heuristic $H_2$ would replace the two paths $p(u)$ and $p(v)$ by disjoint paths only incident to $p(u) \setminus u$ and
The path cannot have labels $\tilde{1}|\tilde{2}$ since heuristic $H_1$ would give an improvement. For the same reason, paths of type $222$ and $232$ must have a label.

(3b) A black vertex $v$ labeled 2 or 3 lies on $q(v)$ containing a gray vertex $w$ labeled 3, 2 or 5. If another black vertex $v'$ labeled 2 or 3 lies on $q(v')$ containing a gray vertex $w'$ labeled 3, 2 or 5, with $p(w) = p(w')$, heuristic $H_2$ will improve this configuration unless the vertices $v$ and $v'$ lie on the same $P$-path labeled $212, 222, 232, 242, 213, 223, 313$ or $323$. The path cannot have labels $212, 213$ or $313$, since heuristic $H_2$ would give an improvement. For the same reason, the other five paths cannot have a label.

(3c) Each $Q$-path with three vertices labeled 4 can have at most two black vertices. If all three were black, heuristic $H_3$ would improve the packing.

**Heuristic $H_3$**

To derive additional constraints for $H_3$, we impose more stringent bounds on black vertices labeled 2 or 3 by starting with the number of corresponding paths labeled +, and carefully avoiding overcounts. Similarly, to bound the number of vertices labeled 2, we consider $q(v)$ as in (3a) containing a gray vertex $w$ labeled 3 and carefully analyze the case whether $p(w)$ contains a gray vertex 4 or not to arrive at a second constraint.

The following additional constraints hold:

$2n(\tilde{2}) + 2n(\tilde{3}) - 3(|222^*| + |232^*| + |323^*| + |223| + |242|) \leq \frac{1}{2}n(\tilde{5})$ \hfill (4a)

$2|P|^+ - (|255^+| + |355^+| + |203^+| + |234^+| + |033^+| + |133^+| - 2|333^*| + n(2) + n(3) + \frac{1}{2}n(5)) \leq 4n(4) + 2|333^+| - 2|334^+| - 4|344^+| - 4|434^+| - |343^+| - |334^+|$ \hfill (4b)

The LP defined by the preceding constraints achieves its optimum $|Q|/|P| = \frac{1}{2}$ for $x(212) = \frac{1}{2}, x(414) = \frac{1}{2}$, and $x(434^+) = \frac{1}{4}$. We justify the additional constraints as follows:

(4a) Like in the argument for (3b), black vertex $v$ labeled 2 or 3 lies on $q(v)$ together with a vertex $w$ labeled 2, 3, or 5. Now, either $p(w)$ is a $P$-path with one label 2, 3, or 5, or $p(w)$ has more than one such label. In the former case, we find a contribution to $2n(\tilde{2}) + 2n(\tilde{3})$ of $+2$ on the left-hand side of (4a), a contribution to $2|P|^+$ of $+2$ on the right-hand side, and no further negative contribution to the right-hand side.

Note that if another short path exists with black vertex $v'$ and gray vertex $w'$, with $p(w) = p(w')$, vertices $v$ and $v'$ lie on the same $P$-path labeled $222, 232, 242, 223$ or $323$, as in the configuration left by heuristic $H_3$. Hence this can explain the second occurrence of a gray node labeled 2, 3, or 55 on $p(w)$, but not a third occurrence.

If $p(w)$ has two gray vertices labeled 2, 3 or 55, both stemming from a black short path, the above argument shows that we have a contribution to $2n(\tilde{2}) + 2n(\tilde{3})$ of $+2$ on the left-hand side of (4a), and a contribution of $-3$ for the path connecting $v$ and $v'$. On the right-hand side we have a contribution to $2|P|^+$ of $+2$ and also a contribution of $-1$ for the path connecting $w$ and $w'$.

If $p(w)$ has more than one gray vertex labeled 2, 3 or 55, and the black short path with vertex $v$ is the only one to intersect the path $p(w)$, then there must be a vertex $z$ on $p(w)$, with $z$ having label 2, 3 or 5, so that $q(z)$ is not a black short path.

Now $q(z)$ either contains a white vertex 2 or 3, or a white pair of 5’s, or it contains a gray vertex 2 or 3, or a gray pair of 5’s, and therefore connects $p(w)$ with another path $p(w')$ with a label.
In the first case we have an extra contribution of $+1$ to the term $n(2) + n(3) + \frac{1}{2} n(5)$, so the total contribution to the right-hand side becomes $2 = 2 - 1 + 1$.

In the latter case, the corresponding black vertices $v$ and $v'$ must be connected by a mutual $P$-path $p(v) = p(v')$ labeled $222^-$, $232^-$, $242$, $223$ or $323^-$, otherwise heuristic $H_4$ can improve upon this solution. We then count on the left-hand side a contribution to $2n(2) + 2n(3) + 4 + 4$ and a contribution of $-3$ for the path connecting $v$ and $v'$. On the right-hand side we have a contribution to $2P^+$ of $+4$ and also a contribution of $-1$ for the paths connecting $w$ and $w'$.

If the $\ddot{4}$ labeled path involved contains more than two gray vertices 2, 3, or 5, and hence is labeled $333^+$, it contains an extra gray vertex $\ddot{4}$ not lying on a black short path. It follows that the counterpart of $\ddot{4}$ on $q(z')$ is a white vertex. This adds a contribution of $+1$ to the term $n(2) + n(3) + \frac{1}{2} n(5)$, and of $-2$ to $-2/|333^+|$, thus compensating for the lost contribution in $-(|255^-| + |355^-| + |032^-| + |234^+| + |033^-| + |334^+|)$. Fig. 3 gives an overview of the possible configurations around a black short path. Each configuration satisfies the constraint (with inequality).

(4b) Note that adding inequalities (3a) and (3c) and multiplying the sum by 2 yields $2n(2) + 2n(4) - 2\{2(222^-+232^-+242^-+223^-)+232^-\} = 4n(4) + 2|\{3+k^+\}|$, which is stronger than inequality (4b), unless we take into account $P^+$-paths containing both a 3 and a 4 labeled vertex. Such a path cannot contain a vertex labeled 1, 2 or 5, and hence at least one vertex labeled 3 is part of a black short path.

Consider a black vertex $v$ labeled 2. As in the configuration left by heuristic $H_3$ it lies on a short path $q(w)$ with $w$ labeled 3 and with $p(w)$ having a $\ddot{1}$ label. Possibly there is another vertex $v'$ labeled 2 on a short path with $w'$, with $p(w') = p(w)$, in which case $v$ and $v'$ have a common $P$-path labeled $222^-, 232^-$ or $242$.

If $p(w)$ contains a single 3 and no 4, we neglect the labels on $p(w) \setminus w$ and have $v$ contribute 2 to $n(2)$ on the left-hand side, and $p(w)$ contribute 2 to $2|\{3+k^+\}|$ on the right-hand side of (4b).

If $p(w)$ contains a single 3 lying on a black short path and a single 4, that is, $p(w)$ has type $34^+, 043^+, 334^+$ or $343^+$, then vertex $v$ contributes 2 to the left-hand side, and $p(w)$ has type $34^+$ or $334^+$ contributes $-2$ to the right-hand side.

If the vertex $z$ labeled 4 lies on $q(z)$ with three gray vertices, we let $z$ contribute 1 to $n(4)$ on the right-hand side, in which case the inequality becomes $2 \leq 4 - 1 - 1$.

If $q(z)$ contains one black vertex $\ddot{z}'$ labeled 4, we have $\ddot{z}'$ contribute $\frac{1}{2}$ to $n(4)$ on the left-hand side, and let $z$ contribute 1 to $n(4)$ on the right-hand side, by which the inequality becomes $2 + 2 - \frac{1}{2} \leq 4 - 1$.

If $z$ is the only gray vertex of $q(z)$, with other nodes $\ddot{z}'$ and $\ddot{z}''$, then $p(v) = p(z')$ or $p(v) = p(z'')$, otherwise $H_4$ would have improved this solution, namely, by dropping $p(w)$, rearranging $p(v)$, $p(z')$, and $p(z'')$, and adding $q(v)$ and $q(z)$. Assume that $p(v) = p(z')$ contains a third vertex $v'$ labeled 1. Then $H_1$ would have improved the solution, by rearranging $p(z'')$, dropping $p(v)$ and $p(w)$, and adding $q(v)$, $q(z)$, $q(v')$. It follows that $p(v)$ must be $v-v'-z'$ of type 224 (with either a $+\ddot{1}$ label or a $-\ddot{1}$ label). In this case the contributions of the vertices labeled 4 cancel out in the inequality, whereas $v$ contributes 2 to the left-hand side, $p(v)$ contributes $-3$ to the left-hand side, and $p(w)$ contributes $-1$ or 0 to the right-hand side of (4b).

What remains is the case (33*) where $p(w)$ contains two gray vertices labeled 3, both lying on a black short path, and the case (344), where $p(w)$ is of type 344$^+$ or 434$^-$. In the case (33*), as described before, the black short paths are connected by a path $p(v)$ of type 222$^-$, 232$^-$ or 242, which contributes $-2$ to the left-hand side. Should the path $p(w)$ contain a gray vertex $z$ labeled 4, we let the vertex contribute to (4b). We do this in the same way as above, with a contribution of 4 to the right-hand side, and potentially a contribution of 1 to $2n(4)$ on the left-hand side. Note that $z$ cannot be the only gray vertex on $q(z)$. If that were the case, the above reasoning would force $p(v)$ to intersect $q(z)$ and have type 224.

In the case (344) we would have $p(w)$ contain gray vertices $z$ and $z'$, say, both labeled 4. It follows from the above reasoning that at most one of $z, z'$ can be the single gray vertex on $q(z)$ or $q(z')$ respectively, as $p(v)$ cannot intersect both $q(z)$ and $q(z')$. Note that $p(w)$ contributes $-4$ to the right-hand side, which is balanced by the gray nodes labeled 4, that contribute $+8$.

Hence in all cases, a vertex $z$ labeled 4 contributes 4 to the right-hand side if $q(z)$ consists of three gray vertices; it contributes 1 to the left-hand side and 4 to the right-hand side if $q(z)$ has one black vertex; and it contributes 4 to both left-hand side and right-hand side if $q(z)$ contains two black vertices. We have to do so to prevent double counting. For instance, a $Q$-path labeled 444 can be adjacent to two distinct black short paths.

Fig. 4 depicts in principle each possible configuration around a black vertex labeled 2. In order to avoid double counting, for the inequality we only count gray and black nodes as far as they lie within the given boundaries. Gray and black nodes labeled 4 that are not accounted for in the picture must lie on (part of) a
Fig. 4. More configurations around a vertex labeled 2 for $H_4$.

$Q$-path with three vertices labeled 4, in which at least one third of the nodes is gray. Each configuration satisfies the constraint, and almost all do so with equality.

Acknowledgments

We are grateful for the comments of the area editor and the referees, which helped us to improve the presentation.

References