Approximation of Minimum Cost Homomorphisms

Pavol Hell[1], Monaldo Mastrolilli[2], Mayssam Mohammadi Nevisi[1], and Arash Rafiey[2,3]

[1] School of Computing Science, SFU, Burnaby, Canada, `pavol,maysamm@sfu.ca`⋆
[2] IDSIA, Lugano, Switzerland `monaldo,arash@idsia.ch`⋆⋆
[3] Informatics Department, University of Bergen, Norway, `arash.rafiey@ii.uib.no`⋆⋆⋆

Abstract. Let H be a fixed graph without loops. We prove that if H is a co-circular arc bigraph then the minimum cost homomorphism problem to H admits a polynomial time constant ratio approximation algorithm; otherwise the minimum cost homomorphism problem to H is known to be not approximable. This solves a problem posed in an earlier paper. For the purposes of the approximation, we provide a new characterization of co-circular arc bigraphs by the existence of min ordering. Our algorithm is then obtained by derandomizing a two-phase randomized procedure. We show a similar result for graphs H in which all vertices have loops: if H is an interval graph, then the minimum cost homomorphism problem to H admits a polynomial time constant ratio approximation algorithm, and otherwise the minimum cost homomorphism problem to H is not approximable.

1 Introduction

We study the approximability of the minimum cost homomorphism problem, introduced below. A *c-approximation algorithm* produces a solution of cost at most c times the minimum cost. A *constant ratio* approximation algorithm is a c-approximation algorithm for some constant c. When we say a problem has a c-approximation algorithm, we mean a polynomial time algorithm. We say that a problem is *not approximable* if it there is no polynomial time approximation algorithm with a multiplicative guarantee unless $P = NP$.

The minimum cost homomorphism problem was introduced in [8]. It consists of minimizing a certain cost function over all homomorphisms of an input graph G to a fixed graph H. This offers a natural and practical way to model many optimization problems. For instance, in [8] it was used to model a problem of minimizing the cost of a repair and maintenance schedule for large machinery. It generalizes many other problems such as list homomorphism problems

⋆ supported by NSERC Canada; IRMACS facilities gratefully acknowledged
⋆⋆ supported by the Swiss National Science Foundation project 200020-122110/1 "Approximation Algorithms for Machine Scheduling III",
⋆⋆⋆ supported by ERC advanced grant PREPROCESSING, 267959.

(see below), retraction problems [6], and various optimum cost chromatic partition problems [10, 15–17]. (A different kind of the minimum cost homomorphism problem was introduced in [1].) Certain minimum cost homomorphism problems have polynomial time algorithms [7–9, 14], but most are NP-hard. Therefore we investigate the approximability of these problems. Note that we approximate the cost over real homomophisms, rather than approximating the maximum weight of satisfied constraints, as in, say, MAXSAT.

We call a graph *reflexive* if every vertex has a loop, and *irreflexive* if no vertex has a loop. An *interval graph* is a graph that is the intersection graph of a family of real intervals, and a *circular arc graph* is a graph that is the intersection graph of a family of arcs on a circle. We interpret the concept of an intersection graph literally, thus any intersection graph is automatically reflexive, since a set always intersects itself. A bipartite graph whose complement is a circular arc graph, will be called a *co-circular arc bigraph.* When forming the complement, we take all edges that were not in the graph, including loops and edges between vertices in the same colour. In general, the word *bigraph* will be reserved for a bipartite graph with a fixed bipartition of vertices; we shall refer to *white* and *black* vertices to reflect this fixed bipartition. Bigraphs can be conveniently viewed as directed bipartite graphs with all edges oriented from the white to the black vertices. Thus, by definition, interval graphs are reflexive, and co-circular arc bigraphs are irreflexive. Despite the apparent differences in their definition, these two graph classes exhibit certain natural similarities [2, 3]. There is also a concept of an *interval bigraph* H, which is defined for two families of real intervals, one family for the white vertices and one family for the black vertices: a white vertex is adjacent to a black vertex if and only if their corresponding intervals intersect.

A reflexive graph is a *proper interval graph* if it is an interval graph in which the defining family of real intervals can be chosen to be inclusion-free. A bigraph is a *proper interval bigraph* if it is an interval bigraph in which the defining two families of real intervals can be chosen to be inclusion-free. It turns out [11] that proper interval bigraphs are a subclass of co-circular arc bigraphs.

A *homomorphism* of a graph G to a graph H is a mapping $f : V(G) \to V(H)$ such that for any edge xy of G the pair $f(x)f(y)$ is an edge of H.

Let H be a fixed graph.

The *list homomorphism problem* to H, denoted ListHOM(H), seeks, for a given input graph G and lists $L(x) \subseteq V(H), x \in V(G)$, a homomorphism f of G to H such that $f(x) \in L(x)$ for all $x \in V(G)$. It was proved in [3] that for irreflexive graphs, the problem ListHOM(H) is polynomial time solvable if H is a co-circular arc bigraph, and is NP-complete otherwise. It was shown in [2] that for reflexive graphs H, the problem ListHOM(H) is polynomial time solvable if H is an interval graph, and is NP-complete otherwise.

The *minimum cost homomorphism problem* to H, denoted MinHOM(H), seeks, for a given input graph G and vertex-mapping costs $c(x, u), x \in V(G), u \in V(H)$, a homomorphism f of G to H that minimizes total cost $\sum_{x \in V(G)} c(x, f(x))$. It was proved in [9] that for irreflexive graphs, the problem MinHOM(H) is polynomial time solvable if H is a proper interval bigraph, and it is NP-complete otherwise. It was also shown there that for reflexive graphs H, the problem MinHOM(H) is polynomial time solvable if H is a proper interval graph, and it is NP-complete otherwise.

In [20], the authors have shown that MinHOM(H) is not approximable if H is a graph that is not bipartite or not a co-circular arc graph, and gave randomized 2-approximation algorithms for MinHOM(H) for a certain subclass of co-circular arc bigraphs H. The authors have asked for the exact complexity classification for these problems. We answer the question by showing that the problem MinHOM(H) in fact has a $|V(H)|$-approximation algorithm for **all** co-circular arc bigraphs H. Thus for an irreflexive graph H the problem MinHOM(H) has a constant ratio approximation algorithm if H is a co-circular arc bigraph, and is not approximable otherwise. We also prove that for a reflexive graph H the problem MinHOM(H) has a constant ratio approximation algorithm if H is an interval graph, and is not approximable otherwise. We use the method of randomized rounding, a novel technique of randomized shifting, and then a simple derandomization.

A *min ordering* of a graph H is an ordering of its vertices $a_1, a_2, \ldots, a_n$, so that the existence of the edges $a_i a_j, a_{i'} a_{j'}$ with $i < i', j' < j$ implies the existence of the edge $a_i a_{j'}$. A *min-max ordering* of a graph H is an ordering of its vertices $a_1, a_2, \ldots, a_n$, so that the existence of the edges $a_i a_j, a_{i'} a_{j'}$ with $i < i', j' < j$ implies the existence of the edges $a_i a_{j'}, a_{i'} a_j$. For bigraphs, it is more convenient to speak of two orderings, and we define a *min ordering* of a bigraph H to be an ordering $a_1, a_2, \ldots, a_p$ of the white vertices and an ordering $b_1, b_2, \ldots, b_q$ of the black vertices, so that the existence of the edges $a_i b_j, a_{i'} b_{j'}$ with $i < i', j' < j$ implies the existence of the edge $a_i b_{j'}$; and a *min-max ordering* of a bigraph H to be an ordering of $a_1, a_2, \ldots, a_p$ of the white vertices and an ordering $b_1, b_2, \ldots, b_q$ of the black vertices, so that the existence of the edges $a_i b_j, a_{i'} b_{j'}$ with $i < i', j' < j$ implies the existence of the edges $a_i b_{j'}, a_{i'} b_j$. (Both are instances of a general definition of min ordering for directed graphs [13].)

In Section 2 we prove that co-circular arc bigraphs are precisely the bigraphs that admit a min ordering. In the realm of reflexive graphs, such a result is known about the class of interval graphs (they are precisely the reflexive graphs that admit a min ordering) [12]. In Section 3 we discuss a linear program that computes a solution to MinHOM(H) when H has a min-max ordering. In [9], the authors used a network flow problem equivalent to this linear program, to solve to MinHOM(H) when H admits a min-max ordering. In Section 4 we recall that MinHOM(H) is not approximable when H does not have min ordering, and describe a $|V(H)|$-approximation algorithm when H is a bigraph that admits a min ordering. Finally, in Section 5 we extend our results to reflexive graphs and suggest some future work.

2 Co-circular Bigraphs and Min Ordering

A reflexive graph has a min ordering if and only if it is an interval graph [12]. In this section we prove a similar result about bigraphs. Two auxiliary concepts from [3, 5] are introduced first.

An *edge asteroid* of a bigraph H consists of $2k+1$ disjoint edges $a_0b_0, a_1b_1, \ldots, a_{2k}b_{2k}$ such that each pair a_i, a_{i+1} is joined by a path disjoint from all neighbours of $a_{i+k+1}b_{i+k+1}$ (subscripts modulo $2k+1$).

An *invertible pair* in a bigraph H is a pair of white vertices a, a' and two pairs of walks $a = v_1, v_2, \ldots, v_k = a'$, $a' = v'_1, v'_2, \ldots, v'_k = a$, and $a' = w_1, w_2, \ldots, w_m = a$, $a = w'_1, w'_2, \ldots, w'_m = a'$ such that v_i is not adjacent to v'_{i+1} for all $i = 1, 2, \ldots, k$ and w_j is not adjacent to w'_{j+1} for all $j = 1, 2, \ldots, m$.

Theorem 1. *A bigraph H is a co-circular arc graph if and only if it admits a min ordering.*

Proof. Consider the following statements for a bigraph H:

1. H has no induced cycles of length greater than three and no edge asteroids
2. H is a co-circular-arc graph
3. H has a min ordering
4. H has no invertible pairs

$1 \Rightarrow 2$ is proved in [3].

$2 \Rightarrow 3$ is seen as follows: Suppose H is a co-circular arc bigraph; thus the complement $\overline{H}$ is a circular arc graph that can be covered by two cliques. It is known for such graphs that there exist two points, the *north pole* and the *south pole*, on the circle, so that the white vertices u of H correspond to arcs A_u containing the north pole but not the south pole, and the black vertices v of H correspond to arcs A_v containing the south pole but not the north pole. We now define a min ordering of H as follows. The white vertices are ordered according to the clockwise order of the corresponding clockwise extremes, i.e., u comes before u' if the clockwise end of A_u precedes the clockwise end of $A_{u'}$. The same definition, applied to the black vertices v and arcs A_v, gives an ordering of the black vertices of H. It is now easy to see from the definitions that if $uv, u'v'$ are edges of H with $u < u'$ and $v > v'$, then A_u and $A_{v'}$ must be disjoint, and so uv' is an edge of H.

$3 \Rightarrow 4$ is easy to see from the definitions (see, for instance [5]).

$4 \Rightarrow 1$ is checked as follows: If C is an induced cycle in H, then C must be even, and any two of its opposite vertices together with the walks around the cycle form an invertible pair of H. In an edge-asteroid $a_0b_0, \ldots, a_{2k}b_{2k}$ as defined above, it is easy to see that, say, a_0, a_k is an invertible pair. Indeed, there is, for any i, a walk from a_i to a_{i+1} that has no edges to the walk $a_{i+k}, b_{i+k}, a_{i+k}, b_{i+k}, \ldots, a_{i+k}$ of the same length. Similarly, a walk a_{i+1}, b_{i+1}, a_{i+1}, $b_{i+1}, \ldots, a_{i+1}$ has no edges to a walk from a_{i+k} to a_{i+k+1} implied by the definition of an edge-asteroid. By composing such walks we see that a_0, a_k is an invertible pair. $\square$

We note that it can be decided in time polynomial in the size of H, whether a graph H is a (co-)circular arc bigraph [19].

3 An Exact Algorithm

If H is a fixed bigraph with a min-max ordering, there is an exact algorithm for the problem MinHOM(H). Suppose H has the white vertices ordered $a_1, a_2, \cdots,$ a_p, and the black vertices ordered $b_1, b_2, \cdots, b_q$. Define $\ell(i)$ to be the smallest subscript j such that b_j is a neighbour of a_i (and $\ell'(i)$ to be the smallest subscript j such that a_j is a neighbour of b_i) with respect to the ordering. Suppose G is a bigraph with white vertices u and black vertices v. We seek a minimum cost homomorphism of G to H that preserves colours, i.e., maps white vertices of G to white vertices of H and similarly for black vertices.

We define a set of variables $x_{u,i}, x_{v,j}$ for all vertices u and v of G and all $i = 1, 2, \ldots, p+1, j = 1, 2, \ldots, q+1$, and the following linear system $\mathcal{S}$.

For all vertices u (respectively v) in G and $i = 1, \ldots, p$ (respectively $j = 1, \ldots, q$)

- $x_{u,i} \geq 0$ (respectively $x_{v,j} \geq 0$)
- $x_{u,1} = 1$ (respectively $x_{v,1} = 1$)
- $x_{u,p+1} = 0$ (respectively $x_{v,q+1} = 0$)
- $x_{u,i+1} \leq x_{u,i}$ (respectively $x_{v,j+1} \leq x_{v,j}$).

For all edges uv of G and $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, q$

- $x_{u,i} \leq x_{v,\ell(i)}$
- $x_{v,j} \leq x_{u,\ell'(j)}$

Theorem 2. *There is a one-to-one correspondence between homomorphisms of G to H and integer solutions of $\mathcal{S}$. Furthermore, the cost of the homomorphism is equal to* $\sum_{u,i} c(u,i)(x_{u,i} - x_{u,i+1}) + \sum_{v,j} c(v,j)(x_{v,j} - x_{v,j+1})$.

Proof. If $f : G \to H$ is a homomorphism, we set the value $x_{u,i} = 1$ if $f(u) = a_t$ for some $t \geq i$, otherwise we set $x_{u,i} = 0$; and similarly for $x_{v,j}$. Now all the variables are non-negative, we have all $x_{u,1} = 1$, $x_{u,p+1} = 0$, and $x_{u,i+1} \leq x_{u,i}$; and similarly for $x_{v,j}$. It remains to show that $x_{u,i} \leq x_{v,\ell(i)}$ for any edge uv of G and any subscript i. (The proof of $x_{v,j} \leq x_{u,\ell'(j)}$ is analogous.) Suppose for a contradiction that $x_{u,i} = 1$ and $x_{v,\ell(i)} = 0$, and let $f(u) = a_r$, $f(v) = b_s$. This implies that $x_{u,r} = 1, x_{u,r+1} = 0$, whence $i \leq r$; and that $x_{v,s} = 1$, whence $s < \ell(i)$. Since both $a_i b_{\ell(i)}, a_r b_s$ are edges of H, the fact that we have a min ordering implies that $a_i b_s$ must also be an edge of H, contradicting the definition of $\ell(i)$.

Conversely, if there is an integer solution for $\mathcal{S}$, we define a homomorphism f as follows: we let $f(u) = a_i$ when i is the largest subscript with $x_{u,i} = 1$ (and similarly, $f(v) = b_j$ when j is the largest subscript with $x_{v,j} = 1$). Clearly, every vertex of G is mapped to some vertex of H, of the same colour. We prove that this is indeed a homomorphism by showing that every edge of G is mapped to an edge of H. Let $e = uv$ be an edge of G, and assume $f(u) = a_r, f(v) = b_s$.

We will show that $a_r b_s$ is an edge of H. Observe that $1 = x_{u,r} \leq x_{v,\ell(r)} \leq 1$ and $1 = x_{v,s} \leq x_{u,\ell'(s)} \leq 1$, so we must have $x_{u,\ell'(s)} = x_{v,\ell(r)} = 1$. Also observe that $x_{u,i} = 0$ for all $i > r$, and $x_{v,j} = 0$ for all $j > s$. Thus, $\ell(r) \leq s$ and $\ell'(s) \leq r$. Since $a_r b_{\ell(r)}$ and $a_{\ell'(s)} b_s$ are edges in H, we must have the edge $a_r b_s$, as we have a min-max ordering.

Furthermore, $f(u) = a_i$ if and only if $x_{u,i} = 1$ and $x_{u,i+1} = 0$, so, $c(u, i)$ contributes to the sum if and only if $f(u) = a_i$ (and similarly, if $f(v) = b_j$). □

We have translated the minimum cost homomorphism problem to an integer linear program: minimize the objective function in Theorem 2 over the linear system $\mathcal{S}$. In fact, this linear program corresponds to a minimum cut problem in an auxiliary network, and can be solved by network flow algorithms [9, 20]. We shall enhance the above system $\mathcal{S}$ to obtain an approximation algorithm for the case H is only assumed to have a min ordering.

4 An Approximation Algorithm

In this section we describe our approximation algorithm for MinHOM(H) in the case the fixed bigraph H has a min ordering, i.e., is a co-circular arc bigraph, cf. Theorem 1. We recall that if H is not a co-circular arc bigraph, then the list homomorphism problem ListHOM(H) is NP-complete [3], and this implies that MinHOM(H) is not approximable for such graphs H [20]. By Theorem 1 we conclude the following.

Theorem 3. *If a bigraph H has no min ordering, then MinHOM(H) is not approximable.*

Our main result is the following converse: if H has a min ordering (is a co-circular arc bigraph), then there exists a constant ratio approximation algorithm. (Since H is fixed, $|V(H)|$ is a constant.)

Theorem 4. *If H is a bigraph that admits a min ordering, then MinHOM(H) has a $|V(H)|$-approximation algorithm.*

Proof. Suppose H has a min ordering with the white vertices ordered $a_1, a_2, \cdots, a_p$, and the black vertices ordered $b_1, b_2, \cdots, b_q$. Let E' denote the set of all pairs $a_i b_j$ such that $a_i b_j$ is not an edge of H, but there is an edge $a_i b_{j'}$ of H with $j' < j$ and an edge $a_{i'} b_j$ of H with $i' < i$. Let $E = E(H)$ and define H' to be the graph with vertex set $V(H)$ and edge set $E \cup E'$. (Note that E and E' are disjoint sets.)

Observation 1. The ordering $a_1, a_2, \cdots, a_p$, and $b_1, b_2, \cdots, b_q$ is a min-max ordering of H'.

We show that for every pair of edges $e = a_i b_{j'}$ and $e' = a_{i'} b_j$ in $E \cup E'$, with $i' < i$ and $j' < j$, both $f = a_i b_j$ and $f' = a_{i'} b_{j'}$ are in $E \cup E'$.

If both e and e' are in E, $f \in E \cup E'$ and $f' \in E$.

If one of the edges, say e, is in E', there is a vertex $b_{j''}$ with $a_i b_{j''} \in E$ and $j'' < j'$, and a vertex $a_{i''}$ with $a_{i''} b_{j'} \in E$ and $i'' < i$. Now, $a_{i'} b_j$ and $a_i b_{j''}$ are both in E, so $f \in E \cup E'$. We may assume that $i'' \neq i'$, otherwise $f' = a_{i''}b_{j'} \in E$. If $i'' < i'$, then $f' \in E \cup E'$ because $a_{i'}b_{j''} \in E$; and if $i'' > i'$, then $f' \in E$ because $a_{i'}b_j \in E$.

If both edges e, e' are in E', then the earlier neighbours of a_i and b_j in E imply that $f \in E \cup E'$, and the earlier neighbours of $a_{i'}$ and $b_{j'}$ in E imply that $f' \in E \cup E'$.

Observation 2. Let $e = a_i b_j \in E'$. Then a_i is not adjacent in E to any vertex after b_j, or b_j is not adjacent in E to any vertex after a_i.

This easily follows from the fact that we have a min ordering.

Our algorithm first constructs the graph H' and then proceeds as follows. Consider an input bigraph G. Since H' has a min-max ordering, we can form the system $\mathcal{S}$ of linear inequalities for H'. By Theorem 2, homomorphisms of G to H' are in a one-to-one correspondence with integer solutions of $\mathcal{S}$. However, we are interested in homomorphisms of G to H, not H'. Therefore we shall add further inequalities to $\mathcal{S}$ to ensure that we only admit homomorphisms of G to H, i.e., avoid mapping edges of G to the edges in E'.

For every edge $e = a_i b_j \in E'$ and every edge $uv \in E(G)$, two of the following inequalities will be added to $\mathcal{S}$.

- if a_s is the first neighbour of b_j after a_i, we add the inequality
  $x_{v,j} \leq x_{u,s} + \sum_{a_t b_j \in E,\ t<i} (x_{u,t} - x_{u,t+1})$
- else if b_j has no neighbours after a_i, we add the inequality
  $x_{v,j} \leq x_{v,j+1} + \sum_{a_t b_j \in E,\ t<i} (x_{u,t} - x_{u,t+1})$
- if b_s is the first neighbour of a_i after b_j, we add the inequality
  $x_{u,i} \leq x_{v,s} + \sum_{a_i b_t \in E,\ t<j} (x_{v,t} - x_{v,t+1})$
- else if a_i has no neighbour after b_j, we add the inequality
  $x_{u,i} \leq x_{u,i+1} + \sum_{a_i b_t \in E,\ t<j} (x_{v,t} - x_{v,t+1})$.

Claim: There is a one-to-one correspondence between homomorphisms of G to H and integer solutions of the expanded system $\mathcal{S}$.

The correspondence between the integer solutions and the homomorphisms is defined as before. Thus we have a homomorphism of G to H' if and only if the old inequalities are satisfied. We shall show that the additional inequalities are also satisfied if and only if each edge of G is mapped to an edge in E, i.e., we have a homomorphism to H.

Suppose f is a homomorphism of G to H', obtained from an integer solution for $\mathcal{S}$, and, for some edge uv of G, let $f(u) = a_i, f(v) = b_j$. We have $x_{u,i} = 1$, $x_{u,i+1} = 0$, $x_{v,j} = 1$, $x_{v,j+1} = 0$, and for all $a_t b_j \in E$ with $t < i$ we have $x_{u,t} - x_{u,t+1} = 0$. If a_s is the first neighbour of b_j after a_i, then we will also have $x_{u,s} = 0$, and so the first inequality fails. Else if b_j is not adjacent to any vertex after a_i, and the second inequality fails. The remaining two other cases are similar.

Conversely, suppose f is a homomorphism of G to H (i.e., f maps the edges of G to the edges in E). For a contradiction, assume that the first inequalities fails (the other inequalities are similar). This means that for some edge $uv \in E(G)$ and some edge $a_ib_j \in E'$, we have $x_{v,j} = 1$, $x_{u,s} = 0$, and the sum of $(x_{u,t} - x_{u,t+1}) = 0$, summed over all $t < i$ such that a_t is a neighbour of b_j. The latter two facts easily imply that $f(u) = a_i$. Since b_j has a neighbour after a_i, Observation 2 tells us that a_i has no neighbours after b_j, whence $f(v) = b_j$ and thus $a_ib_j \in E$, contradicting the fact that $a_ib_j \in E'$. This proves the Claim.

At this point, our algorithm will minimize the cost function over $\mathcal{S}$ in polynomial time using a linear programming algorithm. This will generally result in a fractional solution. (Even though the original system $\mathcal{S}$ is known to be totally unimodular [20] and hence have integral optima, we have added inequalities, and hence lost this advantage.) We will obtain an integer solution by a randomized procedure called *rounding*. We choose a random variable $X \in [0,1]$, and define the rounded values $x'_{u,i} = 1$ when $x_{u,i} \geq X$, and $x'_{u,i} = 0$ otherwise; and similarly for $x'_{v,j}$. It is easy to check that the rounded values satisfy the original inequalities, i.e., correspond to a homomorphism f of G to H'.

Now the algorithm will once more modify the solution f to become a homomorphism of G to H, i.e., to avoid mapping edges of G to the edges in E'. This will be accomplished by another randomized procedure, which we call *shifting*. We choose another random variable $Y \in [0,1]$, which will guide the shifting. Let F denote the set of all edges in E' to which some edge of G is mapped by f. If F is empty, we need no shifting. Otherwise, let a_ib_j be an edge of F with maximum sum $i+j$ (among all edges of F). By the maximality of $i+j$, we know that a_ib_j is the last edge of F from both a_i and b_j. Since $F \subseteq E'$, Observation 2 implies that $e = a_ib_j$ is also the last edge of E from a_i or from b_j. Suppose e is the last edge of E from a_i. (The shifting process is similar in the other case.) So a_i does not have any edges of F or of E after a_ib_j. (There could be edges of $E' - F$, but since no edge of G is mapped to such edges, they don't matter.) We now consider, one by one, vertices u in G such that $f(u) = a_i$ and u has a neighbour v in G with $f(v) = b_j$. (Such vertices u exist by the definition of F.) For such a vertex u, consider the set of all vertices a_t with $t < i$ such that $a_tb_j \in E$. This set is not empty, since e is in E' because of two edges of E. Suppose the set consists of a_t with subscripts t ordered as $t_1 < t_2 < \ldots t_k$. The algorithm now selects one vertex from this set as follows. Let $P_{u,t} = \frac{x_{u,t} - x_{u,t+1}}{P_u}$, where

$$P_u = \sum_{a_tb_j \in E,\ t<i} (x_{u,t} - x_{u,t+1}).$$

Then a_{t_q} is selected if $\sum_{p=1}^{q} P_{u,t_p} < Y \leq \sum_{p=1}^{q+1} P_{u,t_p}$. Thus a concrete a_t is selected with probability $P_{u,t}$, which proportional to the difference of the fractional values $x_{u,t} - x_{u,t+1}$.

When the selected vertex is a_t, we shift the image of the vertex u from a_i to a_t. This modifies the homomorphism f, and hence the corresponding values of the variables. Namely, $x'_{u,t+1}, \ldots, x'_{u,i}$ are reset to 0, keeping all other values the same. Note that these modified values still satisfy the original constraints, i.e., the modified mapping is still a homomorphism.

We repeat the same process for the next u with these properties, until $a_i b_j$ is no longer in F (because no edge of G maps to it). This ends the iteration on $a_i b_j$, and we proceed to the next edge $a_{i'} b_{j'}$ with the maximum $i' + j'$ for the next iteration. Each iteration involves at most $|V(G)|$ shifts. After at most $|E'|$ iterations, the set F is empty and we no longer need to shift.

We now claim that because of the randomization, the cost of this homomorphism is at most $|V(H)|$ times the minimum cost of a homomorphism. We denote by w the value of the objective function with the fractional optimum $x_{u,i}, x_{v,j}$, and by w' the value of the objective function with the final values $x'_{u,i}, x'_{v,j}$, after the rounding and all the shifting. We also denote by w^* the minimum cost of a homomorphism of G to H. Obviously we have $w \leq w^* \leq w'$.

We now show that the expected value of w' is at most a constant times w. We focus on the contribution of one summand, say $x'_{u,t} - x'_{u,t+1}$, to the calculation of the cost. (The other case, $x'_{v,s} - x'_{v,s+1}$, is similar.)

In any integer solution, $x'_{u,t} - x'_{u,t+1}$ is either 0 or 1. The probability that $x'_{u,t} - x'_{u,t+1}$ contributes to w' is the probability of the event that $x'_{u,t} = 1$ and $x'_{u,t+1} = 0$. This can happen in the following situations.

1. u is mapped to a_t by rounding, and is not shifted away. In other words, we have $x'_{u,t} = 1$ and $x'_{u,t+1} = 0$ after rounding, and these values don't change by shifting.
2. u is first mapped to some $a_i, i > t$, by rounding, and then re-mapped to a_t by shifting. This happens if there exist j and v such that uv is an edge of G mapped to $a_i b_j \in F$, and then the image of u is shifted to a_t, where $a_t b_j \in E$. In other words, we have $x'_{u,i} = x'_{v,j} = 1$ and $x'_{u,i+1} = x'_{v,j+1} = 0$ after rounding; and then u is shifted from a_i to a_t.

For the situation in 1, we compute the expectation as follows. The values $x'_{u,t} = 1$, $x'_{u,t+1} = 0$ are obtained by rounding if $x_{u,t+1} < X \leq x_{u,t}$, i.e., with probability $x_{u,t} - x_{u,t+1}$. The probability that they are not changed by shifting is at most 1, whence this situation occurs with probability at most $x_{u,t} - x_{u,t+1}$, and the expected contribution is at most $c(u,t)(x_{u,t} - x_{u,t+1})$.

For the situation in 2, we first compute the contribution for a fixed i (for which there exist j and v as described above). The values $x'_{u,i} = x'_{v,j} = 1$ and $x'_{u,i+1} = x'_{v,j+1} = 0$ are obtained by rounding if X satisfies $\max\{x_{u,i+1}, x_{v,j+1}\} < X \leq \min\{x_{u,i}, x_{v,j}\}$, i.e., with probability $\min\{x_{u,i}, x_{v,j}\} - \max\{x_{u,i+1}, x_{v,j+1}\} \leq x_{v,j} - x_{u,i+1} \leq x_{v,j} - x_{u,s} \leq P_u$. In the last two inequalities above we have assumed that a_s is the first neighbour of b_j after a_i, and used the first inequality added above the Claim. If b_j has no neighbours after a_i, the proof is analogous, using the second added inequality. When uv maps to $a_i b_j$, we shift u to a_t with probability $P_{u,t} = \frac{(x_{u,t} - x_{u,t+1})}{P_u}$, so the overall probability is also at most $x_{u,t} - x_{u,t+1}$, and the expected contribution for a fixed i (with its j and v) is also at most $c(u,t)(x_{u,t} - x_{u,t+1})$.

Let r denote the number of vertices of H, of the same colour as a_t, that are incident with some edges of E'. Clearly the situation in 2 can occur at for at most r different values of i. Therefore a fixed u in G contributes at most $(1+r)c(u,t)(x_{u,t} - x_{u,t+1})$ to the expected value of w'. Thus the expected value of w' is at most

$$(1+r)\ (\sum_{u,i} c(u,i)(x_{u,i} - x_{u,i+1}) + \sum_{v,j} c(v,j)(x_{v,j} - x_{v,j+1})) \leq (1+r)w.$$

Since we have $w \leq w^*$, this means that the expected value of w' is at most $(1+r)w^*$. Note that $1+r \leq 1+|E'|$, and also $1+r < |V(H)|$ because a_1 (and b_1) are not incident with any edges of E' by definition.

At this point we have proved that our two-phase randomized procedure produces a homomorphism whose expected cost is at most $(1+r)$ times the minimum cost. It can be transformed to a deterministic algorithm as follows. There are only polynomially many values $x_{u,t}$ (at most $|V(G)||V(H)|$). When X lies anywhere between two such consecutive values, all computations will remain the same. Thus we can derandomize the first phase by trying all these values of X and choosing the best solution. Similarly, there are only polynomially many values of the partial sums $\sum_{p=1}^{q} P_{u,t_p}$ (again at most $|V(G)||V(H)|$), and when Y lies between two such consecutive values, all computations remain the same. Thus we can also derandomize the second phase by trying all possible values and choosing the best. Since the expected value is at most $(1+r)$ times the minimum cost, this bound also applies to this best solution. □

Corollary 1. *Let H be a co-circular arc bigraph in which at most r vertices of either colour are incident to edges of E', and let $c \geq 1+r$ be any constant.*

Then the problem MinHOM(H) has a c-approximation algorithm.

Note that c can be taken to be $|V(H)|$, or $1+|E'|$, as noted above. For $c = 1+|E'|$, we have an approximation with best bound when E' is small, in particular, an exact algorithm when E' is empty.

Finally, we conclude the following classification for the complexity of approximation of minimum cost homomorphism problems.

Corollary 2. *Let H be an irreflexive graph.*

Then the problem MinHOM(H) has a constant ratio approximation algorithm if H is a co-circular arc bigraph, and is not approximable otherwise.

5 Extensions and Future Work

Interestingly, all our steps can be repeated verbatim for the case of reflexive graphs. If H is a reflexive graph with min-max ordering $a_1, a_2, \ldots, a_n$, we again define $\ell(i)$ as the smallest subscript j, with respect to this ordering, such that a_j is a neighbour of a_i. For an input graph G, we again define variables $x_{u,i}$, for $u \in V(G), i = 1, 2, \ldots, n$, and a the same system of linear inequalities $\mathcal{S}$

(restricted to the u's), and obtain an analogue of Theorem 2. Provided H has a min ordering, we can again add edges E' as above to produce a reflexive graph H' with a min-max ordering, with the analogous properties expressed in Observations 1 and 2. We can add the corresponding inequalities to $\mathcal{S}$ as above, and there will again be a one-to-one correspondence between homomorphisms of G to H and the integer solutions to the system. Finally, we can define the approximation by the same sequence of rounding and shifting. Everything works exactly as before because it only depends on the definition of min (and min-max) ordering, which are the same. We leave the details to the reader. Finally, we use the fact that a reflexive graph has a min ordering if and only if it is an interval graph [2, 12], and the fact that the list homomorphism problem ListHOM(H) is NP-complete if the reflexive graph H is not an interval graph [2]. The last facts implies, as in [20], that the problem MinHOM(H) is not approximable if H is a reflexive graph that is not an interval graph.

Theorem 5. *Let H be a reflexive graph.*

The problem MinHOM(H) has a $|V(H)|$-approximation algorithm if H is an interval graph, and is not approximable otherwise.

We leave open the problem of approximability of MinHOM problems for general graphs, i.e., graphs with loops allowed (some vertices have loops while other don't). It should be noted that the complexity of both ListHOM and MinHOM problems for general graphs has been classified in [4, 9] respectively.

We also leave open the problem of approximability of MinHOM problems for directed graphs. The complexity of ListHOM and MinHOM problems for directed graphs has been classified in [13, 14] respectively.

It would be particularly interesting to see (in both of the open cases) whether the complexity classification for constant ratio approximability again coincides with the complexity classification for list homomorphisms.

The most interesting open question is whether the approximation ratio can be bounded by a constant independent of H. Our algorithm is both a $|V(H)|$-approximation and a $1 + |E'|$-approximation algorithm. These are constants independent of the input G, but very much dependent on the fixed graph H. For many bipartite graphs H (including the bipartite tent, net, or claw), one can choose $|E'| = 1$, thus obtaining a 2-approximation algorithm. With a bit more effort it can be shown that a 2-approximation algorithm exists for the so-called doubly convex bigraphs. We have not excluded the possibility that there exist polynomial time 2-approximation algorithms (or k-approximation algorithms, for some absolute constant k) for all co-circular arc bigraphs H. Until such a possibility is excluded, there is not much interest in making slight improvements to the approximation ratio. However, we do have a more complicated d-approximation algorithm for co-circular arc bigraphs (and reflexive interval graphs) H with maximum degree d. We have not included it here but will be happy to communicate it to interested readers upon request.

We have recently learned that Benoit Larose and Adrian Lemaitre have also characterized bipartite graphs with a min ordering [18].

References

1. G. Aggarwal, T. Feder, R. Motwani, A. Zhu, Channel assignment in wireless networks and classiffcation of minimum graph homomorphisms, Electronic Colloq. on Comput. Complexity (ECCC) TR06-040, 2006.
2. T. Feder, P. Hell List Homomorphism to reflexive graphs, J. Combin. Theory B 72 (1998) 236 – 250.
3. T. Feder, P. Hell, and J. Huang, List homomorphisms and circular arc graphs, Combinatorica 19 (1999) 487 – 505.
4. T. Feder, P. Hell, and J. Huang, Bi-arc graphs and the complexity of list homomorphisms, J. Graph Th. 42 (2003) 61– 80.
5. T. Feder, P. Hell, J.Huang and A. Rafiey, Interval graphs, adjusted interval digraphs, and reflexive list homomorphisms, Discrete Appl. Math., 2011.
6. T. Feder, P. Hell, P. Jonsson, A. Krokhin, G. Nordh, Retractions to pseudo-forests, SIAM J. on Discrete Math. 24 (2010) 101 – 112.
7. A. Gupta, P. Hell, M. Karimi, and A. Rafiey, Minimum cost homomorphisms to reflexive digraphs, LATIN 2008 (E.S. Laber et al. Eds.), LNCS 4957(2008) 182 – 193.
8. G. Gutin, A. Rafiey, A. Yeo, M. Tso Level of repair analysis and minimum cost homomorphisms of graphs, Discrete Appl. Math. 154 (2006) 881 – 889.
9. G. Gutin, P. Hell, A. Rafiey and A. Yeo, A dichotomy for minimum cost graph homomorphisms, European J. Combin. 29 (2008) 900 – 911.
10. M.M. Halldórsson, G. Kortsarz, H. Shachnai, Minimizing average completion of dedicated tasks and interval graphs, in M.X. Goemans, K. Jansen, J.D.P. Rolim, and L. Trevisan (eds.), LNCS 2129 (2001) 114 – 126.
11. P. Hell and J. Huang, Interval bigraphs and circular arc graphs, J. Graph Theory 46 (2004) 313 – 327.
12. P. Hell and J. Nešetřil, Graphs and homomorphisms. Oxford University Press, 2004.
13. P. Hell and A. Rafiey, The dichotomy of list homomorphisms for digraphs, SODA 2011.
14. P. Hell and A. Rafiey, Duality for min-max orderings and dichotomy for minimum cost homomorphisms, arXiv:0907.3016v1 [cs.DM].
15. K. Jansen, Approximation results for the optimum cost chromatic partition problem, J. Algorithms 34 (2000) 54 – 89.
16. T. Jiang, D.B. West, Coloring of trees with minimum sum of colors, J. Graph Theory 32 (1999) 354 – 358.
17. L.G. Kroon, A. Sen, H. Deng, A. Roy, The optimal cost chromatic partition problem for trees and interval graphs, in F. D'Amore, A. Marchetti-Spaccamela, P.G. Franciosa (eds.), WG 1996, LNCS 1197 (1997) 279 – 292.
18. B. Larose, A. Lemaitre, List-homomorphism problems on graphs and arc consistency, manuscript 2012.
19. R.M. McConnell, Linear-time recognition of circular-arc graphs, Algorithmica , 37 (2003), 93-147.
20. M. Mastrolilli, A. Rafiey, *On the approximation of minimum cost homomorphism to bipartite graphs*, Discrete Applied Mathematics, in press, available online 22 June 2011, http://dx.doi.org/10.1016/j.dam.2011.05.002 .