Computing minimum distortion embeddings into a path for bipartite permutation graphs and threshold graphs*

Pinar Heggernes† Daniel Meister† Andrzej Proskurowski‡

Abstract

The problem of computing minimum distortion embeddings of a given graph into a line (path) was introduced in 2004 and has quickly attracted significant attention with subsequent results appearing in recent STOC and SODA conferences. So far, all such results concern approximation algorithms or exponential-time exact algorithms. We give the first polynomial-time algorithms for computing minimum distortion embeddings of graphs into a path when the input graphs belong to specific graph classes. In particular, we solve this problem in polynomial time for bipartite permutation graphs and threshold graphs. For both graph classes, the distortion can be arbitrarily large. The graphs that we consider are unweighted.

1 Introduction

A metric space is defined by a set of points and a distance function between pairs of points. Given two metric spaces (U, d) and (U', d'), an *embedding* of the first into the second is a mapping $f : U \to U'$. The embedding has *distortion* k if for all $x, y \in U$, $d(x, y) \leq d'(f(x), f(y)) \leq k \cdot d(x, y)$. Low distortion embeddings between metric spaces are well-studied and have a long history. Embeddings of finite metric spaces into low-dimensional geometric spaces have applications in various areas of computer science, like computer vision [23] and computational chemistry (see [12, 13] for an introduction and a list of applications). Traditionally, *combinatorial* problems on low distortion embeddings have been subject to extensive study. Results in this direction give bounds on the distortion within which a metric space of a given class can be embedded into a metric space of another class. The study of *algorithmic* problems on low distortion embeddings is more recent, and it concerns computing a minimum (or low) distortion embedding of a given metric space to another (or a class of) given metric space(s).

Minimum distortion embeddings are difficult to compute. It is NP-hard even to approximate by a ratio better than 3 a bijective minimum distortion embedding between two given finite 3-dimensional metric spaces [19].

Every finite metric space can be represented by a matrix whose entries are the distances between pairs of points, and hence corresponds to a graph. Kenyon et al. [14] initiated the study

*This work is supported by the Research Council of Norway. A preliminary version was presented at SWAT 2008.

†Department of Informatics, University of Bergen, Norway. Emails: pinar.heggernes@ii.uib.no, daniel.meister@ii.uib.no

‡Department of Information and Computer Science, University of Oregon, USA. Email: andrzej@cs.uoregon.edu

of computing a minimum distortion embedding of a given graph onto[1] another given graph, and they gave a parametrised algorithm for computing a minimum distortion embedding between an arbitrary unweighted graph and a bounded-degree tree. Subsequently, Badoiu et al. [3] gave a constant-factor approximation algorithm for computing minimum distortion embeddings of arbitrary unweighted graphs into trees.

Since then, computing a minimum distortion embedding for a given graph on n vertices into a path was identified as a fundamental problem. This is exactly the problem that we study in this paper. Bădoiu et al. [2] showed that this problem is hard to approximate within a constant factor. They gave an exponential-time exact algorithm and a polynomial-time $\mathcal{O}(n^{1/2})$-approximation algorithm for arbitrary unweighted input graphs, along with a polynomial-time $\mathcal{O}(n^{1/3})$-approximation algorithm for unweighted trees. In another paper, Bădoiu et al. [1] showed that the problem is hard to approximate by a factor polynomial in n, even for weighted trees. They also gave a better polynomial-time approximation algorithm for general weighted graphs, along with a polynomial-time algorithm that approximates the minimum distortion embedding of a weighted tree into a path by a factor that is polynomial in the distortion. Finally, Fellows et al. [7] showed that whether a general unweighted input graph can be embedded into a path with distortion at most d is fixed-parameter tractable when parametrised by d. They also showed that for weighted input graphs, the problem is NP-hard for every fixed d.

We initiate the study of designing polynomial-time algorithms for exact computation of minimum distortion embeddings into a path for input graphs of specific graph classes. In particular, we give polynomial-time algorithms for the solution of this problem on bipartite permutation graphs and on threshold graphs. Bipartite permutation graphs are bipartite graphs, and threshold graphs are split graphs. Deciding whether a bipartite graph or a split graph can be embedded into a path with distortion at most d is NP-hard [11]. Thus, the results of this paper complement the hardness results and narrow the gap between known tractable and intractable cases. Our input graphs are unweighted, and this restriction is necessary as otherwise the results would extend to arbitrary weighted complete graphs, which can encode arbitrary finite metric spaces. It is important to note that the minimum distortion required to embed an unweighted bipartite permutation or threshold graph into a path is unbounded and it can be $\Theta(n)$. All previous algorithms for exact computation of minimum distortion into a path, mentioned above, are practical only when distortion is bounded.

Minimum distortion into a path is very closely related to the widely known and extensively studied graph parameter bandwidth. The only difference between the two parameters is that a minimum distortion embedding has to be non-contractive, meaning that the distance in the embedding between two vertices of the input graph has to be at least their original distance, whereas there is no such restriction for bandwidth. Bandwidth is known to be one of the hardest graph problems; it is NP-hard even for very simple graphs like caterpillars of hair-length at most 3 [18], and it is hard to approximate by a constant factor even for trees [4]. Polynomial-time algorithms for the exact computation of bandwidth are known for very few graph classes, including bipartite permutation graphs [10] and threshold graphs (that are interval graphs) [15, 22]. However, simple examples exist to show that these bandwidth algorithms cannot be used to generate minimum distortion embeddings into a path for these graph classes. In fact, there exist very simple bipartite permutation graphs, like $K_{3,4}$, for which no optimal bandwidth layout corresponds to a minimum distortion embedding into a path. It should be noted that the bandwidth and the minimum distortion into a path of a graph can be very different. For

[1] They study a more restricted version of the problem where both graphs have the same number of vertices.

example, it is common knowledge that a cycle of length n has bandwidth 2, whereas its minimum distortion into a path is $\Omega(n)$. In this paper, we also prove that the latter is exactly $n-1$.

The running times of the algorithms we present in this paper are $\mathcal{O}(n^2)$ for bipartite permutation graphs and $\mathcal{O}(n)$ for threshold graphs. We would like to mention that our algorithms operate significantly different than known (non-trivial) bandwidth algorithms. Most algorithms for bandwidth take as input a graph and an integer k, and decide whether the bandwidth of the input graph is at most k. The bandwidth of the graph can afterwards be computed by binary search on possible values of k. As opposed to this approach, both of the algorithms that we present in this paper compute the minimum distortion into a path of a graph directly.

This paper is organised as follows. In the next section we give the necessary definitions and notations. In Section 3 we give the first preliminary results on simple graphs, like cycles. Sections 4 and 5 present the polynomial-time algorithms for threshold graphs and bipartite permutation graphs, respectively.

2 Definitions and notations

We study simple finite undirected unweighted graphs that are connected. A graph is denoted by $G = (V, E)$, where V is the vertex set and E is the edge set of G. Usually we refer to $|V|$ as n. The set of neighbours of a vertex v is denoted by $N_G(v)$, and $N_G[v] = N_G(v) \cup \{v\}$. Similarly, for $S \subseteq V$, $N_G[S] = \bigcup_{v \in S} N_G[v]$. A vertex u of G with $N_G[u] = V$ is called *universal*. The *degree* of a vertex v is $d_G(v) = |N_G(v)|$. We will omit the subscripts when the graph is clear from the context. Two non-adjacent vertices u and v are called *false twins* if $N(u) = N(v)$. The subgraph of G induced by the vertices in S is denoted by $G[S]$. For any $v \in V$, $G{-}v$ denotes $G[V \setminus \{v\}]$. A *u, v-path* is a path between u and v, including u and v. The *distance* $\mathrm{d}_G(u, v)$ between two vertices u and v in G is the number of edges in a shortest u, v-path in G. For any mapping f from V to (a subset of) $\mathbb{Z}$, the *distance* $\mathrm{d}_f(u, v)$ between u and v in f is $|f(u) - f(v)|$. We write $u \prec_f v$ when $f(u) < f(v)$. For a vertex v of G, every vertex u with $u \prec_f v$ is *to the left* of v, and every vertex w with $v \prec_f w$ is *to the right* of v in f. We will also informally write *leftmost* and *rightmost* vertex accordingly.

An *embedding into a path (line)* for a graph $G = (V, E)$ is a mapping $\mathcal{E} : V \to \mathbb{Z}$. In the rest of this paper we use simply *embedding* to mean an embedding into a path. An embedding $\mathcal{E}$ is *non-contractive* if $\mathrm{d}_{\mathcal{E}}(u, v) \geq \mathrm{d}_G(u, v)$ for every pair of vertices $u, v \in V$. Note that this condition can only be satisfied by connected graphs. The *distortion* $\mathrm{D}(G, \mathcal{E})$ of a non-contractive embedding $\mathcal{E}$ for G is defined to be the smallest k such that $\mathrm{d}_{\mathcal{E}}(u, v) \leq k \cdot \mathrm{d}_G(u, v)$ for every pair of vertices $u, v \in V$. Since we consider only unweighted graphs, it is easy to see that $\mathrm{D}(G, \mathcal{E})$ is the smallest k such that $\mathrm{d}_{\mathcal{E}}(u, v) \leq k$ for every edge uv of G (see also [14]). A *minimum distortion embedding* is a non-contractive embedding for G of smallest possible distortion. In this paper, the *distortion* of G, denoted by $\mathrm{D}(G)$, is the distortion of a minimum distortion embedding for G. Hence, our purpose is to compute $\mathrm{D}(G)$ when G is a bipartite permutation graph or a threshold graph.

Each integer (*position*) between the smallest and the largest integers that are mapped to in an embedding will be called a *slot* of that embedding. Exactly n slots of a non-contractive embedding are occupied by the vertices of G, and the rest are called *empty slots*. For a given vertex v, we refer to the rightmost vertex to the left of v of a certain property by *the close vertex to the left of* v of that property (*close vertex to the right* is defined symmetrically). For two vertices u, v, where $u \prec_{\mathcal{E}} v$, a vertex w is *between* u and v in $\mathcal{E}$ if $\mathcal{E}(u) \leq \mathcal{E}(w) \leq \mathcal{E}(v)$. In particular, w can be equal to u or v. The *vertex ordering underlying* $\mathcal{E}$, denoted by $\text{ord}(\mathcal{E})$, is an ordered list of the n vertices occupying the non-empty slots of $\mathcal{E}$ in increasing order of positions.

In general, a *vertex ordering* for $G = (V, E)$ is a mapping $\sigma : V \to \{1, 2, \dots, |V|\}$, thus a special kind of embedding. Since every ordering can be considered as a permutation of V, we will also give an ordering as an ordered list of vertices $\sigma = \langle x_1, x_2, \dots, x_n \rangle$. For an integer $k \geq 0$, we call σ a *k-ordering* for G if for every edge uv of G, $\text{d}_\sigma(u, v) \leq k$. The *bandwidth* of G, $\text{bw}(G)$, is the smallest k such that G has a k-ordering. Note that for a minimum distortion embedding $\mathcal{E}$ for G, $\text{ord}(\mathcal{E})$ is not necessarily a minimum bandwidth ordering for G. Similarly, adding a minimum number of empty slots to a minimum bandwidth ordering to achieve a non-contractive embedding does not necessarily result in a minimum distortion embedding for G. A simple example is C_n, the cycle on n vertices, for which minimum distortion embeddings are without empty slots (as we will show in the next section), but no minimum bandwidth ordering is a minimum distortion embedding.

Each of the graph classes studied in this paper will be introduced in the section that presents results on it. All graph classes mentioned in this paper can be recognised in linear time [5, 9].

3 Preliminary results on distortion

3.1 Minimum distortion embeddings of arbitrary graphs

In this subsection we present results on minimum distortion embeddings that will be useful for our proofs later in the paper. We start by showing that in a minimum distortion embedding we can always assume consecutive vertices to have the same distance in the embedding as they have in the graph.

Lemma 3.1 *Let G be a connected graph, and let $\mathcal{E}$ be an embedding for G with* $\text{ord}(\mathcal{E}) = \langle x_1, \dots, x_n \rangle$*. If* $\text{d}_\mathcal{E}(x_i, x_{i+1}) \geq \text{d}_G(x_i, x_{i+1})$ *for every* $1 \leq i < n$ *then $\mathcal{E}$ is non-contractive.*

Proof. Assume for a contradiction that $\text{d}_\mathcal{E}(x_i, x_{i+1}) \geq \text{d}_G(x_i, x_{i+1})$ for every $1 \leq i < n$, but that $\mathcal{E}$ is not non-contractive. Then, there is a pair u, v of vertices of G such that $\text{d}_\mathcal{E}(u, v) < \text{d}_G(u, v)$. Among all such pairs we choose u and v with smallest $\text{d}_\mathcal{E}(u, v)$. Without loss of generality, we can assume that u appears to the left of v in $\mathcal{E}$. If $u = x_i$ and $v = x_{i+1}$ for some $1 \leq i < n$ then $\text{d}_\mathcal{E}(x_i, x_{i+1}) < \text{d}_G(x_i, x_{i+1})$, which is a contradiction to our assumption about $\mathcal{E}$. So, there is a vertex w between u and v in $\mathcal{E}$, $w \neq u, v$, and by the choice of u and v, $\text{d}_\mathcal{E}(u, w) \geq \text{d}_G(u, w)$ and $\text{d}_\mathcal{E}(w, v) \geq \text{d}_G(w, v)$. However, $\text{d}_\mathcal{E}(u, v) = \text{d}_\mathcal{E}(u, w) + \text{d}_\mathcal{E}(w, v)$ and $\text{d}_G(u, v) \leq \text{d}_G(u, w) + \text{d}_G(w, v)$ contradict the choice of u and v. ■

Corollary 3.2 *Every connected graph G has a minimum distortion embedding $\mathcal{E}$ with* $\text{ord}(\mathcal{E}) = \langle x_1, \dots, x_n \rangle$ *such that* $\text{d}_\mathcal{E}(x_i, x_{i+1}) = \text{d}_G(x_i, x_{i+1})$ *for every* $1 \leq i < n$*.*

Proof. Let $\mathcal{F}$ be a minimum distortion embedding for G, and let $\text{ord}(\mathcal{F}) = \langle x_1, \dots, x_n \rangle$. Obtain $\mathcal{E}$ by placing x_1 in the slot at position 1 and x_{i+1} at distance $\text{d}_G(x_i, x_{i+1})$ to the right of x_i for every $1 \leq i < n$. Informally spoken, $\mathcal{E}$ is obtained from $\text{ord}(\mathcal{F})$ by adding the minimum number of necessary empty slots between consecutive vertices. Then, $\mathcal{E}$ satisfies the condition of Lemma 3.1, thus is non-contractive. It holds that $\text{d}_\mathcal{E}(x_i, x_{i+1}) \leq \text{d}_\mathcal{F}(x_i, x_{i+1})$, $1 \leq i < n$, so that $\text{d}_\mathcal{E}(u, v) \leq \text{d}_\mathcal{F}(u, v)$ for every pair u, v of adjacent vertices. Thus, $\text{D}(G) \leq \text{D}(G, \mathcal{E}) \leq \text{D}(G, \mathcal{F})$, and $\mathcal{E}$ is a minimum distortion embedding for G. ■

About the above result, note in particular that there are no empty slots between consecutive vertices in $\mathcal{E}$ that are adjacent in G. We say that an embedding *does not contain unnecessary empty slots* if it satisfies the distance condition of Corollary 3.2, i.e., consecutive vertices in the embedding are at distance exactly their distance in the graph.

A *bipartite graph* is a graph whose vertex set can be partitioned into two independent sets. We denote such a graph by $G = (A, B, E)$ where $A \cup B$ is the vertex set of G, and A and B are independent sets, also called *colour classes*. If G is a connected bipartite graph, then the partition of the vertex set into the two colour classes is unique.

Lemma 3.3 *The distortion of a connected bipartite graph is an odd integer.*

Proof. Let $G = (A, B, E)$ be a connected bipartite graph, and let $\mathcal{E}$ be a minimum distortion embedding for G. Let $\mathrm{ord}(\mathcal{E}) = \langle x_1, \ldots, x_n \rangle$. According to Corollary 3.2, we can choose $\mathcal{E}$ such that $\mathrm{d}_{\mathcal{E}}(x_i, x_{i+1}) = \mathrm{d}_G(x_i, x_{i+1})$. Then, x_i and x_{i+1} belong to the same colour class if and only if $\mathrm{d}_{\mathcal{E}}(x_i, x_{i+1})$ is even. By induction, it can be shown that the vertices at even distance from x_i in $\mathcal{E}$ are exactly the vertices from the colour class of x_i. Hence, u and v belong to the same colour class of G if and only if $\mathrm{d}_{\mathcal{E}}(u, v)$ is even. Since adjacent vertices of G belong to different colour classes, every edge joins two vertices at odd distance in $\mathcal{E}$. Thus, $\mathrm{D}(G, \mathcal{E})$ is odd. ▪

Lemma 3.4 *For every connected graph G,* $\mathrm{D}(G) \geq \mathrm{bw}(G)$.

Proof. Let $\mathcal{E}$ be a minimum distortion embedding for G with $\mathrm{ord}(\mathcal{E}) = \langle x_1, \ldots, x_n \rangle$. For every pair x_i, x_{i+r} of adjacent vertices of G, $\mathrm{d}_{\mathcal{E}}(x_i, x_{i+r}) \geq r$. Thus, $\mathrm{ord}(\mathcal{E})$ is a $\mathrm{D}(G, \mathcal{E})$-ordering and $\mathrm{bw}(G) \leq \mathrm{D}(G, \mathcal{E}) = \mathrm{D}(G)$. ▪

In some of our proofs, we will identify a subgraph of a given graph and use the distortion of the subgraph as a lower bound for the distortion of the given graph. For this reason, we need the following lemmas. We say that a subgraph H of G is *distance-preserving* if $\mathrm{d}_H(u, v) \leq \mathrm{d}_G(u, v)$ for all $u, v \in V(H)$. It follows directly that distances in H and G are then equal, since every path in H is a path in G. In particular, distance-preserving subgraphs are induced subgraphs.

Lemma 3.5 *Let H be a subgraph of a graph G. If H is a distance-preserving subgraph of G then* $\mathrm{D}(G) \geq \mathrm{D}(H)$.

Proof. Let $\mathcal{E}$ be a minimum distortion embedding for G, and let $\mathcal{F}$ be obtained from $\mathcal{E}$ by removing all vertices that are not in H. Let u and v be vertices of H. Clearly, $\mathrm{d}_{\mathcal{F}}(u, v) = \mathrm{d}_{\mathcal{E}}(u, v)$ and $\mathrm{D}(H, \mathcal{F}) \leq \mathrm{D}(G, \mathcal{E})$. Since H is distance-preserving and $\mathcal{E}$ is non-contractive for G, we obtain $\mathrm{d}_H(u, v) = \mathrm{d}_G(u, v) \leq \mathrm{d}_{\mathcal{E}}(u, v) = \mathrm{d}_{\mathcal{F}}(u, v)$. Hence, $\mathcal{F}$ is a non-contractive embedding for H, and thus $\mathrm{D}(H) \leq \mathrm{D}(G)$. ▪

For applying Lemma 3.5, the main task is to identify distance-preserving subgraphs. We give sufficient conditions for two easy situations.

Lemma 3.6 *Let u and v be two false twin vertices of a graph G. Let H be a connected subgraph of G that contains u and v. If $H - v$ is a distance-preserving subgraph of G then H is a distance-preserving subgraph of G.*

Proof. Let $H{-}v$ be distance-preserving. Let a and b be two vertices of H. If $a \neq v$ and $b \neq v$ then $\mathrm{d}_H(a,b) \leq \mathrm{d}_{H-v}(a,b)$ since adding vertices does not increase distances. Now, let $a = v$. If $b = u$ then u and v have a common neighbour in H (since H is connected) and G, and thus $\mathrm{d}_H(v,u) = \mathrm{d}_G(v,u) = 2$. If $b \neq u$ then $\mathrm{d}_G(u,b) = \mathrm{d}_G(v,b)$. Let $(w_0, w_1, \ldots, w_s)$ be a shortest u,b-path in $H{-}v$. By $H{-}v$ being distance-preserving, $\mathrm{d}_G(u,b) = s$. Then, $(v, w_1, \ldots, w_s)$ is a v,b-path in H, so that v and b are at distance at most $s = \mathrm{d}_G(v,b)$ in H. Hence, H is a distance-preserving subgraph of G. ▪

Lemma 3.7 *Let u and v be two vertices of a graph G such that $N_G(v) \subseteq N_G[u]$. Then, $G{-}v$ is a distance-preserving subgraph of G.*

Proof. Let a, b be vertices of $G{-}v$, and let P be a shortest a,b-path in G. If P does not contain v then $\mathrm{d}_{G-v}(a,b) = \mathrm{d}_G(a,b)$. Otherwise, if P contains v, obtain P' by replacing v with u. If P' is a simple path, which means that no vertex appears more than once on P', P' is a path in $G{-}v$, and we conclude $\mathrm{d}_{G-v}(a,b) = \mathrm{d}_G(a,b)$. Suppose now that P' is not a simple path. Then, u occurs twice on P'. We obtain P'' from P' by cutting the piece from the first occurrence of u on P' until before the second occurrence of u. Then, P'' is an a,b-path in G of shorter length than P, which contradicts the choice of P. ▪

3.2 Graph classes with easy minimum distortion embeddings

As a warm-up before we start with the more involved algorithms in the next sections, and as interesting independent results on their own, we present combinatorial results on the minimum distortion of proper interval graphs, cycles, complete bipartite graphs and complete split graphs. The result on complete bipartite graphs is heavily needed for our results on bipartite permutation graphs.

A graph is an *interval graph* if sets of consecutive integers (intervals) can be assigned to its vertices such that two vertices are adjacent if and only if their intervals have a non-empty intersection. An interval graph is a *proper interval graph* if intervals can be assigned such that no interval is a subset of another. Proper interval graphs are equivalent to *unit interval graphs* meaning that there is an assignment with all intervals of the same length [20]. The vertex ordering by the smallest (or equivalently largest) element of the assigned intervals is called a *proper interval ordering.*

Theorem 3.8 *For every connected proper interval graph G,* $\mathrm{D}(G) = \mathrm{bw}(G)$ *.*

Proof. Let $G = (V, E)$ be a connected proper interval graph with proper interval ordering $\langle x_1, \ldots, x_n \rangle$. Let $\mathcal{E}$ be the non-contractive embedding without unnecessary empty slots with underlying vertex ordering $\langle x_1, \ldots, x_n \rangle$. Since G is connected, $x_i x_{i+1} \in E$ for every $1 \leq i < n$, so that there are no empty slots between the vertices in $\mathcal{E}$. For every pair x_i, x_j of adjacent vertices, where $i < j$, the set $\{x_i, x_{i+1}, \ldots, x_j\}$ is a clique in G [16, 9]. Consequently, the maximum distance of two adjacent vertices is $\omega(G) - 1 = \mathrm{bw}(G)$, which shows $\mathrm{D}(G) \leq \mathrm{bw}(G)$. Equality then follows with Lemma 3.4. ▪

The following three theorems show that the distortion of cycles, complete bipartite graphs and complete split graphs only depend on the number of vertices in these graphs. The chordless cycle on n vertices for $n \geq 3$ is denoted by C_n.

Theorem 3.9 $\mathrm{D}(C_n) = n-1$ *for* $n \geq 3$.

Proof. Let $(v_1, v_2, \ldots, v_n)$ be a cycle in C_n. Vertex ordering $\sigma = \langle v_1, v_2, \ldots, v_n \rangle$ is a non-contractive embedding for C_n of distortion $\mathrm{d}_\sigma(v_1, v_n) = n-1$. Thus, $\mathrm{D}(C_n) \leq n-1$.

For the lower bound, let $\mathcal{E}$ be a minimum distortion embedding for C_n with the smallest number of pairs of non-adjacent consecutive vertices. Let $\mathrm{ord}(\mathcal{E}) = \langle x_1, \ldots, x_n \rangle$. For $1 \leq i < n$, we call position $\mathcal{E}(x_i)$ a *gap position* if $x_i x_{i+1} \notin E$. If $\mathcal{E}$ has no gap positions then $x_i x_{i+1} \in E$ for all $1 \leq i < n$, and $(x_1, \ldots, x_n)$ is a path in C_n. Then, $x_1 x_n \in E$ and $\mathrm{D}(C_n) = \mathrm{D}(C_n, \mathcal{E}) = \mathrm{d}_\mathcal{E}(x_1, x_n) = n-1$. Now, assume that there is a gap position in $\mathcal{E}$. We construct a non-contractive embedding for C_n with a smaller number of gap positions and without increasing the distortion. Let $\mathcal{E}(x_j)$ be a gap position of $\mathcal{E}$. The number of empty slots between x_j and x_{j+1} in $\mathcal{E}$ can be assumed to be $\mathrm{d}_{C_n}(x_j, x_{j+1}) - 1$. Let P be a shortest x_j, x_{j+1}-path in C_n. We obtain $\mathcal{F}$ from $\mathcal{E}$ by moving the vertices in P that are different from x_j and x_{j+1} into the empty slots between x_j and x_{j+1} respecting their order in P. Clearly, $\mathcal{F}$ is non-contractive. We determine the distortion of $\mathcal{F}$. Moved vertices are at distance 1 to their two neighbours, so that it holds for every pair u, v of adjacent vertices at distance more than 1 in $\mathcal{F}$ that $\mathcal{F}(u) = \mathcal{E}(u)$ and $\mathcal{F}(v) = \mathcal{E}(v)$, thus $\mathrm{d}_\mathcal{F}(u, v) = \mathrm{d}_\mathcal{E}(u, v)$. Hence, $\mathrm{D}(G, \mathcal{F}) \leq \mathrm{D}(G, \mathcal{E})$. We consider the number of pairs of non-adjacent consecutive vertices in $\mathcal{F}$. Let x_i and x_{i+1} be adjacent in $\mathcal{E}$. Note that x_i is moved if and only if x_{i+1} is moved. Then, x_i and x_{i+1} appear consecutively in $\mathcal{F}$. Thus, the number of pairs of non-adjacent consecutive vertices in $\mathcal{F}$ is at most the number in $\mathcal{E}$. However, since x_j and the close vertex to the right of x_j in $\mathcal{F}$ are adjacent, the number of pairs of consecutive non-adjacent vertices in $\mathcal{F}$ is smaller than the number in $\mathcal{E}$, which contradicts the choice of $\mathcal{E}$. Consequently, $\mathcal{E}$ does not contain a gap position. ■

A bipartite graph $G = (A, B, E)$ is a *complete bipartite graph* if every vertex in A is adjacent to every vertex in B. Such a graph is denoted by $K_{n,m}$, where $n = |A|$ and $m = |B|$.

Theorem 3.10 *Let* n *and* m *be integers satisfying* $1 \leq n \leq m$. *If* $n+m$ *is odd then* $\mathrm{D}(K_{n,m}) = n+m-2$, *and if* $n+m$ *is even then* $\mathrm{D}(K_{n,m}) = n+m-1$.

Proof. Let A and B be the two colour classes of $K_{n,m}$ with $|A| = n$ and $|B| = m$.

First we prove a lower bound on the distortion of $K_{n,m}$. Clearly, $\mathrm{D}(K_{1,1}) = 1$. Assume in the following that $m \geq 2$. Let $\mathcal{E}$ be a non-contractive embedding for $K_{n,m}$. The distance between consecutive vertices from the same colour class is at least 2. Denote by a and a' the respectively leftmost and rightmost vertex in $\mathcal{E}$, and denote by b and b' the respectively leftmost and rightmost vertex from B. It holds that $\mathrm{d}_\mathcal{E}(a, a') \geq 2n-2$ and $\mathrm{d}_\mathcal{E}(b, b') \geq 2m-2$, and $\mathrm{D}(K_{n,m}, \mathcal{E}) = \max\{\mathrm{d}_\mathcal{E}(a, b'), \mathrm{d}_\mathcal{E}(b, a')\}$. We distinguish two cases. If there is a vertex from A to the left of b or to the right of b' then the distortion of $\mathcal{E}$ is at least $2m-1 \geq m+n-1$. Now, let all vertices from A be between b and b'. Note that $\mathrm{d}_\mathcal{E}(b, a') = \mathrm{d}_\mathcal{E}(b, a) + \mathrm{d}_\mathcal{E}(a, a')$. So, $\mathrm{d}_\mathcal{E}(b, a') + \mathrm{d}_\mathcal{E}(a, b') = \mathrm{d}_\mathcal{E}(b, a') + \mathrm{d}_\mathcal{E}(a, a') + \mathrm{d}_\mathcal{E}(a', b') = \mathrm{d}_\mathcal{E}(a, a') + \mathrm{d}_\mathcal{E}(b, b')$. A lower bound on this sum is $2n-2+2m-2 = 2(n+m-2)$. Hence, $\mathrm{D}(K_{n,m}, \mathcal{E}) \geq n+m-2$, which already gives the lower bound in the case $n+m$ odd. Let $n+m$ be even. If $\mathrm{d}_\mathcal{E}(b, a') \leq n+m-2$ then there are at most $\frac{1}{2}(n+m-2)$ vertices from B to the left of a', and at least $m - \frac{1}{2}(n+m-2) = \frac{1}{2}(m-n+2)$ vertices from B are to the right of a'. Hence, $\mathrm{d}_\mathcal{E}(a, b') = \mathrm{d}_\mathcal{E}(a, a') + \mathrm{d}_\mathcal{E}(a', b') \geq 2n-2+(m-n+2)-1 = n+m-1$. This completes the proof of the lower bound.

We prove an upper bound on the distortion by defining an embedding $\mathcal{E}$. Lay out the vertices from B in any order with exactly one empty slot between consecutive vertices. Denote by b and b' the respectively leftmost and rightmost vertex in $\mathcal{E}$. Let $p =_{\mathrm{def}} \mathcal{E}(b') - (n+m-2)$ or

$p =_{\text{def}} \mathcal{E}(b') - (n + m - 1)$ depending on whether $n+m$ is odd or even, respectively. Note that the slot at position p is empty in $\mathcal{E}$. Starting in the slot at position p and continuing towards the right, place the vertices from A in any order with one slot between consecutive vertices. This completes the definition of $\mathcal{E}$. Observe that $\mathcal{E}$ is a proper embedding. Furthermore, $\mathcal{E}$ is non-contractive, since vertices of the same colour class are at distance at least 2 from each other, and vertices from different colour classes are adjacent. Denote by a and a' the respectively leftmost and rightmost vertex from A in $\mathcal{E}$. It holds that $\mathrm{d}_{\mathcal{E}}(a, a') = 2n - 2$ and $\mathrm{d}_{\mathcal{E}}(b, b') = 2m - 2$. Then, $\mathrm{d}_{\mathcal{E}}(b, a') = \mathrm{d}_{\mathcal{E}}(b, b') - \mathrm{d}_{\mathcal{E}}(a, b') + \mathrm{d}_{\mathcal{E}}(a, a') \leq 2m - 2 - (n + m - 2) + 2n - 2 = n + m - 2$. Thus, if $n + m$ is odd then $\mathrm{D}(K_{n,m}, \mathcal{E}) = n + m - 2$, if $n + m$ is even then $\mathrm{D}(K_{n,m}, \mathcal{E}) = n + m - 1$. ∎

A graph is a *split graph* if its vertices can be partitioned into a clique X and an independent set I. We call such a partition a *split partition* and denote it by (X, I). Generally, a split graph can have more than one split partition. A split graph $G = (V, E)$ with split partition (X, I) is also denoted by (X, I, E). We refer to the vertices in X and I as *X-vertices* and *I-vertices*, respectively. We call a split graph a *complete split graph* if it has a split partition (X, I) such that all X-vertices are adjacent to all I-vertices, and we denote it by $S_{n,m}$, where n is the number of X-vertices and m is the number of I-vertices. Note that $S_{1,m}$ coincides with $K_{1,m}$ and that $S_{n,1}$ is a complete graph.

Theorem 3.11 *Let n and m be natural numbers where $n \geq 2$ and $m \geq 2$. Then, $\mathrm{D}(S_{n,m}) = n + m - 2$.*

Proof. Let (X, I) be a split partition of $S_{n,m}$ with $|X| = n$ and $|I| = m$. Note that each of the X-vertices is adjacent to each of the I-vertices.

First we prove a lower bound on the distortion of $S_{n,m}$. Let $\mathcal{E}$ be a minimum distortion embedding for $S_{n,m}$ with the smallest number of I-vertices between X-vertices. The leftmost and rightmost vertex in $\mathcal{E}$ are at distance at least $n + m - 1$. If one of these two vertices is an X-vertex then the two vertices are adjacent and the distortion of $\mathcal{E}$ is at least $n + m - 1$. Now, let the leftmost and rightmost vertex be I-vertices, denoted as b and b', respectively. Denote by a and a' the respectively leftmost and rightmost X-vertex in $\mathcal{E}$. It holds that $\mathrm{D}(S_{n,m}, \mathcal{E}) = \max\{\mathrm{d}_{\mathcal{E}}(b, a'), \mathrm{d}_{\mathcal{E}}(a, b')\}$. Furthermore, with $s =_{\text{def}} \mathrm{d}_{\mathcal{E}}(b, a') + \mathrm{d}_{\mathcal{E}}(a, b') = \mathrm{d}_{\mathcal{E}}(b, b') + \mathrm{d}_{\mathcal{E}}(a, a')$, it holds that $\mathrm{D}(S_{n,m}, \mathcal{E}) \geq \frac{1}{2}s$. We distinguish three cases. First, let there be no I-vertex between X-vertices in $\mathcal{E}$. Then, $\mathrm{d}_{\mathcal{E}}(b, b') \geq 2m - 2 + n - 1$ and $\mathrm{d}_{\mathcal{E}}(a, a') \geq n - 1$. For the second case, let there be exactly one I-vertex between a and a' in $\mathcal{E}$. Then, $\mathrm{d}_{\mathcal{E}}(b, b') \geq 2m - 2 + n - 2$ and $\mathrm{d}_{\mathcal{E}}(a, a') \geq n$. In both cases, we obtain $s \geq 2m - 2 + 2n - 2$, thus $\mathrm{D}(S_{n,m}, \mathcal{E}) \geq n + m - 2$. For the third case, assume that there are at least two I-vertices between a and a' in $\mathcal{E}$. Let c and c' be the close I-vertex to the right of a and to the left of a', respectively. We obtain $\mathcal{F}$ from $\mathcal{E}$ by removing c and c', moving all vertices to the left of c one position further to the right, all vertices to the right of c' one position further to the left, placing c at distance 2 to the left of b and c' at distance 2 to the right of b'. Since I-vertices are at distance at least 2 from each other in $\mathcal{F}$, $\mathcal{F}$ is a non-contractive embedding for $S_{n,m}$. Furthermore, $\mathrm{d}_{\mathcal{F}}(c, a') = \mathrm{d}_{\mathcal{E}}(b, a')$ and $\mathrm{d}_{\mathcal{F}}(a, c') = \mathrm{d}_{\mathcal{E}}(a, b')$, so that $\mathrm{D}(S_{n,m}, \mathcal{F}) = \mathrm{D}(S_{n,m}, \mathcal{E})$. Since the number of I-vertices between X-vertices in $\mathcal{F}$ is smaller than the number in $\mathcal{E}$, we obtain a contradiction to the choice of $\mathcal{E}$. This completes the proof of the lower bound.

We prove the upper bound on the distortion by defining an embedding. We distinguish two cases. Let m be even. Let $\mathcal{E}$ be a non-contractive embedding without unnecessary empty slots with underlying vertex ordering of the following form: first $\frac{m}{2}$ I-vertices, then all X-vertices, then the remaining $\frac{m}{2}$ I-vertices. It clearly holds that $\mathrm{D}(S_{n,2}, \mathcal{E}) = n$ and $\mathrm{D}(S_{n,m}, \mathcal{E}) =$

$2 \cdot (\frac{m}{2} - 1) + 1 + n - 1$ for $m \geq 4$. In the case where m is odd, we define embedding $\mathcal{E}$ as: take the above defined embedding for $S_{n,m-1}$ and place the last I-vertex between two X-vertices. Then, $\mathcal{E}$ is a non-contractive embedding for $S_{n,m}$ of distortion $\mathrm{D}(S_{n,m-1}) + 1 = n + m - 2$. ■

4 Distortion of threshold graphs

Threshold graphs are split graphs, and they have various characterisations [5, 9]. For our purposes, the following characterisation will serve as a definition. A graph is a *threshold graph* if and only if it is split and the vertices of the independent set can be ordered by neighbourhood inclusion, for any split partition for it [17]. Equivalently, the vertices of the clique can be ordered by neighbourhood inclusion [17]. Hence, for any split partition (X, I) for a threshold graph G, the X-vertices can be ordered as $a_1, a_2, \ldots, a_n$ such that $N(a_1) \supseteq N(a_2) \supseteq \cdots \supseteq N(a_n)$, and the I-vertices can be ordered as $b_1, b_2, \ldots, b_m$ such that $N(b_1) \subseteq N(b_2) \subseteq \cdots \subseteq N(b_m)$. In particular, this means that I-vertices of the same degree have exactly the same neighbourhood, and the same for X-vertices. Therefore, the given orderings correspond to a non-increasing degree order for the X-vertices and a non-decreasing degree order for the I-vertices. For simplicity, we say *decreasing* instead of non-increasing and *increasing* instead of non-decreasing. Every connected threshold graph has a universal vertex, which is a vertex that is adjacent to every other vertex of the graph. Thus, every pair of vertices in a connected threshold graph is at distance at most 2. In threshold graph $G = (X, I, E)$, if there is no X-vertex without a neighbour in I, there is an I-vertex b that is adjacent to all X-vertices. Then, $(X \cup \{b\}, I \setminus \{b\})$ is also a split partition for G. In the following, we assume for split partitions that an X-vertex of smallest degree has no neighbours outside X. In particular, the threshold graphs that we consider here contain at least three vertices and at least two X-vertices.

In this section, we give an efficient algorithm for computing the distortion of threshold graphs. The algorithm is based on a structural result about minimum distortion embeddings for threshold graphs that we prove first. We show that a minimum distortion embedding can be assumed to list the X-vertices in decreasing degree order. When we say in the following that we "remove a vertex from the embedding" we mean that the slot containing the vertex becomes an empty slot. Note that every embedding for a threshold graph can be partitioned into three sections: I-vertices to the left of all X-vertices, I-vertices to the right of all X-vertices and all other vertices in between, that are between X-vertices.

Lemma 4.1 *Let $G = (X, I, E)$ be a connected threshold graph. There is a minimum distortion embedding for G without empty slots between X-vertices.*

Proof. Let $\mathcal{E}$ be a minimum distortion embedding for G without unnecessary empty slots and with the smallest number of empty slots between X-vertices. In particular, pairs of consecutive vertices are at distance at most 2 in $\mathcal{E}$. We show that $\mathcal{E}$ satisfies the lemma. Let a and b be the respectively leftmost and rightmost X-vertex in $\mathcal{E}$. Assume for a contradiction that there is an empty slot at position p between a and b in $\mathcal{E}$. Let x and y be the vertices occupying the slots at position $p-1$ and $p+1$, respectively. Since $\mathrm{d}_{\mathcal{E}}(x, y) = 2 = \mathrm{d}_G(x, y)$, it follows that at least one of these two vertices is an I-vertex. Assume that y is an I-vertex, and if x is also an I-vertex then assume that $d_G(x) \geq d_G(y)$; otherwise, we repeat the arguments on the reverse of $\mathcal{E}$. Obtain embedding $\mathcal{F}$ from $\mathcal{E}$ by removing y and moving all vertices to the left of y two positions to the right. Observe that the slot at position $\mathcal{F}(x) + 1 = \mathcal{E}(y) + 1$ in $\mathcal{F}$ is either empty or occupied by an X-vertex that is adjacent to x. Note that the latter is particularly true for x and y both I-vertices, since every neighbour of y is a neighbour of x. Thus, $\mathcal{F}$ is non-contractive for $G-y$. Let u be a universal vertex in G, and let a^* and b^* be the respectively leftmost and rightmost vertex in $\mathcal{F}$ (and thus in $\mathcal{E}$ because of $y \prec_{\mathcal{E}} b$). We obtain $\mathcal{F}'$ from $\mathcal{F}$ as follows:

– if $y \prec_{\mathcal{E}} u$ then place y at distance 2 to the left of a^*

– if $u \prec_{\mathcal{E}} y$ then place y at distance 2 to the right of b^* .

Then, $\mathcal{F}'$ is a non-contractive embedding for G. Furthermore, $\mathrm{D}(G,\mathcal{F}') \leq \mathrm{D}(G,\mathcal{E})$, since $\max\{\mathrm{d}_{\mathcal{F}'}(y,u),\, \mathrm{d}_{\mathcal{F}'}(a^*,u),\, \mathrm{d}_{\mathcal{F}'}(b^*,u)\} \leq \max\{\mathrm{d}_{\mathcal{E}}(a^*,u),\, \mathrm{d}_{\mathcal{E}}(b^*,u)\}$. Thus, $\mathcal{F}'$ is a minimum distortion embedding for G with fewer empty slots between a and b than $\mathcal{E}$, contradicting the choice of $\mathcal{E}$. ■

Note that non-contractive embeddings for threshold graphs that have no empty slots between X-vertices do not contain two or more consecutive I-vertices between two X-vertices.

Lemma 4.2 *Let $G = (X, I, E)$ be a connected threshold graph. There is a minimum distortion embedding for G without empty slots between X-vertices such that the X-vertices appear in decreasing degree order.*

Proof. Let $\mathcal{E}$ be a minimum distortion embedding for G without empty slots between X-vertices and without unnecessary empty slots; such an embedding exists due to Lemma 4.1. Let u be the leftmost universal vertex in $\mathcal{E}$. Without loss of generality, we can assume that there is an X-vertex of smallest degree to the right of u in $\mathcal{E}$; otherwise we use the reverse of $\mathcal{E}$ instead of $\mathcal{E}$. Let v be the rightmost vertex in $\mathcal{E}$ among the X-vertices of smallest degree. Remember that v has no neighbour in I. Denote by a and b the respectively leftmost and rightmost X-vertex in $\mathcal{E}$. Note that $\mathrm{D}(G,\mathcal{E}) \geq \mathrm{d}_{\mathcal{E}}(a,b)$. Without loss of generality, we can assume that all I-vertices to the left of a appear in increasing degree order and all I-vertices to the right of b appear in decreasing degree order (ordering the I-vertices in this way does not increase $\mathrm{D}(G,\mathcal{E})$). This assumption is of importance only for making later arguments shorter. Based on $\mathcal{E}$, we will define a new embedding that satisfies the conditions of the lemma. Before, we collect helpful properties.

Let M be the set of I-vertices to the right of b that are at distance more than $\mathrm{D}(G,\mathcal{E})$ to a in $\mathcal{E}$. Note that no vertex in M is adjacent to a. Furthermore, the vertices in M appear consecutively in $\mathcal{E}$, and if M is non-empty then the rightmost vertex in $\mathcal{E}$ is contained in M. Let M be non-empty and let w' be the leftmost vertex in M; clearly $\mathrm{D}(G,\mathcal{E})+1 \leq \mathrm{d}_{\mathcal{E}}(a,w') \leq \mathrm{D}(G,\mathcal{E})+2$. In the following, we distinguish between the two cases $\mathrm{d}_{\mathcal{E}}(a,w') = \mathrm{D}(G,\mathcal{E})+1$ and $\mathrm{d}_{\mathcal{E}}(a,w') = \mathrm{D}(G,\mathcal{E})+2$ as the “short” and the “long” case, respectively. The *working interval* is the interval of slots between positions $\mathcal{E}(a)$ and $\mathcal{E}(b)$, potentially extended by the positions:

– $\mathcal{E}(b)+1$ if the slot at this position is non-empty and not occupied by w'

– $\mathcal{E}(a)-1$ if the slot at this position is non-empty and M is non-empty and we are in the short case and there are two X-vertices between a and u at distance 1 from each other.

Denote by a^* and b^* the respectively leftmost and rightmost vertex in the working interval in $\mathcal{E}$. We show the following auxiliary result. Let M be non-empty, let $y' \in M$, let d' be the leftmost neighbour of y' in $\mathcal{E}$ and let l be the number of vertices between w' and y' in $\mathcal{E}$. Then, the following holds:

1) $\mathrm{d}_{\mathcal{E}}(a^*, d') \geq \mathrm{d}_{\mathcal{E}}(w', y') + 1 = 2l - 1$

2) if the slot at position $\mathcal{E}(a) - 1$ in $\mathcal{E}$ is non-empty then $\mathrm{d}_{\mathcal{E}}(a^*, d') \geq \mathrm{d}_{\mathcal{E}}(w', y') + 2 = 2l$.

Note that $\mathrm{d}_{\mathcal{E}}(w', y') = 2l - 2$, since pairs of consecutive vertices between w' and y' in $\mathcal{E}$ are at distance 2. Hence, the first statement directly follows with the definition of M:

$$\mathrm{d}_{\mathcal{E}}(a^*, d') = \mathrm{d}_{\mathcal{E}}(a^*, w') + \mathrm{d}_{\mathcal{E}}(w', y') - \mathrm{d}_{\mathcal{E}}(d', y') \geq \mathrm{D}(G, \mathcal{E}) + 1 + \mathrm{d}_{\mathcal{E}}(w', y') - \mathrm{D}(G, \mathcal{E}).$$

The second statement holds with similar arguments in the long case and in case $a^* \prec_{\mathcal{E}} a$. So, consider the short case where $a^* = a$. By the definition of the working interval, all pairs of consecutive X-vertices between a and u are at distance 2. In particular, there is an I-vertex between every pair of consecutive X-vertices between a and u. Thus, $\mathrm{d}_{\mathcal{E}}(d', y') \leq \mathrm{D}(G, \mathcal{E}) - 1$, since the slot at distance $\mathrm{D}(G, \mathcal{E})$ to the left of y' is occupied by an I-vertex, and the correctness of the second statement follows.

Let A be the set of I-vertices in the working interval. For $S \subseteq A$, an ordering for $X \cup S$ is *good* if the X-vertices are ordered by decreasing degree and each I-vertex is between two neighbours. Note that the two neighbours of an I-vertex naturally are X-vertices. Let $B \subseteq A$ be of largest cardinality among all subsets of A such that $X \cup B$ has a good ordering; let β be a good ordering for $X \cup B$ such that no I-vertex in β can appear further right without changing the order of the I-vertices. Without loss of generality, we can assume that u and v are the respectively leftmost and rightmost vertex in β. Denote by $n(x)$ the number of I-vertices to the right of X-vertex x in β. We determine a lower bound on the value of $n(x)$. Assign I-vertices to X-vertices in the following way. For $y \in A$:

- if $v \prec_{\mathcal{E}} y$ then assign y to the close vertex to the left
- if $u \prec_{\mathcal{E}} y \prec_{\mathcal{E}} v$ then assign y to the close vertex to the right
- if $a^* = a$: if $y \prec_{\mathcal{E}} u$ then assign y to the close vertex to the left
- if $a^* \prec_{\mathcal{E}} a$: let z be the leftmost X-vertex such that the close vertex to the right is an X-vertex: if $y \prec_{\mathcal{E}} z$ then assign y to the close vertex to the right, if $z \prec_{\mathcal{E}} y \prec_{\mathcal{E}} u$ then assign y to the close vertex to the left.

Note that every vertex from A is assigned to an X-vertex, u and v have no assigned I-vertex (particularly since v has no I-vertex neighbours) and no X-vertex has two assigned I-vertices (particularly since the close vertex to the right of z is an X-vertex). Let x be an X-vertex satisfying $u \prec_{\beta} x \prec_{\beta} v$, and let x be assigned an I-vertex y. If the close vertex to the left of x in β is an X-vertex then y is to the right of x in β; otherwise y could be placed between x and the close vertex to the left, thus obtaining an ordering of the desired form with another I-vertex in the ordering or an I-vertex further to the right. Hence, $n(x)$ for an arbitrary X-vertex x is at least the number of X-vertices to the right of x in β that are assigned an I-vertex. In particular, $n(u)$ is equal to $|A|$, which shows that $B = A$ and β is an ordering for $X \cup A$, i.e., for all vertices in the working interval.

Denote by I_l and I_r the set of I-vertices respectively to the left and right of the working interval in $\mathcal{E}$. Note that $M \subseteq I_r$. We define an embedding $\mathcal{F}$ for G. We specify the underlying vertex ordering of the embedding; the actual embedding is obtained by adding the necessary (but no unnecessary) empty slots: place the vertices in $I_l \cup M$ ordered increasingly by degree where, for convenience reasons, vertices in I_l preserve their $\mathcal{E}$-order and vertices in M appear in their reverse $\mathcal{E}$-order, then place the vertices from the working interval in order according to β, then place the vertices in $I_r \setminus M$ in their $\mathcal{E}$-order. By definition, $\mathcal{F}$ is non-contractive, and there are no empty slots between u and v in $\mathcal{F}$ due to the definition of β. In the following, we determine the distortion of $\mathcal{F}$. Since the working interval in $\mathcal{E}$ does not contain empty slots, $\mathrm{d}_{\mathcal{F}}(u,v) = \mathrm{d}_{\mathcal{E}}(a^*,b^*)$. Furthermore, the slot at position $\mathcal{F}(u)-1$ in $\mathcal{F}$ is non-empty if $I_l \cup M$ is non-empty, and the slot at position $\mathcal{F}(v)+1$ in $\mathcal{F}$ is empty. As the first case, we consider the vertices to the right of u in $\mathcal{F}$. Since u is universal, thus the leftmost neighbour of every vertex, it suffices to consider the distance between u and the rightmost vertex in $\mathcal{F}$. Let w be the rightmost vertex in $\mathcal{E}$ that is not contained in M. Let $\mathrm{d}_{\mathcal{E}}(b,w) \geq 2$. Then, w is the rightmost vertex in $\mathcal{F}$. It holds either $\mathrm{D}(G,\mathcal{E}) = \mathrm{d}_{\mathcal{E}}(a,w)$ (which also means $a^* = a$) or $\mathrm{D}(G,\mathcal{E}) \geq \mathrm{d}_{\mathcal{E}}(a,w)+1$. Thus, $\mathrm{d}_{\mathcal{F}}(u,w) \leq \mathrm{D}(G,\mathcal{E})$. Let $\mathrm{d}_{\mathcal{E}}(b,w) \leq 1$. Then, $\mathcal{E}(w)$ belongs to the working interval and v is the rightmost vertex in $\mathcal{F}$. If $\mathrm{D}(G,\mathcal{E}) \geq \mathrm{d}_{\mathcal{E}}(a,w)+1$ then $\mathrm{d}_{\mathcal{F}}(u,v) \leq \mathrm{D}(G,\mathcal{E})$. Let $\mathrm{D}(G,\mathcal{E}) = \mathrm{d}_{\mathcal{E}}(a,w)$. If $\mathrm{d}_{\mathcal{E}}(w,w') = 2$ then $a^* = a$ and $\mathrm{d}_{\mathcal{F}}(u,v) \leq \mathrm{D}(G,\mathcal{E})$. Let $\mathrm{d}_{\mathcal{E}}(w,w') = 1$. Then, $w = b = b^*$ and $bw' \in E$ and $aw' \notin E$ and therefore $d_G(a) < d_G(b)$ and therefore the slot at position $\mathcal{E}(a)-1$ is empty. Consequently, $a^* = a$, and thus $\mathrm{d}_{\mathcal{F}}(u,v) = \mathrm{d}_{\mathcal{E}}(a,b)$.

As the second case, we consider the vertices to the left of u in $\mathcal{F}$. We define a "correction value" s. If $a^* = a$ then $s =_{\mathrm{def}} 0$, if $a^* \prec_{\mathcal{E}} a$ then $s =_{\mathrm{def}} -1$. Let $y \in I_l \cup M$, and let ℓ be the number of vertices from M between y and u in $\mathcal{F}$. Suppose there is no vertex from I_l between y and u in $\mathcal{F}$. In particular, $y \in M$ and $\mathrm{d}_{\mathcal{F}}(y,u) = 2\ell - 1$. Let d be the rightmost neighbour of y in $\mathcal{F}$, and let d' be the leftmost neighbour of y in $\mathcal{E}$. We determine $\mathrm{d}_{\mathcal{F}}(d,v)$. All X-vertices to the left of d' in $\mathcal{E}$ have degree smaller than $d_G(d')$ and $d_G(d)$ and thus are to the right of d in $\mathcal{F}$. By the result about the value of $n(d)$ it holds that $n(d)+s$ is not smaller than the number of I-vertices to the left of d' in the working interval in $\mathcal{E}$. Remember that $s=0$ implies that no I-vertex to the left of d' in $\mathcal{E}$ is assigned to d'. Thus, $\mathrm{d}_{\mathcal{F}}(d,v) \geq \mathrm{d}_{\mathcal{E}}(a^*,d') + s \geq 2\ell - 1$ due to the auxiliary result, and hence

$$\mathrm{d}_{\mathcal{F}}(y,d) = \mathrm{d}_{\mathcal{F}}(y,u) + \mathrm{d}_{\mathcal{F}}(u,v) - \mathrm{d}_{\mathcal{F}}(d,v) \leq 2\ell - 1 + \mathrm{d}_{\mathcal{F}}(u,v) - 2\ell + 1 = \mathrm{d}_{\mathcal{F}}(u,v)\,.$$

Now, let there be a vertex from I_l between y and u in $\mathcal{F}$; let y^* be the leftmost vertex from I_l between y and u in $\mathcal{F}$. It holds that $\mathrm{d}_{\mathcal{F}}(y,u) \leq \mathrm{d}_{\mathcal{E}}(y^*,a^*) + 2\ell + s$. Let c be the rightmost neighbour of y in $\mathcal{F}$. We determine $\mathrm{d}_{\mathcal{F}}(c,v)$. Let c^* be the rightmost neighbour of y^* in $\mathcal{F}$, and let c' be the rightmost neighbour of y^* in $\mathcal{E}$. Note that y^* is adjacent to c. All X-vertices to the right of c' in $\mathcal{E}$ have degree smaller than $d_G(c')$ and $d_G(c^*)$ and thus are to the right of c^* in $\mathcal{F}$. If $v \prec_{\mathcal{E}} c'$ and the close vertex to the right of c' in $\mathcal{E}$ is an I-vertex then $n(c^*)$ is at least the number of I-vertices to the right of c' in the working interval in $\mathcal{E}$ minus 1. Since v is to the right of c in $\mathcal{F}$ it follows that $\mathrm{d}_{\mathcal{F}}(c^*,v) \geq \mathrm{d}_{\mathcal{E}}(c',b^*)$. If $c' \prec_{\mathcal{E}} v$ then $\mathrm{d}_{\mathcal{F}}(c^*,v) \geq \mathrm{d}_{\mathcal{E}}(c',b^*)$ due to the result about the value of $n(c^*)$. Thus, if $\ell = 0$ then $y = y^*$ and $c = c^*$ and $\mathrm{d}_{\mathcal{F}}(y,c) = \mathrm{d}_{\mathcal{F}}(y,v) - \mathrm{d}_{\mathcal{F}}(c,v) \leq \mathrm{d}_{\mathcal{E}}(y,b^*) - \mathrm{d}_{\mathcal{E}}(c',b^*) = \mathrm{d}_{\mathcal{E}}(y,c')$. So, let $\ell \geq 1$. Let y^{**} be the leftmost vertex from M between y and u in $\mathcal{F}$. With the results shown above it follows that $\mathrm{d}_{\mathcal{F}}(c,v) \geq \mathrm{d}_{\mathcal{E}}(c^*,b^*) + 2\ell + s$. Here, it is important to note that the X-vertices that contribute to this number are to the left of u for y^{**} and to the right of u for y^* in $\mathcal{E}$, thus disjoint sets. Therefore, we obtain:

$$\begin{array}{rcl} \mathrm{d}_{\mathcal{F}}(y,c) & = & \mathrm{d}_{\mathcal{F}}(y,u) + \mathrm{d}_{\mathcal{F}}(u,v) - \mathrm{d}_{\mathcal{F}}(c,v) \\ & \leq & \mathrm{d}_{\mathcal{E}}(y^*,a^*) + 2\ell + s + \mathrm{d}_{\mathcal{E}}(a^*,b^*) - \mathrm{d}_{\mathcal{E}}(c^*,b^*) - 2\ell - s \;\; = \;\; \mathrm{d}_{\mathcal{E}}(y^*,c^*)\,. \end{array}$$

Hence, $\mathrm{D}(G, \mathcal{F}) \leq \mathrm{D}(G, \mathcal{E})$, and $\mathcal{F}$ is a minimum distortion embedding for G of the desired form. This completes the proof. ∎

The structural result of Lemma 4.2 leads to a simple algorithm for computing the distortion of threshold graphs. The algorithm finds an embedding of smallest distortion among all non-contractive embeddings where the X-vertices appear in decreasing degree order. Lemma 4.2 then shows that this actually is a minimum distortion embedding. Let $G = (X, I, E)$ be a connected threshold graph and let $\mathcal{E}$ be an embedding for G where the X-vertices appear in decreasing degree order. Let u be the leftmost X-vertex in $\mathcal{E}$. Note that u is universal. Denote by $R(\mathcal{E})$ the distance in $\mathcal{E}$ between u and the rightmost vertex, and denote by $L(\mathcal{E})$ the maximum taken over all distances between a vertex to the left of u and its rightmost neighbour in $\mathcal{E}$. If u is the leftmost vertex in $\mathcal{E}$ then $L(\mathcal{E}) = 0$. It holds that $\mathrm{D}(G, \mathcal{E}) = \max\{L(\mathcal{E}), R(\mathcal{E})\}$. The following algorithm computes the distortion of connected threshold graphs. It iteratively decreases the distortion of an initial embedding by moving vertices.

Algorithm `thrg-distortion`
Input connected threshold graph $G = (X, I, E)$ and
increasing degree ordering $\langle y_1, \ldots, y_{|I|} \rangle$ of the I-vertices, i.e., such that $d_G(y_1) \leq \cdots \leq d_G(y_{|I|})$

begin
  let $\mathcal{E}_0 =$ `start-embedding`; **let** u be the leftmost vertex in $\mathcal{E}_0$; **let** $i = 0$;
  while $R(\mathcal{E}_i) \geq L(\mathcal{E}_i) + 2$ **and** $i < |I|$ **do**
    set $i = i + 1$; **let** $\mathcal{E}_i =$ `moveleft`$(\mathcal{E}_{i-1}, y_i)$
  end while;
  let v be the close I-vertex to the right of u; **let** $\mathcal{E} =$ `moveright`$(\mathcal{E}_i, v)$;
  return $\min\{\mathrm{D}(G, \mathcal{E}), \mathrm{D}(G, \mathcal{E}_i)\}$ and the corresponding embedding
end.

To complete the definition of `thrg-distortion`, we explain three operations. These operations define embeddings. For ease of description, we only define the underlying vertex orderings; the actual embeddings are non-contractive and without unnecessary empty slots.

`start-embedding`
The X-vertices appear in decreasing degree order, and the I-vertices are added as follows, iteratively processed in order $y_{|I|}, \ldots, y_1$: y_i is placed rightmost between two neighbours if possible, and if not possible it is placed at the right end, particularly to the right of the rightmost X-vertex. The result for a sample graph is depicted in Figure 1.

`moveleft`$(\mathcal{E}_{i-1}, y_i)$
The result is obtained from $\mathcal{E}_{i-1}$ by moving y_i and making it the close vertex to the left of u.

`moveright`$(\mathcal{E}_i, v)$
If v is undefined then $\mathcal{E} = \mathcal{E}_i$; otherwise move v to the right and place it as the rightmost vertex, particularly to the right of the rightmost X-vertex.

For the correctness of the algorithm, the following observations are important. There are no empty slots between X-vertices in the start embedding. The start embedding has smallest distortion among all non-contractive embeddings with the leftmost vertex a universal vertex. After application of operation `moveleft`, the distance between u and the rightmost vertex decreases by 1 or 2 depending on whether y_i is between neighbours in $\mathcal{E}_{i-1}$ or to the right of the rightmost X-vertex.

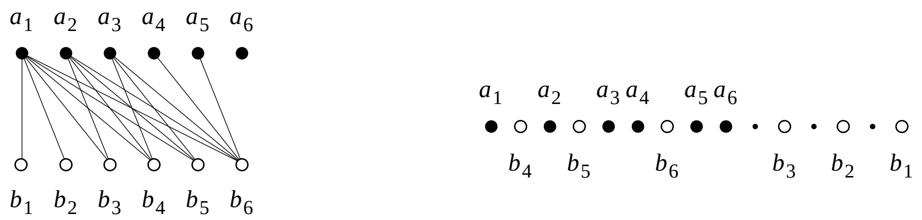

Figure 1: The left hand side shows a threshold graph. The X-vertices are represented by full circles and the I-vertices are represented by empty circles. The edges between X-vertices are omitted. The right hand side shows the result of the `start-embedding` procedure when applied to the graph.

A *succinct representation* of a threshold graph lists the vertices and their degrees. This representation is unique for a threshold graph.

Theorem 4.3 *There is an $\mathcal{O}(n)$-time algorithm that computes the distortion of a connected threshold graph on n vertices and outputs a minimum distortion embedding. The graph is given in succinct representation.*

Proof. We prove that Algorithm `thrg-distortion` satisfies the theorem. Let $G = (X, I, E)$ be a connected threshold graph where $y_1, \ldots, y_{|I|}$ are the I-vertices in increasing degree order. Apply `thrg-distortion` to G. Let r be the number of **while** loop executions and let embeddings $\mathcal{E}_0, \ldots, \mathcal{E}_r, \mathcal{E}$ and vertex u be defined according to `thrg-distortion`. Note that $r \geq 1$ since $L(\mathcal{E}_0) = 0$ and $R(\mathcal{E}_0) \geq 2$. We show that $\mathcal{E}$ or $\mathcal{E}_r$ has smallest distortion among all non-contractive embeddings for G with the X-vertices appearing in decreasing degree order. Lemma 4.2 then shows that $\mathcal{E}$ or $\mathcal{E}_r$ is a minimum distortion embedding for G.

We begin by studying $\mathcal{E}_0, \ldots, \mathcal{E}_r$. Let $1 \leq i \leq r$. It clearly holds that $L(\mathcal{E}_{i-1}) + 1 \leq L(\mathcal{E}_i)$ and $R(\mathcal{E}_{i-1}) - 2 \leq R(\mathcal{E}_i) \leq R(\mathcal{E}_{i-1}) - 1$. Furthermore, the rightmost neighbour of y_i is at distance at most $R(\mathcal{E}_i)$ in $\mathcal{E}_i$ since the rightmost X-vertex has no I-vertex neighbour, so that $L(\mathcal{E}_i) \leq \max\{L(\mathcal{E}_{i-1}) + 2, R(\mathcal{E}_i)\}$. Together with the **while** loop condition, we obtain $L(\mathcal{E}_i) \leq R(\mathcal{E}_{i-1})$. Thus, $\mathrm{D}(G, \mathcal{E}_i) \leq \mathrm{D}(G, \mathcal{E}_{i-1})$ and $L(\mathcal{E}_i) - R(\mathcal{E}_i) \leq 2$. We distinguish between two cases with respect to the value of r.

Case A: $r = |I|$

This means that there is no I-vertex to the right of u in $\mathcal{E}_r$. Note that for every non-contractive embedding $\mathcal{F}$ for G with the X-vertices appearing in decreasing degree order and an I-vertex to the right of some X-vertex, it holds that $R(\mathcal{F}) \geq R(\mathcal{E}_{r-1}) \geq R(\mathcal{E}_r) + 1 = |X|$. Since $L(\mathcal{E}_{r-1}) < R(\mathcal{E}_{r-1})$, the above shown inequalities prove that $\mathrm{D}(G, \mathcal{F}) \geq R(\mathcal{F}) \geq R(\mathcal{E}_{r-1}) = \mathrm{D}(G, \mathcal{E}_{r-1}) \geq \mathrm{D}(G, \mathcal{E}_r)$. Thus, $\mathcal{E}_r$ is of minimum distortion among all non-contractive embeddings for G with the X-vertices appearing in decreasing degree order.

Case B: $r < |I|$

This means that there is at least one I-vertex to the right of u in $\mathcal{E}_r$. By the inequalities from the second paragraph and the **while** loop condition, it holds that $-1 \leq L(\mathcal{E}_r) - R(\mathcal{E}_r) \leq 2$. Assume there is a non-contractive embedding for G with the X-vertices appearing in decreasing degree order without empty slots between X-vertices and of distortion smaller than $\mathrm{D}(G, \mathcal{E}_r)$. Choose $\mathcal{F}$ to be such an embedding with the smallest number of I-vertices to the left of all X-vertices. Without loss of generality, we can assume that u is the leftmost X-vertex in $\mathcal{F}$, and I-vertices to the left of u in $\mathcal{F}$ appear in increasing degree order and have degree not larger than any I-vertex to the right of u. We will show that $\mathrm{D}(G,\mathcal{F}) = \mathrm{D}(G,\mathcal{E})$.

There are at most $|I| - r$ I-vertices to the right of u in $\mathcal{F}$, as can be seen as follows: $L(\mathcal{E}_{r-1}) < R(\mathcal{E}_{r-1})$ by the **while** loop condition and $R(\mathcal{G}) \geq R(\mathcal{E}_{r-1})$ for all non-contractive embeddings $\mathcal{G}$ with X-vertices appearing in decreasing degree order, starting with u, and at least $|I| - r + 1$ I-vertices to the right of u. The second property directly follows from the definition of the start embedding. Thus, by the assumption about the I-vertices in $\mathcal{F}$, we can assume that $y_1, \ldots, y_r$ are to the left of u in $\mathcal{F}$.

Let y be the leftmost I-vertex in $\mathcal{E}_r$ such that $\mathrm{d}_{\mathcal{E}_r}(y,b) = L(\mathcal{E}_r)$ for b the rightmost neighbour of y. Then, y is the leftmost neighbour of b in $\mathcal{E}_r$. With exchanging vertices of the same degree, if necessary, we can assume that b is the rightmost neighbour of y and y is the leftmost neighbour of b also in $\mathcal{F}$. If $L(\mathcal{F}) > L(\mathcal{E}_r)$ then $\mathrm{D}(G,\mathcal{F}) \geq L(\mathcal{E}_r) + 1 \geq R(\mathcal{E}_r)$, i.e., $\mathrm{D}(G,\mathcal{F}) \geq \mathrm{D}(G,\mathcal{E}_r)$ as a contradiction to the choice of $\mathcal{F}$. Thus, $L(\mathcal{F}) \leq L(\mathcal{E}_r)$, and in particular, $\mathrm{d}_{\mathcal{F}}(y,b) \leq \mathrm{d}_{\mathcal{E}_r}(y,b)$. We consider the number of I-vertices between y and b in $\mathcal{E}_r$ and $\mathcal{F}$. Suppose for a contradiction that there are at least as many I-vertices between y and b in $\mathcal{F}$ as in $\mathcal{E}_r$. Since the number of X-vertices between y and b is equal in $\mathcal{E}_r$ and $\mathcal{F}$, the partition of the I-vertices between y and b in $\mathcal{E}_r$ into vertices to the left of u and vertices to the right of u is uniquely defined. For this argument, it is important to remember that there are no empty slots between X-vertices, to that every I-vertex to the left of u contributes 2 to $\mathrm{d}_{\mathcal{E}_r}(u,b)$ and every I-vertex to the right of u contributes 1. It therefore follows the same for $\mathcal{F}$, so that $\mathrm{d}_{\mathcal{F}}(y,b) = \mathrm{d}_{\mathcal{E}_r}(y,b)$ and $L(\mathcal{F}) = L(\mathcal{E}_r)$. Now, observe that the same number of I-vertices to the left of b in $\mathcal{E}_r$ and $\mathcal{F}$ directly implies the same number of I-vertices to the right of b in $\mathcal{E}_r$ and $\mathcal{F}$. Since $R(\mathcal{F}) < R(\mathcal{E}_r)$ must hold due to the choice of $\mathcal{F}$, there is an I-vertex between two neighbours to the right of b in $\mathcal{F}$ that is to the right of all X-vertices in $\mathcal{E}_r$. This, however, contradicts the definition of the start embedding. Hence, we conclude that the number of I-vertices between y and b in $\mathcal{F}$ is smaller than between y and b in $\mathcal{E}_r$. And since there are exactly r I-vertices to the left of u in $\mathcal{E}_r$ and at least r I-vertices to the left of u in $\mathcal{F}$, the number of I-vertices between u and b in $\mathcal{F}$ is smaller than the number in $\mathcal{E}_r$.

We have seen that there are I-vertices between u and b in $\mathcal{E}_r$ that are not between u and b in $\mathcal{F}$. In particular, $u \neq b$. Let M be the set of I-vertices between u and b in $\mathcal{E}_r$ that are not between u and b in $\mathcal{F}$. Without loss of generality, we can assume that no vertex in M is between X-vertices in $\mathcal{F}$. This follows from the definition of the start embedding (otherwise, there would be an empty slot between X-vertices). Note that $|M| \geq 1$. As a first case, assume that $|M| \geq 2$. Observe that the definition of M implies the existence of (at least) two pairs of consecutive X-vertices in $\mathcal{F}$. Let a and a' be two vertices from M where at least one of them, say a, is to the right of u in $\mathcal{F}$. Obtain $\mathcal{F}'$ from $\mathcal{F}$ as follows, where unnecessary empty slots are deleted and necessary empty slots are inserted:

- if a' is to the left of u in $\mathcal{F}$ then move a and a' between two pairs of consecutive X-vertices between u and b
- if a' is to the right of u in $\mathcal{F}$ then move a and a' as in the previous case and additionally move the close vertex to the left of u in $\mathcal{F}$ to the right end.

For an illustration of the two cases, see Figure 2. It holds that $L(\mathcal{F}') = L(\mathcal{F})$ and $R(\mathcal{F}') = R(\mathcal{F})$. Since $\mathcal{F}'$ contains fewer vertices to the left of u than $\mathcal{F}$, this is a contradiction to the choice of $\mathcal{F}$. We conclude for the case $|M| \geq 2$ that all vertices in M are to the left of u in $\mathcal{F}$.

Figure 2: The two figures show modifications of embeddings, that are used in the proof of Theorem 4.3. In both figures, the arrows describe the places where the vertices are moved to. Necessary slots are inserted implicitly.

Let p be the number of vertices to the left of u in $\mathcal{F}$. Suppose that $p \geq r+1$. Without loss of generality, we can assume that $y_1, \ldots, y_p$ are the vertices to the left of u in $\mathcal{F}$, and $y_{r+1}, \ldots, y_p$ are to the right of y in $\mathcal{F}$. Thus, $\mathrm{d}_{\mathcal{F}}(y,u) = \mathrm{d}_{\mathcal{E}_r}(y,u) + 2(p-r)$. From $\mathrm{d}_{\mathcal{F}}(y,b) \leq \mathrm{d}_{\mathcal{E}_r}(y,b)$, it follows that $|M| \geq 2(p-r)$, which particularly means $|M| \geq 2$. Then, all vertices in M are between y and u in $\mathcal{F}$, so that $p-r \geq |M|$. Thus, $|M| \geq 2|M|$, which yields a contradiction for $|M| \geq 1$. We conclude that $p = r$ and $|M| = 1$. Consider $\mathcal{E}$, that is obtained from $\mathcal{E}_r$ by moving the close vertex to the right of u to the right end. Note that the moved vertex is between u and b, since $|M| = 1$. Then, $\mathrm{d}_{\mathcal{E}}(y,b) = \mathrm{d}_{\mathcal{F}}(y,b) = L(\mathcal{E}_r) - 1$ and $R(\mathcal{E}) = R(\mathcal{E}_r) + 1$. By minimality of $R(\mathcal{E}_r)$, $R(\mathcal{F}) = R(\mathcal{E})$. Consequently, $\mathrm{D}(G,\mathcal{F}) = \mathrm{D}(G,\mathcal{E})$. We conclude that `thrg-distortion` correctly computes the distortion of G and outputs a minimum distortion embedding.

For the running time of `thrg-distortion`, observe first that the leftmost and rightmost neighbour of every I-vertex in a decreasing degree order of the X-vertices can be computed in $\mathcal{O}(n)$ time from the succinct representation. Therefore, it is sufficient to describe how L- and R-values are computed efficiently. Clearly, R-values are obtained by simple subtraction, since it suffices to remember whether an I-vertex to the right of u is between two neighbours (whose moving results in decreasing the value by 1) or at the end (whose moving results in decreasing the value by 2). For I-vertices, we have to distinguish two cases. If the moved I-vertex is at the right end then the L-value increases by at least 2 and is the maximum over the distance of the moved vertex to its rightmost neighbour and the increased previous L-value. If the moved I-vertex is between neighbours then some distances increase by 2 and some distances increase by 1. Here, we have to find the leftmost I-vertex with rightmost neighbour at maximum distance. This information can be computed in a preprocessing step, when there is no I-vertex between X-vertices and this is only a neighbourhood cardinality problem. Hence, `thrg-distortion` has an $\mathcal{O}(n)$-time implementation. ■

The main structural result about minimum distortion embeddings in this section is given in Lemma 4.2. A natural question is whether a similar result holds also for I-vertices. As a complementary result, we show that this is indeed the case. Note, however, that this may require empty slots between X-vertices, since I-vertices may appear consecutively between X-vertices.

Proposition 4.4 *Let $G = (X, I, E)$ be a connected threshold graph. There is a minimum distortion embedding for G such that the X-vertices appear in decreasing degree order and the I-vertices appear in increasing degree order.*

Proof. Let $\mathcal{E}$ be a minimum distortion embedding for G without empty slots between X-vertices such that the X-vertices appear in decreasing degree order; $\mathcal{E}$ exists due to Lemma 4.2. Denote by a and b the respectively leftmost and rightmost X-vertex in $\mathcal{E}$. Since the vertex degree corresponds to neighbourhood inclusion, a simple vertex exchange argument shows that we can

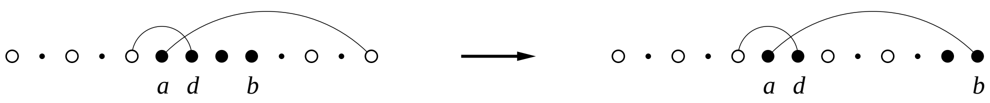

Figure 3: An illustration of the central embedding modification in the proof of Proposition 4.4 for obtaining a minimum distortion embedding where X-vertices are ordered by decreasing degree and I-vertices are ordered by increasing degree.

assume without loss of generality that no I-vertex to the left of a has degree larger than any I-vertex to the right of a in $\mathcal{E}$. Assume that there is an I-vertex to the right of b in $\mathcal{E}$. Let d be the rightmost neighbour of the close I-vertex to the left of a. Then, every I-vertex to the right of a is also adjacent to d, particularly the close I-vertex to the right of b. Let $\mathcal{F}$ be the non-contractive embedding without unnecessary empty slots with underlying vertex ordering the following: modify $\mathrm{ord}(\mathcal{E})$ by moving the I-vertices to the right of b between d and the close vertex to the right of d. See Figure 3 for an illustration of the construction. By the construction it is clear that $L(\mathcal{F}) = L(\mathcal{E})$ and $R(\mathcal{F}) \leq R(\mathcal{E})$. Thus, $\mathcal{F}$ is a minimum distortion embedding for G without I-vertices to the right of b. To obtain an embedding that satisfies the statement, it remains to exchange pairs of I-vertices to make them appear in increasing degree order. Here, it is important to note that no new empty slots are required between a and b since the close X-vertex to the left of an I-vertex between a and b is a neighbour. ■

Note that Proposition 4.4 does not result in a straightforward algorithm for computing the distortion. The main reason is that it does not speak about the positions of the I-vertices that are placed between X-vertices in the initial minimum distortion embedding $\mathcal{E}$.

5 Distortion of bipartite permutation graphs

Bipartite permutation graphs are permutation graphs that are bipartite. For the definition and properties of permutation graphs, we refer to [5]. Let $G = (A, B, E)$ be a bipartite graph. A *strong ordering* for G is a pair of orderings (σ_A, σ_B) on respectively A and B such that for every pair of edges ab and $a'b'$ in E with $a, a' \in A$ and $b, b' \in B$, $a \prec_{\sigma_A} a'$ and $b' \prec_{\sigma_B} b$ implies that ab' and $a'b$ are in E. If we denote by $(\sigma_A, \sigma_B)^R$ the pair of the reverse of σ_A and σ_B then $(\sigma_A, \sigma_B)^R$ is also a strong ordering for G. The following characterisation of bipartite permutation graphs is the only property that we will need in this section, and thus we use it as a definition.

Theorem 5.1 ([21]) *A bipartite graph is a bipartite permutation graph if and only if it has a strong ordering.*

Spinrad et al. give a linear-time recognition algorithm for bipartite permutation graphs that produces a strong ordering if the input graph is bipartite permutation [21]. It follows from the definition of a strong ordering that if $G = (A, B, E)$ is a connected bipartite permutation graph then any strong ordering (σ_A, σ_B) satisfies the following. For every vertex a in A, the neighbours of a appear consecutively in σ_B. Furthermore, if $N(a) \subseteq N(a')$ for two vertices $a, a' \in A$ then a is adjacent to the leftmost or rightmost neighbour of a' in σ_B.

We show two main results about distortion of bipartite permutation graphs. We give a fast algorithm for computing the distortion of bipartite permutation graphs and we give a complete characterisation of bipartite permutation graphs of bounded distortion by forbidden induced subgraphs. Before that, we consider the relationship of bandwidth and distortion for bipartite permutation graphs. For each vertex u of a bipartite permutation graph, we denote by $cc(u)$ the colour class of u and by $\overline{cc}(u)$ the other colour class, i.e., A and B.

5.1 Relationship to bandwidth

As already mentioned, bandwidth and distortion do not always coincide on bipartite permutation graphs, not even on the restricted subclass of complete bipartite graphs. As an example, $\mathrm{bw}(K_{3,4}) = 4$ (two vertices of the second colour class are placed first, followed by all three vertices of the first colour class, followed by the last two vertices of the second colour class) and $\mathrm{D}(K_{3,4}) = 5$. The question arises whether the difference between bandwidth and distortion can be arbitrarily large, like for cycles. We answer this question completely in this subsection. We show that distortion is an approximation by ratio 2 of bandwidth of connected bipartite permutation graphs.

Let $G = (A, B, E)$ be a connected bipartite permutation graph with strong ordering (σ_A, σ_B). We say that a vertex ordering β for G is *normalised* (with respect to (σ_A, σ_B)) if it satisfies the following two conditions:

(C1) for every pair a, a' of vertices in A: $a \prec_{\sigma_A} a'$ implies $a \prec_\beta a'$,
for every pair b, b' of vertices in B: $b \prec_{\sigma_B} b'$ implies $b \prec_\beta b'$

(C2) for every triple u, v, w of vertices of G where $u \prec_\beta v \prec_\beta w$ and $uw \in E$:
$uv \in E$ or $vw \in E$.

Condition (C1) requires that β respects the two given orderings. Orderings that respect condition (C2) are called *cocomparability orderings*; hence, condition (C2) requires β to be a cocomparability ordering for G.

As a corollary of a theorem by Fishburn et al. [8], the following normalisation result for optimal bandwidth orderings can be obtained.

Theorem 5.2 ([10]) *Let $G = (A, B, E)$ be a connected bipartite permutation graph with strong ordering (σ_A, σ_B), and let $k \geq 0$ be an integer. If G has a k-ordering then G has a k-ordering that is normalised with respect to (σ_A, σ_B).*

Theorem 5.3 *Let G be a connected bipartite permutation graph. Then,* $\mathrm{D}(G) \leq 2 \cdot \mathrm{bw}(G) - 1$.

Proof. Let (σ_A, σ_B) be a strong ordering for $G = (A, B, E)$. Let β be a $\mathrm{bw}(G)$-ordering for G. By Theorem 5.2 we can assume that β is normalised with respect to (σ_A, σ_B). Let $\mathcal{E}$ be the non-contractive embedding for G without unnecessary empty slots and underlying vertex ordering β. We determine $\mathrm{D}(G, \mathcal{E})$ by showing for every pair u, v of adjacent vertices of G that $\mathrm{d}_\mathcal{E}(u, v) \leq 2 \cdot \mathrm{d}_\beta(u, v) - 1$. We prove the claim by induction over the distances in β between adjacent vertices. Let u, v be a pair of adjacent vertices such that $\mathrm{d}_\beta(u, v) = 1$; then $\mathrm{d}_\mathcal{E}(u, v) = 1$. Suppose that the claim holds for each pair of adjacent vertices at distance at most s in β. Let u, v be a pair of adjacent vertices such that $u \prec_\beta v$ and $\mathrm{d}_\beta(u, v) = s + 1$. From condition (C2), it follows for the vertices between u and v that all vertices from $cc(u)$ are adjacent to v and all vertices from $cc(v)$ are adjacent to u. Hence, vertices of the same colour class are at distance 2 in G and vertices from different colour classes are at distance 1 or 3 in G. We distinguish two cases. First, let there be no pair of consecutive vertices between u and v in $\mathcal{E}$ at distance 3 in G. Then, pairs of consecutive vertices between u and v are at distance at most 2 in $\mathcal{E}$. Furthermore, since u and v are from different colour classes, there is a pair of consecutive vertices from different colour classes between u and v, that are adjacent due to the normalisation conditions and the properties of strong orderings. Thus, $\mathrm{d}_{\mathcal{E}}(u,v) \leq 2 \cdot \mathrm{d}_{\beta}(u,v) - 1$. For the other case, let x, y be a pair of consecutive vertices between u and v such that $x \prec_{\beta} y$ and $\mathrm{d}_{\mathcal{E}}(x,y) = 3$. Note that $x \neq u$ and $y \neq v$ by the above observation and x and y are from different colour classes. Since $cc(u) = cc(x)$ and condition (C2) imply $xy \in E$ and therefore a contradiction to $\mathrm{d}_{\mathcal{E}}(x,y) > 1$, it holds that $cc(u) = \overline{cc}(x) = cc(y) = \overline{cc}(v)$. By condition (C2), $ux \in E$ and $yv \in E$, and since $\mathrm{d}_{\beta}(u,x) \leq \mathrm{d}_{\beta}(u,v) - 2$ and $\mathrm{d}_{\beta}(y,v) \leq \mathrm{d}_{\beta}(u,v) - 2$, we know $\mathrm{d}_{\mathcal{E}}(u,x) \leq 2 \cdot \mathrm{d}_{\beta}(u,x) - 1$ and $\mathrm{d}_{\mathcal{E}}(y,v) \leq 2 \cdot \mathrm{d}_{\beta}(y,v) - 1$ by induction hypothesis. Consequently,

$$\begin{aligned}\mathrm{d}_{\mathcal{E}}(u,v) = \mathrm{d}_{\mathcal{E}}(u,x) + 3 + \mathrm{d}_{\mathcal{E}}(y,v) &\leq 2 \cdot \mathrm{d}_{\beta}(u,x) - 1 + 2 \cdot \mathrm{d}_{\beta}(x,y) + 1 + 2 \cdot \mathrm{d}_{\beta}(y,v) - 1 \\ &= 2 \cdot \mathrm{d}_{\beta}(u,v) - 1\,.\end{aligned}$$

Thus, $\mathrm{D}(G) \leq \mathrm{D}(G,\mathcal{E}) \leq 2 \cdot \mathrm{bw}(G) - 1$. ∎

The bandwidth upper bound on the distortion of connected bipartite permutation graphs in Theorem 5.3 is tight. The star graphs $K_{1,m}$ for $m \geq 2$ and m even have bandwidth $\frac{m}{2}$: a minimum bandwidth ordering is obtained by placing the centre vertex in the middle of the ordering; the distortion of $K_{1,m}$ is $m-1$ due to Theorem 3.10. Note that the bandwidth of bipartite permutation graphs can be computed in polynomial time [10].

5.2 Lower bound on the distortion of bipartite permutation graphs

Our main results on distortion of bipartite permutation graphs are an efficient computation algorithm and a forbidden induced subgraph characterisation of bipartite permutation graphs of bounded distortion. Both results are obtained simultaneously and presented in the next subsection. Both results rely on the properties of special bipartite permutation graphs, that we study in this subsection. We identify a class of bipartite permutation graphs that are distance-preserving as subgraphs and for which we can give a highly non-trivial lower bound on the distortion.

A *clawpath* is a tree that is obtained from a chordless path by attaching a leaf to every vertex of the path. Thus, clawpaths have even number of vertices and every vertex of the path has degree 3 except the end vertices of the path, that have degree 2. The number of edges on the path is called the *length* of the clawpath. Note that the smallest clawpath is $K_{1,1}$, of length 0, and one of the two vertices is chosen to form the path. (Clawpaths are thus caterpillars where every vertex that is not a leaf has exactly one neighbour that is a leaf.)

Definition 5.4 *A* thick clawpath *is a graph obtained from a clawpath by replacing each vertex by a (non-empty) independent set of new vertices.*

When replacing a vertex v by a set of new vertices $v_1, \ldots, v_\ell$ with $\ell \geq 1$, we give each v_i the same neighbourhood as v had. Thus we can view this process as iteratively adding to the graph new false twins of chosen vertices. The *underlying clawpath* of a thick clawpath is the clawpath from which the graph was obtained according to Definition 5.4. The *length* of a thick clawpath is the length of its underlying clawpath. An example is given in Figure 4.

Thick clawpaths are both bipartite and permutation. Thus, they form a subclass of bipartite permutation graphs. Furthermore, they are connected and contain at least one edge. Following Definition 5.4, any thick clawpath of length r can be represented by a pair $(x_0, \ldots, x_r)$ and

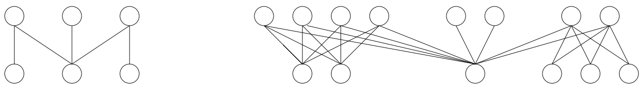

Figure 4: A clawpath G of length 2 on the left, and a thick clawpath of length 2 that has G as its underlying clawpath on the right.

$((C_0, D_0), \dots, (C_r, D_r))$ where $x_0, \dots, x_r$ are the path vertices of the underlying clawpath, C_i is the set of vertices path vertex x_i was replaced with, and D_i is the set of vertices the single leaf neighbour of x_i was replaced with. It is in fact sufficient to specify only $((C_0, D_0), \dots, (C_r, D_r))$, which we will call the *sequence representation*. Thus, every thick clawpath has a sequence representation.

Lemma 5.5 *Let G be a bipartite permutation graph. Every induced subgraph of G that is a thick clawpath is distance-preserving.*

Proof. We obtain the result in several steps. Let $H = (A, B, E)$ be an induced subgraph of G that is a thick clawpath. Let $((C_0, D_0), \dots, (C_r, D_r))$ be a sequence representation for H. If $r = 0$ then H is a complete bipartite graph and clearly a distance-preserving subgraph of G. So, let $r \geq 1$. According to Lemma 3.6, it suffices to consider the case when $C_0, D_0, \dots, C_r, D_r$ all contain exactly one vertex each, i.e., we can restrict to clawpaths. For ease of notation, we denote these vertices as $c_0, d_0, \dots, c_r, d_r$. Note that $c_0, c_1, \dots, c_r$ correspond to the path vertices. Let (σ_A, σ_B) be a strong ordering for G. Let σ be the union of σ_A and σ_B, so that we do not have to distinguish between colour classes. Without loss of generality, we can assume that $d_0 \prec_\sigma c_1$; otherwise we use $(\sigma_A, \sigma_B)^R$ as strong ordering.

Note the following observation: for (u_1, u_2, u_3, u_4) an induced path in G, it is not difficult to see that $cc(u_1) = \overline{cc}(u_2) = cc(u_3) = \overline{cc}(u_4)$, and $u_1 \prec_\sigma u_3$ if and only if $u_2 \prec_\sigma u_4$ by the properties of strong orderings. So, since (d_0, c_0, c_1, d_1) is an induced path in G, the assumption $d_0 \prec_\sigma c_1$ implies $c_0 \prec_\sigma d_1$. Assume for $1 \leq i < r$ that we have already shown $c_0 \prec_\sigma d_1 \prec_\sigma \cdots \prec_\sigma e_i$ and $d_0 \prec_\sigma c_1 \prec_\sigma \cdots \prec_\sigma f_i$, where $e_i, f_i \in \{c_i, d_i\}$ appropriately. Note that $(d_{i-1}, c_{i-1}, c_i, c_{i+1})$, $(c_{i-1}, c_i, c_{i+1}, d_{i+1})$ and $(d_i, c_i, c_{i+1}, d_{i+1})$ are induced paths in H and therefore in G. Applying the observation to the paths and the assumption $d_{i-1} \prec_\sigma c_i$, we obtain $c_{i-1} \prec_\sigma c_{i+1}$, $c_i \prec_\sigma d_{i+1}$ and $d_i \prec_\sigma c_{i+1}$. Thus, c- and d-vertices are ordered by index in the two colour classes.

We show that H is a distance-preserving subgraph of G by induction over distances in G. Since H is an induced subgraph of G, pairs of non-adjacent vertices in H are non-adjacent in G. Let $s \geq 2$ and assume that $\mathrm{d}_H(x, y) = \mathrm{d}_G(x, y)$ for all pairs x, y of vertices of H where $\mathrm{d}_G(x, y) \leq s - 1$. Let x, y be a pair of vertices of H such that $\mathrm{d}_G(x, y) = s$. Let $P = (u_0, \dots, u_s)$ be an x, y-path in G of length s. Without loss of generality, we can assume that $x \prec_\sigma y$ or $x \prec_\sigma u_{s-1}$ depending on whether x and y belong to the same colour class or to different colour classes. By iterative application of the above observation, we obtain that $u_0 \prec_\sigma u_2 \prec_\sigma u_4 \prec_\sigma \cdots$ and $u_1 \prec_\sigma u_3 \prec_\sigma \cdots$. Let $x \in \{c_j, d_j\}$ for some $0 \leq j \leq r$. Observe that the x, y-path in H contains c_{j+1} and that $y \in \{c_{j'}, d_{j'}\}$ for some $j' \geq j + 1$. We distinguish between two cases.

For the first case, let $x = d_j$. If $y = d_{j+1}$ then $\mathrm{d}_H(x, y) = 3$, and $\mathrm{d}_G(x, y) \geq 3$ since x and y belong to different colour classes and are not adjacent. Now, let $y \neq d_{j+1}$. If $d_{j+1}u_2 \in E$ then $(d_{j+1}, u_2, u_3, \dots, u_s)$ shows that there is a d_{j+1}, y-path in G of length at most $s - 1$. We apply the induction hypothesis and obtain $\mathrm{d}_H(d_{j+1}, y) = \mathrm{d}_G(d_{j+1}, y)$. Since $\mathrm{d}_H(d_{j+1}, y) = \mathrm{d}_H(d_j, y) - 1$, we conclude that $\mathrm{d}_H(d_j, y) = s$. Now, let $d_{j+1}u_2 \notin E$. Since $c_{j+1}d_{j+1} \in E$, $u_2 \neq c_{j+1}$. We claim that $u_1 \prec_\sigma d_{j+1}$. To see this, observe that if $u_1 = d_{j+1}$ then d_{j+1} and u_2 are adjacent in contradiction to our assumption, and if $d_{j+1} \prec_\sigma u_1$ then $u_0 d_{j+1} \in E$ because of $d_j \prec_\sigma c_{j+1}$ and $c_{j+1}d_{j+1} \in E$ and the properties of strong orderings. Thus, $u_1 \prec_\sigma d_{j+1}$. Furthermore, note that $cc(c_{j+1}) = cc(u_2)$. If $c_{j+1} \prec_\sigma u_2$ then $u_1 \prec_\sigma d_{j+1}$ and $\{u_1u_2, c_{j+1}d_{j+1}\} \subseteq E$ imply $d_{j+1}u_2 \in E$, a contradiction. So, $u_2 \prec_\sigma c_{j+1}$, and thus, $x \prec_\sigma u_2 \prec_\sigma c_{j+1}$. Independent of whether $u_1 \prec_\sigma c_j$ or $u_1 = c_j$ or $c_j \prec_\sigma u_1$, we have that $u_2c_j \in E$. Consequently, $(c_j, u_2, \ldots, u_s)$ shows that there is a c_j, y-path in G of length at most $s-1$. We apply the induction hypothesis and conclude with $\mathrm{d}_H(c_j, y) = \mathrm{d}_H(d_j, y) - 1$ that $\mathrm{d}_H(d_j, y) = s$. This completes the first case.

For the second case, let $x = c_j$. If $y = d_{j+1}$ then c_{j+1} is a common neighbour of x and y in H and G and thus $\mathrm{d}_H(x, y) = \mathrm{d}_G(x, y) = 2$. Let $y \neq d_{j+1}$. If $u_1 \prec_\sigma c_{j+1}$ then $u_2c_{j+1} \in E$ because of $x \prec_\sigma u_2$ and the properties of strong orderings. Then, $(c_{j+1}, u_2, u_3, \ldots, u_s)$ shows that there is a c_{j+1}, y-path in G of length at most $s-1$. Induction hypothesis and $\mathrm{d}_H(c_{j+1}, y) = \mathrm{d}_H(c_j, y) - 1$ yield $\mathrm{d}_H(c_j, y) = s$. Finally, let $c_{j+1} \prec_\sigma u_1$ or $c_{j+1} = u_1$. Then, u_1 and d_{j+1} are adjacent (remember that $c_j \prec_\sigma d_{j+1}$), and $(d_{j+1}, u_1, \ldots, u_s)$ shows that there is a d_{j+1}, y-path in G of length at most s. Thus, $\mathrm{d}_G(d_{j+1}, y) \leq s$, and by induction hypothesis and the first case, we obtain $\mathrm{d}_H(d_{j+1}, y) = \mathrm{d}_G(d_{j+1}, y) \leq s$. Since $\mathrm{d}_H(d_{j+1}, y) = \mathrm{d}_H(c_j, y)$, we conclude $\mathrm{d}_H(c_j, y) = s$. This completes the second case and the proof. ∎

Lemma 5.6 *Let $G = (V, E)$ be a thick clawpath of length r. Let $k \geq 1$ be an odd integer. If $|V| \geq \frac{1}{2}(rk + r + 2k + 6)$ then $\mathrm{D}(G) \geq k + 2$.*

Proof. We show the lemma by induction over the length of the thick clawpath. First, let G be a thick clawpath of length $r = 0$. Then, G is a complete bipartite graph. If G has at least $\frac{1}{2}(2k + 6) = k + 3$ vertices, which is an even number, we obtain $\mathrm{D}(G) \geq k + 2$ by applying Theorem 3.10. Now, let $r \geq 1$, and assume that the lemma holds for all thick clawpaths of length at most $r - 1$. We show the lemma for thick clawpaths of length r by induction over the number of vertices in set D_r. Let G be a thick clawpath of length r with sequence representation $((C_0, D_0), \ldots, (C_r, D_r))$ and let G have at least $\frac{1}{2}(rk + r + 2k + 6)$ vertices. Let $|D_r| \leq \frac{k+1}{2}$. Then, $G[V \setminus D_r]$ is a thick clawpath of length $r - 1$ on at least $\frac{1}{2}((r-1)k + (r-1) + 2k + 6)$ vertices. By induction hypothesis, $\mathrm{D}(G[V \setminus D_r]) \geq k + 2$. Since $G[V \setminus D_r]$ is a distance-preserving subgraph of G due to Lemma 5.5, we obtain $\mathrm{D}(G) \geq k + 2$ by Lemma 3.5.

Let n_b be an integer such that $n_b \geq \frac{k+1}{2}$. Assume that the lemma holds for all thick clawpaths of length r with at most n_b vertices in their D_r-set. Let G be a thick clawpath of length r on at least $\frac{1}{2}(rk + r + 2k + 6)$ vertices with sequence representation $((C_0, D_0), \ldots, (C_r, D_r))$ where $|D_r| = n_b + 1$. We determine $\mathrm{D}(G)$. Let $\mathcal{E}$ be a minimum distortion embedding for G. We say that a vertex x from D_r has the *compact property* in $\mathcal{E}$ if the close vertex to the left and right of x are both from $D_r \cup C_r$. Denote by c and c' the respectively leftmost and rightmost vertex from $C_r \cup D_{r-1}$ in $\mathcal{E}$. We distinguish two main cases.

Case A

Let all vertices from D_r between c and c' have the compact property. Let d and d' denote the respectively leftmost and rightmost vertex from D_r in $\mathcal{E}$. Since the vertices in D_r are pairwise non-adjacent, $\mathrm{d}_{\mathcal{E}}(d, d') \geq k + 1$. If a vertex from C_r is to the left of d or to the right of d' in $\mathcal{E}$, we directly obtain $\mathrm{D}(G, \mathcal{E}) = \mathrm{D}(G) \geq k + 2$. Now, let no vertex from C_r be to the left of d or to the right of d', i.e., all vertices from C_r are between d and d' in $\mathcal{E}$. If $c \prec_{\mathcal{E}} d$ or $d' \prec_{\mathcal{E}} c'$ then d or d' does not have the compact property in contradiction to our assumption. Thus, $d \prec_{\mathcal{E}} c$ and $c' \prec_{\mathcal{E}} d'$. Denote by a and a' the respectively leftmost and rightmost vertex from C_r in $\mathcal{E}$.

If $\mathrm{d}_\mathcal{E}(d,a') \geq k+2$ or $\mathrm{d}_\mathcal{E}(a,d') \geq k+2$ then $\mathrm{D}(G) \geq k+2$. Now, let $\mathrm{d}_\mathcal{E}(d,a') \leq k+1$ and $\mathrm{d}_\mathcal{E}(a,d') \leq k+1$. We determine the number of vertices in $D_r \cup C_r \cup D_{r-1}$. For a pair u, v of consecutive vertices from $D_r \cup C_r \cup D_{r-1}$ in $\mathcal{E}$, note the following:

- if one of u, v is from D_r and the other from C_r then $\mathrm{d}_\mathcal{E}(u,v) \geq 1$
- if one of u, v is from D_r and the other from D_{r-1} then $\mathrm{d}_\mathcal{E}(u,v) \geq 3$
- in all other cases, $\mathrm{d}_\mathcal{E}(u,v) \geq 2$.

From the assumptions, it follows that

$$\begin{aligned}\mathrm{d}_\mathcal{E}(d,d') &= \mathrm{d}_\mathcal{E}(d,a') + \mathrm{d}_\mathcal{E}(a,d') - \mathrm{d}_\mathcal{E}(a,a') \leq 2k+2-\mathrm{d}_\mathcal{E}(a,a')\\ \mathrm{d}_\mathcal{E}(d,d') &= \mathrm{d}_\mathcal{E}(d,c) + \mathrm{d}_\mathcal{E}(c,c') + \mathrm{d}_\mathcal{E}(c',d')\,,\end{aligned}$$

which gives

$$\frac{\mathrm{d}_\mathcal{E}(d,c) + \mathrm{d}_\mathcal{E}(c,c') + \mathrm{d}_\mathcal{E}(c',d') + \mathrm{d}_\mathcal{E}(a,a')}{2} \leq k+1\,.$$

It follows from the compact property that all vertices from D_r that are between c and c' are between a and a'. Let p be the number of vertices from D_r that are between c and c'. If $p < |C_r|$ then $\mathrm{d}_\mathcal{E}(a,a') \geq 2|C_r|-2$, if $p \geq |C_r|$ then $\mathrm{d}_\mathcal{E}(a,a') \geq 2|C_r|-2+2(p-|C_r|+1) = 2p$. If there is a vertex from D_{r-1} between a and a', both lower bounds increase by 2 since vertices from D_{r-1} are non-adjacent to vertices from D_r as well as C_r. We distinguish two cases with respect to c, c'. First, let $c \in D_{r-1}$ or $c' \in D_{r-1}$. Then, $\mathrm{d}_\mathcal{E}(d,c) + \mathrm{d}_\mathcal{E}(c',d') \geq 2(|D_r|-p)$, and $\mathrm{d}_\mathcal{E}(c,c') \geq 2|C_r \cup D_{r-1}|-2$ or $\mathrm{d}_\mathcal{E}(c,c') \geq 2|C_r \cup D_{r-1}|-2+2(p-|C_r|+1)$. Remember that $|D_r| = n_b+1$. So, for two cases, we obtain with the above inequality:

- if $p \leq |C_r|-2$ then
  $k+1 \geq n_b+1-p+|C_r|-1+|C_r \cup D_{r-1}|-1$, i.e.,
  $k+1 \geq n_b+|C_r|-p-1+|C_r \cup D_{r-1}| \geq n_b+1+|C_r \cup D_{r-1}| = |D_r \cup C_r \cup D_{r-1}|$
- if $p \geq |C_r|$ then
  $k+1 \geq n_b+1-p+p+|C_r \cup D_{r-1}|-1+p-|C_r|+1$, i.e.,
  $k+1 \geq n_b+p-|C_r|+1+|C_r \cup D_{r-1}| \geq n_b+1+|C_r \cup D_{r-1}| = |D_r \cup C_r \cup D_{r-1}|$.

The case when $p = |C_r|-1$ requires a more careful analysis. If there is a vertex between a and a' that is not from $D_r \cup C_r$ then there is also an empty slot between a and a' (because of $p = |C_r|-1$). Thus, $\mathrm{d}_\mathcal{E}(a,a') \geq 2|C_r|$, and we can conclude $|D_r \cup C_r \cup D_{r-1}| \leq k+1$ similar to the case $p \leq |C_r|-2$ above. If $c, c' \in D_{r-1}$ then $\mathrm{d}_\mathcal{E}(d,c) + \mathrm{d}_\mathcal{E}(c',d') \geq 2(n_b+1-p)+2$, and again we obtain $|D_r \cup C_r \cup D_{r-1}| \leq k+1$ similar to the cases above. Now, let there be only vertices from $D_r \cup C_r$ between a and a' and assume that $c \in C_r$. Note that this means $a = c$ and $c' \in D_{r-1}$. We determine the cardinality of $D_r \cup C_r \cup D_{r-1}$ by partitioning the set into two sets: the set of vertices from D_r to the left of a' and the other vertices. All vertices from $C_r \cup D_{r-1}$ are from a on to the right in $\mathcal{E}$. With $\mathrm{d}_\mathcal{E}(d,a') \leq k+1$, it follows that there are at most $\frac{k+1}{2}$ vertices from D_r to the left of a'. For the other set, observe that there are two slots between c' and its close vertex from D_r to the right. So, the number of vertices in the second set is at most $\lfloor \frac{k+2}{2} \rfloor = \frac{k+1}{2}$. Hence, $|D_r \cup C_r \cup D_{r-1}| \leq k+1$. The case when $c' \in C_r$ is symmetric.

Now, let $c, c' \in C_r$, i.e., $c = a$ and $c' = a'$. Then, $\mathrm{d}_\mathcal{E}(a,a') = \mathrm{d}_\mathcal{E}(c,c')$ and $\mathrm{d}_\mathcal{E}(d,c)+\mathrm{d}_\mathcal{E}(c',d') \geq 2(n_b+1-p)-2$. We analyse analogously to the cases above:

– if $p \le |C_r| - 2$ then
$k+1 \ge n_b + 1 - p - 1 + 2(|C_r \cup D_{r-1}| - 1)$, i.e.,
$k+1 \ge n_b + |C_r| - p + |C_r \cup D_{r-1}| - 1 \ge n_b + 1 + |C_r \cup D_{r-1}| = |D_r \cup C_r \cup D_{r-1}|$

– if $p = |C_r| - 1$ then $\mathrm{d}_{\mathcal{E}}(a, a') \ge 2(|C_r \cup D_{r-1}| - 1) + 1$ since a vertex from D_r between a and a' has an empty slot to its left or right. So,
$k+1 \ge n_b + 1 - p - 1 + 2(|C_r \cup D_{r-1}| - 1) + 1$, i.e.,
$k+1 \ge n_b + |C_r| - p + |C_r \cup D_{r-1}| \ge n_b + 1 + |C_r \cup D_{r-1}| = |D_r \cup C_r \cup D_{r-1}|$

– if $p \ge |C_r|$ then
$k+1 \ge n_b + 1 - p - 1 + 2(|C_r \cup D_{r-1}| - 1 + p - |C_r| + 1)$, i.e.,
$k+1 \ge n_b + p - |C_r| + |D_{r-1}| + |C_r \cup D_{r-1}| \ge n_b + 1 + |C_r \cup D_{r-1}| = |D_r \cup C_r \cup D_{r-1}|$.

We have shown $|D_r \cup C_r \cup D_{r-1}| \le k+1$. If $r \ge 2$ then $G[V \setminus (D_r \cup C_r \cup D_{r-1})]$ is a thick clawpath of length $r-2$ on at least $\frac{1}{2}((r-2)k + (r-2) + 2k + 6) = \frac{1}{2}(rk + r + 2k + 6) - (k+1)$ vertices. Applying the induction hypothesis, $\mathrm{D}(G[V \setminus (D_r \cup C_r \cup D_{r-1})]) \ge k+2$, which gives $\mathrm{D}(G) \ge k+2$ due to Lemmata 5.5 and 3.5. Let $r = 1$. Since $|D_r \cup C_r \cup D_{r-1}| \le k+1$ and $|D_r| \ge \frac{k+1}{2}$, $|D_{r-1}| \le \frac{k+1}{2}$, and $G[V \setminus D_{r-1}]$ is a complete bipartite graph on at least $k+3$ vertices. Applying Theorem 3.10 and Lemmata 5.5 and 3.5, we obtain $\mathrm{D}(G) \ge k+2$.

Case B

Let there be a vertex x from D_r between c and c' that does not have the compact property. We define a new thick clawpath from G and $\mathcal{E}$ with fewer vertices in D_r and without increasing the distortion. Let w be the close vertex to the left or to the right of x in $\mathcal{E}$ such that $w \notin D_r \cup C_r$. Let y be a vertex from D_{r-1}. Observe that $\mathrm{d}_{\mathcal{E}}(w, x) \ge \mathrm{d}_G(w, y) + 1$, since every shortest w,x-path in G contains a vertex from C_{r-1}, that is at distance 2 from x in G and at distance 1 from y. For $z \in D_r \cup C_r$, $z \ne x$, it holds that $\mathrm{d}_G(x, z) = \mathrm{d}_{\mathcal{E}}(y, z) - 1$. We obtain graph H from G by deleting x as a D_r-vertex and making it a D_{r-1}-vertex. This particularly means that the sequence representation of H is the following: $(C_0, D_0), \ldots, (C_{r-2}, D_{r-2}), (C_{r-1}, D_{r-1} \cup \{x\}), (C_r, D_r \setminus \{x\})$. Thus, H is a thick clawpath with n_b vertices in its D_r-set. Obtain embedding $\mathcal{F}$ for H from $\mathcal{E}$ by moving x by one position towards w. Due to the distance observations above, $\mathcal{F}$ is a non-contractive embedding for H. We determine $\mathrm{D}(H, \mathcal{F})$. For u, v vertices of H where $u, v \ne x$, $\mathrm{d}_{\mathcal{F}}(u, v) = \mathrm{d}_{\mathcal{E}}(u, v)$. For $u \in N_H(x) = C_{r-1}$, if u is to the left of x in $\mathcal{E}$ (and $\mathcal{F}$) then $\mathrm{d}_{\mathcal{F}}(u, x) < \mathrm{d}_{\mathcal{E}}(u, c')$, if u is to the right of x then $\mathrm{d}_{\mathcal{F}}(x, u) < \mathrm{d}_{\mathcal{E}}(c, u)$. Since c and c' are adjacent to u in G, it directly follows that $\mathrm{D}(H, \mathcal{F}) \le \mathrm{D}(G, \mathcal{E})$. Applying the induction hypothesis, $k+2 \le \mathrm{D}(H) \le \mathrm{D}(H, \mathcal{F})$, which gives $\mathrm{D}(G) \ge k+2$ by the choice of $\mathcal{E}$. This completes the proof. ■

Corollary 5.7 *Let G be a connected bipartite permutation graph, and let H be an induced subgraph of G that is a thick clawpath of length $r \ge 0$. Let $k \ge 1$ be an odd integer. If H contains at least $\frac{1}{2}(rk + r + 2k + 6)$ vertices then* $\mathrm{D}(G) \ge k+2$.

Proof. The result directly follows from Lemmata 5.6, 5.5 and 3.5. ■

5.3 Upper bound on the distortion of bipartite permutation graphs

We give an efficient algorithm for computing the distortion of bipartite permutation graphs. This algorithm works in a vertex-incremental manner, by computing the distortion for a sequence of induced subgraphs of the input graph. Correctness of our algorithm partially relies on Corollary 5.7.

The main idea of the algorithm is to take a special minimum distortion embedding for a smaller graph, to add a new vertex and to improve the embedding by moving vertices. We specify properties of the special embeddings and the moving operations in the following. Let $G = (A, B, E)$ be a bipartite permutation graph with strong ordering (σ_A, σ_B). Let a be the leftmost A-vertex in σ_A. An embedding $\mathcal{E}$ for G is called *normalised with respect to* (σ_A, σ_B) if it satisfies the following two conditions:

(D1) $\mathrm{ord}(\mathcal{E})$ is normalised with respect to (σ_A, σ_B), i.e., satisfies conditions (C1) and (C2)

(D2) for every A-vertex x, $\mathrm{d}_{\mathcal{E}}(a, x)$ is even; and for every B-vertex x, $\mathrm{d}_{\mathcal{E}}(a, x)$ is odd.

The slots of a normalised embedding can be partitioned into even slots and odd slots, the former ones will only contain A-vertices, the latter ones will only contain B-vertices. The even slots will also be called $cc(a)$*-slots* and the odd slots will also be called $\overline{cc}(a)$*-slots*. The partition into the two slot classes is not a strong restriction on an embedding for a bipartite graph, but it will simplify the description of our algorithms. It is a simple but important observation that $\mathcal{E}$ is normalised with respect to (σ_A, σ_B) if and only if the reverse of $\mathcal{E}$ is normalised with respect to $(\sigma_A, \sigma_B)^R$. We will show that every connected bipartite permutation graph has a minimum distortion embedding that is normalised with respect to a given strong ordering. Thus, a result analogous to Theorem 5.2 also holds for distortion embeddings.

Our algorithm is based solely on moving vertices. Vertex moving will appear in three different forms, depending on which vertices are moved into which direction. The corresponding three operations are called `RightMove`, `LeftMove` and `DeleteTwo`. The latter operation, `DeleteTwo`, receives an embedding $\mathcal{E}$ and a position p as input and "deletes" the slots at positions p and $p+1$ in $\mathcal{E}$, by moving all vertices that are to the right of position p by two positions to the left. Note that the result is a proper embedding if the slots at position p and $p+1$ are empty. When we apply `DeleteTwo`, these two positions are empty.

We give the definition of operation `RightMove` in pseudocode. For the definition, we introduce the following notation. For an embedding $\mathcal{E}$, a vertex u and a position p, $\mathcal{E} - u$ denotes the embedding obtained from $\mathcal{E}$ by removing u (which leaves an empty slot) and $\mathcal{E} + (u \to p)$ is the embedding obtained from $\mathcal{E}$ by placing vertex u in the slot at position p (to obtain a proper embedding, we assume that u is not placed in $\mathcal{E}$ and that the slot at position p in $\mathcal{E}$ is empty). Operation `RightMove` mainly executes a right-shift for vertices of one of the two colour classes (if the input embedding is normalised for a bipartite permutation graph). It receives an embedding $\mathcal{E}$ and a vertex u as input and is defined as

Procedure `RightMove`
begin
    let $p = \mathcal{E}(u) + 2$; **set** $\mathcal{E} = \mathcal{E} - u$;
    while position p in $\mathcal{E}$ is occupied **do**
        let x be the vertex at position p in $\mathcal{E}$;
        set $\mathcal{E} = (\mathcal{E} - x) + (u \to p)$; **set** $u = x$; **set** $p = p + 2$
    end while;
    return $\mathcal{E} + (u \to p)$
end.

An example of a `RightMove` operation is given in Figure 5. The two colour classes are depicted as white and grey circles. In the example, `RightMove` moves u by two positions to the

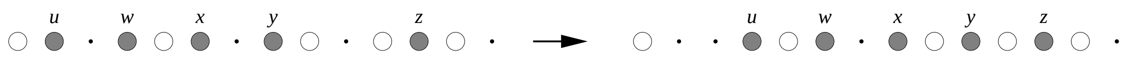

Figure 5: The `RightMove` operation illustrated, applied to vertex u. The small dots indicate empty slots. Vertices are coloured according to the colour classes they belong to.

right; this operation also moves vertices w, x, y, and the positions of the other vertices remain unchanged. The resulting embedding is shown on the right side.

Operation `LeftMove` can be considered the counterpart of `RightMove`. It receives an embedding $\mathcal{E}$ and a vertex u as input. The result is the reverse of the result obtained from applying `RightMove` to the reverse of $\mathcal{E}$ and u. The following lemma shows that the three operations are compatible with the notion of normalised embedding.

Lemma 5.8 *Let $G = (A, B, E)$ be a connected bipartite permutation graph, and let $\mathcal{E}$ be a normalised non-contractive embedding for G.*

1. *Let u be a vertex that has a neighbour to its right in $\mathcal{E}$. Let v be the rightmost neighbour of u. Let there be an empty $cc(u)$-slot between u and v in $\mathcal{E}$.*
   Then, `RightMove`$(\mathcal{E}, u)$ *is a normalised non-contractive embedding for G.*

2. *Let v be a vertex that has a neighbour to its left in $\mathcal{E}$. Let u be the leftmost neighbour of v. Let there be an empty $cc(v)$-slot between u and v in $\mathcal{E}$.*
   Then, `LeftMove`$(\mathcal{E}, v)$ *is a normalised non-contractive embedding for G.*

3. *Let u be a vertex such that all $\overline{cc}(u)$-vertices to its right in $\mathcal{E}$ are adjacent to u.*
   Then, `RightMove`$(\mathcal{E}, u)$ *is a normalised non-contractive embedding for G.*

4. *Let the slots at position p and $p + 1$ in $\mathcal{E}$ be empty.*
   Then, `DeleteTwo`$(\mathcal{E}, p)$ *is a normalised embedding for G.*

Proof. First note that the result in all cases is a proper embedding, meaning that every slot is occupied by at most one vertex. Furthermore, vertices that change position move exactly two positions, so that the distance between any pair of vertices from the same colour class is even and between any pair of vertices from different colour classes is odd. Thus, all embeddings satisfy condition (D2). The correctness of statement 4 then is immediate, since vertices are not deleted and the vertex ordering underlying the resulting embedding is equal to $\mathrm{ord}(\mathcal{E})$. For statements 1, 2, 3, the vertex ordering underlying the resulting embedding satisfies condition (C1), since vertices of the same colour class do not change order. We show that also condition (C2) is satisfied. Let $\mathcal{F} =_{\mathrm{def}}$ `RightMove`$(\mathcal{E}, u)$. Let a, b, c be three vertices of G where $a \prec_{\mathcal{F}} b \prec_{\mathcal{F}} c$, and let $ac \in E$. If $a \prec_{\mathcal{E}} b \prec_{\mathcal{E}} c$ then $ab \in E$ or $bc \in E$, since $\mathrm{ord}(\mathcal{E})$ satisfies condition (C2). Otherwise, $b \prec_{\mathcal{E}} a \prec_{\mathcal{E}} c$ or $a \prec_{\mathcal{E}} c \prec_{\mathcal{E}} b$, depending on whether $cc(b) = cc(c)$ or $cc(a) = cc(b)$. (Note that every vertex moves at most two positions for the construction of $\mathcal{F}$, which means it can change its relative order with at most one vertex.) In the former case, b is a $cc(u)$-vertex, a is a $\overline{cc}(u)$-vertex and $u \prec_{\mathcal{E}} a \prec_{\mathcal{E}} v$. Thus, $ua \in E$. And since $u = b$ or $u \prec_{\mathcal{E}} b \prec_{\mathcal{E}} a$, $ba \in E$. In the latter case, b is a $\overline{cc}(u)$-vertex, $u \prec_{\mathcal{E}} b \prec_{\mathcal{E}} v$ and $ub \in E$, and so $bc \in E$ because of $u = c$ or $u \prec_{\mathcal{E}} c \prec_{\mathcal{E}} b$. Consequently, $\mathcal{F}$ is normalised. For the non-contractiveness condition, let w be a $\overline{cc}(u)$-vertex between u and v in $\mathcal{F}$. Then, w is between u and v also in $\mathcal{E}$ and therefore $uw \in E$ by condition (C2), and thus w is adjacent to all $cc(u)$-vertices between u and w in $\mathcal{E}$. Hence, the close $cc(u)$-vertex to the left of w in $\mathcal{F}$ is a neighbour, and the close $cc(u)$-vertex x to the right of w in $\mathcal{F}$ is a neighbour or $\mathrm{d}_{\mathcal{F}}(w,x) \geq \mathrm{d}_{\mathcal{E}}(w,x)$. For vertices to the left of u or to the right of v in $\mathcal{E}$, nothing has changed in $\mathcal{F}$. Thus, $\mathcal{F}$ is non-contractive. The correctness of statement 2 immediately follows from the correctness of statement 1.

For statement 3, we distinguish cases with respect to the number of $\overline{cc}(u)$-vertices to the right of u in $\mathcal{E}$. Let $\mathcal{F} =_{\mathrm{def}}$ `RightMove`$(\mathcal{E}, u)$. If there is no $\overline{cc}(u)$-vertex to the right of u in $\mathcal{E}$, then all vertices to the right of u in $\mathcal{E}$ are $cc(u)$-vertices and $\mathrm{ord}(\mathcal{F}) = \mathrm{ord}(\mathcal{E})$, and $\mathcal{F}$ is clearly a normalised non-contractive embedding for G. Let there be exactly one $\overline{cc}(u)$-vertex to the right of u in $\mathcal{E}$, say v, and let $\mathrm{d}_{\mathcal{E}}(u,v) = 1$ and let the slot at position $\mathcal{E}(u) + 2$ in $\mathcal{E}$ be empty. Then, $\mathcal{F}$ differs from $\mathcal{E}$ only in the position of u, and $\mathcal{F}$ is non-contractive. Note that non-contractiveness here relies on the properties of condition (D2), since u cannot be placed at distance 1 to a $cc(u)$-vertex. For satisfaction of condition (C2), it suffices to observe that u and v are adjacent and consecutive in $\mathcal{E}$ and that they changed their order to obtain $\mathrm{ord}(\mathcal{F})$. Thus, $\mathcal{F}$ is normalised. If there are at least two $\overline{cc}(u)$-vertices to the right of u in $\mathcal{E}$, then $\mathcal{F}$ is the result of at most three consecutive applications of `RightMove` with the following vertices: the $cc(u)$-vertex at distance 1 to the right of the rightmost neighbour of u, then the vertex at distance 1 to the left of the rightmost neighbour of u and finally to u. The last case is captured by statement 1. ■

We will always apply the three operations to normalised non-contractive embeddings. Statement 4 of Lemma 5.8 cannot be extended by an unconditional statement about non-contractiveness. However, in all cases when we apply `DeleteTwo`, the two consecutive vertices around the deleted positions never violate the distance condition. Therefore, we assume throughout the subsection that the result of any application of the three operations is a normalised non-contractive embedding, and we will not mention this explicitly again.

To give a first outline, our algorithm for computing the distortion of bipartite permutation graphs iteratively takes a minimum distortion embedding for a connected induced subgraph, adds a new vertex to this embedding and determines on this basis the distortion of the extended graph. The new vertex is not an arbitrary vertex but one with special properties. This process defines a vertex ordering for the given graph, that we formalise in the following. Let $G = (A, B, E)$ be a connected bipartite permutation graph on at least two vertices with strong ordering (σ_A, σ_B). We say that a vertex ordering $\sigma = \langle x_1, \ldots, x_n \rangle$ for G is *competitive* if it has the following properties:

- σ satisfies condition (C1), at the beginning of Subsection 5.1
- x_1 is the leftmost A-vertex in σ_A and x_2 is the leftmost B-vertex in σ_B
- for $3 \leq i \leq n$, $N(x_i) \cap \{x_1, \ldots, x_{i-1}\} \subseteq N(w)$ where w is the $cc(x_i)$-vertex preceding x_i in σ_A or σ_B.

Observe that competitive vertex orderings exist for all connected bipartite permutation graphs and given strong orderings: if the rightmost A-vertex has a neighbour that is not a neighbour of the previous A-vertex then this neighbour has degree 1. Without loss of generality, this neighbour can be chosen as the last B-vertex. And since G is connected the last A-vertex is adjacent to the last two B-vertices, from which follows that all neighbours of the last B-vertex are neighbours of the previous B-vertex. Iteration proves the existence. The following lemma is important for the correctness of the approach of our algorithm. Note that a competitive ordering defines a strong ordering for a connected bipartite permutation graph.

Lemma 5.9 *Let $G = (A, B, E)$ be a connected bipartite permutation graph with competitive ordering $\sigma = \langle x_1, \ldots, x_n \rangle$. Then, $G[\{x_1, \ldots, x_i\}]$ is connected for every $1 \leq i \leq n$.*

Proof. Suppose the contrary. Let $i \geq 2$ be the smallest value such that $G[\{x_1, \ldots, x_i\}]$ is not connected, which means that x_i has no neighbour among $x_1, \ldots, x_{i-1}$. Note that $i \geq 3$, since x_2 is adjacent to x_1. Since G is connected, x_i has a leftmost neighbour y in σ, and $x_i \prec_\sigma y$. Let v be the $cc(y)$-vertex preceding y in σ. Since $v \prec_\sigma y$, v is not adjacent to x_i by the definition of y. Then, however, the third condition for competitive orderings is violated by y, which is a contradiction. Hence, $G[\{x_1, \ldots, x_i\}]$ is connected. ∎

We give the first step of our algorithm. We take an induced subgraph and a minimum distortion embedding and extend both by adding a new vertex, which is picked according to a competitive ordering. For a graph $G = (V, E)$, an embedding $\mathcal{E}$ for G and an integer $k \geq 0$ we say that a vertex x is *$(G, \mathcal{E}, k)$-bad* if x has a neighbour y in G where $y \prec_\mathcal{E} x$ such that $\mathrm{d}_\mathcal{E}(x, y) > k$. In particular, if x is a $(G, \mathcal{E}, k)$-bad vertex then its leftmost neighbour in $\mathcal{E}$ is at distance more than k in $\mathcal{E}$. If the context is clear we write "$(\mathcal{E}, k)$-bad vertex" or simply "k-bad vertex".

Lemma 5.10 *Let $G = (A, B, E)$ be a connected bipartite permutation graph on at least three vertices with competitive ordering σ. Let x be the rightmost vertex in σ. Let c be the $cc(x)$-vertex preceding x in σ, and let d be the leftmost neighbour of x in σ. Let $\mathcal{E}$ be a normalised minimum distortion embedding for $G{-}x$, and let $k =_{\mathrm{def}} \mathrm{D}(G{-}x, \mathcal{E})$.*

1. *Let $c \prec_\mathcal{E} d$ and $\mathcal{F} =_{\mathrm{def}} \mathcal{E} + (x \to \mathcal{E}(d) + 1)$.*
   Then, $\mathcal{F}$ is a normalised minimum distortion embedding for G of distortion k.

2. *Let $d \prec_\mathcal{E} c$ and $\mathcal{F} =_{\mathrm{def}} \mathcal{E} + (x \to \mathcal{E}(c) + 2)$.*
   Then, $\mathcal{F}$ is a normalised non-contractive embedding for G of distortion k or $k + 2$, and if there is an $(\mathcal{F}, k)$-bad vertex then it is x.

Proof. Note that in either case $\mathcal{F}$ is a normalised embedding: x occupies a $cc(x)$-slot in $\mathcal{F}$ (at odd distance to d or even distance to c) that is empty in $\mathcal{E}$; therefore, $\mathcal{F}$ satisfies condition (D2). Condition (C1) is satisfied by $\mathrm{ord}(\mathcal{F})$ since x is rightmost among all $cc(x)$-vertices in σ and $\mathcal{F}$. Now, let u, v, w be three vertices of G where $u \prec_\mathcal{F} v \prec_\mathcal{F} w$ and let $uw \in E$. If $u = x$ then $xv \in E$ since $d \prec_\mathcal{F} x$ and all $\overline{cc}(x)$-vertices to the right of d are neighbours of x. If $v = x$ then $xw \in E$ by the same argument. Let $w = x$. If v is a $\overline{cc}(x)$-vertex then $vx \in E$, since $d \prec_\mathcal{F} v$. Let v be a $cc(x)$-vertex. By definition of σ and since x is rightmost vertex in σ, $N_G(x) \subseteq N_G(c)$. If $v = c$ then $uv \in E$. If $d \prec_\mathcal{F} v \prec_\mathcal{F} c$ then $uv \in E$, since $dc \in E$ and $cc(v) = cc(c)$ and $\mathcal{E}$ satisfies condition (C2). If $u, v, w \neq x$ then $uv \in E$ or $vw \in E$, since $\mathcal{E}$ satisfies condition (C2). Therefore, $\mathcal{F}$ satisfies condition (C2), and $\mathcal{F}$ is a normalised embedding for G. For non-contractiveness, note that all vertices to the right of x in $\mathcal{F}$ are neighbours of x and the close vertex to the left of x is a neighbour at distance 1 (cases 1 and 2) or a non-neighbour, namely c, at distance 2 and c and x have a common neighbour.

It remains to consider the distortion of $\mathcal{F}$. For a neighbour y of x such that $x \prec_\mathcal{F} y$, it holds that $\mathrm{d}_\mathcal{F}(x, y) \leq \mathrm{d}_\mathcal{F}(c, y) - 2 = \mathrm{d}_\mathcal{E}(c, y) - 2 \leq k$. As the first case, let $c \prec_\mathcal{E} d$. By construction of $\mathcal{F}$, x has exactly one neighbour to the left, and this neighbour is d, at distance 1. Thus, $\mathrm{D}(G, \mathcal{F}) \leq k$. Due to Lemma 3.7, $G{-}x$ is a distance-preserving subgraph of G, so that $\mathrm{D}(G{-}x) \leq \mathrm{D}(G)$ due to Lemma 3.5. Thus, $\mathrm{D}(G) = \mathrm{D}(G, \mathcal{F})$ and $\mathcal{F}$ is a minimum distortion embedding for G. As the second case, let $d \prec_\mathcal{E} c$. Since d is a neighbour also of c, $\mathrm{d}_\mathcal{F}(d, x) =$

$d_{\mathcal{F}}(d,c)+2 = d_{\mathcal{E}}(d,c)+2 \le k+2$. Consequently, $D(G,\mathcal{F}) = k$ because of $D(G-x,\mathcal{E}) = k$ or $D(G,\mathcal{F}) = k+2$ because of $d_{\mathcal{F}}(d,x) = k+2$. Note that $D(G,\mathcal{F}) \neq k+1$ since edges join vertices on positions of different parity by condition (D2). And since $\mathcal{E}$ and $\mathcal{F}$ coincide on all vertices of $G-x$, x can be the only $(\mathcal{F},k)$-bad vertex. ∎

In the following, we want to solve the question that is raised by the second case of Lemma 5.10, namely we want to decide whether the distortion of the graph in this case is at most k or exactly $k+2$. Remember that $k+1$ is not a possible value of distortion for a bipartite graph, due to Lemma 3.3. The main subroutine in our algorithm will answer exactly this question but requires an input embedding of a special form. The next result shows that this form can be achieved by few modifications or it is easy to decide the distortion question already by looking at a small part of the given embedding. For a connected bipartite permutation graph $G = (A,B,E)$, an integer $k \ge 1$ and a normalised non-contractive embedding $\mathcal{E}$ for G, we say that $\mathcal{E}$ has a *nice beginning* if, for b_l and b_r the respectively leftmost and rightmost $(G,\mathcal{E},k)$-bad vertex in $\mathcal{E}$ and a_r the leftmost neighbour of b_r, all $(G,\mathcal{E},k)$-bad vertices are $cc(b_r)$-vertices, $d_{\mathcal{E}}(b_l,b_r) \le k-1$, there is no empty $cc(b_r)$-slot between a_r and b_r, and there is an empty $\overline{cc}(b_r)$-slot between a_r and b_l in $\mathcal{E}$. Note that $a_r \prec_{\mathcal{E}} b_l$ by the distance conditions.

Lemma 5.11 *Let $G = (A,B,E)$ be a connected bipartite permutation graph on at least three vertices with competitive ordering σ. Let $\mathcal{E}$ be a normalised non-contractive embedding for G of distortion $k+2$, and let there be exactly one $(G,\mathcal{E},k)$-bad vertex x. Let x be the rightmost $cc(x)$-vertex in σ. Then, one of the following cases is true:*

1. $D(G) \le k$, *which is certified by a normalised non-contractive embedding for G*
2. $D(G) = k+2$, *which is certified by a normalised non-contractive embedding for G of distortion $k+2$ and an induced subgraph that is complete bipartite on $k+3$ vertices*
3. $D(G) \le k+2$, *which is certified by a normalised non-contractive embedding for G of distortion $k+2$ and with a nice beginning.*

There is an $\mathcal{O}(n)$-time algorithm that identifies a true case and outputs the certificates.

Proof. Let y be the rightmost $\overline{cc}(x)$-vertex in $\mathcal{E}$. If $x \prec_{\mathcal{E}} y$ and there is an empty $\overline{cc}(x)$-slot between x and y then `LeftMove`$(\mathcal{E},y)$ is a normalised non-contractive embedding for G that satisfies the assumptions of the lemma. Repeated application deletes all empty $\overline{cc}(x)$-slots between x and y. So, we can assume in the following that there are no empty $\overline{cc}(x)$-slots between x and the rightmost vertex in $\mathcal{E}$. Let d be the leftmost neighbour of x in $\mathcal{E}$, and let $\mathcal{F} =_{\rm def}$ `RightMove`$(\mathcal{E},d)$. If there is no $(\mathcal{F},k)$-bad vertex then $D(G) \le k$, which is certified by normalised non-contractive embedding $\mathcal{F}$. Now, suppose there is an $(\mathcal{F},k)$-bad vertex. Note that, by the definition of $\mathcal{F}$, x is not $(\mathcal{F},k)$-bad and no other $cc(x)$-vertex is $(\mathcal{F},k)$-bad. Let w be the rightmost $(\mathcal{F},k)$-bad vertex in $\mathcal{F}$. Since w must be a moved vertex, w is between d and y.

Case A

Let $x \prec_{\mathcal{F}} w$. Since w is a moved vertex, there is no empty $\overline{cc}(x)$-slot between d and w in $\mathcal{E}$, and thus there is no empty $\overline{cc}(x)$-slot between d and y in $\mathcal{E}$. In particular, all $\overline{cc}(x)$-vertices to the right of d moved for the definition of $\mathcal{F}$. Let c be the leftmost neighbour of w in $\mathcal{F}$. First, let there be an empty $cc(x)$-slot between c and x in $\mathcal{F}$ and $\mathcal{E}$. Note that, by the choice of w and the definition of c, no $cc(x)$-vertex to the right of c has a right neighbour at distance more than

$k-2$ in $\mathcal{E}$. Let $\mathcal{E}' =_{\text{def}}$ `LeftMove`$(\mathcal{E}, x)$. Since $d \prec_\mathcal{E} c \prec_\mathcal{E} x$, $\mathcal{E}'$ is normalised and non-contractive, and $\mathrm{D}(G, \mathcal{E}') = k$. Hence, $\mathrm{D}(G) \leq k$.

For the other case, let there be no empty $cc(x)$-slot between c and x. Denote by C the $cc(x)$-vertices between c and x and denote by D the $\overline{cc}(x)$-vertices between d and w. By the properties of strong orderings, all vertices in C are adjacent to all vertices in D, which means that $G[C \cup D]$ is a complete bipartite graph. We determine the number of vertices in $C \cup D$ based on $\mathcal{E}$. Remember that $\mathrm{d}_\mathcal{E}(d, x) = k+2$ and $\mathrm{d}_\mathcal{E}(c, w) = k$. If $w \prec_\mathcal{E} x$ then $C \cup D$ is the set of vertices between d and x in $\mathcal{E}$, hence, $|C \cup D| = k+3$. Now, let $x \prec_\mathcal{E} w$. From D there are $\frac{k+3}{2}$ vertices between d and x, $\frac{k+1}{2}$ vertices between c and w and $\frac{1}{2}\mathrm{d}_\mathcal{E}(c, x)$ vertices between c and x (that have been counted twice), and there are $\frac{1}{2}\mathrm{d}_\mathcal{E}(c, x) + 1$ vertices in C. We sum up and obtain:

$$\frac{k+3+k+1+\mathrm{d}_\mathcal{E}(c,x)+2-\mathrm{d}_\mathcal{E}(c,x)}{2} = \frac{2k+6}{2} = k+3$$

vertices in $C \cup D$. Applying Theorem 3.10, $G[C \cup D]$ has distortion $k+2$. And since $G[C \cup D]$ is a distance-preserving subgraph of G due to Lemma 5.5, G has distortion at least $k+2$ according to Lemma 3.5. Since $\mathrm{D}(G) \leq \mathrm{D}(G, \mathcal{E})$, we conclude $\mathrm{D}(G) = k+2$.

Case B

Let $w \prec_\mathcal{F} x$. All $(\mathcal{F}, k)$-bad vertices are between d and x, at distance at most $k-1$ to d in $\mathcal{F}$. If the slot at position $\mathcal{F}(d) - 1$ in $\mathcal{F}$ is not occupied, the two slots at position $\mathcal{F}(d) - 2$ and $\mathcal{F}(d) - 1$ in $\mathcal{F}$ are not occupied. (Remember that d occupies the slot at position $\mathcal{F}(d) - 2$ in $\mathcal{E}$.) We obtain a normalised non-contractive embedding $\mathcal{F}'$ for G as `DeleteTwo`$(\mathcal{F}, \mathcal{F}(d) - 2)$. Since all leftmost neighbours of $(\mathcal{F}, k)$-bad vertices are to the left of d in $\mathcal{F}$, $\mathrm{D}(G, \mathcal{F}') = k$, and thus $\mathrm{D}(G) \leq k$. Now, let the slot at position $\mathcal{F}(d) - 1$ in $\mathcal{F}$ be occupied, say by vertex a.

Let there be no empty $cc(x)$-slot between a and x in $\mathcal{E}$. If there is an empty $\overline{cc}(x)$-slot between d and x in $\mathcal{E}$ then $\mathcal{E}$ is an embedding with a nice beginning. Otherwise, if there is no empty $\overline{cc}(x)$-slot between d and x, let vertex z occupy position $\mathcal{E}(x) - 1$ in $\mathcal{E}$. Note that $z \neq d$. According to the properties of $\mathcal{F}$, `RightMove`$(\mathcal{E}, z)$ is a normalised non-contractive embedding for G of distortion $k+2$ with a nice beginning.

Let there be an empty $cc(x)$-slot between a and x in $\mathcal{E}$. Let v be the leftmost $cc(x)$-vertex such that there is no empty $cc(x)$-slot between v and x in $\mathcal{E}$. Let $\mathcal{G} =_{\text{def}}$ `LeftMove`$(\mathcal{E}, x)$. If there is no $(\mathcal{G}, k)$-bad vertex then $\mathcal{G}$ is a normalised non-contractive embedding certifying $\mathrm{D}(G) \leq k$. So, let there be a $(\mathcal{G}, k)$-bad vertex. Let u be the leftmost $(\mathcal{G}, k)$-bad vertex in $\mathcal{G}$. Since x is not $(\mathcal{G}, k)$-bad, all $(\mathcal{G}, k)$-bad vertices are $\overline{cc}(x)$-vertices and $x \prec_\mathcal{G} u$ and $x \prec_\mathcal{E} u$ (the second relationship follows from the fact that $\mathrm{d}_\mathcal{E}(a, x) = k+1$ and $\mathrm{d}_\mathcal{E}(v, x) \leq k-3$) and $\mathrm{d}_\mathcal{G}(x, u) \geq 5$. If there is an empty $\overline{cc}(x)$-slot between v and u in $\mathcal{G}$ then `LeftMove`$(\mathcal{G}, y)$ is a normalised non-contractive embedding of distortion k for G. Remember that there is no empty $\overline{cc}(x)$-slot between u and y in $\mathcal{E}$ by the discussion at the beginning of the proof. If there is no empty $\overline{cc}(x)$-slot between v and u in $\mathcal{G}$ then $\mathcal{G}$ is a normalised non-contractive embedding with a nice beginning, particularly since there is an empty $cc(x)$-slot between x and u in $\mathcal{G}$. ■

After just two more definitions we will be ready for presenting the central subroutine of our algorithm. Let $G = (A, B, E)$ be a bipartite permutation graph and let $\mathcal{E}$ be a normalised embedding for G. We call a pair (v, w) of vertices for v a $\overline{cc}(w)$-vertex a *blocking pair* if $v \prec_\mathcal{E} w$, $\mathrm{d}_\mathcal{E}(v, w) = 3$ and $vw \notin E$. Let d and x be vertices of G from different colour classes where $d \prec_\mathcal{E} x$. We call a $\overline{cc}(x)$-vertex w for $d \prec_\mathcal{E} w \prec_\mathcal{E} x$ a *breakpoint vertex* between d and x if (v, w) is a blocking pair for some vertex v, there is no empty $cc(x)$-slot between d and v, and no empty $\overline{cc}(x)$-slot between w and x in $\mathcal{E}$. The algorithm of the main subroutine is then the following:

Algorithm `RepairAndDecide`
Input An embedding $\mathcal{E}$ and an integer k
Output Acceptance if $\mathcal{E}$ can be repaired into an embedding $\mathcal{F}$ of distortion at most k; rejection otherwise.

```
begin
    while there is an (E, k)-bad vertex do
        let x be the rightmost k-bad vertex in E;
        let d be the leftmost neighbour of x in E;
        if there is no empty cc̄(x)-slot between d and x in E then reject end if;
        let F = RightMove(E, d);
        if slot at position F(d) − 1 is not occupied in F then accept end if;
        if there is no breakpoint vertex between d and x in F and
            there is an empty cc(x)-slot between d and x in F then accept end if;
        set E = F
    end while;
    accept
end.
```

The input of the above algorithm is a normalised non-contractive embedding of distortion $k+2$ with a nice beginning. With the results of Lemma 5.8 it is clear that all embeddings during the execution of `RepairAndDecide` are normalised non-contractive. If the execution of the **while** loop stops since there is no k-bad vertex in $\mathcal{E}$, $\mathcal{E}$ has distortion at most k, and the algorithm accepts correctly. In the following, we show that the algorithm always stops with the correct answer, which means that it accepts if the distortion of the input graph is at most k and it rejects if the distortion of the input graph is at least $k+2$. This correctness proof is partitioned into two lemmata. We begin with properties about the intermediate embeddings. An iteration of the **while** loop is called a *round* of the algorithm.

Lemma 5.12 *Let $G=(A,B,E)$ be a connected bipartite permutation graph with normalised non-contractive embedding $\mathcal{G}$ of distortion $k+2$ with a nice beginning. Apply* `RepairAndDecide` *to $(\mathcal{G},k)$. Let $\mathcal{E}$, $\mathcal{F}$, c and x have the values according to* `RepairAndDecide` *at the end of a round, where we assume that there is an empty $\overline{cc}(x)$-slot between d and x in $\mathcal{E}$. Denote by x_l and x_r the respectively leftmost and rightmost $(\mathcal{F},k)$-bad vertex.*

(W1) $\mathrm{D}(G,\mathcal{F}) \le k+2$

(W2) *$d \prec_{\mathcal{F}} x_l$ or $d = x_l$, and $x_r \prec_{\mathcal{F}} x$*

(W3) *the slot at position $\mathcal{F}(d)-2$ in $\mathcal{F}$ is empty*

(W4) *all $(\mathcal{F},k)$-bad $cc(x_r)$-vertices are to the right of all $(\mathcal{F},k)$-bad $\overline{cc}(x_r)$-vertices*

(W5) *if there is an empty $cc(x)$-slot between d and x in $\mathcal{F}$ then there is an empty $cc(x)$-slot between d and the leftmost $(\mathcal{F},k)$-bad $cc(x)$-vertex in $\mathcal{F}$.*

Proof. We prove satisfaction of the conditions by induction over the number of rounds. If the current round is the first round, $\mathcal{E}$ is an embedding with a nice beginning. If the current round is not the first round, we assume that $\mathcal{E}$ satisfies the conditions. Let u be the rightmost $\overline{cc}(x)$-vertex such that there is no empty $\overline{cc}(x)$-slot between d and u in $\mathcal{E}$; note that $u \prec_{\mathcal{E}} x$ by the empty slot assumption of the lemma. Then, the $\overline{cc}(x)$-vertices between d and u are exactly the vertices that have different positions in $\mathcal{E}$ and $\mathcal{F}$. It follows that all $(\mathcal{F},k)$-bad $cc(x)$-vertices are

$(\mathcal{E},k)$-bad, since they are not moved and their leftmost neighbours are not moved (the leftmost neighbours are to the left of d). An $(\mathcal{F},k)$-bad $\overline{cc}(x)$-vertex is $(\mathcal{E},k)$-bad or is between d and u.

Claim u is at distance at least 3 to the left of the leftmost $(\mathcal{E},k)$-bad vertex in $\mathcal{E}$.

Proof. For $\mathcal{E}$ in the first round, this is clear from the fact that there is an empty $\overline{cc}(x)$-slot between d and the leftmost k-bad vertex by definition of nice beginning. Let the current round not be the first round. Then, $\mathcal{E}$ is the result of a `RightMove` operation, applied to some vertex d'. By assumption, $\mathcal{E}$ satisfies condition (W3), so that the slot at position $\mathcal{E}(d')-2$ in $\mathcal{E}$ is empty. If d' is a $\overline{cc}(x)$-vertex then u is clearly to the left of d' at distance at least 4 and no k-bad vertex is to the left of d' in $\mathcal{E}$ by condition (W2). For the other case, let d' be a $cc(x)$-vertex. We show that there is no empty $\overline{cc}(x)$-slot between d' and x. Let $\mathcal{E}'$ be the input embedding to the previous round, and let x' be the rightmost $(\mathcal{E}',k)$-bad vertex. Note that x' is a $\overline{cc}(x)$-vertex. Since $d'x' \in E$ and $dx \in E$, all $\overline{cc}(x)$-vertices between d and x' are adjacent to all $cc(x)$-vertices between d' and x. Consequently, there is no vertex w between d' and x in $\mathcal{E}$ such that (v,w) for some vertex v is a blocking pair. Here, it is important to note that $\mathrm{d}_{\mathcal{E}}(d',x) \leq k-1$, so that $d \prec_{\mathcal{E}} w$ and $\mathrm{d}_{\mathcal{E}}(d,w) \geq 3$ for all vertices w between d' and x in $\mathcal{E}$. Therefore, there is no breakpoint vertex between d' and x in $\mathcal{E}$. If there is an empty $cc(x')$-slot between d' and x' then the algorithm would have accepted in the previous round. Therefore, there can be no empty $\overline{cc}(x)$-slot between d' and x in $\mathcal{E}$. Thus, all empty $\overline{cc}(x)$-slots between d and x are to the left of d', and since there exists an empty $\overline{cc}(x)$-slot due to assumption of the lemma, u is at distance at least 3 to d' in $\mathcal{E}$. □

(W1)
No $\overline{cc}(x)$-vertex between d and u is $(\mathcal{E},k)$-bad. Therefore, moved vertices have left neighbours at distance at most k, and so no $(\mathcal{F},k)$-bad vertex has a neighbour at distance more than $k+2$ in $\mathcal{F}$. This means $\mathrm{D}(G,\mathcal{F}) \leq k+2$.

(W2)
Since no vertex to the right of x is moved for defining $\mathcal{F}$ according to the claim or is $(\mathcal{E},k)$-bad, and since d is the leftmost neighbour of x, all left neighbours of x in $\mathcal{F}$ are at distance at most k. Thus, $x_r \prec_{\mathcal{F}} x$. For x_l, it follows from the claim that no $(\mathcal{F},k)$-bad vertex is to the left of d in $\mathcal{F}$.

(W3)
This is immediately clear from the fact that d is the leftmost moved vertex.

(W4)
Vertices that are $(\mathcal{F},k)$-bad but not $(\mathcal{E},k)$-bad are $\overline{cc}(x)$-vertices between d and u. Since $\mathcal{E}$ satisfies condition (W4), which is clear for $\mathcal{G}$ by the definition of nice beginning, no $(\mathcal{F},k)$-bad $\overline{cc}(x)$-vertex is to the right of an $(\mathcal{F},k)$-bad $cc(x)$-vertex in $\mathcal{F}$.

(W5)
For this condition, we partition the sequence of rounds into intervals. A new interval always starts when x changes colour class with respect to the previous round, and the first interval starts with the first round. Note that during the rounds of a single interval, new bad vertices are from the same colour class. So, it suffices to consider only first rounds of intervals. Consider the first round, which is the first round of the first interval. By definition of nice beginning, there is no empty $cc(x)$-slot between d and x in $\mathcal{G}$, thus there is no empty $cc(x)$-slot between d and x in $\mathcal{F}$. Now, consider the beginning of an arbitrary but later interval. Let $\mathcal{E}$ be the input embedding of the first round of the interval, and denote by b_l and b_r the respectively leftmost

and rightmost $(\mathcal{E},k)$-bad vertex. Let $\mathcal{E}'$ be the input embedding of the previous round, which is the last round of the previous interval. Denote by x' the rightmost $(\mathcal{E}',k)$-bad vertex in $\mathcal{E}'$ and denote by d' its leftmost neighbour. Then, $d' = b_l$ or $d' \prec_{\mathcal{E}} b_l$ according to condition (W2). And the slot at position $\mathcal{E}(d') - 2$ is empty in $\mathcal{E}$. And since the leftmost neighbour of b_r, denoted as d_r, is at distance $k+2$ to the left of b_r in $\mathcal{E}$, which means at distance at least 3 to the left of d' in $\mathcal{E}$, there is an empty $cc(b_r)$-slot between d_r and b_l in $\texttt{RightMove}(\mathcal{E}, d_r)$. This completes the proof. ∎

Let $G = (A,B,E)$ be a bipartite permutation graph and let $\mathcal{E}$ be an embedding for G. Let b, x be two vertices of G of the same colour class where $b \prec_{\mathcal{E}} x$. Let H be an induced subgraph of G that is a thick clawpath. We say that H *has a proper connection on* (b,x) if H and (b,x) satisfy the following conditions in $\mathcal{E}$:

(P1) $x \in V(H)$, the slot at position $\mathcal{E}(x) - 1$ is occupied, say by vertex c, and $bc \in E$

(P2) H contains no $cc(x)$-vertex to the left of b and no $\overline{cc}(x)$-vertex to the left of c

(P3) no $\overline{cc}(x)$-vertex to the left of x has a neighbour in H to the right of x

(P4) the $cc(x)$-vertices between b and x in H correspond to a last path vertex of the clawpath underlying H.

We use such thick clawpaths to extend them on their proper connections.

Lemma 5.13 *Let $G = (A,B,E)$ be a connected bipartite permutation graph with normalised non-contractive embedding $\mathcal{E}$ of distortion $k+2$ with a nice beginning. Apply* $\texttt{RepairAndDecide}$ *to $(\mathcal{E},k)$. If the algorithm accepts then $\mathrm{D}(G) \leq k$, if the algorithm rejects then G contains a thick clawpath of length r on $\frac{1}{2}(rk + r + 2k + 6)$ vertices as induced subgraph.*

Proof. We show the lemma by induction over the number of rounds of $\texttt{RepairAndDecide}$. We begin with the first round; note that there is a first round. Let x and d be the vertices chosen according to the algorithm. By definition of nice beginning, there is an empty $\overline{cc}(x)$-slot between d and x in $\mathcal{E}$ and the slot at position $\mathcal{E}(d) + 1$ is occupied, say by vertex u. Let $\mathcal{F} =_{\mathrm{def}} \texttt{RightMove}(\mathcal{E}, d)$. Let c be any $\overline{cc}(x)$-vertex between d and x in $\mathcal{F}$ such that there is no empty $\overline{cc}(x)$-slot between d and c in $\mathcal{F}$. Then, G contains a thick clawpath of length 0 on $\frac{k+5}{2}$ vertices as induced subgraph with a proper connection on (d,c), as we show in the following. According to the properties of nice beginning, there is no empty $cc(x)$-slot between u and x in $\mathcal{F}$. Let b be the vertex occupying the slot at position $\mathcal{F}(c) - 1$ in $\mathcal{F}$. Note that b exists, since b is a $cc(x)$-vertex between u and x in $\mathcal{E}$. Then, $bc \in E$ due to non-contractiveness of $\mathcal{F}$. Let $H_{d,c}$ be the subgraph of G induced by the $\overline{cc}(x)$-vertices between d and c and the $cc(x)$-vertices between b and x in $\mathcal{F}$. To show satisfaction of the conditions (P1–4), it remains to show satisfaction of condition (P4); the other conditions are clearly satisfied by the definition of $H_{d,c}$. Since $dx \in E$ and $bc \in E$, all $cc(x)$-vertices in $H_{d,c}$ are adjacent to all $\overline{cc}(x)$-vertices in $H_{d,c}$ by the properties of normalised embeddings and strong orderings. Thus, $H_{d,c}$ is a complete bipartite graph, i.e., a thick clawpath of length 0, and the $cc(c)$-vertices correspond to a last path vertex of the underlying clawpath. For the number of vertices in $H_{d,c}$, note that $\mathrm{d}_{\mathcal{F}}(d,x) = k$ and there are $\frac{1}{2}\mathrm{d}_{\mathcal{F}}(d,c) + 1$ many $cc(c)$-vertices and $\frac{1}{2}\mathrm{d}_{\mathcal{F}}(b,x) + 1$ many $\overline{cc}(c)$-vertices in $H_{d,c}$. This sums up to $\frac{k+5}{2}$ vertices, since $\mathrm{d}_{\mathcal{F}}(d,c) + \mathrm{d}_{\mathcal{F}}(b,x) = k+1$.

We now consider an arbitrary but later round. Let $\mathcal{E}$, x and d be defined according to the algorithm. We assume that there is a $cc(x)$-vertex b, where $b = x$ or $b \prec_{\mathcal{E}} x$, such that there is no empty $cc(x)$-slot between b and x in $\mathcal{E}$ and (b,x) is a proper connection for a thick clawpath $H_{b,x}$ of length r on $|V(H_{b,x})| = \frac{1}{2}(rk+r+k+5)$ vertices. We consider cases according to `RepairAndDecide`.

No empty slot

Suppose there is no empty $\overline{cc}(x)$-slot between d and x in $\mathcal{E}$. Let c be the vertex occupying position $\mathcal{E}(x)-1$ in $\mathcal{E}$. Since $dx \in E$ and $bc \in E$ (according to condition (P1)), all $\overline{cc}(x)$-vertices between d and x are adjacent to all $cc(x)$-vertices between b and x. Because of conditions (P3–4), subgraph H of G induced by the $\overline{cc}(x)$-vertices between d and x and $V(H_{b,x})$ is a thick clawpath of length r. We determine the number of vertices of H. There are $\frac{k+3}{2}$ $\overline{cc}(x)$-vertices between d and x in $\mathcal{E}$ and at least all $\frac{k+1}{2}$ $\overline{cc}(x)$-vertices to the left of c are not contained in $H_{b,x}$ due to condition (P2). Hence, $|V(H)| \geq \frac{1}{2}(rk+r+2k+6)$.

For the other cases, let there be an empty $\overline{cc}(x)$-slot between d and x in $\mathcal{E}$. Let $\mathcal{F} =_{\text{def}}$ `RightMove`$(\mathcal{E}, d)$.

Position $\mathcal{F}(d)-1$ not occupied

Let the slot at position $\mathcal{F}(d)-1$ not be occupied in $\mathcal{F}$. Then, the slots at position $\mathcal{F}(d)-2$ and $\mathcal{F}(d)-1$ are not occupied in $\mathcal{F}$. Let $\mathcal{G} =_{\text{def}}$ `DeleteTwo`$(\mathcal{F}, \mathcal{F}(d)-2)$. Due to Lemma 5.12, all $(\mathcal{F},k)$-bad vertices are between d and x in $\mathcal{F}$. And since $\mathrm{d}_{\mathcal{F}}(d,x) = k$, the leftmost neighbour of every $(\mathcal{F},k)$-bad vertex is to the left of d in $\mathcal{F}$. If there is no vertex to the left of d in $\mathcal{F}$, then there are no $(\mathcal{F},k)$-bad vertices, and $\mathrm{D}(G,\mathcal{F}) = \mathrm{D}(G,\mathcal{G}) = k$. Otherwise, let w be the close vertex to the left of d in $\mathcal{F}$. Then, w is the close vertex to the left of d also in $\mathcal{E}$, and $\mathrm{d}_{\mathcal{G}}(w,d) = \mathrm{d}_{\mathcal{F}}(w,d) - 2 = \mathrm{d}_{\mathcal{E}}(w,d)$. Thus, $\mathcal{G}$ is a normalised non-contractive embedding for G. And since $\mathrm{D}(G,\mathcal{F}) \leq k+2$, it follows that $\mathrm{D}(G,\mathcal{G}) = k$; equality is shown by $\mathrm{d}_{\mathcal{G}}(d,x) = k$.

Position $\mathcal{F}(d)-1$ occupied

Let u be the vertex occupying position $\mathcal{F}(d)-1$ in $\mathcal{F}$. As the first case, let there be no empty $cc(x)$-slot between d and x. Let c be a $\overline{cc}(x)$-vertex between d and x such that there is no empty $\overline{cc}(x)$-slot between d and c in $\mathcal{F}$, and let b be the vertex occupying the slot at position $\mathcal{F}(c)-1$ in $\mathcal{F}$. Analogous to the beginning of the proof, the $\overline{cc}(x)$-vertices between d and c and the $cc(x)$-vertices between c and x define a thick clawpath of length 0 on $\frac{k+5}{2}$ vertices with a proper connection on (d,c). As the second case, let there be an empty $cc(x)$-slot between d and x in $\mathcal{F}$; let p be the position of the leftmost empty $cc(x)$-slot between d and x. Let there be no breakpoint vertex between d and x in $\mathcal{F}$. We want to move u two positions to the right to obtain an embedding without vertex occupying the slot at position $\mathcal{F}(d)-1$. Suppose there is a blocking pair (v,w) such that v is a $cc(x)$-vertex and w is a $\overline{cc}(x)$-vertex and $d \prec_{\mathcal{F}} w$ and $\mathcal{F}(v) < p$. When we move u then v has to move and would come too close to the non-neighbour w. Note that $vw \notin E$ implies that no vertex to the right of w is adjacent to v. In particular, no $\overline{cc}(x)$-vertex between w and x has a left neighbour at distance more than k. Remember that $\mathrm{d}_{\mathcal{F}}(u,x) = k+1$. Since w is no breakpoint vertex, there is an empty $\overline{cc}(x)$-slot between w and x. Since $wx \in E$ due to condition (C2), `RightMove`$(\mathcal{F}, w)$ is a normalised non-contractive embedding without (v,w) being a blocking pair. If there are further blocking pairs with vertices to the left of position p, repeat the described procedure. If there are no (further) blocking pairs, $\mathcal{F}' =_{\text{def}}$ `RightMove`$(\mathcal{F}, u)$ is a normalised non-contractive embedding of distortion at most $k+2$ with $(\mathcal{F}',k)$-bad vertices only between d and x. We obtain a normalised non-contractive embedding of distortion at most k by deleting the two empty slots to the left of d, similar to the case above.

Finally, let there be a breakpoint vertex w between d and x; let v be the vertex such that (v, w) is a blocking pair. By definition, there is no empty $cc(x)$-slot between d and v and there is no empty $\overline{cc}(x)$-slot between w and x. Note that $v \prec_{\mathcal{F}} b$ by condition (W5) of Lemma 5.12. Let a be an $(\mathcal{F}, k)$-bad $\overline{cc}(x)$-vertex that is not $(\mathcal{E}, k)$-bad. This particularly means that there are no empty $\overline{cc}(x)$-slots between d and a in $\mathcal{F}$. Observe that $a \prec_{\mathcal{F}} w$. Let c be the vertex occupying the slot at position $\mathcal{F}(a) - 1$. Now, let $H_{d,a}$ be the subgraph of G induced by $V(H_{b,x})$ and the $\overline{cc}(x)$-vertices between d and a and between w and x and the $cc(x)$-vertices between c and v. Similar to the beginning of the proof, all $\overline{cc}(x)$-vertices between d and x are adjacent to all $cc(x)$-vertices between b and x. And since $dx \in E$ and $au \in E$, all $cc(x)$-vertices between u and v are adjacent to all $\overline{cc}(x)$-vertices between d and a. And no $cc(x)$-vertex between u and v is adjacent to a vertex from w on to the right in $\mathcal{F}$. Thus, $H_{d,a}$ is a thick clawpath with a proper connection on (d, a) of length $r + 1$. It remains to determine the number of vertices in $V(H_{d,a}) \setminus V(H_{b,x})$:

- $\frac{1}{2}\mathrm{d}_{\mathcal{F}}(c, v) + 1$ $cc(x)$-vertices between c and v
- $\frac{1}{2}\mathrm{d}_{\mathcal{F}}(d, a) + 1$ $\overline{cc}(x)$-vertices between d and a
- $\frac{1}{2}(\mathrm{d}_{\mathcal{F}}(w, x) - 1)$ $\overline{cc}(x)$-vertices between w and x, where the vertex occupying the slot at position $\mathcal{F}(x) - 1$ in $\mathcal{F}$ is not counted,

which sums up to $\frac{1}{2}(\mathrm{d}_{\mathcal{F}}(c, v) + \mathrm{d}_{\mathcal{F}}(d, a) + \mathrm{d}_{\mathcal{F}}(w, x) - 1) + 2$ new vertices. With the definition of the selected vertices, it holds that $\mathrm{d}_{\mathcal{F}}(c, a) = 1$ and $\mathrm{d}_{\mathcal{F}}(v, w) = 3$, so that

$$\mathrm{d}_{\mathcal{F}}(c, v) + \mathrm{d}_{\mathcal{F}}(d, a) + \mathrm{d}_{\mathcal{F}}(w, x) = k + 1 - 3 = k - 2\,.$$

Thus, $H_{d,a}$ contains $|V(H_{b,x})| + \frac{k+1}{2} \geq \frac{1}{2}((r + 1)k + (r + 1) + k + 6)$ vertices.

We have seen that in case `RepairAndDecide` stops during a round then the decision is correct with respect to our definitions; and if it does not stop then every k-bad vertex is associated with a thick clawpath of special properties. This completes the proof. ∎

So far, there is a third possible case for `RepairAndDecide` that is not covered by Lemma 5.13, namely that the algorithm might not terminate on an input. However, we have actually already proven that this cannot happen, as condition (W2) in Lemma 5.12: in every round of the algorithm, the number of vertices to the right of k-bad vertices increases. Now, we are ready for presenting the two main results of this section.

Theorem 5.14 *Let $G = (A, B, E)$ be a connected bipartite permutation graph, and let $k \geq 1$ be an odd integer. Then, $\mathrm{D}(G) \leq k$ or G contains a thick clawpath of length r on $\frac{1}{2}(rk + r + 2k + 6)$ vertices as induced subgraph.*

Proof. We show the statement by induction over the number of vertices of G. If G contains at most two vertices then $\mathrm{D}(G) \leq 1$. So, let G have $n \geq 3$ vertices. Assume that the claim holds for all graphs on at most $n - 1$ vertices. Let σ be a competitive ordering for G, and let x be the last vertex in σ. If $\mathrm{D}(G - x) \geq k + 2$ then $G - x$ contains a thick clawpath of length r on $\frac{1}{2}(rk + r + 2k + 6)$ vertices as induced subgraph, and thus G. Now, let $\mathrm{D}(G - x) \leq k$, and let $\mathcal{F}$ be the embedding obtained as in Lemma 5.10 on input $\mathcal{E}$, σ and x. Assume that $\mathrm{D}(\mathcal{F}) = k + 2$. Then, Lemma 5.11 can be applied to $\mathcal{F}$, and in connection with Lemma 5.13, we obtain the claim. ∎

Corollary 5.15 *A connected bipartite permutation graph G has distortion at most k for $k \geq 1$ an odd integer if and only if G does not contain a thick clawpath of length r on $\frac{1}{2}(rk+r+2k+6)$ vertices as induced subgraph.*

Proof. The statement directly follows from Theorem 5.14 and Corollary 5.7. ∎

Note that Corollary 5.15 also gives a lower bound on the number of vertices of graphs of high distortion.

With the result of Theorem 5.14, we can conclude that Lemmata 5.10, 5.11 and 5.13 readily give an algorithm for computing the distortion of a connected bipartite permutation graph directly. We summarize this as our main algorithm `bpg-distortion`, which we call `bpg-distortion`:

Algorithm `bpg-distortion`
Input connected bipartite permutation graph G
Output $k = D(G)$, a corresponding embedding $\mathcal{E}$ with $D(G, \mathcal{E}) = k$ and
an induced thick clawpath subgraph H such that $D(H) > k - 2$

begin
  if G consists of a single vertex x **then** **return** $k = 0$, $\mathcal{E} = \langle x \rangle$, $H = G$ **end if**;
  Compute a competitive ordering $\langle x_1, x_2, \ldots, x_n \rangle$ for G;
  let $\mathcal{E} = \langle x_1, x_2 \rangle$; **let** $k = 1$; **let** $H = G[\{x_1, x_2\}]$;
  for $i = 3$ **to** n **do**
    let $G_i = G[\{x_1, x_2, \ldots x_i\}]$;
    Apply Lemma 5.10 on G_i and $\mathcal{E}$ to obtain embedding $\mathcal{F}$; **set** $\mathcal{E} = \mathcal{F}$;
    if $D(G_i, \mathcal{E}) > k$ **then**
      Apply Lemma 5.11 on G_i and $\mathcal{E}$ to obtain a true case and a certificate
      embedding $\mathcal{F}$ corresponding to this case; **set** $\mathcal{E} = \mathcal{F}$;
      if Case 2 of Lemma 5.11 **then**
        set $k = k + 2$;
        set H = induced complete bipartite subgraph returned by this case
      end if;
      if Case 3 of Lemma 5.11 **then**
        Apply Algorithm `RepairAndDecide` to $(\mathcal{E}, k)$;
        if `RepairAndDecide` rejects **then**
          set $k = k + 2$;
          Apply the algorithm in the proof of Lemma 5.13 to compute an induced
          thick clawpath subgraph H
        else
          set $\mathcal{E}$ = output embedding $\mathcal{F}$ of Algorithm `RepairAndDecide`
        end if
      end if
    end if
  end for;
  return k, $\mathcal{E}$, H
end.

Theorem 5.16 *There is an $\mathcal{O}(n^2)$-time algorithm that computes the distortion of a connected bipartite permutation graph on n vertices. The algorithm certifies the computed distortion by a normalised non-contractive embedding as an upper bound and an induced thick clawpath subgraph as a lower bound.*

Proof. We show that Algorithm `bpg-distortion` is such an algorithm. Let $G = (A, B, E)$ be a connected bipartite permutation graph with a competitive ordering $\langle x_1, x_2, \ldots, x_n \rangle$, and let $G_i = G[\{x_1, \ldots, x_i\}]$. Minimum distortion embeddings for G_1 and G_2 are trivial. For G_1, there is no certifying induced subgraph, and G_2 is a thick clawpath of length 0 on two vertices, thus clearly $D(G_2) = 1$. The correctness of the **for** loop of Algorithm `bpg-distortion` follows from Lemmata 5.10, 5.11, 5.13 and Corollary 5.15. Notice that if $D(G, \mathcal{E}) \leq k$ after the application of Lemma 5.11, then the loop continues to next value of i, and none of the remaining commands inside the loop are executed. Similarly, the application of Lemma 5.11 returns exactly one true case, and if this is Case 1 ($D(G) \leq k$) then the loop continues to next value of i with the same k value. If the value of k does not change from one iteration to the next then the induced thick clawpath H does not change either, since this is still a certificate of $D(G_i) > k - 2$ for the next i value as well.

For the running time, observe first that a competitive ordering for G can be computed in linear time. Furthermore, we see that the algorithms of Lemmata 5.10 and 5.11 are executed at most n times, which sums up to $\mathcal{O}(n^2)$ time. Algorithm `RepairAndDecide` is applied at most n times, so that it remains to consider the running time of a single `RepairAndDecide` application. Observe that there are at most two empty slots between consecutive vertices in a normalised embedding for a connected bipartite permutation graph. Thus, the distance between leftmost and rightmost vertex in such an embedding is at most $3n$. With the definition of nice beginning, it also follows that no further slots are needed during a computation. Every vertex is moved at most once. For every slot, we store the number of vertices of the two colour classes to its right in the embedding. Existence of empty slots can be decided from the difference of these numbers for two positions. When vertices are moved, the number information has to be updated, which takes time linear in the number of moved vertices. The existence of a breakpoint vertex can be checked straightforward since a breakpoint vertex is not moved (anymore), and thus a vertex has to be checked for being breakpoint vertex at most once. Finally, the next bad vertex is found by simply checking the vertices to the left of the previous bad vertex in (the reverse of) their order in the current embedding. This follows from condition (W2) in Lemma 5.12. Thus, `RepairAndDecide` has an $\mathcal{O}(n)$-time implementation, and the total computation running time is $\mathcal{O}(n^2)$. It remains to consider the time for computing the certificates. Modifications on the embedding in case `RepairAndDecide` accepts can be executed in $\mathcal{O}(n)$ time, since they require only some move operations. In case `RepairAndDecide` rejects a thick clawpath has to be found. The proof of Lemma 5.13 gives a recursive algorithm for doing this, that has an $\mathcal{O}(n)$-time implementation. ∎

6 Final remarks

We gave an $\mathcal{O}(n^2)$-time implementation of an algorithm for computing the distortion of connected bipartite permutation graphs. In our implementation of `RepairAndDecide` the input embedding is expected to be arbitrary. However, the actual embedding given to `RepairAndDecide` by Algorithm `bpg-distortion` is of a specific form. Is it possible to give a linear-time implementation of Algorithm `bpg-distortion` using the information about the embedding gained during previous iterations of the main loop?

Acknowledgments

The authors thank Fedor V. Fomin for suggesting the study of minimum distortion embeddings of specific graph classes. The authors also thank Dieter Kratsch for preliminary discussions

on the topic. The authors thank anonymous referees for the hints on the improvement of the presentation of the results.

References

[1] M. Bădoiu, J. Chuzhoy, P. Indyk, A. Sidiropoulos. Low-distortion embeddings of general metrics into the line. *STOC 2005*, pp. 225–233, ACM, 2005.

[2] M. Bădoiu, K. Dhamdhere, A. Gupta, Y. Rabinovich, H. Räcke, R. Ravi, A. Sidiropoulos. Approximation algorithms for low-distortion embeddings into low-dimensional spaces. *SODA 2005*, pp. 119–128, ACM and SIAM, 2005.

[3] M. Bădoiu, P. Indyk, A. Sidiropoulos. Approximation algorithms for embedding general metrics into trees. *SODA 2007*, pp. 512–521, ACM and SIAM, 2007.

[4] G. Blache, M. Karpinski, J. Wirtgen. On approximation intractability of the bandwidth problem. Technical report TR98-014, University of Bonn, 1997.

[5] A. Brandstädt, V. B. Le, J. P. Spinrad. *Graph Classes: A Survey.* SIAM Monographs on Discrete Mathematics and Applications, 1999.

[6] V. Chvátal and P. L. Hammer. Set-packing and threshold graphs. University of Waterloo Research Reports, CORR 73–21, 1973.

[7] M. R. Fellows, F. V. Fomin, D. Lokshtanov, E. Losievskaja, F. A. Rosamond, S. Saurabh. Distortion Is Fixed Parameter Tractable. *ICALP 2009*, Springer LNCS, 5555:463–474, 2009.

[8] P. Fishburn, P. Tanenbaum, A. Trenk. Linear discrepancy and bandwidth. *Order*, 18:237–245, 2001.

[9] M.C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs*. Second edition. Annals of Discrete Mathematics 57, Elsevier, 2004.

[10] P. Heggernes, D. Kratsch, D. Meister. Bandwidth of bipartite permutation graphs in polynomial time. *Journal of Discrete Algorithms*, 7:533–544, 2009.

[11] P. Heggernes and D. Meister. Hardness and approximation of minimum distortion embeddings. *Information Processing Letters*, 110:312–316, 2010.

[12] P. Indyk. Algorithmic applications of low-distortion geometric embeddings. *FOCS 2001*, pp. 10–35, IEEE, 2005.

[13] P. Indyk and J. Matousek. Low-distortion embeddings of finite metric spaces. *Handbook of Discrete and Computational Geometry*, second edition, pp. 177–196, CRC press, 2004.

[14] C. Kenyon, Y. Rabani, A. Sinclair. Low distortion maps between point sets. *STOC 2004*, pp. 272–280, ACM, 2004.

[15] D. J. Kleitman and R. V. Vohra. Computing the bandwidth of interval graphs. *SIAM Journal on Discrete Mathematics*, 3:373–375, 1990.

[16] P. J. Looges and S. Olariu. Optimal greedy algorithms for indifference graphs. *Computers & Mathematics with Applications*, 25:15–25, 1993.

[17] N. Mahadev and U. Peled. *Threshold graphs and related topics.* Annals of Discrete Mathematics 56. North Holland, 1995.

[18] B. Monien. The Bandwidth-Minimization Problem for Caterpillars with Hair Length 3 is NP-Complete. *SIAM Journal on Algebraic and Discrete Methods*, 7:505–512, 1986.

[19] C. Papadimitriou and S. Safra. The complexity of low-distortion embeddings between point sets. *SODA 2005*, pp. 112–118, ACM and SIAM, 2005.

[20] F. S. Roberts. Indifference graphs. In F. Harary (Ed.), *Proof techniques in graph theory*, pp. 139–146, Academic Press, New York, 1969.

[21] J. Spinrad, A. Brandstädt, L. Stewart. Bipartite permutation graphs. *Discrete Applied Mathematics*, 18:279–292, 1987.

[22] A. P. Sprague. An $O(n \log n)$ algorithm for bandwidth of interval graphs. *SIAM Journal on Discrete Mathematics*, 7:213–220, 1994.

[23] J. B. Tenenbaum, V. de Silva, J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, 2000.