On counting homomorphisms to directed acyclic graphs*

Martin Dyer [†] Leslie Ann Goldberg[‡] Mike Paterson[§]

October 13, 2006

Abstract

It is known that if P and NP are different then there is an infinite hierarchy of different complexity classes which lie strictly between them. Thus, unless the P≠NP? question can be answered, there will be problems in NP whose precise complexity cannot be resolved. This situation has led to attempts to identify smaller classes of problems within NP where *dichotomy* results may hold: every problem is either in P or is NP-complete. A similar situation exists for *counting* problems. If P≠#P, there is an infinite hierarchy in between and it is important to identify subclasses of #P where dichotomy results hold. Graph homomorphism problems are a fertile setting in which to explore dichotomy theorems. Indeed, Feder and Vardi have shown that a dichotomy theorem for the problem of deciding whether there is a homomorphism to a fixed directed acyclic graph would resolve their long-standing dichotomy conjecture for all constraint satisfaction problems. In this paper we give a dichotomy theorem for the problem of counting homomorphisms to directed acyclic graphs. Let H be a fixed directed acyclic graph. The problem is, given an input digraph G, determine how many homomorphisms there are from G to H. We give a graph-theoretic classification, showing that for some digraphs H, the problem is in P and for the rest of the digraphs H the problem is #P-complete. An interesting feature of the dichotomy, which is absent from previously-known dichotomy results, is that there is a rich supply of tractable graphs H with complex structure.

1 Introduction

It has long been known [11] that, if P and NP are different, there is an infinite hierarchy of different complexity classes which lie strictly between them. Thus, unless the P≠NP? question can be answered, there will be problems in NP whose precise complexity cannot be resolved. This unsatisfactory situation has led to attempts to identify smaller classes of problems within NP where *dichotomy* results may hold: every problem is either in P or is NP-complete. The first such result was due to Schaeffer [14], for generalised Boolean satisfiability problems, and there has been much subsequent work. A similar situation exists for *counting* problems. The proof of Ladner's theorem [11] is easily modified to show that, if P≠#P,

*Partially supported by the EPSRC grant Discontinuous Behaviour in the Complexity of Randomized Algorithms. Some of the work was done while the authors were visiting the Mathematical Sciences Research Institute in Berkeley. A preliminary version of this paper appeared in the Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP 2006).

[†]School of Computing, University of Leeds, Leeds LS2 9JT, United Kingdom.

[‡]Department of Computer Science, University of Liverpool, Liverpool L69 3BX, United Kingdom.

[§]Department of Computer Science, University of Warwick, Coventry, CV4 7AL, United Kingdom.

there is an infinite hierarchy in between. In consequence, problem classes where counting dichotomies may exist are of equal interest. The first such result was proved by Creignou and Hermann [3], again for Boolean satisfiability, and others have followed. The theorem presented here is of this type: a dichotomy for the class of counting functions determined by the number of homomorphisms from an input digraph to a fixed directed acyclic graph.

A *homomorphism* from a (directed) graph $G = (V, E)$ to a (directed) graph $H = (\mathcal{V}, \mathcal{E})$ is a function from V to $\mathcal{V}$ that preserves (directed) edges. That is, the function maps every edge of G to an edge of H.

Hell and Nešetřil[8] gave a dichotomy theorem for the *decision* problem for undirected graphs H. In this case, H is an undirected graph (possibly with self-loops). The input, G, is an undirected simple graph. The question is "Is there a homomorphism from G to H?". Hell and Nešetřil [8] showed that the decision problem is in P if the fixed graph H has a loop, or is bipartite. Otherwise, it is NP-complete. Dyer and Greenhill [4] established a dichotomy theorem for the corresponding *counting* problem in which the question is "How many homomorphisms are there from G to H?". They showed that the problem is in P if every component of H is either a complete graph with all loops present or a complete bipartite graph with no loops present[1]. Otherwise, it is #P-complete. Bulatov and Grohe [1] extended the counting dichotomy theorem to the case in which H is an undirected *multigraph*. Their result will be discussed in more detail below.

In this paper, we study the corresponding counting problem for *directed* graphs. First, consider the decision problem: H is a fixed digraph and, given an input digraph G, we ask "Is there a homomorphism from G to H?". It is conjectured [9, Conjecture 5.12] that there is a dichotomy theorem for this problem, in the sense that, for every H, the problem is either polynomial-time solvable or NP-complete. Currently, there is no graph-theoretic conjecture stating what the two classes of digraphs will look like. Obtaining such a dichotomy may be difficult. Indeed, Feder and Vardi [7, Theorem 13] have shown that the resolution of the dichotomy conjecture for *layered* (or *balanced*) digraphs, which are a small subset of *directed acyclic graphs*, would resolve their long-standing dichotomy conjecture for all *constraint satisfaction problems*. There are some known dichotomy classifications for restricted classes of digraphs. However, the problem is open even when H is restricted to oriented trees [9], which are a small subset of layered digraphs.

The corresponding dichotomy is also open for the *counting* problem in general digraphs, although some partial results exist [2, 1]. Note that, even if the dichotomy question for the existence problem were resolved, this would not necessarily imply a dichotomy for counting, since the reductions for the existence question may not be parsimonious.

In this paper, we give a dichotomy theorem for the counting problem in which H can be any directed acyclic graph. An interesting feature of this problem, which is different from any previous dichotomy theorem for counting, is that there is a rich supply of tractable graphs H with complex structure.

The formal statement of our dichotomy is given below. Here is an informal description. First, the problem is #P-complete unless H is a *layered* digraph, meaning that the vertices of H can be arranged in levels, with edges going from one level to the next. We show (see Theorem 6.1 for a precise statement) that the problem is in P for a layered digraph H if the following condition is true (otherwise it is #P-complete). The condition is that, for every pair of vertices x and x' on level i and every pair of vertices y and y' on level $j > i$, the product of the graphs $H_{x,y}$ and $H_{x',y'}$ is isomorphic to the product of the graphs $H_{x,y'}$ and $H_{x',y}$. The precise definition of $H_{x,y}$ is given below, but the reader can think of it as the subgraph between vertex x and vertex y. The details of the product that we use (from [5])

[1] The graph with a singleton isolated vertex is taken to be a (degenerate) complete bipartite graph with no loops.

are given below. The notion of isomorphism is the usual (graph-theoretic) one, except that certain short components are dropped, as described below. Some fairly complex graphs H satisfy this condition (see, for example, Figure 5), so for these graphs H the counting problem is in P.

Our algorithm for counting graph homomorphisms for tractable digraphs H is based on *factoring.* A difficulty is that the relevant algebra lacks unique factorisation. We deal with this by introducing "preconditioners". See Section 6.

Before giving precise definitions and proving our dichotomy theorem, we note that our proof relies on two fundamental results of Bulatov and Grohe [1] and Lovász [12]. These will be introduced below in Section 3.

2 Notation and definitions

Let $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$. For $m, n \in \mathbb{N}_0$, we will write $[m, n] = \{m, m + 1, \ldots, n - 1, n\}$ and $[n] = [1, n]$. We will generally let $H = (\mathcal{V}, \mathcal{E})$ denote a fixed "colouring" digraph, and $G = (V, E)$ an "input" digraph. We denote the *empty digraph* $(\emptyset, \emptyset)$ by $\mathbf{0}$.

2.1 Homomorphisms

Let $G = (V, E)$, $H = (\mathcal{V}, \mathcal{E})$. If $f : V \to \mathcal{V}$, and $e = (v, v') \in E$, we write $f(e) = (f(v), f(v'))$. Then f is a *homomorphism* from G to H (or an *H-colouring* of G) if $f(E) \subseteq \mathcal{E}$. We will denote the number of distinct homomorphisms from G to H by $\#H(G)$. Note that $\#H(\mathbf{0}) = 1$ for all H.

Let f be a homomorphism from $H_1 = (\mathcal{V}_1, \mathcal{E}_1)$ to $H_2 = (\mathcal{V}_2, \mathcal{E}_2)$. If f is also injective, it is a *monomorphism.* Then $|\mathcal{E}_1| = |f(\mathcal{E}_1)| \leq |\mathcal{E}_2|$. If there exist monomorphisms f from H_1 to H_2 and f' from H_2 to H_1, then f is an *isomorphism* from H_1 to H_2. Then $|\mathcal{E}_1| = |f(\mathcal{E}_1)| \leq |\mathcal{E}_2| = |f'(\mathcal{E}_2)| \leq |\mathcal{E}_1|$, so $|f(\mathcal{E}_1)| = |\mathcal{E}_2|$ which implies $f(\mathcal{E}_1) = \mathcal{E}_2$. If there is an isomorphism from H_1 to H_2, we write $H_1 \cong H_2$ and say H_1 is *isomorphic* to H_2. The relation $\cong$ is easily seen to be an equivalence. We will usually use H_1 and H_2 to denote equivalence classes of isomorphic graphs, and write $H_1 = H_2$ rather than $H_1 \cong H_2$.

In this paper, we consider the particular case where $H = (\mathcal{V}, \mathcal{E})$ is a *directed acyclic graph* (DAG). Thus, in particular, H has no self-loops, and $\#H(G) = 0$ if G is not a DAG.

2.2 Layered graphs

A DAG $H = (\mathcal{V}, \mathcal{E})$ is a *layered digraph* [2] with ℓ *layers* if $\mathcal{V}$ is partitioned into $(\ell + 1)$ *levels* $\mathcal{V}_i$ $(i \in [0, \ell])$ such that $(u, u') \in \mathcal{E}$ only if $u \in \mathcal{V}_{i-1}, u' \in \mathcal{V}_i$ for some $i \in [\ell]$. We will allow $\mathcal{V}_i = \emptyset$. We will call $\mathcal{V}_0$ the *top* and $\mathcal{V}_\ell$ the *bottom.* Nodes in $\mathcal{V}_0$ are called *sources* and nodes in $\mathcal{V}_\ell$ are called *sinks.* (Note that the usage of the words *source* and *sink* varies. In this paper a vertex is called a source only if it is in $\mathcal{V}_0$. A vertex in $\mathcal{V}_i$ for some $i \neq 0$ is not called a source, even if it has in-degree 0, and similarly for sinks.) Layer i is the edge set $\mathcal{E}_i \subseteq \mathcal{E}$ of the subgraph $H^{[i-1,i]}$ induced by $\mathcal{V}_{i-1} \cup \mathcal{V}_i$. More generally we will write $H^{[i,j]}$ for the subgraph induced by $\bigcup_{k=i}^{j} \mathcal{V}_k$.

Let $\mathcal{G}_\ell$ be the class of all layered digraphs with ℓ layers and let $\mathcal{C}_\ell$ be the subclass of $\mathcal{G}_\ell$ in which every connected component spans all $\ell + 1$ levels. If $H \in \mathcal{C}_\ell$ and $G = (V, E) \in \mathcal{C}_\ell$, with V_i denoting level i $(i \in [0, \ell])$ and E_i denoting layer i $(i \in [\ell])$, then any homomorphism from

[2] This is called a *balanced digraph* in [7, 9]. However, "balanced" has other meanings in the study of digraphs.

G to H is a sequence of functions $f_i : V_i \to \mathcal{V}_i$ ($i \in [0, \ell]$) which induce a mapping from E_i into $\mathcal{E}_i$ ($i \in [\ell]$).

We use $\mathcal{C}_\ell$ to define an equivalence relation on $\mathcal{G}_\ell$. In particular, for $H_1, H_2 \in \mathcal{G}_\ell$, $H_1 \equiv H_2$ if and only if $\widehat{H_1} = \widehat{H_2}$, where $\widehat{H_i} \in \mathcal{C}_\ell$ is obtained from H_i by deleting every connected component that spans fewer than $\ell + 1$ levels.

2.3 Sums and products

If $H_1 = (\mathcal{V}_1, \mathcal{E}_1)$, $H_2 = (\mathcal{V}_2, \mathcal{E}_2)$ are disjoint digraphs, the *union* $H_1 + H_2$ is the digraph $H = (\mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E}_1 \cup \mathcal{E}_2)$. Clearly $\mathbf{0}$ is the additive identity and $H_1 + H_2 = H_2 + H_1$. If G is connected then $\#(H_1 + H_2)(G) = \#H_1(G) + \#H_2(G)$, and if $G = G_1 + G_2$ then $\#H(G) = \#H(G_1)\#H(G_2)$.

The *layered cross product* [5] $H = H_1 \times H_2$ of layered digraphs $H_1 = (\mathcal{V}_1, \mathcal{E}_1), H_2 = (\mathcal{V}_2, \mathcal{E}_2) \in \mathcal{G}_\ell$ is the layered digraph $H = (\mathcal{V}, \mathcal{E}) \in \mathcal{G}_\ell$ such that $\mathcal{V}_i = \mathcal{V}_{1i} \times \mathcal{V}_{2i}$ ($i \in [0, \ell]$), and we have $\big((u_1, u_2), (u'_1, u'_2)\big) \in \mathcal{E}$ if and only if $(u_1, u'_1) \in \mathcal{E}_1$ and $(u_2, u'_2) \in \mathcal{E}_2$. We will usually write $H_1 \times H_2$ simply as H_1H_2. It is clear that H_1H_2 is connected only if both H_1 and H_2 are connected. The converse is not necessarily true. See Figure 1.

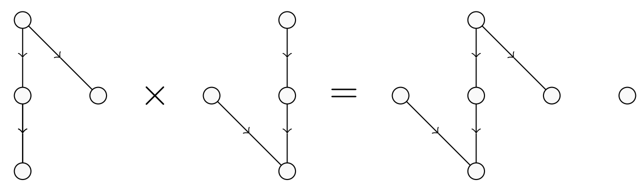

Figure 1: A disconnected product

Nevertheless, we have the following lemma.

Lemma 2.1. *If $H_1, H_2 \in \mathcal{C}_\ell$ and both of these graphs contain a directed path from every source to every sink then exactly one component of H_1H_2 spans all $\ell + 1$ levels. In each other component level 0 and level ℓ are empty.*

Proof. There is a directed path from every source of H_1H_2 to every sink. Thus, the sources and sinks are all in the same connected component. □

Note that $H_1H_2 = H_2H_1$, using the isomorphism $(u_1, u_2) \mapsto (u_2, u_1)$.

If $G, H_1, H_2 \in \mathcal{C}_\ell$ then any homomorphism $f : G \to H_1H_2$ can be written as a product $f_1 \times f_2$ of homomorphisms $f_1 : G \to H_1$ and $f_2 : G \to H_2$, and any such product is a homomorphism. Thus $\#H_1H_2(G) = \#H_1(G)\,\#H_2(G)$. Observe that the directed path P_ℓ of length ℓ gives the multiplicative identity $\mathbf{1}$ and that $\mathbf{0}H = H\mathbf{0} = \mathbf{0}$ for all H. It also follows easily that $H(H_1 + H_2) = HH_1 + HH_2$, so $\times$ distributes over $+$. The algebra $\mathcal{A} = (\mathcal{G}_\ell, +, \times, \mathbf{0}, \mathbf{1})$ is a *commutative semiring*. The $+$ operation is clearly cancellative[3]. We will show in Lemma 3.6 that $\times$ is also cancellative, at least for $\mathcal{C}_\ell$. In many respects, this algebra resembles arithmetic on $\mathbb{N}_0$, but there is an important difference. In $\mathcal{A}$ we do not have *unique factorisation into primes*. A *prime* is any $H \in \mathcal{G}_\ell$ which has only the *trivial factorisation* $H = \mathbf{1}H$. Here we may have $H = H_1H_2 = H'_1H'_2$ with H_1, H_2, H'_1, H'_2 prime and no pair equal, even if all the graphs are connected.

[3]This means that $H + H_1 = H + H_2$ implies $H_1 = H_2$. Similarly for $\times$.

Example 2.2. Let $\vec{K}_{1,m}$ be the usual undirected bipartite clique $K_{1,m}$, but with all edges directed from the root vertex, and let $\vec{K}_{m,1}$ be $\vec{K}_{1,m}$ with all edges reversed. The graphs H_1, H_2, H_1', H_2' will all have three layers. Each has top layer $\vec{K}_{1,m_1}$, middle layer a disjoint union of $K_{1,m}$'s, and bottom layer $\vec{K}_{m_3,1}$. We specify the subgraphs in each layer in the following table. We show H_1 in Figure 2, where all edges are directed downwards.

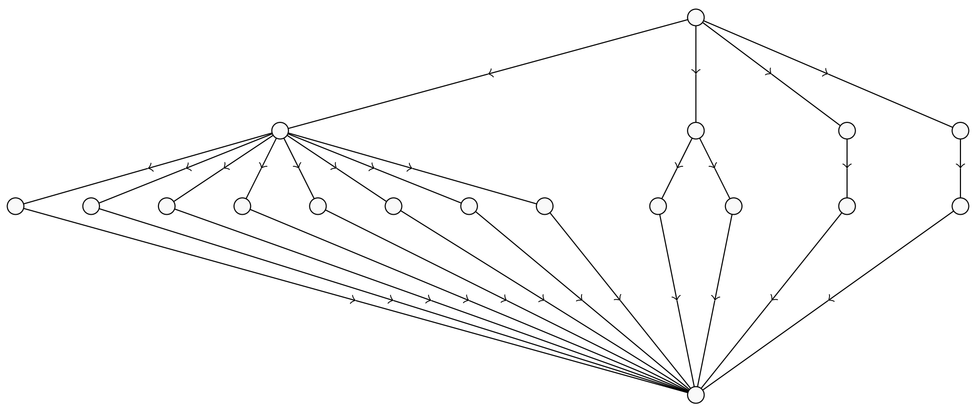

Figure 2: The graph H_1 in Example 2.2

Graph	Layer 1	Layer 2	Layer 3
H_1	$\vec{K}_{1,4}$	$\vec{K}_{1,8}+\vec{K}_{1,2}+\vec{K}_{1,1}+\vec{K}_{1,1}$	$\vec{K}_{12,1}$
H_2	$\vec{K}_{1,9}$	$\vec{K}_{1,8}+\vec{K}_{1,4}+\vec{K}_{1,2}+\vec{K}_{1,1}+\vec{K}_{1,1}+\vec{K}_{1,1}+\vec{K}_{1,1}+\vec{K}_{1,1}+\vec{K}_{1,1}$	$\vec{K}_{20,1}$
H_1'	$\vec{K}_{1,6}$	$\vec{K}_{1,8}+\vec{K}_{1,4}+\vec{K}_{1,1}+\vec{K}_{1,1}+\vec{K}_{1,1}+\vec{K}_{1,1}$	$\vec{K}_{16,1}$
H_2'	$\vec{K}_{1,6}$	$\vec{K}_{1,8}+\vec{K}_{1,2}+\vec{K}_{1,2}+\vec{K}_{1,1}+\vec{K}_{1,1}+\vec{K}_{1,1}$	$\vec{K}_{15,1}$

It is clear that H_1, H_2, H_1', H_2' are connected, and it is not difficult to show that none of them has a nontrivial factorisation. However, it is easy to verify that $H_1H_2 = H_1'H_2'$.

The layered cross product was defined in [5] in the context of interconnection networks. It is similar to the (non-layered) *direct product* [10], which also lacks unique factorisation, but they are not identical. In general, they have different numbers of vertices and edges.

3 Fundamentals

Our proof relies on two fundamental results of Bulatov and Grohe [1] and Lovász [12].

First we give the basic result of Lovász [12]. (See also [9, Theorem 2.11].) If $H = (\mathcal{V}, \mathcal{E})$, $G = (V, E)$ are DAGs, we denote the number of monomorphisms from G to H by $\Diamond H(G)$. The following is essentially a special case of Lovász [12, Theorem 3.6], though stated rather differently. We give a proof since it yields additional information.

Theorem 3.1 (Lovász)**.** *If $\#H_1(G) = \#H_2(G)$ for all G, then $H_1 = H_2$.*

Proof. Let f be any homomorphism from G to H. Then f^{-1} induces a partition I of V into disjoint sets $S_{I,1}, \ldots, S_{I,k_I}$ such that each $S_{I,i}$ ($i \in [k_I]$) is independent in G. Each partition I

fixes subsets $S_{I,i} \subseteq V$ which map to the same $u_i \in \mathcal{V}$. Let $\mathcal{I}$ be the set of all such partitions. With the relation $I \preceq I'$ whenever I is a refinement of I', $\mathsf{P}_G = (\mathcal{I}, \preceq)$ is a poset. Note that P_G depends only on G. We will write $\perp$ for the partition into singletons, so $\perp \preceq I$ for any $I \in \mathcal{I}$. Let G/I be the digraph obtained from G by identifying all vertices in $S_{I,i}$ $(i \in [k_I])$. Then we have

$$\#H(G) = \#H(G/\perp) = \sum_{I \in \mathcal{I}} \Diamond H(G/I) = \sum_{I \in \mathcal{I}} \Diamond H(G/I)\zeta(\perp, I),$$

where $\zeta(I, I') = 1$ if $I \preceq I'$, and $\zeta(I, I') = 0$ otherwise, defines the ζ-function of P_G. More generally, the same reasoning gives

$$\#H(G/I) = \sum_{I \preceq I'} \Diamond H(G/I') = \sum_{I' \in \mathcal{I}} \Diamond H(G/I')\zeta(I, I').$$

Now Möbius inversion for posets [13, Ch. 25] implies that the matrix $\zeta : \mathcal{I} \times \mathcal{I} \to \{0, 1\}$ has a unique inverse $\mu : \mathcal{I} \times \mathcal{I} \to \mathbb{Z}$. It follows directly that

$$\Diamond H(G) = \sum_{I \in \mathcal{I}} \#H(G/I)\mu(\perp, I).$$

Using the assumption of the theorem, for every G we now have

$$\Diamond H_1(G) = \sum_{I \in \mathcal{I}} \#H_1(G/I)\mu(\perp, I) = \sum_{I \in \mathcal{I}} \#H_2(G/I)\mu(\perp, I) = \Diamond H_2(G). \quad (1)$$

In particular this implies $\Diamond H_2(H_1) = \Diamond H_1(H_1) > 0$, so there is a monomorphism f from H_1 into H_2. By symmetry, there is also a monomorphism f' from H_2 into H_1. □

Remark 3.2. In this paper, the digraph H is always a DAG, so it has no self-loops. However, if we generalise to the situation in which H_1 and H_2 can have both looped and unlooped vertices (but every vertex in G is unlooped), as in the usual definition of the general H-colouring problem [4], the above proof is no longer valid. The reason is that $S_{I,i}$ must be an independent set if u_i is unlooped, but can be arbitrary if u_i is looped. Thus P_G no longer depends only on G. However, if H_1 and H_2 have *all* their vertices looped, an obvious modification of the proof goes through. Whether the theorem remains true for H with both looped and unlooped vertices is, as far as we know, an open question.

Note that, if $G \in \mathcal{C}_\ell$, the poset P_G is a *lattice* since it possesses a (unique) element $\top$ such that $I \preceq \top$ for all $I \in \mathcal{I}$, with $G/\top = P_\ell$. The following variant of Theorem 3.1 restricts H_1, H_2 and G to C_ℓ.

Theorem 3.3. *If $H_1, H_2 \in \mathcal{C}_\ell$ and $\#H_1(G) = \#H_2(G)$ for all $G \in \mathcal{C}_\ell$, then $H_1 = H_2$.*

The proof follows from the proof of Theorem 3.1 since the only instances of G used in that proof are of the form H_1/I or H_2/I and these are in $\mathcal{C}_\ell$. Similar reasoning gives the following corollaries.

Corollary 3.4. *Suppose $H_k = (\mathcal{V}_k, \mathcal{E}_k)$ $(k = 1, 2)$. If there is any G with $\#H_1(G) \neq \#H_2(G)$, there is a G such that $0 < |V| \leq \max_{k=1,2} |\mathcal{V}_k|$.*

Proof. Clearly G must be non-empty. Assume $|\mathcal{V}_1| \leq |\mathcal{V}_2|$. If $|\mathcal{V}_1| < |\mathcal{V}_2|$, taking $G_0 = (\{v_0\}, \emptyset)$ gives $\#H_1(G_0) = |\mathcal{V}_1| \neq |\mathcal{V}_2| = \#H_2(G_0)$, so we may take $|V| = 1 \leq |\mathcal{V}_2|$. Otherwise, since $H_1 \neq H_2$, either $\Diamond H_2(H_1) \neq \Diamond H_1(H_1)$ or $\Diamond H_2(H_2) \neq \Diamond H_1(H_2)$. In the former case, we can use Equation 1 with $G = H_1$ to see that one of the H_1/I must be such that $\#H_1(H_1/I) \neq \#H_2(H_1/I)$. In the latter case, one of the H_2/I must be such that $\#H_2(H_2/I) \neq \#H_1(H_2/I)$. But all the graphs H_1/I, H_2/I have at most $\max_{i=1,2} |\mathcal{V}_i|$ vertices. □

Corollary 3.5. *Suppose $H_1 = (\mathcal{V}_1, \mathcal{E}_1), H_2 = (\mathcal{V}_2, \mathcal{E}_2) \in \mathcal{C}_\ell$. If there is any $G \in \mathcal{C}_\ell$ with $\#H_1(G) \neq \#H_2(G)$, then there is such a G such that $0 < |V| \leq \max_{k=1,2} |\mathcal{V}_k|$.*

Therefore we can find a witness to the predicate $\exists G : \#H_1(G) \neq \#H_2(G)$, if one exists, among the graphs of the form H_1/I and H_2/I.

Lemma 3.6. *If $H_1H = H_2H$ for $H_1, H_2, H \in \mathcal{C}_\ell$, then $H_1 = H_2$.*

Proof. Suppose $G \in \mathcal{C}_\ell$. Since $H \in \mathcal{C}_\ell$, $\#H(G) \neq 0$. Also, since H_1 and H_2 are in $\mathcal{C}_\ell$ and $H_1H = H_2H$, $\#H_1(G)\#H(G) = \#H_2(G)\#H(G)$ so $\#H_1(G) = \#H_2(G)$. Now use Theorem 3.3. □

Here is another similar lemma that we will need. Recall that $\equiv$ denotes the equivalence relation on G_ℓ which ignores "short" components.

Lemma 3.7. *If $H_1, H_2, H \in \mathcal{C}_\ell$ and each of these contains a directed path from every source to every sink and $H_1H \equiv H_2H$ then $H_1 = H_2$.*

Proof. Suppose $G \in \mathcal{C}_\ell$. Since $H \in \mathcal{C}_\ell$, $\#H(G) \neq 0$. Let $\widehat{H_1}$ be the single full component of H_1H from Lemma 2.1. Similarly, let $\widehat{H_2}$ be the single full component of H_2H. Then since $G \in \mathcal{C}_\ell$,

$$\#\widehat{H_1}(G) = \#H_1H(G) = \#H_1(G)\#H(G)$$

and

$$\#\widehat{H_2}(G) = \#H_2H(G) = \#H_2(G)\#H(G).$$

So since $\#\widehat{H_1}(G) = \#\widehat{H_2}(G)$ by the equivalence in the statement of the lemma, we have $\#H_1(G) = \#H_2(G)$. Now use Theorem 3.3. □

The second fundamental result is a theorem of Bulatov and Grohe [1, Theorem 1], which provides a powerful generalisation of a theorem of Dyer and Greenhill [4].

Let $A = (A_{ij})$ be a $k \times k$ matrix of non-negative rationals. We view A as a weighted digraph such that there is an edge (i, j) with weight A_{ij} if $A_{ij} > 0$. Given a digraph $G = (V, E)$, $\textsc{Eval}(A)$ is the problem of computing the *partition function*

$$Z_A(G) = \sum_{\sigma: V \to \{1,\dots,k\}} \prod_{(u,v) \in E} A_{\sigma(u)\sigma(v)}. \tag{2}$$

In particular, if A is the adjacency matrix of a digraph H, $Z_A(G) = \#H(G)$. Thus $\textsc{Eval}(A)$ has the same complexity as $\#H$. If A is *symmetric*, corresponding to a weighted *undirected* graph, the following theorem characterises the complexity of $\textsc{Eval}(A)$.

Theorem 3.8 (Bulatov and Grohe)**.** *Let A be a non-negative rational symmetric matrix.*

(1) If A is connected and not bipartite, then $\textsc{Eval}(A)$ is in polynomial time if the row rank of A is at most 1; otherwise $\textsc{Eval}(A)$ is #P-complete.
(2) If A is connected and bipartite, then $\textsc{Eval}(A)$ is in polynomial time if the row rank of A is at most 2; otherwise $\textsc{Eval}(A)$ is #P-complete.
(3) If A is not connected, then $\textsc{Eval}(A)$ is in polynomial time if each of its connected components satisfies the corresponding condition stated in (1) or (2); otherwise $\textsc{Eval}(A)$ is #P-complete.

4 Reduction from acyclic H to layered H

Let $H = (\mathcal{V}, \mathcal{E})$ be a DAG. Clearly #H is in #P. We will call H *easy* if #H is in P and *hard* if #H is #P-complete. We will show that H is hard unless it can be represented as a *layered* digraph. Essentially, we do this using a "gadget" consisting of two opposing directed k-paths to simulate the edges of an undirected graph and then apply Theorem 3.8. To this end, let $N_k(u, u')$ be the number of paths of length k from u to u' in H. Say that vertices $u, u' \in \mathcal{V}$ are *k-compatible* if, for some vertex w, there is a length-k path from u to w and from u' to w. We say that H is *k-good* if, for every k-compatible pair (u, u'), there is a rational number λ such that $N_k(u, v) = \lambda N_k(u', v)$ $(\forall v \in \mathcal{V})$.

Lemma 4.1. *If there is a k such that H is not k-good then #H is #P-complete.*

Proof. Fix k, u, and u' such that u and u' are k-compatible, but there is no λ. Let A be the adjacency matrix of H and let $A' = A^k (A^k)^T$. Note that $(A^k)_{u,w} = N_k(u, w)$.

First, we show that $\textsc{Eval}(A')$ is #P-hard. Note that A' is symmetric with non-negative rational entries. Let A'' be the square sub-matrix of A' corresponding to the connected component containing u and u'. Note that u and u' are in the same connected component since they are k-compatible. Also, there are loops on vertices u and u', so A'' is not bipartite. To show that $\textsc{Eval}(A')$ is #P-complete, we need only show that the rank of A'' is bigger than 1. To do this, we just need to find a 2×2 submatrix that is non-singular, i.e., with nonzero determinant. Take the principal submatrix indexed by rows u and u' and columns u and u'. The determinant is

$$\left(\sum_w N_k(u, w)^2\right) \left(\sum_w N_k(u', w)^2\right) - \left(\sum_w N_k(u, w) N_k(u', w)\right)^2.$$

By Cauchy-Schwartz, the determinant is non-negative, and is zero only if λ exists, which we have assumed not to be the case. Thus $\textsc{Eval}(A')$ is #P-complete.

Now we use the hardness of $\textsc{Eval}(A')$ to show that $\textsc{Eval}(A)$ is #P-hard. To do this, take an undirected graph G which is an instance of $\textsc{Eval}(A')$. Construct a digraph G' by taking every edge $\{v, v'\}$ of G and replacing it with a digraph consisting of a directed length-k path P_k from v to a new vertex w and a directed length-k path P_k from v' to w. Note that $\textsc{Eval}(A)$ on G' is the same as $\textsc{Eval}(A')$ on G. Thus $\textsc{Eval}(A)$ is #P-complete, and #H has the same complexity. □

Remark 4.2. We have used the path P_k as a *gadget* in the above reduction, in order to simulate an edge of an undirected graph. We can use any other DAG G having a single source and single sink in the same way, and we do that below.

Remark 4.3. The statement of Lemma 4.1 is not symmetrical with respect to the direction of edges in H. However, let us define the digraph H^R to be that obtained from H by reversing every edge. Then #H^R and #H have the same complexity. To see this, simply observe that $\#H^R(G^R) = \#H(G)$ for all G.

We are now in a position to prove the main result of this section.

Lemma 4.4. *If H is a DAG, but it cannot be represented as a layered digraph, then #H is #P-hard.*

Proof. Suppose that H contains two paths of different lengths from u to u'. Choose $k > 0$, the length of the short path as small as possible. Choose $k' > k$, the length of the long path, as large as possible subject to the choice of k. Suppose that k edges of the long path take us from u to b and that k edges of the long path take us from a to u'. We claim that H has no length-k path from a to b. Since u and a are k-compatible, Lemma 4.1 will then give the conclusion.

If $k' \geq 2k$, the claim follows from the fact that H has no cycles. (Either $a = b$ or there is a path from b to a on the long path.) If $k' < 2k$ then the claim follows from the choice of k'. If there were a length-k path from a to b then we could go from u to a following the long path, from a to b on a k-edge path and from b to u' again following the long path. The concatenation of these paths would have length greater than k'. □

5 A structural condition for hardness

We can now formulate a sufficient condition for hardness of a layered digraph $H = (\mathcal{V}, \mathcal{E}) \in \mathcal{G}_\ell$. Suppose $s \in \mathcal{V}_i$ and $t \in \mathcal{V}_j$ for $i < j$. If there is a directed path in H from s to t, we let H_{st} be the subgraph of H induced by s, t, and all components of $H^{[i+1,j-1]}$ to which both s and t are incident. Otherwise, we let $H_{st} = \mathbf{0}$.

Lemma 5.1. *If there exist $x, x' \in \mathcal{V}_0$, $y, y' \in \mathcal{V}_\ell$ such that $H_{xy}H_{x'y'} \not\equiv H_{xy'}H_{x'y}$, and at most one of $H_{xy}, H_{xy'}, H_{x'y}, H_{x'y'}$ is $\mathbf{0}$, then $\#H$ is #P-complete.*

Proof. If exactly one of $H_{xy}, H_{xy'}, H_{x'y}, H_{x'y'}$ is $\mathbf{0}$ then Lemma 4.1 applies. Suppose that none of $H_{xy}, H_{xy'}, H_{x'y}, H_{x'y'}$ is $\mathbf{0}$. Note that x, x', y and y' are all in the same component of H and that this component is in $\mathcal{C}_\ell$. By Lemma 2.1, $H_{xy}H_{x'y'}$ contains a single component $H_{xy;x'y'}$ that spans all $\ell+1$ levels. So if $G \in \mathcal{C}_\ell$, $\#H_{xy}H_{x'y'}(G) = \#H_{xy;x'y'}(G)$. Similarly, $H_{xy'}H_{x'y}$ contains a single component $H_{xy';x'y}$ that spans all $\ell+1$ levels and $\#H_{xy'}H_{x'y}(G) = \#H_{xy';x'y}(G)$. By the assumption in the lemma, $H_{xy;x'y'} \neq H_{xy';x'y}$. So, by Theorem 3.3, there is a $\Gamma \in \mathcal{C}_\ell$ such that

$$\#H_{xy;x'y'}(\Gamma) \neq \#H_{xy';x'y}(\Gamma). \tag{3}$$

We can assume without loss of generality that Γ has a single source and sink since $H_{xy;x'y'}$ and $H_{xy';x'y}$ do. By Corollary 3.5, we can also ensure that Γ has at most $|\mathcal{V}|$ vertices. Equation (3) implies

$$\#H_{xy}(\Gamma)\#H_{x'y'}(\Gamma) \neq \#H_{xy'}(\Gamma)\#H_{x'y}(\Gamma). \tag{4}$$

We now follow the proof of Lemma 4.1, replacing the path gadget with Γ as indicated in Remark 4.2. The matrix A is indexed by sources u of H and sinks w. A_{uw} is $\#H_{uw}(\Gamma)$. Then $A' = AA^T$ so $A'_{uu'} = \sum_w A_{uw}A_{u'w}$.

As in the proof of Lemma 4.1, we first show that $\textsc{Eval}(A')$ is #P-hard. Let A'' be the square submatrix corresponding to the connected component containing x and x'. Note that they are in the same connected component since $\Gamma \in \mathcal{C}_\ell$ and $H_{x,y} \neq \mathbf{0}$ and $H_{x'y} \neq \mathbf{0}$ so $A_{xy} \neq 0$ and $A_{x'y} \neq 0$. Consider the principal submatrix indexed by rows x and x' and columns x and x'. The determinant is

$$\left(\sum_w A_{xw}^2\right)\left(\sum_w A_{x'w}^2\right) - \left(\sum_w A_{xw}A_{x'w}\right)^2.$$

As before, this is zero only if there is a λ such that $A_{xw} = \lambda A_{x'w}$ for all w and this is false by (4) which says $A_{xy}A_{x'y'} \neq A_{xy'}A_{x'y}$. Thus, the rank of A'' is bigger than 1 and $\textsc{Eval}(A')$ is #P-complete.

Now let B be the adjacency matrix of H. Reduce $\textsc{Eval}(A')$ to $\textsc{Eval}(B)$ as follows. Take an undirected graph G which is an instance of $\textsc{Eval}(A')$. Construct a digraph G' by taking every edge $\{v, v'\}$ of G and replacing it with a digraph consisting of a copy of Γ with source v

and a copy of Γ with source v' with the sinks identified. Note that $\textsc{Eval}(B)$ on G' is the same as $\textsc{Eval}(A')$ on G. Thus $\textsc{Eval}(B)$ is #P-complete, and $\#H$ has the same complexity. □

Clearly checking the condition of the Lemma and carrying out the search for the gadget Γ both require only constant time (since the size of H is a constant). Note that if x, x', y, y' are not all in the same component of H then at least two of $H_{xy}, H_{x'y'}, H_{xy'}, H_{x'y}$ are $\mathbf{0}$, so Lemma 5.1 has no content.

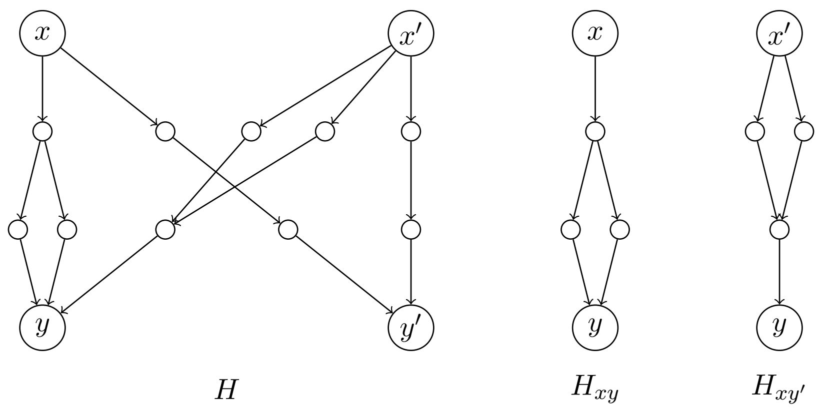

Figure 3: The graph of Example 5.2

Example 5.2. Consider the H in Figure 3. We have $H_{xy}H_{x'y'} = H_{xy}$ and $H_{xy'}H_{x'y} = H_{x'y}$. Clearly H_{xy} and $H_{x'y}$ are not isomorphic, so $\#H$ is #P-complete. A suitable gadget is $\Gamma = H_{xy}$. The following table gives $\#H_{xy}(\Gamma)$, $\#H_{xy'}(\Gamma)$, $\#H_{x'y}(\Gamma)$ and $\#H_{x'y'}(\Gamma)$. Since this matrix has rank 2 and is indecomposable, we can prove #P-completeness using Γ as a gadget.

	y	y'
x	4	1
x'	2	1

We may generalise Lemma 5.1 as follows.

Lemma 5.3. *If there exist $x, x' \in \mathcal{V}_i$, $y, y' \in \mathcal{V}_j$ ($0 \le i < j \le \ell$) such that $H_{xy}H_{x'y'} \not\equiv H_{xy'}H_{x'y}$, and at most one of $H_{xy}, H_{xy'}, H_{x'y}, H_{x'y'}$ is* $\mathbf{0}$*, then $\#H$ is #P-complete.*

Proof. If $\ell' = j - i$, consider any G having ℓ' layers. For $k = 0, 1, \ldots, \ell - \ell'$, let H_k be the subgraph of H induced by $\mathcal{V}_k \cup \mathcal{V}_{k+1} \cdots \cup \mathcal{V}_{k+\ell'}$. Then

$$\#H(G) = \sum_{k=0}^{\ell-\ell'} \#H_k(G),$$

so colouring with H is equivalent to colouring with the graph $H' = (\mathcal{V}', \mathcal{E}') = \sum_{k=0}^{\ell-\ell'} H_k \in \mathcal{G}'_{\ell}$. But x, x' are in $\mathcal{V}'_0$, and c, d in $\mathcal{V}'_{\ell'}$. The result now follows from Lemma 5.1. □

Note that the "N" of Bulatov and Dalmau [2] is the special case of Lemma 5.3 in which $j = i + 1$, $H_{xy} = H_{xy'} = H_{x'y'} = \mathbf{1}$, and $H_{x'y} = \mathbf{0}$. More generally, any structure with $H_{xy}, H_{xy'}, H_{x'y'} \neq \mathbf{0}$ and $H_{x'y} = \mathbf{0}$ is a special case of Lemma 5.3, so is sufficient to prove #P-completeness. Such a structure is equivalent to the existence of paths from x to y, x to

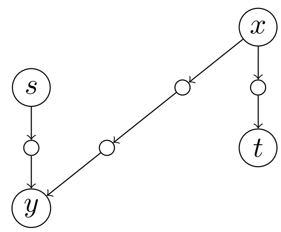

Figure 4: An easy H with no st path

y' and x' to y', when no path from x' to y exists. We call such a structure an "N", since it generalises Bulatov and Dalmau's [2] construction.

Lemma 5.4. *If $H \in \mathcal{C}_\ell$ is connected and not hard, then there exists a directed path from every source to every sink.*

Proof. Clearly, by Lemma 5.3, H must be N-free. Since it is connected, there is an undirected path from any $x \in \mathcal{V}_0$ to any $y \in \mathcal{V}_\ell$. Suppose that the shortest undirected path from x to y is not a directed path. Then some part of it induces an N in H, giving a contradiction. □

Lemma 5.4 cannot be generalised by replacing "source" with "node (at any level) with indegree 0" and replacing "sink" similarly, as the graph in Figure 4 illustrates.

We will call four vertices x, x', y, y' in H, with $x, x' \in \mathcal{V}_i$ and $y, y' \in \mathcal{V}_j$ $(0 \leq i < j \leq \ell)$, a *Lovász violation*[4] if at most one of $H_{xy}, H_{xy'}, H_{x'y}, H_{x'y'}$ is $\mathbf{0}$ and $H_{xy}H_{x'y'} \not\equiv H_{xy'}H_{x'y}$. A graph H with no Lovász violation will be called *Lovász-good*. We show next that this property is preserved under the layered cross product.

Lemma 5.5. *If $H, H_1, H_2 \in \mathcal{C}_\ell$ and $H = H_1H_2$ then H is Lovász-good if and only if both H_1 and H_2 are Lovász-good.*

Proof. Let $H_k = (\mathcal{V}_k, \mathcal{E}_k)$ with levels $\mathcal{V}_{k,i}$ $(k = 1, 2;\ i \in [0, \ell])$ and $H = (\mathcal{V}, \mathcal{E})$ with levels $\mathcal{V}_i$ $(i \in [0, \ell])$. Suppose $x_k, x'_k \in \mathcal{V}_{k,i},\ y_k, y'_k \in \mathcal{V}_{k,j}$ $(k = 1, 2;\ 0 \leq i < j \leq \ell)$, not necessarily distinct. We will write, for example, $x_1x_2 \in \mathcal{V}_i$ for product vertices and $H^1_{x,y}$ for $(H_1)_{x,y}$. Then

$$H_{x_1x_2,y_1y_2}H_{x'_1x'_2,y'_1y'_2} = H^1_{x_1,y_1}H^2_{x_2,y_2}H^1_{x'_1,y'_1}H^2_{x'_2,y'_2} \tag{5}$$

$$\text{and} \quad H_{x_1x_2,y'_1y'_2}H_{x'_1x'_2,y_1y_2} = H^1_{x_1,y'_1}H^2_{x_2,y'_2}H^1_{x'_1,y_1}H^2_{x'_2,y_2} \tag{6}$$

Now, if $H^1_{x_1,y_1}H^1_{x'_1,y'_1} \equiv H^1_{x_1,y'_1}H^1_{x'_1,y_1}$ and $H^2_{x_2,y_2}H^2_{x'_2,y'_2} \equiv H^2_{x_2,y'_2}H^2_{x'_2,y_2}$, then (5) and (6) imply $H_{x_1x_2,y_1y_2}H_{x'_1x'_2,y'_1y'_2} \equiv H_{x_1x_2,y'_1y'_2}H_{x'_1x'_2,y_1y_2}$. This if H_1 and H_2 are Lovász-good, so is H.

Conversely, suppose without loss of generality that H_1 is not Lovász-good, and that

$$H^1_{x_1,y_1}H^1_{x'_1,y'_1} \not\equiv H^1_{x_1,y'_1}H^1_{x'_1,y_1}.$$

Taking $x'_2 = x_2,\ y'_2 = y_2$ for any vertices $x_2 \in \mathcal{V}_{2,i},\ y_2 \in \mathcal{V}_{2,j}$ such that $H^2_{x_2,y_2} \neq \mathbf{0}$, from (5) and (6) we have

$$H_{x_1x_2,y_1y_2}H_{x'_1x_2,y'_1y_2} = H^1_{x_1,y_1}H^1_{x'_1,y'_1}(H^2_{x_2,y_2})^2$$

[4]The name derives from the isomorphism theorem (Theorem 3.1) of Lovász.

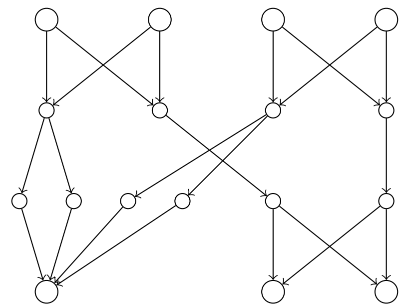

Figure 5: A Lovász-good H

and

$$H_{x_1x_2,y_1y_2}H_{x_1'x_2,y_1'y_2} = H^1_{x_1,y_1'}H^1_{x_1',y_1}(H^2_{x_2,y_2})^2.$$

First, suppose that none of $H^1_{x_1,y_1}, H^1_{x_1',y_1'}, H^1_{x_1,y_1'}, H^1_{x_1',y_1}$ is $\mathbf{0}$. Let Z_1 be the full component of $H^1_{x_1,y_1}H^1_{x_1',y_1'}$ according to Lemma 2.1. Similarly, let Z_2 be the full component of $H^1_{x_1,y_1'}H^1_{x_1',y_1}$. Let $Z = (H^2_{x_2,y_2})^2$. To show that H is not Lovász-good, we wish to show $Z_1Z \not\equiv Z_2Z$. By Lemma 3.7, this follows from $Z_1 \neq Z_2$.

Finally, suppose that exactly one of $H^1_{x_1,y_1}, H^1_{x_1',y_1'}, H^1_{x_1,y_1'}, H^1_{x_1',y_1}$ is $\mathbf{0}$. Then exactly one of $H_{x_1x_2,y_1y_2}H_{x_1'x_2,y_1'y_2}$ and $H_{x_1x_2,y_1y_2}H_{x_1'x_2,y_1'y_2}$ is $\mathbf{0}$, so they are not equivalent under $\equiv$ and H is not Lovász-good. □

The requirement of H being Lovász-good is essentially a "rank 1" condition in the algebra $\mathcal{A}$ of Section 2.3, and therefore resembles the conditions of [1, 4]. However, since $\mathcal{A}$ lacks unique factorisation, difficulties arise which are not present in the analyses of [1, 4]. But a more important difference is that, whereas the conditions of [1, 4] permit only trivial easy graphs, Lovász-good graphs can have complex structure. See Figure 5 for a small example.

6 Main theorem

We can now state the dichotomy theorem for counting homomorphisms to directed acyclic graphs.

Theorem 6.1. *Let H be a directed acyclic graph. Then $\#H$ is in P if H is layered and Lovász-good. Otherwise $\#H$ is #P-complete.*

The proof of Theorem 6.1 will use the following lemma, which we prove later.

Lemma 6.2. *Suppose $H \in \mathcal{C}_\ell$ is connected, with a single source and sink, and is Lovász-good. There is a polynomial-time algorithm for the following problem. Given a connected $G \in \mathcal{C}_\ell$ with a single source and sink, compute $\#H(G)$.*

Proof of Theorem 6.1. We have already shown in Lemma 4.4 that any non-layered H is hard. We have also shown in Lemma 5.3 that H is hard if it is not Lovász-good. Suppose $H \in \mathcal{G}_\ell$ is Lovász-good. We will show how to compute $\#H(G)$.

First, we may assume that G is connected since, as noted in Section 2.3, if $G = G_1 + G_2$ then $\#H(G) = \#H(G_1)\#H(G_2)$. We can also assume that H is connected since, for connected G, $\#(H_1 + H_2)(G) = \#H_1(G) + \#H_2(G)$, but H_1 and H_2 are Lovász-good if $H_1 + H_2$ is.

So we can now assume that $H \in \mathcal{C}_\ell$ is connected and G is connected. If G has more than $\ell+1$ non-empty levels then $\#H(G) = 0$. If G has fewer than ℓ non-empty levels then decompose H into component subgraphs H_1, H_2, ... as in the Proof of Lemma 5.3, and proceed with each component separately. So we can assume without loss of generality that both H and G are connected and in $\mathcal{C}_\ell$.

Now we just add a new level at the top of H with a single vertex, adjacent to all sources of H and a new level at the bottom of H with a single vertex, adjacent to all sinks of H. We do the same to G. Then we use Lemma 6.2. □

Before proving Lemma 6.2 we need some definitions. Suppose H is a connected graph in $\mathcal{C}_\ell$. For a subset S of sources of H, let $H_S^{[0,j]}$ be the subgraph of $H^{[0,j]}$ induced by those vertices from which there is an (undirected) path to S in $H^{[0,j]}$. We say that H is *top-j disjoint* if, for every pair of distinct sources s, s', $H_{\{s\}}^{[0,j]}$ and $H_{\{s'\}}^{[0,j]}$ are disjoint. We say that H is *bottom-j disjoint* if the reversed graph H^R from Remark 4.3 is top-j disjoint. Finally, We say that H is *fully disjoint* if it is top-$(\ell-1)$ disjoint and bottom-$(\ell-1)$ disjoint.

We will say that (Q, U, D) is a *good factorisation* of H if Q, U and D are connected Lovász-good graphs in $\mathcal{C}_\ell$ such that

- $QH \equiv UD$,
- Q has a single source and sink,
- U has a single sink, and
- D has a single source.

Remark 6.3. The presence of the "preconditioner" Q in the definition of a good factorisation is due to the absence of unique factorisation in the algebra $\mathcal{A}$. Our algorithm for computing homomorphisms to a Lovász-good H works by factorisation. However, it is possible to have a non-trivial Lovász-good H which is prime. A simple example can be constructed from the graphs H_1, H_2, H_1', H_2' of Example 2.2, by identifying the sources in H_1, H_1' and in H_2, H_2', and the sinks in H_1, H_2' and in H_1', H_2. The resulting 2-source, 2-sink graph has no nontrivial factorisation.

We use the following operations on a Lovász-good connected digraph $H \in \mathcal{C}_\ell$.

Local Multiplication: Suppose that U is a connected Lovász-good single-sink graph in $\mathcal{C}_j$ on levels $0, \dots, j$ for $j \le \ell$. Let C be a Lovász-good connected component in $H^{[0,j]}$ with no empty levels. Then $\mathrm{Mul}(H, C, U)$ is the graph constructed from H by replacing C with the full component of UC. (Note that there is only one full component, by Lemma 5.4 and 2.1.)

Local Division: Suppose $S \subseteq \mathcal{V}_0$, and that (Q, U, D) is a good factorisation of $H_S^{[0,j]}$. Then $\mathrm{Div}(H, Q, U, D)$ is the graph constructed from H by replacing $H_S^{[0,j]}$ with D.

We can now state our main structural lemma.

Lemma 6.4. *If $H \in \mathcal{C}_\ell$ is connected, and Lovász-good, then it has a good factorisation (Q, U, D).*

We prove Lemma 6.4 below in Section 7. In the course of the proof, we give an algorithm for constructing (Q, U, D). We now describe how we use Lemma 6.4 (and the algorithm) to prove Lemma 6.2.

Proof of Lemma 6.2. The proof is by induction on ℓ. The base case is $\ell = 2$. (Note that calculating $\#H(G)$ is easy in this case.) For the inductive step, suppose $\ell > 2$. Let H' denote the part of H excluding levels 0 and ℓ and let G' denote the part of G excluding levels 0 and ℓ. Using reasoning similar to that in the proof of Theorem 6.1, we can assume that G' is connected and then that H' is connected. Since H is Lovász-good, so is H'. Now by Lemma 6.4 there is a good factorisation (Q', U', D') of H'.

Let $S \subseteq \mathcal{V}_1$ be the nodes in level 1 of H that are adjacent to the source and $T \subseteq \mathcal{V}_{\ell-1}$ be the nodes in level $\ell - 1$ of H that are adjacent to the sink. Note that $\mathcal{V}_1$ is the top level of U' and $\mathcal{V}_{\ell-1}$ is the bottom level of D'.

Construct Q from Q' by adding a new top and bottom level with a new source and sink. Connect the new source and sink to the old ones. Construct D from D' by adding a new top and bottom level with a new source and sink. Connect the new source to the old one and the new sink to T. Finally, construct U from U' by adding a new top and bottom level with a new source and sink. Connect the new source to S and the new sink to the old one. See Figure 6. Note that (Q, U, D) is a good factorisation of H. To see that $QH \equiv UD$, consider the component of $Q'H'$ that includes sources and sinks. (There is just one of these. Since H' is Lovász-good, it has a directed path from every source to every sink by Lemma 5.4. So does Q'. Then use Lemma 2.1.) This is isomorphic to the corresponding component in $D'U'$ since (Q', U', D') is a good factorisation of H'. The isomorphism maps S in H' to a corresponding S in U' and now note that the new top level is appropriate in QH and DU. Similarly, the new bottom level is appropriate.

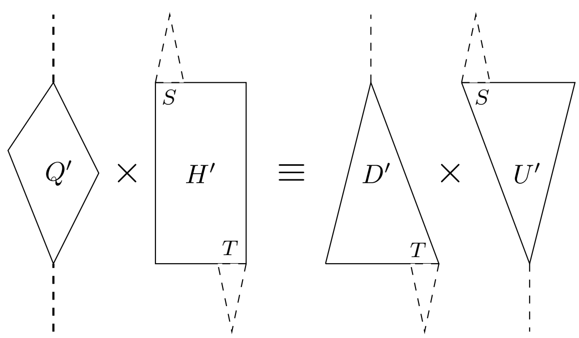

Figure 6: A good factorisation of H

Now let's consider how to compute $\#Q(G)$. In any homomorphism from G to Q, every node in level 1 of G gets mapped to the singleton in level 1 of Q. Thus, we can collapse all level 1 nodes of G into a single vertex without changing the problem. At this point, the top level of G and Q are not doing anything, so they can be removed, and we have a sub-problem with fewer levels. So $\#Q(G)$ can be computed recursively. The same is true for $\#D(G)$ and $\#U(G)$.

Since $G \in \mathcal{C}_\ell$, $\#QH(G) = \#Q(G)\#H(G)$. Also, since components without sources and sinks cannot be used to colour G (which has the full ℓ layers), this is equal to $\#D(G)\#U(G)$. Thus, we can output $\#H(G) = \#D(G)\#U(G)/\#Q(G)$.

□

That concludes the proof of Theorem 6.2, so it only remains to prove Lemma 6.4. The proof will be by induction, and we will need the following technical lemmas in the induction.

Lemma 6.5. *Suppose that H is Lovász-good, top-$(j-1)$ disjoint, $S \subseteq \mathcal{V}_0$ and $H_S^{[0,j]}$ is connected. If (Q, U, D) is a good factorisation of $H_S^{[0,j]}$, then* $\mathrm{Div}(H, Q, U, D)$ *is Lovász-good.*

Proof. Suppose to the contrary that $\widetilde{H} = \mathrm{Div}(H, Q, U, D)$ contains a Lovász violation x, x', y, y', with x, x' in level i, and y, y' in level k. Since D is Lovász-good, we may assume without loss that x is in D, with $i < j$, and that $k > j$. Thus y, y' are not in D. Let $\widetilde{Q}$, $\widetilde{U}$ be $Q^{[i,j]}$, $U^{[i,j]}$, both extended downwards by a single path to level k. There are two cases.

If x' is in D, Lemma 5.5 gives us a Lovász violation in $\widetilde{H}\widetilde{U}^{[i,k]}$. (By the proof of Lemma 5.5, this involves nodes at level i and k.) This gives us a Lovász violation in $H^{[i,k]}\widetilde{Q}^{[i,k]}$ since these graphs are isomorphic except for components that don't extend down to y and y'. Finally, by Lemma 5.5, we find a Lovász violation in $H^{[i,k]}$, and hence in H, which gives a contradiction.

If x' is not in D, then let H^* be $\mathrm{Mul}(\widetilde{H}, D, U)$. Let w be some vertex on level i of $\widetilde{U}$ which has a directed path to the sink z of $\widetilde{U}$. Now by construction

$$H^*_{xw,y} H^*_{x',y'} \equiv \widetilde{H}_{xy} \widetilde{H}_{x'y'} (\widetilde{U}_{wz}),$$

$$H^*_{xw,y'} H^*_{x',y} \equiv \widetilde{H}_{xy'} \widetilde{H}_{x'y} (\widetilde{U}_{wz}).$$

Because these graphs all have single sources and sinks (at levels i and k), we can apply Lemma 3.7 to cancel and find a Lovász violation xw, y, x', y' in H^*.

Then do a local multiplication multiplying the component of x' by Q and again, in the same way, we find a Lovász violation $xw, y, x'w', y'$ in the resulting graph, H^{**}

Now since $DU \equiv QH_S^{[0,j]}$ and since xw does have a path down to level j of H^{**}, the component containing xw in $H^{**[0,j]}$ is isomorphic to the corresponding component in $QH^{[0,j]}$. Thus, the same Lovász violation exists in $\widetilde{Q}H$ (where now we extend the tail of Q all the way down to level ℓ of H.

Now to get a Lovász violation in H itself, we use the reasoning from the proof of Theorem 5.5. Say the Loász violation is $x_Q x_H, y_Q y_H, x'_Q x'_H, y'_Q y'_H$. Then

$$(\widetilde{Q}H)^{[i,k]}_{x_Q x_H y_Q y_H} (\widetilde{Q}H)^{[i,k]}_{x'_Q x'_H y'_Q y'_H} \equiv Q_{x_Q y_Q} H_{x_H y_H} Q_{x'_Q y'_Q} H_{x'_H y'_H},$$

$$(\widetilde{Q}H)^{[i,k]}_{x_Q x_H y'_Q y'_H} (\widetilde{Q}H)^{[i,k]}_{x'_Q x'_H, y_Q y_H} \equiv Q_{x_Q y'_Q} H_{x_H y'_H} Q_{x'_Q y_Q} H_{x'_H y_H}.$$

And since Q is Lovász-good, the first of these is equivalent to $Q_{x_Q y'_Q} H_{x_H y_H} Q_{x'_Q y_Q} H_{x'_H y'_H}$. Now if we had

$$H_{x_H y_H} H_{x'_H y'_H} \equiv H_{x_H y'_H} H_{x'_H y_H}$$

we could multiply both sides by $Q_{x_Q y'_Q} Q_{x'_Q y_Q}$ to get

$$(\widetilde{Q}H)^{[i,k]}_{x_Q x_H y_Q y_H} (\widetilde{Q}H)^{[i,k]}_{x'_Q x'_H y'_Q y'_H} \equiv (\widetilde{Q}H)^{[i,k]}_{x_Q x_H y'_Q y'_H} (\widetilde{Q}H)^{[i,k]}_{x'_Q x'_H, y_Q y_H}$$

which is a contradiction since this was a Lovász violation.

□

Lemma 6.6. *Suppose that H and U are Lovász-good. Then* $\mathrm{Mul}(H, C, U)$ *is Lovász-good.*

Proof. Suppose $M = \mathrm{Mul}(H, C, U)$ is not Lovász-good. By Lemma 5.5, the full component of UC is Lovász-good, so there is a Lovász violation in levels i and k of M with $i \le j$ and $k > j$. Also, one of the relevant vertices in level i is in the component from UC in the local multiplication.

First, suppose that both of the relevant vertices in level i are in this component. Then there is a Lovász violation (the same one) in $\widehat{U}H$, where $\widehat{U}$ extends U with a path. Now restricting attention to levels i–k, Lemma 5.5 shows that there is a Lovász violation in H or $\widehat{U}$ (hence U), which is a contradiction.

Second, suppose that only one of the relevant vertices in level i is in the component from CU in the local multiplication. Then the Lovász violation can be written

$$M^{[i,k]}_{U_x C_x, y} M^{[i,k]}_{x', y'} \not\equiv M^{[i,k]}_{U_x C_x, y'} M^{[i,k]}_{x', y}$$

and expanding the product, letting $\widehat{U}$ denote the extension of U downwards with a path to node s on level k, this is

$$\widehat{U}^{[i,k]}_{U_x, s} H^{[i,k]}_{C_x, y} H^{[i,k]}_{x', y'} \not\equiv \widehat{U}^{[i,k]}_{U_x, s} H^{[i,k]}_{C_x, y'} H^{[i,k]}_{x', y}.$$

But this gives us a Lovász violation in H, which is a contradiction. □

7 Proof of Lemma 6.4

A *top-dangler* is a component in $H^{[1,\ell-1]}$ that is incident to a source but not to a sink. Similarly, a *bottom-dangler* is a component in $H^{[1,\ell-1]}$ that is incident to a sink but not to a source. (Note that a bottom-dangler in H is a top-dangler in H^R.)

The proof of Lemma 6.4 will be by induction. The base case will be $\ell = 1$, where it is easy to see that a connected Lovász-good H must be a complete bipartite graph. The ordering for the induction will be lexicographic on the following criteria (in order).

1. the number of levels,
2. the number of sources,
3. the number of top-danglers,
4. the number of sinks,
5. the number of bottom-danglers.

Thus, for example, if H' has fewer levels than H then H' precedes H in the induction. If H' and H have the same number of levels, the same number of sources and the same number of top-danglers but H' has fewer sinks then H' precedes H in the induction.

The inductive step will be broken into five cases. The cases are exhaustive but not mutually exclusive – given an H we will apply the first applicable case.

Case 1: H is top-$(j-1)$ disjoint and has a top-dangler with depth at most $j-1$.

Case 2: For $j < \ell$, H is top-$(j-1)$ disjoint, but not top-j disjoint, and has no top-dangler with depth at most $j-1$.

Case 3: H is top-$(\ell-1)$ disjoint and has no top-dangler and is bottom-$(j-1)$ disjoint and has a bottom-dangler with height at most $(j-1)$.

Case 4: For $j < \ell$, H is top-$(\ell-1)$ disjoint and has no top-dangler and is bottom-$(j-1)$ disjoint, but not bottom-j disjoint, and has no bottom-dangler with height at most $(j-1)$.

Case 5: H is fully disjoint. and has no top-danglers or bottom-danglers.

7.1 Case 1: *H is top-$(j-1)$ disjoint and has a top-dangler with depth at most $j-1$.*

Let R be a top-dangler with depth j' where $j' < j$ (meaning that it is a component in $H^{[1,\dots,\ell-1]}$ that is incident to a source but not to a sink, and that levels $j'+1,\dots,\ell-1$ are empty and level j' is non-empty). Note that since H is top-$(j-1)$ disjoint, R must be adjacent to a single source, v, in H. This follows from the definition of top-$(j-1)$ disjoint, and from the fact that R has depth at most $j-1$.

Construct H' from H by removing R. Note that H' is connected. By construction (from H), H' is Lovász-good, and has no empty levels. It precedes H in the induction order since it has the same number of levels, the same number of sources and one fewer top-dangler. By induction, it has a good factorisation (Q', U', D') so $Q'H' \equiv U'D'$.

Construct $\widehat{D}$ as follows. On layers $1,\dots,j'$, $\widehat{D}$ is identical to D'. On layers $j'+2,\dots,\ell$, $\widehat{D}$ is a path. Every node in level j' is connected to the singleton vertex in level $j'+1$. Then clearly $\widehat{D}Q'H' \equiv \widehat{U}D'$ where $\widehat{U}$ is the single full component of $\widehat{D}U'$. (There is just one of these. Since $\widehat{D}$ is Lovász-good, it has a directed path from every source to every sink by Lemma 5.4. So does U'. Then use Lemma 2.1.) Note that $\widehat{U}$ has a single sink.

Let R' be the graph obtained from R by adding the source v. Let R'' be $Q'^{[0,j']}R'$. Form U'' from R'' and $\widehat{U}$ by identifying v with the appropriate source of $\widehat{U}$. (Note that $\widehat{U}$ has the same sources as H). Then $\widehat{D}Q'H \equiv U''D'$.

Thus, we have a good factorisation (Q, U, D') of H by taking Q to be the full component of $\widehat{D}Q'$ and U to be the full component of U''. To see that it is a good factorisation, use Lemma 5.5 to show that Q and U are Lovász-good.

7.2 Case 2: *For $j < \ell$, H is top-$(j-1)$ disjoint, but not top-j disjoint, and has no top-dangler with depth at most $j-1$.*

Partition the sources of H into equivalence classes $S_1,\dots,S_k$ so that the graphs $H_{S_i}^{[0,j]}$ are connected and pairwise disjoint. See Figure 7. Since H is not top-j disjoint, some equivalence class, say S_1, contains more than one source.

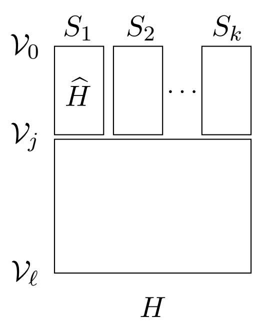

Figure 7: H

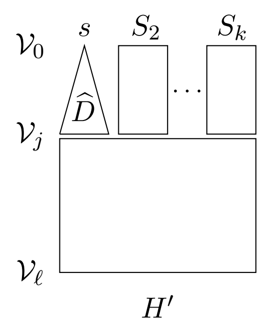

Figure 8: H'

Let $\widehat{H}$ denote $H_{S_1}^{[0,j]}$. $\widehat{H}$ is shorter than H, so it comes before H in the induction order. It is connected by construction of the equivalence classes, and it is Lovász-good by virtue of being a subgraph of H. By induction we can construct a good factorisation $(\widehat{Q}, \widehat{U}, \widehat{D})$ of $\widehat{H}$. Let

$H' = \mathrm{Div}(H, \widehat{Q}, \widehat{U}, \widehat{D})$. See Figure 8.

H' comes before H in the induction order because it has the same number of levels, but fewer sources. To see that H' is connected, note that H is connected and $\widehat{H}$ is connected. Since $(\widehat{Q}, \widehat{U}, \widehat{D})$ is a good factorisation), we know $\widehat{D}$ is connected, so H' is connected. By Lemma 6.5, H' is Lovász-good. By induction, we can construct a good factorisation (Q', U', D') of H'.

Let s be the (single) source of $\widehat{D}$. By construction, the sources of U' are $\{s\} \cup S_2 \cup \cdots \cup S_k$. See Figure 9.

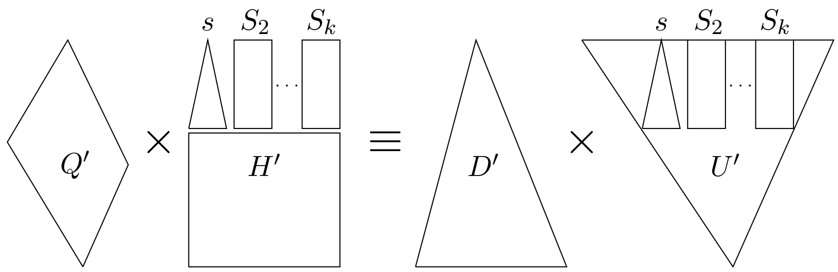

Figure 9: $Q'H' \equiv D'U'$

Let $C_1, \ldots, C_z$ be the connected components of ${U'}^{[0,j]}$. Let C_1 be the component containing s. Since the $H_{S_i}^{[0,j]}$ are connected and pairwise disjoint, and Q' has a single source, there is a single connected component of $(Q'H')^{[0,j]}$ containing all of S_i (and no other sources) so (see Figure 9) there is a single connected component of ${U'}^{[0,j]}$ containing all of S_i (and no other sources). For convenience, call this C_i.

If $z > k$ then components $C_{k+1}, \ldots, C_z$ do not contain any sources. (They are due to danglers in U', which in this case are nodes that are not descendants of a source.)

Now consider $U_1 = \mathrm{Mul}(U', C_1, \widehat{U})$. U_1 is the graph constructed from U' by replacing C_1 with the full component of $C_1\widehat{U}$. For $i \in \{2, \ldots, z\}$, let $U_i = \mathrm{Mul}(U_{i-1}, C_i, \widehat{Q})$. U_z is the graph constructed from U' by replacing C_1 with the full component of $C_1\widehat{U}$ and replacing every other C_i with the full component of $C_i\widehat{Q}$.

Let $\widetilde{Q}$ extend $\widehat{Q}$ down to level ℓ with a single path. We claim that

$$Q'\widetilde{Q}H \equiv U_z D'. \tag{7}$$

To establish Equation (7), note that on levels $[j, \ldots, \ell]$ the left-hand side is

$$(Q'\widetilde{Q}H)^{[j,\ell]} \equiv (Q'H')^{[j,\ell]}.$$

Any components of $Q'H'$ that differ from $U'D'$ do not include level ℓ, so this is equivalent to

$$(U'D')^{[j,\ell]} \equiv (U_z D')^{[j,\ell]},$$

which is the right-hand side. So focus on levels $0, \ldots, j$.

From the left-hand side, look at the component $(Q'\widetilde{Q}H)_{S_1}^{[0,j]}$. Note that it is connected. It is

$$({Q'}^{[0,j]}\widehat{Q}\widehat{H})_{S_1}^{[0,j]} \equiv ({Q'}^{[0,j]}\widehat{U}\widehat{D})_{S_1}^{[0,j]} \equiv (({D'}^{[0,j]}{U'}^{[0,j]})_{\{s\}}\widehat{U})_{S_1}^{[0,j]} \equiv (D'U_z)_{S_1}^{[0,j]},$$

which is the right-hand side.

Then look at the component S_2.

$$(Q'\widetilde{Q}H)_{S_2}^{[0,j]} \equiv (Q'^{[0,j]}\widehat{Q}H^{[0,j]})_{S_2} \equiv ((D'^{[0,j]}U'^{[0,j]})_{S_2}\widehat{Q})_{S_2}^{[0,j]} \equiv (D'U_z)_{S_2}^{[0,j]}.$$

The other components containing sources (which are the only components that we care about) are similar.

Having established (7), we observe that (Q, U, D') is a good factorisation of H where Q is the full component of $Q'\widetilde{Q}$ and U is the full component of U_z. Use Lemma 5.5 to show Q is Lovász-good and Lemma 6.6 to show that U_z is.

7.3 Case 3: *H is top-$(\ell - 1)$ disjoint and has no top-dangler and is bottom-$(j - 1)$ disjoint and has a bottom-dangler with height at most $(j - 1)$.*

We apply an analysis similar to Section 7.1 to the reversed graph from Remark 4.3. Let H^R be an instance in Case 3. Thus H is top-$(j - 1)$ disjoint and has a top-dangler with depth at most $j - 1$. Also, H is bottom-$(\ell - 1)$ disjoint and has no bottom dangler. Apply the transformation in Section 7.1 to H. This produces as inductive instance H'. As we noted in Section 7.1, H' precedes H in the induction order since it has the same number of levels, the same number of sources and one fewer top-dangler. Crucially, H' has the same number of sinks as H and the same number of bottom-danglers as H. (We have not added any.) Thus, ${H'}^R$ precedes H^R in the inductive order. It has one fewer bottom-dangler and everything else is the same. Then the good factoring (Q, U, D) of H that we produce gives us a good factoring (Q^R, D^R, U^R) of H^R.

7.4 Case 4: *For $j < \ell$, H is top-$(\ell - 1)$ disjoint and has no top-dangler and is bottom-$(j - 1)$ disjoint, but not bottom-j disjoint, and has no bottom-dangler with height at most $(j - 1)$.*

We apply an analysis similar to Section 7.2 to the reversed graph from Remark 4.3. Let H^R be an instance in Case 4. The reversed graph H is bottom-$(\ell - 1)$-disjoint with no bottom-dangler. For some $j < \ell$, it is top-$(j-1)$-disjoint, but not top-j disjoint and has no top-dangler with depth at most $j - 1$. Apply the analysis in Case 3. This produces a recursive instance H' with fewer sources. H' has the same number of levels as H. Furthermore, H' has the same number of sinks as H and the same number of bottom-danglers. Thus, ${H'}^R$ has fewer sinks than H^R, but the same number of levels, sources, and top-danglers. So it precedes H^R in the induction order. Then the good factoring (Q, U, D) of H that we produce gives us a good factoring (Q^R, D^R, U^R) of H^R.

7.5 Case 5: *H is fully disjoint and has no top-danglers or bottom-danglers.*

In the fully disjoint case, the subgraphs H_{st} $(s \in \mathcal{V}_0, t \in \mathcal{V}_\ell)$ satisfy

$$H_{st} \cap H_{s't'} = \begin{cases} \{s\}, & \text{if } s = s',\ t \neq t'; \\ \{t\}, & \text{if } s \neq s',\ t = t'; \\ \emptyset, & \text{if } s \neq s',\ t \neq t'. \end{cases} \tag{8}$$

$H_{st} \neq \mathbf{0}$ and $H_{st}H_{s't'} \equiv H_{st'}H_{s't}$, since H is Lovász-good. We assume without loss that $|\mathcal{V}_0| > 1$ and $|\mathcal{V}_\ell| > 1$, since otherwise $(\mathbf{1}, H, \mathbf{1})$ or $(\mathbf{1}, \mathbf{1}, H)$ is a good factorisation. See Figure 10.

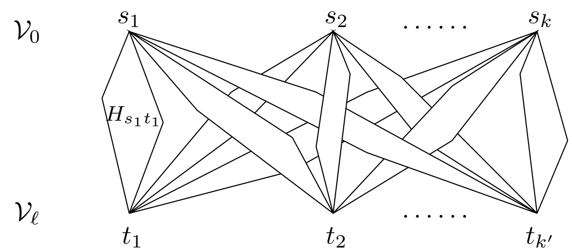

Figure 10: Fully disjoint case

Choose any $s^* \in \mathcal{V}_0, t^* \in \mathcal{V}_\ell$, and let $Q = H_{s^*t^*}$. Note that Q is connected with a single source and sink, and is Lovász-good because H is Lovász-good. Let D be the subgraph $\bigcup_{t\in\mathcal{V}_\ell} H_{s^*t}$ of H, and let U be the subgraph $\bigcup_{s\in\mathcal{V}_0} H_{st^*}$ of H. These are both connected and Lovász-good since H is. Clearly D has a single source and U has a single sink. Also $QH \equiv DU$ follows from (8) and from the fact that there are no top-danglers or bottom-danglers and

$$(DU)_{s^*s,tt^*} = D_{s^*t}U_{st^*} \equiv H_{s^*t}H_{st^*} = H_{s^*t^*}H_{st} = QH_{st} \qquad (s \in \mathcal{V}_0;\ t \in \mathcal{V}_\ell), \tag{9}$$

where we have used the fact that H is Lovász-good. Thus (Q, U, D) is a good factorisation of H.

References

[1] A. Bulatov and M. Grohe, The complexity of partition functions, in *Automata, Languages & Programming: 31st International Colloquium*, Lecture Notes in Computer Science **3142**, pp. 294–306, 2004.

[2] A. Bulatov and V. Dalmau, Towards a dichotomy theorem for the counting constraint satisfaction problem, in *Proc. 44th IEEE Symposium on Foundations of Computer Science*, IEEE, pp. 562–572, 2003.

[3] N. Creignou and M. Hermann, Complexity of generalized satisfiability counting problems, *Information and Computation* **125** (1996), 1–12.

[4] M. Dyer and C. Greenhill, The complexity of counting graph homomorphisms, *Random Structures & Algorithms* **17** (2000), 260–289.

[5] S. Even and A. Litman, Layered cross product: a technique to construct interconnection networks. *Networks* **29** (1997), 219–223.

[6] T. Feder and M. Vardi, Monotone monadic SNP and constraint satisfaction, in *Proc. 25th Annual ACM Symposium on Theory of Computing*, ACM Press, 1993, pp. 612–622.

[7] T. Feder and M. Vardi, The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory, *SIAM J. Comput.* **28** (1998) 57–104.

[8] P. Hell and J. Nešetřil, On the complexity of H-coloring, *Journal of Combinatorial Theory Series B* **48** (1990), 92–110.

[9] P. Hell and J. Nešetřil, *Graphs and homomorphisms*, Oxford University Press, 2004.

[10] W. Imrich and S. Klavžar, *Product graphs: structure and recognition*, Wiley, New York, 2000.

[11] R. Ladner, On the structure of polynomial time reducibility, *Journal of the Association for Computing Machinery* **22** (1975), 155–171.

[12] L. Lovász, Operations with structures, *Acta. Math. Acad. Sci. Hung.*, **18** (1967), 321–328.

[13] J. van Lint and R. Wilson, *A course in combinatorics* (2nd ed.), CUP, Cambridge, 2001.

[14] T. Schaefer, The complexity of satisfiability problems, in *Proc. 10th Annual ACM Symposium on Theory of Computing*, ACM Press, 1978, pp. 216–226.