Electronic Colloquium on Computational Complexity, Report No. 45 (2019)

On the Fine-grained Complexity of Least Weight Subsequence in Graphs

Jiawei Gao*
University of California, San Diego
jiawei@cs.ucsd.edu

February 19, 2019

Abstract

Least Weight Subsequence (LWS) is a type of highly sequential optimization problems with form $F(j) = \min_{i<j}[F(i) + c_{i,j}]$. They can be solved in quadratic time using dynamic programming, but it is not known whether these problems can be solved faster than $n^{2-o(1)}$ time. Surprisingly, each such problem is subquadratic time reducible to a highly parallel, non-dynamic programming problem [KPS17]. In other words, if a "static" problem is faster than quadratic time, so is an LWS problem. For many instances of LWS, the sequential versions are equivalent to their static versions by subquadratic time reductions. The previous result applies to LWS on linear structures, and this paper extends this result to LWS on paths in sparse graphs. When the graph is a multitree (i.e. a DAG where any pair vertices can have at most one path) or when the graph is a DAG whose underlying undirected graph has constant treewidth, we show that LWS on this graph is still subquadratically reducible to their corresponding static problems. For many instances, the graph versions are still equivalent to their static versions.

Moreover, this paper shows that on these graphs, property testing of form $\exists x \exists y (\mathrm{TC}_E(x, y) \wedge P(x, y))$ is subquadratically reducible to property testing of form $\exists x \exists y P(x, y)$, where P is a property checkable in time linear to the sizes of x and y, and TC_E is the transitive closure of relation E. Furthermore, when P is definable by a first-order logic formula with at most one quantified variable, then the above two problems are equivalent to each other by subquadratic reductions.

1 Introduction

1.1 Extending one-dimensional dynamic programming to graphs

The Least Weight Subsequence (LWS) is type of dynamic programming problems introduced by [HL87]: select a set of elements from a linearly ordered set so that the total cost incurred by the adjacent pairs of elements is optimized. LWS is defined as follows: Given elements $x_0, \ldots, x_n$, and an $n \times n$ matrix C of costs $c_{i,j}$, for all pairs of indices $i < j$, compute F on all elements, defined by

$$F(j) = \begin{cases} 0, & \text{for } j = 0 \\ \min_{0 \le i < j}[F(i) + c_{i,j}], & \text{for } j = 1, \ldots, n \end{cases}$$

$F(j)$ is the optimal cost value from the first element up to the j-th element. The Airplane Refueling problem [HL87] is a well known example of LWS: Given the locations of airports on a line, find

*Work supported by a Simons Investigator Award from the Simons Foundation.

ISSN 1433-8092

a subset of the airports for an airplane to add fuel, that minimizes the sum of the cost. The cost of flying from the i-th to the j-th airport is defined by $c_{i,j}$. Other LWS examples include finding a longest chain satisfying some property, such as Longest Increasing Subsequence [Fre75] and Longest Subset Chain [KPS17]; breaking a linear structure into blocks, such as Pretty Printing [KP81]; variations of Subset Sum such as the Coin Change problem, and the Knapsack problem. These problems have $O(n^2)$ time algorithms using dynamic programming, and in many special cases it can be improved: when the cost satisfies quadrangle inequality or some other properties, there are near linear time algorithms (e.g. [Yao80, Wil88, GP89]). But for the general LWS, it is not known whether these problems can be solved faster than $n^{2-o(1)}$ time.

A general approach to understanding the fine-grained complexity of these problems was initiated in [KPS17]. Many LWS problems have succinct representations of $c_{i,j}$. Taking problems defined in [KPS17] as examples, in LowRankLWS, $c_{i,j} = \langle \mu_i, \sigma_j \rangle$, where μ_i and σ_j are boolean vectors of length $d \ll n$ associated with each element. The ChainLWS problem has costs $c_1, \ldots, c_n$ defined a property P so that $c_{i,j}$ equals c_j if $P(i,j)$ is true, and ∞ otherwise. P is computable by data associated with the pair (i,j). (For example, in LongestSubsetChain, $P(i,j)$ is true iff set S_i is contained in set S_j.) So the goal of the problem becomes finding a longest chain of elements so that adjacent elements satisfy property P. When C can be represented succinctly, we can ask whether there exist subquadratic time algorithms for these problems, or try to find subquadratic time reductions between problems. [KPS17] showed that in many LWS problems when C can be succinctly described in the input, in subquadratic time it is reducible to a corresponding problem, which is called a StaticLWS problem. The problem StaticLWS is: given elements $x_1, \ldots, x_{2n}$, a cost matrix C, and values $F(i)$ on all $i \in \{1, \ldots, n\}$, compute $F(j) = \min_{i \in \{1,\ldots,n\}}[F(i) + c_{i,j}]$ for all $j \in \{n+1, \ldots, 2n\}$. It is a parallel, batch version (with many values of j rather than a single one) of the update rule applied sequentially one index at a time in the standard DP algorithm. The reduction from LWS to StaticLWS implies that a highly sequential problem can be reducible to a highly parallel one. If a StaticLWS problem can be solved faster than quadratic time, so can the LWS problem. Apart from one-directional reductions from LWS to StaticLWS, [KPS17] also proved subquadratic time equivalence between some concrete problems: LowRankLWS is equivalent to MinInnerProduct, NestedBoxes is equivalent to VectorDomination, LongestSubsetChain is equivalent to OrthogonalVectors, and ChainLWS is equivalent to Selection.

Some of the LWS problems can be naturally extended from lines to DAGs. For example, on a road map, we wish to find a path for a vehicle, along which we wish to find a sequence of cities where the vehicle can rest and add fuel so that the cost is minimized. The cost of traveling between cities x and y is defined by cost $c_{x,y}$. Connections between cities could be a general graph, not just a line. Works about algorithms for LWS problems on graphs include [AST94, Sch98, CWHL11, LjLW12].

Using a similar approach as [KPS17], this paper extend the Least Weight Subsequence problems to the Least Weight Subpath (LWSP) problem whose objective is to find a least weight subsequence on a path of a given directed acyclic graph $G = (V, E)$. Let there be a set V_0 containing vertices that can be the starting point of a sequence. The optimum value on each vertex is defined by:

$$F(v) = \begin{cases} \min(0, \min_{u \rightsquigarrow v}[F(u) + c_{u,v}]), \text{ for } v \in V_0 \\ \min_{u \rightsquigarrow v}[F(u) + c_{u,v}], \text{ for } v \notin v_0 \end{cases}$$

where $u \rightsquigarrow v$ means u is reachable to v. The goal of LWSP is to compute $F(v)$ on all vertices $v \in V$. Examples of LWSP problems will be given in Appendix A. LWSP can be solved in time $O(|V| \cdot |E|)$ by doing reversed depth/breadth first search from each vertex, and update the F value on the vertex accordingly. It is not known whether it has faster algorithms, even for Longest Increasing Subsequence, which is an LWS instance solvable in $O(n \log n)$ time on linear structures. If C is succinctly describable in similar ways as LowRankLWS, NestedBoxes, SubsetChain or ChainLWS, we wish to study if there are subquadratic time algorithms or subquadratic time reductions between problems.

In this paper we will always reduce between problems with the same C. Therefore we omit the C from the input parameter list of LWS, StaticLWS and LWSP. Because the matrix C is either fixed in the problem or given succinctly in the input, we consider that every vertex has some additional data so that $c_{x,y}$ can be computed by the data contained in x and y. Let the size of additional data associated with each vertex v be its weight $w(v)$. The weight of a vertex can be defined in different ways according to the problems. For example, in LowRankLWS, the weight of an element can be defined as the length of its associated vector; and in SubsetChain, the weight of an element is the size of its corresponding subset. We use $m = |E|$ as the number of graph edges. Let n be the number of vertices. Let the total weight of all vertices be N. We use $M = \max(m, N)$ as the size of the input. In this paper we will see that if we can improve the algorithm for StaticLWS to be subquadratic time, then on some interesting classes of graphs we can solve LWSP faster than $M^{2-o(1)}$ time.

1.2 Fine-grained complexity preliminaries

Fine-grained complexity studies the exact-time reductions between problems, and the completeness of problems in classes under exact-time reductions. These reductions have established conditional lower bounds for many interesting problems. The Orthogonal Vectors problem (OV) is a well-studied problem solvable in quadratic time. If the *Strong Exponential Time Hypothesis (SETH)* [IP01, IPZ01] is true, then OV does not have truly-subquadratic time algorithms[Wil05]. The problem OV is defined as follows: Given n boolean vectors of dimension $d = \omega(\log n)$, and decide whether there is a pair of vectors whose inner product is zero. The best algorithm is in time $n^{k-\Omega(1/\log(d/\log n))}$ [AWY15, CW16]. The *Moderate-dimension OV conjecture (MDOVC)* states that for all $\epsilon > 0$, there is no $O(n^{2-\epsilon}\mathsf{poly}(d))$ time algorithm that solves OV with dimension d. If this conjecture is true, then many interesting problems would get conditional lower bounds, including dynamic programming problems such as Longest Common Subsequence [ABW15, BK15], Edit Distance [BI15, AHWW16], Fréchet distance [Bri14, BK17, BM15], Local Alignment [AWW14], CFG Parsing and RNA Folding [ABBK17], Regular Expression Matching [BI16, BGL17] and Subset Sum [ABHS19], and also many graph problems [RVW13, AWW16, BRS$^+$18]. There are also conditional hardness results about graph problems based on the hardness of All Pair Shortest Path [WW10, AGW14, AR16, LWW18].

The *fine-grained reduction* was introduced in [WW10], which can preserve polynomial saving factors in the running time between problems. The statements for fine-grained complexity are usually like this: if there is some $\epsilon_2 > 0$ such that problem Π_2 is in $\mathsf{TIME}((T_2(m))^{1-\epsilon_2})$, then problem Π_1 is in $\mathsf{TIME}((T_1(m))^{1-\epsilon_1})$ for some ϵ_1. If T_1 and T_2 are both $O(m^2)$ then this reduction is called a subquadratic reduction. Furthermore, the *exact-complexity reduction* is a more strict version that can preserve sub-polynomial savings factors between problems. We use $(\Pi_1, T_1(m)) \leq_{\text{EC}} (\Pi_2, T_2(m))$ to denote that if problem Π_2 is in $\mathsf{TIME}(T_2(m))$, then problem Π_1 is in $\mathsf{TIME}(T_1(m))$.

1.3 Introducing reachability to first-order model checking

Introducing reachability to first-order property problems is analogous to extending LWS to paths in graphs, which makes parallel problems become sequential. The first-order property (or first-order model checking) problem is to decide whether an input structure satisfies a fixed first-order logic formula φ. Even if model checking for input formulas is PSPACE-complete [Sto74, Var82], when φ

is fixed by the problem, it is solvable in polynomial time. The sparse version of OV is one of these problems, defined by formula $\exists u \exists v \forall i \in [d](\neg One(u,i) \vee (\neg One(v,i)))$, where relation $One(u,i)$ is true iff the i-th coordinate of vector u is one.

If φ has k quantifiers ($k \geq 2$), then on input structures of n elements and m tuples of relations, it can be solved in time $O(n^{k-2}m)$ [GIKW17]. On dense graphs where $k \geq 9$, it can be solved in time $O(n^{k-3+\omega})$, where ω is the matrix multiplication exponent [Wil14]. Here we study the case where the input structure is sparse, i.e. $m = n^{1+o(1)}$, and ask whether a three-quantifier first-order formula can be model checked in time faster than $m^{2-o(1)}$. The *first-order property conjecture (FOPC)* states that there exists integer $k \geq 2$, so that first-order model checking for $(k+1)$-quantifier formulas cannot be solved in time $O(m^{k-\epsilon})$ for any $\epsilon > 0$. This conjecture is equivalent to MDOVC, since OV is proven to be a complete problem in the class of first-order model checking problems; in other words, any model checking of 3 quantifier formulas on sparse graphs is subquadratic time reducible to OV [GIKW17]. This means from improved algorithms for OV we can get improved algorithms for first-order model checking.

The first-order property problems are highly parallelizable. If we introduce the transitive closure (TC) operation on the relations, then these problems will become sequential. The transitive closure of a binary relation E can be considered as the reachability relation by edges of E in a graph. In a sparse structure, the TC of a relation may be dense. So it can be considered as a dense relation succinctly described in the input. In finite model theory, adding transitive closure significantly adds to the expressive power of first-order logic (First discovered by Fagin in 1974 according to [Lib13], and then re-discovered by [AU79].) In fine-grained complexity, adding arbitrary transitive closure operations on the formulas strictly increases the hardness of the model checking problem. More precisely, [GI19] shows that The SETH on constant depth circuits, which is a weaker conjecture than the SETH on k-CNF-SAT, implies the model checking for two-quantifier first-order formulas with transitive closure operations cannot be solved in time $O(m^{2-\epsilon})$ for any $\epsilon > 0$. This means this problem may stay hard even if the SETH on k-CNF-SAT is refuted.

However, we will see that for a class of three-quantifier formulas with transitive closure, model checking is no harder than OV under subquadratic time reductions.

We define problem $\mathsf{Selection}_P$ to be the property testing problem for $(\exists x \in X)(\exists y \in Y)P(x,y)$. $P(x,y)$ is a fixed property specified by the problem that can be decided in time $O(w(x) + w(y))$, where $w(x)$ is the size of additional data on element x. For example, OV is $\mathsf{Selection}_P$ where $P(x,y)$ iff x and y are a pair of orthogonal vectors. In this case $w(x)$ is defined as the length of vector x. (If we work on the sparse version of OV, the weight $w(x)$ is defined by the Hamming weight of x.)

On a directed graph $G = (V,E)$, we define Path_P to be the problem of deciding whether $(\exists x \in V)(\exists y \in V)[\mathrm{TC}_E(x,y) \wedge P(x,y)]$, where TC_E is the transitive closure of relation E and $P(x,y)$ is a property on x,y fixed by the problem. That is, whether there exist two vertices x,y not only satisfying property P but also x is reachable to y. We will give an example of Path_P in Appendix A. Also, we define $\mathsf{ListPath}_P$ to be the problem of listing all $x \in V$ such that $(\exists y \in V)[\mathrm{TC}_E(x,y) \wedge P(x,y)]$.

Considering the model checking problems, we let *PathFO$_3$* and *ListPathFO$_3$* denote the class of Path_P and $\mathsf{ListPath}_P$ such that P is of form $\exists z \psi(x,y,z)$ or $\forall z \psi(x,y,z)$, where ψ is a quantifier-free logical formula. Later we will see that problems in *PathFO$_3$* and *ListPathFO$_3$* are no harder than OV. In these model checking problems, the weight of an element is the number of tuples in the structure that the element is contained in.

Trivially, $\mathsf{Selection}_P$ on input size $|X| = N_1, |Y| = N_2$ can be decided in time $O(N_1 N_2)$. Path_P and $\mathsf{ListPath}_P$ on input size M and total vertex weight N are solvable time $O(MN)$ by depth/breadth first search from each vertex. In this paper we will show that on some graphs, if $\mathsf{Selection}_P$ is in truly subquadratic time, so is Path_P and $\mathsf{ListPath}_P$. Interestingly, by applying the same reduction techniques from Path_P to $\mathsf{Selection}_P$, we can get a similar reduction from a dynamic programming problem on a graph to a static problem.

1.4 Main results

This paper works on two classes of graphs, both having some similarities to trees. The first class is where the graph G is a multitree. A *multitree* is a directed acyclic graph where the set of vertices reachable from any vertex form a tree. Or equivalently a DAG is a multitree if and only if on all pairs of vertices u, v, there is at most one path from u to v. In different contexts, multitrees are also called *strongly unambiguous graphs*, *mangroves* or *diamond-free posets* [GLL12]. These graphs can be used to model computational paths in nondeterministic algorithms where there is at most one path connecting any two states [AL96].

The second class of graphs is when we treat G as undirected by replacing all directed edges by undirected edges, the underlying graph has constant treewidth. *Treewidth* [RS86] is an important parameter of graphs that describes how similar they are to trees. If a graph has constant treewidth, it is very similar to a tree. On these classes of graphs, we have the following theorems.

Theorem 1 (Reductions between decision problems.). *Let $t(M) \geq 2^{\Omega(\sqrt{\log M})}$, and let the DAG $G = (V, E)$ satisfy one of the following conditions:*

- *G is a multitree or a multitree of strongly connected components, or*
- *The underlying undirected graph of G has constant treewidth,*

then, the following statements are true:

1. *If $\mathsf{Selection}_P$ is in time $N_1 N_2 / t(\min(N_1, N_2))$, then Path_P is in time $M^2/t(\mathsf{poly}M)$.*[1]
2. *If Path_P is in time $M^2/t(M)$, then $\mathsf{ListPath}_P$ is in time $M^2/t(\mathsf{poly}M)$.*
3. *When $P(x, y)$ is of form $\exists z \psi(x, y, z)$ or $\forall z \psi(x, y, z)$ where ψ is a quantifier-free first-order formula, $\mathsf{Selection}_P$ is in time $N_1 N_2 / t(\min(N_1, N_2))$ iff Path_P is in time $M^2/t(\mathsf{poly}M)$ iff $\mathsf{ListPath}_P$ is in time $M^2/t(\mathsf{poly}M)$.*

This theorem implies that OV is hard for classes *PathFO$_3$* and *ListPathFO$_3$*. By the improved algorithm for OV [AWY15, CW16], we get improved algorithms for *PathFO$_3$* and *ListPathFO$_3$*:

Corollary 1.1 (Improved algorithms.). *Let the graph G be a multitree, or multitree of strongly connected components, or a DAG whose underlying undirected graph has constant treewidth. Then PathFO$_3$ and ListPathFO$_3$ are in time $M^2/2^{\Omega(\sqrt{\log M})}$.*

Next, we consider the dynamic programming problems. If the cost matrix C in LWSP is succinctly describable, we get the following reduction from LWSP to $\mathsf{StaticLWS}$. Except for the reduction from $\mathsf{LongestSubsetChain}$ to OV, we always assume that all vertices have the same weight.[2]

Theorem 2 (Reductions between optimization problems.). *On a multitree graph, or a DAG whose underlying undirected graph has constant treewidth, let $t(N) \geq 2^{\Omega(\sqrt{\log N})}$, then,*

1. *if $\mathsf{StaticLWS}$ of total weight N is in time $N^2/t(N)$, then LWSP on input size M is in time $M^2/t(\mathsf{poly}(M))$.*
2. *if LWSP is in time $M^2/t(M)$, then LWS on input size N is in time $N^2/t(\mathsf{poly}(N))$.*

[1] This reduction also applies to optimization versions of these two problems. Let Path_F be a problem to compute $\min_{x,y \in V, x \rightsquigarrow y} F(x, y)$ and $\mathsf{Selection}_F$ be a problem to compute $\min_{x \in X, y \in Y} F(x, y)$, where F is a function on x, y, instead of a boolean property. Then the same technique gives us a reduction from Path_F to $\mathsf{Selection}_F$. We will leave the details to the full version of the paper.

[2] In $\mathsf{LongestSubsetChain}$, the subsets corresponding to different vertices can have different sizes.

If there is a reduction from a concrete StaticLWS problem to its corresponding LWS problem (e.g. from MinInnerProduct to LowRankLWS, from VectorDomination to NestedBoxes and from OV LongestSubsetChain [KPS17]), then the corresponding LWS, StaticLWS and LWSP problems are subquadratic-time equivalent.

Finally, because VectorDomination is equivalent to OV, we get improved algorithm for problem LongestSubsetChain:

Corollary 1.2 (Improved algorithm)**.** *On a sparse multitree graph or a DAG whose underlying undirected graph has constant treewidth,* LongestSubsetChain *is in time* $M^2/2^{\Omega(\sqrt{\log M})}$*.*

1.5 List of problem definitions

Here we list the main problems in the paper. Some problems will be used later in the proofs.

LWS**:** Given elements $x_1, \dots, x_n$ and value $F(0) = 0$, compute $F(j) = \min_{0 \le i < j}[F(i) + c_{i,j}]$ for all $j \in \{1, \dots, 2n\}$.

StaticLWS**:** Given elements $x_1, \dots, x_{2n}$ and values of $F(i)$ on all $i \in \{1, \dots, n\}$, compute $F(j) = \min_{i \in \{1,\dots,n\}}[F(i) + c_{i,j}]$ for all $j \in \{n+1, \dots, 2n\}$.

LWSP**:** Given DAG $G = (V, E)$ and starting vertex set $V_0 \subseteq V$, compute on each $v \in V$, the value of $F(v)$, where

$$F(v) = \begin{cases} \min(0, \min_{u \rightsquigarrow v}[F(u) + c_{u,v}]), & \text{for } v \in V_0 \\ \min_{u \rightsquigarrow v}[F(u) + c_{u,v}], & \text{for } v \notin v_0 \end{cases}$$

CutLWSP**:** On DAG G with a cut (S, T) where edges only direct from S to T, given the values of function F_S on S, for all $t \in T$ compute $F_T(t) = \min_{s \in S, s \rightsquigarrow t}[F_S(s) + c_{s,t}]$.

Selection$_P$**:** On two sets X, Y, decide whether $(\exists x \in X)(\exists y \in Y)P(x, y)$.

Path$_P$**:** On graph $G = (V, E)$, decide whether $(\exists x \in V)(\exists y \in V)[\mathrm{TC}_E(x, y) \wedge P(x, y)]$.

ListPath$_P$**:** On graph $G = (V, E)$, for all $x \in V$, decide whether $(\exists y \in V)[\mathrm{TC}_E(x, y) \wedge P(x, y)]$.

CutPath$_P$**:** On graph $G = (V, E)$ with cut (S, T) where edges only direct from S to T, decide whether $(\exists x \in S)(\exists y \in T)[\mathrm{TC}_E(x, y) \wedge P(x, y)]$.

1.6 Organization

In Section 2 we prove the first part of Theorem 1, by reduction from Path$_P$ to Selection$_P$. Here we only show the reduction algorithm on multitrees. For DAGs of constant treewidth, the proof will be presented in Section 2.4. Section 3 proves Theorem 2 by showing a reduction from LWSP to StaticLWS. Section 4 proves the second part of Theorem 1 by reduction from ListPath$_P$ to Path$_P$. Section 5 proves the last part of Theorem 1, the subquadratic equivalence of Selection$_P$, Path$_P$ and ListPath$_P$ when P is a first-order property. In Section 6 we talk about open problems. Appendix A shows some problems as examples of LWSP and Path$_P$.

2 From sequential problems to parallel problems

This section will establish the first part of Theorem 1 by showing that if $t(M) \ge 2^{\Omega(\sqrt{\log M})}$, then $(\mathsf{Path}_P, M^2/t(\mathsf{poly} M)) \le_{\mathrm{EC}} (\mathsf{Selection}_P, N_1 N_2 / t(\min(N_1, N_2)))$. We will give the reduction for multitrees and multitrees of strongly connected components. For constant treewidth graphs, the reduction will be left to Appendix 2.4.

2.1 The recursive algorithm

In the algorithm we first remove high degree vertices, then follow a divide and conquer strategy. In the whole process of the reduction from Path_P to $\mathsf{Selection}_P$, for simplicity of description, we will consider each strongly connected component as a single vertex, whose weight equals the total weight of the component. In the following algorithm, whenever querying $\mathsf{Selection}_P$ or exhaustively enumerating pairs of reachable vertices and testing P on them, we will extract all the vertices from a component. Testing P on a pair of vertices (or strongly connected components) of weights N_1, N_2 is in time $O(N_1N_2)$. We use "vertex" to express "vertex or strongly connected component" in the following argument.

Let $\mathsf{CutPath}_P$ be a variation of Path_P. It is the property testing problem for $(\exists x \in S)(\exists y \in T)[TC_E(x,y) \wedge \varphi(x,y)]$, where (S,T) is a cut in the graph, such that all the edges between S and T are directed from S to T. $\mathsf{CutPath}_P$ on input size M and total vertex weight N can be solved in time $O(MN)$ if $P(x,y)$ is decidable in time $O(w(x)+w(y))$: start from each vertex and do depth/breadth first search, and on each pair of reachable vertices decide if P is satisfied.

Lemma 2.1. *For $t(M) \geq 2^{\Omega(\sqrt{\log M})}$, if $\mathsf{Selection}_P(N_1,N_2)$ is in time $N_1N_2/t(\min(N_1,N_2))$ and $\mathsf{CutPath}_P(M)$ is in time $M^2/t(M)$, then $\mathsf{Path}_P(M)$ is in time $M^2/t(\mathsf{poly}(M))$.*

Proof. Let γ be a constant satisfying $0 < \gamma \leq 1/4$. Let $T_\Pi(M)$ be the running time of problem Π on a structure of size M, and let $T_{\mathsf{Selection}_P}(N_1,N_2)$ be the running time of $\mathsf{Selection}_P$ on a pair of sets (X,Y) where the total vertex weight of X is N_1 and of Y is N_2.

We show that there exists a constant c where $0 < c < 1$ so that if $T_{\mathsf{Selection}_P}(N_1,N_2) \leq N_1N_2/t(\min(N_1,N_2))$ and $T_{\mathsf{CutPath}_P}(M) \leq M^2/t(M)$, and $T_{\mathsf{Path}_P}(M')$ is at most $M'^2/t(M'^c)$ for all $M' \leq M$, then $T_{\mathsf{Path}_P}(M) \leq M^2/t(M^c)$. We run the recursive algorithm as shown in Algorithm 1. The intuition is to divide the graph into a cut S,T, recursively compute Path_P on S and T, and deal with paths from S to T. For large-weight vertices, we deal with them separately so that $\mathsf{CutPath}_P$ will not deal with large-weight vertices.

For the vertices of weight more than M^γ, we deal with them separately before the recursive calls. If a vertex has more than M^γ ancestors/descendants, and if $\mathsf{Selection}_P$ on size (N_1,N_2) is in time $O(N_1N_2/t(\min N_1,N_2))$, then the time to deal with a vertex of weight N_i is at most $O(MN_i/t(N_i)) \leq O(MN_i/t(M^\gamma))$. Because all N_i sum to at most M, the total time is $O(M^2/t(M^\gamma))$. If the vertex it has less than M^γ ancestors/descendants, then the exhaustive search time on all such v and all their ancestors/descendants should sum to at most $O(M \cdot M^\gamma)$. After the computation, the vertex becomes an "auxiliary" vertex. In the upcoming steps we will only use auxiliary vertices as intermediate points in the path, but will not include them in calls to $\mathsf{Selection}_P$ or treat them as potential endpoints of a path and check P on them. This can be done by keeping a list of vertices to be ignored.

Let M_S and M_T be the sizes of sets S and T respectively. Assume $M_S \geq M_T$ and let $\Delta = M_S - M_T$. Then we have

$$\begin{aligned} T_{\mathsf{Path}_P}(M) &= T_{\mathsf{Path}_P}(M_S) + T_{\mathsf{Path}_P}(M_T) + T_{\mathsf{CutPath}_P}(M) + O(M^2/t(M^\gamma)) \\ &= T_{\mathsf{Path}_P}(M_T+\Delta) + T_{\mathsf{Path}_P}(M_T) + T_{\mathsf{CutPath}_P}(M) + O(M^2/t(M^\gamma)) \\ &\leq 2T_{\mathsf{Path}_P}(M/2+\Delta) + T_{\mathsf{CutPath}_P}(M) + O(M^2/t(M^\gamma)) \\ &= 2(M/2+\Delta)^2/t((M/2+\Delta)^c) + M^2/t(M) + O(M^2/t(M^\gamma)). \end{aligned}$$

Let d be the constant factor of term $O(M^2/t(M^\gamma)$. We can pick c to be small enough so that $dt(M^c)/t(M^\gamma) = \epsilon$. Thus the term $d \cdot M^2/t(M^\gamma) = \epsilon M^2/t(M^c)$. The term $M^2/t(M)$ is less than $M^2/t(M^\gamma)$, so it is also less than $\epsilon M^2/t(M^c)$. So the running time by the above formula yields to at

Algorithm 1: $\mathsf{Path}_P(G)$

// Reducing Path_P to $\mathsf{Selection}_P$ and $\mathsf{CutPath}_P$

if G has only one vertex **then return** false.
Let M be the size of the problem.
for each vertex v of weight $\geq M^\gamma$ **do**
  if v has at least M^γ ancestors **then**
    Compute $\mathsf{Selection}_P$ on the set of v's ancestors and v.
  else
    Exhaustively search all pairs of vertices on the set of v's ancestors and v, test P on all pairs. If P is true on any pair then **return** true.
  if v has at least M^γ descendants **then**
    Compute $\mathsf{Selection}_P$ on v and the set of v's descendants.
  else
    Exhaustively search all pairs of vertices on v the set of v's descendants, test P on all pairs. If P is true on any pair then **return** true.
  Replace v by an auxiliary vertex of weight 1.
Topological sort all vertices.
Keep adding vertices to S by topological order, until the total weight of S exceeds $M/2$. Let the rest of vertices be T.
Run Path_P on the subgraph induced by S.
Run $\mathsf{CutPath}_P(S, T)$.
Run Path_P on the subgraph induced by T.
if any one of the above three calls returns true **then return** true.

most $2(M/2+\Delta)^2/t((M/2+\Delta)^c)+2\epsilon M^2/t(M^c)$. Because the function $M^2/t(M^c)$ is monotonically increasing, the formula is upper bounded by $2(M/2+\Delta)^2/t((M/2+\Delta)^c)+2\epsilon(M/2+\Delta)^2/t((M/2+\Delta)^c) = (2+2\epsilon)(M/2+\Delta)^2/t((M/2+\Delta)^c)$. When $\Delta \leq M^\gamma$ for $\gamma < 1/4$ and when M is large enough, $M^\gamma \ll M$ so $M/2+\Delta = (1+o(1))M/2$. So $(2+2\epsilon)(M/2+\Delta)^2$ can be significantly less than M^2. Moreover, we can make $(2+2\epsilon)(M/2+\Delta)^2/M^2$ less than $t((M/2+\Delta)^c)/t(M^c)$ because t grows very slow. Thus we get $(2+2\epsilon)(M/2+\Delta)^2/t((M/2+\Delta)^c) \leq M^2/t(M^c)$. □

2.2 A special case that can be exhaustively searched

The following lemma shows that if no vertex has both a lot of ancestors and a lot of descendants, then the total number of reachable pairs of vertices is subquadratic to m. This lemma holds for any DAG, not just for multitrees. We will use this lemma in the next subsection to show that in a subgraph where all vertices have few ancestors and descendants, we can test property P on all pairs of reachable vertices by brute force.

Lemma 2.2. *If in a DAG $G = (V, E)$ of m edges, every vertex has either at most n_1 ancestors or at most n_2 descendants, then there are at most $(m \cdot n_1 \cdot n_2)$ pairs of vertices s, t such that s is reachable to t.*

In a DAG $G = (V, E)$ of m edges, let S, T be two disjoint sets of vertices where edges between S and T only direct from S to T. If every vertex has either at most n_1 ancestors in S or at most n_2 descendants in T, then there are at most $(m \cdot n_1 \cdot n_2)$ pairs of vertices $s \in S$ and $t \in T$ such that s is reachable to t.

Proof. We define the ancestors of an edge $e \in E$ to be the ancestors (or ancestors in S) of its incoming vertex, and its descendants to be the descendants (or descendants in T) of its outgoing vertex. Let the number of its ancestors and descendants be denoted by $anc(e)$ and $des(e)$ respectively.

For each edge e, it belongs to exactly one of the following three types:

Type A: If $anc(e) \le n_1$ but $des(e) > n_2$, then let $count(e)$ be $anc(e)$.

Type B: If $des(e) \le n_2$ but $anc(e) > n_1$, then let $count(e)$ be $des(e)$.

Type C: If $anc(e) \le n_1$ and $des(e) \le n_2$, then let $count(e)$ be $anc(e) \cdot des(e)$.

$\sum_{e \in E} count(e) \le m \cdot n_1 \cdot n_2$ because the *count* value on each edge is bounded by $n_1 \cdot n_2$. We will prove that this value upper bounds the number of reachable pairs of vertices.

For each pair of reachable vertices (u, v) (or (u, v) s.t. $u \in S$ and $v \in T$), let $(e_1, \dots, e_p)$ be the path from u to v. Along the path, *anc* does not decrease, and *dec* does not increase. A path belongs to exactly one of the following three types:

Type a: Along the path $anc(e_1) \le anc(e_2) \le \cdots \le anc(e_p) \le n_1$, and $des(e_1) \ge des(e_2) \ge \cdots \ge des(e_p) > n_2$. That is, all the edges are Type A.

Type b: Along the path $des(e_p) \le des(e_{p-1}) \le \cdots \le des(e_1) \le n_2$, and $anc(e_p) \ge anc(e_{p-1}) \ge \cdots \ge anc(e_1) > n_1$. That is, all the edges are Type B.

Type c: Along the path there is some edge e_i so that $anc(e_i) \le n_1$ and $des(e_i) \le n_2$. That is, it has at least one Type C edge.

There will not be other cases, for otherwise if a Type A edge directly connects to a Type B edge without a Type C edge in the middle, then the vertex joining these two edges would have more than n_1 ancestors and more than n_2 descendants.

If a path from u to v is Type a, then its last edge e_p is Type A. If it is Type b, then its first edge e_1 is Type B. If it is Type c, then there is some edge e_i in the path that is Type C. This means,

1. For each Type A edge e, $count(e)$ is at least the number of all Type a pairs (u, v) whose path has e as its last edge.
2. For each Type B edge e, $count(e)$ is at least the number of all Type b pairs (u, v) whose path has e as its first edge.
3. For each Type C edge e, $count(e)$ is at least the number of all Type c pairs (u, v) whose path contains e.

Therefore each path is counted at least once by the $count(e)$ of some edge e. ☐

2.3 Subroutine: reachability across a cut

Now we will show the reduction from $\mathsf{CutPath}_P$ to $\mathsf{Selection}_P$. The high level idea of $\mathsf{CutPath}_P$ is that we think of the reachability relation on $S \times T$ as an $|S| \times |T|$ boolean matrix whose one-entries correspond to reachable pairs of vertices. If we could partition the matrix into all-one combinatorial rectangles, then we can decide all entries within these rectangles by a query to $\mathsf{Selection}_P$, because in the same rectangle, all pairs are reachable.

Claim 2.1. *Consider the reachability matrix of on sets S and T. Let M_S and M_T be the sizes of S and T. If there is a way to partition the matrix into non-overlapping combinatorial rectangles $(S_1, T_1), \dots, (S_k, T_k)$ of sizes $(r_1, c_1), \dots, (r_k, c_k)$, and if there is some t so that computing each subproblem of size (r_i, c_i) takes time $r_i \cdot c_i / t(\min(r_i, c_i))$, and each r_i and c_i are at least ℓ, then all the computation takes total time $O(M_S \cdot M_T / t(\ell))$.*

Proof. Let the minimum of all r_i be r_{min} and the minimum of all c_i be c_{min}. Then the factor of time saved for computing each combinatorial rectangle is at least $t(\min(r_{min}, c_{min}))$, greater than $t(\ell)$. So the time spent on all rectangles is at most $O((\sum_{i=1}^{t} c_i)(\sum_{i=1}^{t} r_i)/t(\ell))$, also we have

Algorithm 2: $\mathsf{CutPath}_P(S,T)$ on a multitree

Count the number of ancestors $anc(v)$ and descendants $des(v)$ for all vertices.
Insert all vertices with at least M^α ancestors and M^α descendants into linked list L.
while there exists a vertex $v \in L$ **do**
    `// we call` v `a` ***pivot vertex***
    Let A be the set of ancestors of v in S.
    Let B be the set of descendants of v in T.
    Add v to A if $v \in S$, otherwise add v to B.
    Run $\mathsf{Selection}_P$ on (A,B). If it returns true then **return** true.
    for each $a \in A$ **do**
        let $des(a) = des(a) - |B|$.
        if $des(a) < M^\alpha$ and $a \in L$ **then** remove a from L.
    for each $b \in B$ **do**
        let $anc(b) = anc(b) - |A|$.
        if $anc(b) < M^\alpha$ and $b \in L$ **then** remove b from L.
    Remove v from the graph.
for each edge (s,t) crossing the cut(S,T) **do**
    Let A be the set of ancestors of s (including s) in S.
    Let B be the set of descendants of t (including t) in T.
    On all pairs of vertices (a,b) where $a \in A, b \in B$, check property P. If P is true on any pair of (a,b) then **return** true.

$(\sum_{i=1}^t c_i)(\sum_{i=1}^t r_i) \le M_S \times M_T$ because the rectangles are contained inside the matrix of size $M_S \times M_T$ and they do not overlap. So the total time is $O(M_S \cdot M_T / t(\ell))$. ☐

The algorithm $\mathsf{CutPath}_P(S,T)$ is shown in Algorithm 2. It tries to cover the one-entries of the reachability matrix by combinatorial rectangles as many as possible. Finally, for the one-entries not covered, we go through them by exhaustive search, which takes less than quadratic time.

In the beginning, we can count the number of ancestors (or descendants) of all vertices in the DAG in $O(M)$ time by going through all vertices by topological order (or reversed topological order).

In each query to $\mathsf{Selection}_P(A,B)$, all vertices in A are reachable to all vertices in B, because they all go through v. For any pair of reachable vertices $s \in S, t \in T$, if they go through any pivot vertex, then the pair is queried to $\mathsf{Selection}_P$. Otherwise it is left to the end, and checked by exhaustive search on all pairs of reachable vertices.

The calls to $\mathsf{Selection}_P$ correspond to non-overlapping all-one combinatorial rectangles in the reachability matrix. For each call to $\mathsf{Selection}_P$, the rectangle size is at least $M^\alpha \times M^\alpha$. Thus the total time for all the ∃∃P calls is $O(M^2/t(M^\alpha))$ by Claim 2.1.

Each time we remove a pivot vertex v, there will be no more paths from set A to set B, for otherwise there would be two distinct paths connecting the same pair of vertices. Thus, removing a v decreases the total number of pairs of reachable vertices by at least $M^\alpha \cdot M^\alpha = M^{2\alpha}$. There are M^2 pairs of vertices, so the total number of pivot vertices v is at most $M^2/M^{2\alpha} = M^{2-2\alpha}$.

Each time we find a pivot vertex v, we update the number of descendants for all its ancestors, and update the number of ancestors for all its descendants. Because it has at least M^α ancestors and M^α descendants, the value decrease on each affected vertex is at least M^α. So each vertex has decreased its ancestors/descendants values for at most $M/M^\alpha = M^{1-\alpha}$ times. In other words, each vertex can be an ancestor/descendant of at most $M^{1-\alpha}$ pivot vertices. The total time to deal with all ancestors/descendants of all pivot vertices in the while loop is in $O(M \cdot M^{1-\alpha}) = O(M^{2-\alpha})$.

Finally, after the while loop, there are no vertices with both more than M^α ancestors and M^α descendants. In this case, by Lemma 2.2 in Section 2.2, the total number of reachable vertices is bounded by $M \cdot M^\alpha \cdot M^\alpha = M^{1+2\alpha}$. Each vertex has weight at most M^γ. So the total time to deal with these paths is $O(M^{1+2\alpha} \cdot M^\gamma \cdot M^\gamma) = O(M^{1+2\alpha+2\gamma})$.

Thus the total running time is $O(M^2/t(M^\alpha) + M^{2-\alpha} + M^{1+2\alpha+2\gamma})$. By choosing α and γ to be appropriate constants, we get subquadratic running time.

If $t(M) = M^\epsilon$, then by choosing $\alpha = \gamma = 1/(4+\epsilon)$, we get running time $M^{2-\epsilon/(4+\epsilon)}$.

2.4 $\mathsf{CutPath}_P$ for constant treewidth graphs

We prove the first part of Theorem 1 on DAGs whose underlying undirected graphs have constant treewidth. The algorithm Path_P for constant treewidth graphs is the same as the one for multitrees. In this section we will show the reduction algorithm $\mathsf{CutPath}_P$ for constant treewidth graphs on a cut (S, T).

Let $\mathcal{T}$ be the decomposition tree of a graph G. Recall that by the definition of tree decomposition, each node z of the tree corresponds to a set $\mathcal{B}(z)$ which is a subset of vertices of G. Because the treewidth is constant, each set $\mathcal{B}(z)$ has a constant number of vertices. Every vertex of G appears in at least one set of a tree node. Also, for every edge of G, there is at least one tree node whose set contains both its endpoints. And if a vertex v appears both in $\mathcal{B}(z_1)$ and $\mathcal{B}(z_2)$, then along the path from z_1 to z_2, v must appear in all the sets of the tree nodes. Here we consider the decomposition tree as rooted, where all edges are directed from the root to leaves.

We use a similar reduction idea as Section 2.3. In the decomposition tree, each time we find a node z to split the tree into two connected components. We first deal with all the paths that go through the vertices in $\mathcal{B}(z)$. Any other path in the graph must be completely contained in one of the connected components we have created. In the end, all connected components are so small that we can go through all pairs of reachable vertices by exhaustive search. The algorithm is defined in Algorithm 3.

The following claim uses a $1/3 - 2/3$ trick on trees:

Claim 2.2. *In a rooted tree of size n, we can find a connected subgraph of size between $(1/3)n$ and $(2/3)n$ in $O(n)$ time.*

Proof. For each node z in the tree, we will compute the size of the subtree rooted at z, denoted by $f(z)$. We compute $f(z)$ from the leaves up to the root, by a reversed topological order. If z is a leaf then let $size(z) \leftarrow 1$.

On each parent node p, we initially let $f(p) \leftarrow 1$, and then for each child c_i of p, add the value $f(c_i)$ to $f(c_i)$. If before we add the $f(c_i)$ of certain child c_i to $f(p)$, $f(p) < (1/3)n$, and after we add $f(c_i)$ to $f(p)$, $f(p) \geq (1/3)n$, then there are two cases:

If $f(p) \leq (2/3)n$, then the subgraph formed by p and its subtrees $c_1, \ldots, c_i$ is the connected subgraph we want.

If $f(p) > (2/3)n$, then it must be $f(c_i) \geq (2/3)n - (1/3)n = (1/3)n$. That is, the subtree rooted at c_i has size between $(1/3)n$ and $(2/3)n$. But then we should have already returned the subtree rooted c_i instead. So this case would not happen.

After we have added the sizes of all the children of p to $f(p)$, we have finished computing $f(p)$. If $f(p)$ is still less than $1/3$, we will continue to let the next vertex by the reversed topological order be the current parent. ☐

Algorithm 3: $\mathsf{CutPath}_P(S, T)$ on constant treewidth DAG

Compute $\mathcal{T}$, the tree decomposition of the underlying undirected graph.
for each z in $\mathcal{T}$ **do**
    Let $size(z)$ be the number of nodes of $\mathcal{T}$.
while there exists a tree node z in $\mathcal{T}$ so that there is a connected subgraph of $\mathcal{T}$ rooted at z with size between $(1/3)size(z)$ and $(2/3)size(z)$ **do**
    `//` z `can be found in time` $O(size(z))$ `by Claim 2.2.`
    for each $v \in \mathcal{B}(z)$ **do**
        `// Deal with all paths going through` v`.`
        Let A be the set of ancestors of v in S.
        Let B be the set of descendants of v in T.
        Add v to A if $v \in S$, otherwise add v to B.
        if both A and B have at least M^α vertices **then**
            Run $\mathsf{Selection}_P$ on (A, B). If it returns true then **return** true.
        else
            Exhaustively check P on all pairs of $a \in A$ and $b \in B$. If P is true on any (a, b) then **return** true.
        Remove v from the graph, and from the sets of all the tree nodes.
    Remove z from $\mathcal{T}$.
    for each tree node z' who was originally in the same connected component with z **do**
        Update $size(z')$ to be the new size of the connected component z' is in.
for each edge (s, t) crossing the cut(S, T), **do**
    Let A be the set of ancestors of s (including s) in S.
    Let B be the set of descendants of t (including t) in T.
    On all pairs of vertices (a, b) where $a \in A, b \in B$, check property P. If P is true on any pair of (a, b) then **return** true.

Next we will analyze the reduction algorithm. First, if a the treewidth of a graph is constant, then the corresponding decomposition tree can be computed in linear time [Bod96].

Unlike multitree graphs, here the calls to $\mathsf{Selection}_P$ are not non-overlapping rectangles: different v from the same $\mathcal{B}(z)$ may share the same ancestors or descendants. However, each time after removing a z, the connected components of the decomposition tree correspond to non-overlapping rectangles in the reachability matrix, and will not overlap with the rectangles corresponding to the ancestors and descendants for any $v \in \mathcal{B}(z)$. Thus, the overlapping only happens when dealing with the ancestors and descendants of different v from the same $\mathcal{B}(z)$, and these $\mathsf{Selection}_P$ rectangles will not overlap with other $\mathsf{Selection}_P$ rectangles after z is removed. Because in each non-overlapping rectangle corresponding to a connected component, we only computed the $\mathsf{Selection}_P$ for $|\mathcal{B}(z)|$ times, which is a constant. So by Claim 2.1, the total time spent on all the calls to $\mathsf{Selection}_P$ is still $O(M^2/t(M^\alpha))$.

When we remove all vertices $v \in \mathcal{B}(z)$, the graph vertices from sets of different connected components of the decomposition tree are not reachable to each other. Because any path from one connected component to another must go through some vertex in $\mathcal{B}(z)$.

Unlike multitree graphs, this time some vertex v in $\mathcal{B}(z)$ may have fewer than M^α ancestors or descendants. If so, then we do exhaustive search on the sets of v's ancestors and descendants, since calling $\mathsf{Selection}_P$ will not save time. Each time we find a v, the connected component of the decomposition tree that v belongs to loses at least $(1/3)\textit{size}(v)$ of its vertices, thus each vertex can be the ancestor/descendants of at most $O(\log_{3/2} M)$ such v's. There are at most M vertices in the graph, each of which can take part in at most M^α such paths going through each such v. So the total time is $O(M \cdot \log_{3/2} M \cdot M^\alpha) = O(M^{1+\alpha} \cdot \log_{3/2} M)$.

Also, because each vertex can be the ancestor/descendants of at most $O(\log_{3/2} M)$ such v's, the total time for updating *size* for all of them is also bounded by $O(M \cdot \log_{3/2} M)$.

In the end, each remaining vertex has $O(M^\alpha)$ ancestors and $O(M^\alpha)$ descendants. The total running time for the exhaustive search is $O(M \cdot M^\alpha \cdot M^\alpha \cdot M^{2\gamma}) = O(M^{1+2\alpha+2\gamma})$ by Lemma 2.2.

The overall running time is $O(M^2/t(M^\alpha) + M^{1+\alpha} \cdot \log_{3/2} M + M^{1+2\alpha+2\gamma})$. By choosing α and γ to be appropriate small constants, we get subquadratic running time.

3 Application to Least Weight Subsequence

In this section we will prove Theorem 2. The reduction from LWSP to $\mathsf{StaticLWS}$ uses the same structure as the reduction from Path_P to $\mathsf{Selection}_P$ in the proof of Theorem 1 shown in Section 2.

Process $\mathsf{LWSP}(G, F_0)$ computes values of F on initial values F_0 defined on all vertices of G. On a given LWSP problem, we will reduce it to an asymmetric variation of $\mathsf{StaticLWS}$. Process $\mathsf{StaticLWS}(A, B, F_A)$ computes all the values of function F_B defined on domain B, given all the values of F_A defined on domain A, such that $F_B(b) = \min_{a \in A}[F_A(s) + c_{a,b}]$. Let N_A and N_B be the total weight of A and B respectively. It is easy to see that if $\mathsf{StaticLWS}$ on $|N_A| = |N_B|$ is in time $N_A^2/t(N_A)$, then $\mathsf{StaticLWS}$ on general A, B is in time $O(N_A \cdot N_B/t(\min(N_A, N_B)))$.

We also define process $\mathsf{CutLWSP}(S, T, F_S)$, which computes all the values of F_T defined on domain T, given all the values of F_S on domain S, where $F_T(t) = \min_{s \in S, s \rightsquigarrow t}[F_S(s) + c_{s,t}]$.

The reduction algorithm is adapted from the reduction from Path_P to $\mathsf{Selection}_P$. LWSP is analogous to Path_P, $\mathsf{StaticLWS}$ is analogous to $\mathsf{Selection}_P$, and $\mathsf{CutLWSP}$ is analogous to $\mathsf{CutPath}_P$. In Path_P, we divide the graph into two halves, recursively call Path_P on the subgraphs, and use $\mathsf{CutPath}_P$ to deal with paths from one side of the graph to the other side. Similarly in LWSP, we divide the graph into two halves, recursively compute function F on the source side of the graph, then based on these values we call $\mathsf{CutPath}_P$ to compute the initial values of function F on the sink side of the graph, and finally we recursively call LWSP on the sink side of the graph. In $\mathsf{CutPath}_P$, we first identify large all-one rectangles in the reachability matrix, and then use $\mathsf{Selection}_P$ to solve them, and finally we go through all reachable pairs of vertices that are not covered by these rectangles. Similarly, in LWSP, we will use the similar method to identify large all-one rectangles in the reachability matrix and use $\mathsf{StaticLWS}$ to solve them, and finally we go through all reachable pairs of vertices and update F on each of them.

The algorithm LWSP is similar as Path_P, and is defined in Algorithm 4. Initially, we let $F(v) \leftarrow 0$ for all $v \in V_0$, and let $F(v) \leftarrow +\infty$ for all $v \notin V_0$. We run $\mathsf{LWSP}(G, F_0)$ on the whole graph. Here we only consider the case where all vertices have the same weight. (For $\mathsf{SubsetChain}$ the subset associated with each vertex can have different sizes. But by the universe-shrinking self reduction in [GIKW17] we can transform the universe of the sets to be as small as $2^{\Theta(\sqrt{\log n})}$ for problems with n subsets. By expressing the set using a vector of length equal to the size of the small universe, we will make all vertices have the same weight.)

The algorithm $\mathsf{CutLWSP}(S, T, F_S)$ is adapted from $\mathsf{CutPath}_P$, with the following changes:

1. In the beginning, $F_T(t)$ is initialized to ∞ for all $t \in T$.
2. Each query to $\mathsf{Selection}_P(A, B)$ in $\mathsf{CutPath}_P$ is replaced by
   (a) Compute F_B on domain B by $\mathsf{StaticLWS}(A, B, F_S)$.

Algorithm 4: $\mathsf{LWSP}(G = (V, E, V_0), F_0)$

if G has only one vertex v **then**
  if $v \in V_0$ **then**
    return $\min(0, F_0(v))$.
  return F_0 on v.
Let M be the size of the problem.
Topological sort all vertices.
Keep adding vertices to S by topological order, until the total weight of S exceeds $M/2$. Let the rest of vertices be T.
Compute F on domain S, by $F \leftarrow \mathsf{LWSP}(G_S, F_0)$, where G_S is the subgraph of G induced by S.
Let $F_T \leftarrow \mathsf{CutLWSP}(S, T, F)$.
For each $t \in T$, let $F_0(t) \leftarrow \min(F_0(t), F_T(t))$.
Compute F on domain T, by $F \leftarrow \mathsf{LWSP}(G_T, F_0)$, where G_T is the subgraph of G induced by T.
return F on domain V.

(b) For each vertex b in B, let $F_T(b)$ be the minimum of the original $F_T(b)$ and $F_B(b)$.

3. Whenever processing a pair of vertices s, t such that s is reachable to t in either the preprocessing phase or the final exhaustive search phase, we let $F_T(t) \leftarrow F_S(s) + c_{s,t}$ if $F_S(s) + c_{s,t} < F_T(t)$.
4. In the end, the process returns F_T, the target function on domain T.

The proof of correctness will be shown in Appendix **??**. The time complexity of this reduction algorithm follows from the argument of Section 2. Here because all vertices have the same weight and we are dealing with DAGs so there are no strongly connected components. And in Path_P, there will not be the term $M^2/t(M^\gamma)$. The rest of the time analysis is the same as Section 2.

Correctness of CutLWSP.

The correctness of CutLWSP follows from the correctness of $\mathsf{CutPath}_P$. We claim that after running $\mathsf{CutLWSP}(S, T, F_S)$, for any vertex $t \in T$, there is $F_T(t) = \min_{s \in S, s \rightsquigarrow t}[F_S(s) + c_{s,t}]$. Because for any pair $s \in S$, $t \in T$, such that s reachable to t, they are either processed in a query to $\mathsf{StaticLWS}(A, B)$ where $s \in A, t \in B$, or computed separately thus $F_T(t) \leftarrow \min(F_T(t), F(s) + c_{s,t})$.

Correctness of LWSP.

The LWSP algorithm has the following facts:

1. Whenever a process LWSP on domain $V_1 \subseteq V$ returns, the values of F on V_1 are fixed and will not be changed henceforth.

2. Whenever there is an edge from u to v, then the value of F on u is always fixed before the value on v. So the final values of function F on all vertices are fixed by topological order.

3. Each time we call LWSP on a subset of vertices $V_1 \subseteq V$, the F values on all ancestors of any vertex in V_1 that are not in V_1 have been fixed by some previous calls to LWSP.

Assume that when we call LWSP on subgraph with cut (S, T), initially there is

$$F_0(v) = \begin{cases} \min_{u \in R(v) \setminus (S \cup T), u \rightsquigarrow v}[F(u) + c_{u,v}], & \text{if } v \notin V_0 \\ \min(0, \min_{u \in R(v) \setminus (S \cup T), u \rightsquigarrow v}[F(u) + c_{u,v}]), & \text{if } v \in V_0 \end{cases} \tag{1}$$

where $R(v)$ is the set of vertices reachable to v. Then, if $\mathsf{LWSP}(S, F_0)$ is correct, after running $\mathsf{LWSP}(S, F_0)$, for any $s \in S \backslash V_0$, there is $F(s) = \min_{u \in R(s) \backslash T, u \rightsquigarrow s}[F(u) + c_{u,s}]$. And after running $\mathsf{CutLWSP}(S, T, F)$, we have $F_T(t) = \min_{s \in S, s \rightsquigarrow t}[F(s) + c_{s,t}]$. Then after taking $F_0(t) = \min(F_0(t), F_T(t))$ on all t, for any $t \in T \backslash V_0$, we get $F_0(t) = \min_{u \in R(t) \backslash T, u \rightsquigarrow t}[F(u) + c_{u,t}]$. Similarly for any $t \in T \cap V_0$, $F_0(t)$ gets the the minimum of this value and 0. Therefore, on each call of $\mathsf{LWSP}(V_1, F_0)$ on a subset $V_1 \subset V$ with initial values F_0, F_0 keeps the invariant in formula (1).

4 From listing problems to decision problems

In this section we prove the second part of Theorem 1, that $\mathsf{ListPath}_P$ is reducible to Path_P.

Consider a star graph, which is a graph with its vertex set partitioned in X, Y and another single vertex c. Every $x \in X$ is connected to c, and c is connected to every $y \in Y$. Let problem FindX_P be the following problem: on a star graph, find an $x \in X$ satisfying $(\exists y \in Y) P(x, y)$. We will prove that $\mathsf{ListPath}_P$ is reducible to FindX_P and FindX_P is reducible to Path_P.

Lemma 4.1. *Let* $t(M) \geq 2^{\Omega(\sqrt{\log M})}$. $(\mathsf{ListPath}_P, M^2/(t(\mathsf{poly} M))) \leq_{EC} (\mathsf{FindX}_P, M^2/t(M)))$

Proof. We use a grouping reduction technique similar as the trick in [WW10] and [AWW16].

We modify the algorithm for Path_P in Section 2 to get the algorithm for $\mathsf{ListPath}_P$. That is, we divide the graph into two subgraphs and call $\mathsf{ListPath}_P$ recursively in a similar wa as Path_P. Path_P needs to call $\mathsf{Selection}_P$ as queries, and in the counterpart of $\mathsf{ListPath}_P$ we will call FindX_P as queries.

Whenever we need to call $\mathsf{Selection}_P(X, Y)$, we partition X and Y into groups of size at most $\sqrt{M}$. Thus there are $O((|X|/\sqrt{M}) \times (|Y|/\sqrt{M}))$ groups. For each pair of group X_i, Y_j, we construct a star graph and call FindX_P on it. The star graph is constructed as follows: Connect every $x \in X_i$ to a dummy vertex c, and connect c to every $y \in Y_j$. Thus if there exist some satisfying x in X_i, FindX_P will find a satisfying x.

Every time a satisfying vertex x in X_i is found by FindX_P, we mark it and add it into the list of satisfying x, and then call the FindX_P on the same star again with x removed from the graph. We keep calling FindX_P on this graph, ignoring all marked vertices, until either all elements in X_i are marked and removed, or FindX_P cannot find a satisfying x.

Because there are at most M vertices that can be listed, there are at most M calls to FindX_P that returns a satisfying x. Each call has instance size $\sqrt{M}$. The running time is $O(M \cdot (\sqrt{M})^2 / t(\sqrt{M}))$. The total time spent on the rest of the algorithm is the same as the running time of Path_P. □

Lemma 4.2. *Let* $t(M) \geq 2^{\Omega(\sqrt{\log M})}$. $(\mathsf{FindX}_P, M^2/(t(\mathsf{poly} M))) \leq_{EC} (\mathsf{Path}_P, M^2/t(M)))$

Proof. First, we pick an arbitrary element $x_1 \in X$, and construct a graph by letting x_1 connect to all y in Y. Then we call Path_P on this graph. If it returns yes, then we return x_1.

Otherwise, on the star graph we will replace the center vertex c by x_1, remove the original x_1, and call Path_P on this graph. After each call to Path_P, if it returns yes, we divide X in two halves and call Path_P again. Using binary search and shrinking the size of X by half each time, we will finally find a satisfying x. □

Lemmas 4.1 and 4.2 imply the reduction $(\mathsf{ListPath}_P, M^2/(t(\mathsf{poly} M))) \leq_{\mathrm{EC}} (\mathsf{Path}_P, M^2/t(M)))$ for $t(M) \geq 2^{\Omega(\sqrt{\log M})}$.

5 From parallel problems to sequential problems

We prove the third part of Theorem 1, the other direction of the reduction. The reduction from Path_P to $\mathsf{ListPath}_P$ is straightforward.

To reduce from $\mathsf{Selection}_P$ to Path_P, we can construct a graph with dummy vertex c in the middle, such that each x in set X is connected to c, and c is connected to every y in set Y. If P is expressible by first-order logic, then we will let c act like one of the y's when computing $R(x, c)$, and act like of the x's when computing predicates on $P(c, y)$. Let x_1 be an arbitrary element in X, and y_1 be an arbitrary element in Y. We create c by merging x_1 and y_1 into a single element. c has all the relations x_1 and y_1 have. Thus, on any $x \in X, x \neq x_1$, the value of $P(x, c)$ is the same as $P(x, y_1)$. Symmetrically on any $y \in Y, y \neq y_1$, the value of $P(c, y)$ is the same as $P(x_1, y)$. Therefore, there exists x, y such that $P(x, y)$ is true iff $\mathsf{Selection}_P$ on this graph returns true.

In general, if we are allowed to define another property P' such that $P'(x, y) \leftarrow (P(x, y) \wedge (x \neq c) \wedge (y \neq c))$, we have a reduction from $\mathsf{Selection}_P$ to $\mathsf{Path}_{P'}$.

6 Open problems

One open problem is to extend Path_P and LWSP to general DAGs and find subquadratic time reductions and equivalences. Also, we would like to consider the case where the graph is not sparse, where we use $O(MN)$ as the baseline time complexity instead of $O(M^2)$.

It would also be desirable to study the fine-grained complexity of the DAG versions of other quadratic time solvable dynamic programming problems, e.g. the Longest Common Subsequence problem.

Acknowledgements. I sincerely thank my advisor Russell Impagliazzo for his guidance and advice on this paper. Also I would like to thank Marco Carmosino and Jessica Sorrell for helpful comments.

References

[ABBK17] Amir Abboud, Arturs Backurs, Karl Bringmann, and Marvin Künnemann. Fine-grained complexity of analyzing compressed data: Quantifying improvements over decompress-and-solve. In *Foundations of Computer Science (FOCS), 2017 IEEE 58th Annual Symposium on*, pages 192–203. IEEE, 2017.

[ABHS19] Amir Abboud, Karl Bringmann, Danny Hermelin, and Dvir Shabtay. SETH-based lower bounds for subset sum and bicriteria path. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 41–57. SIAM, 2019.

[ABW15] Amir Abboud, Arturs Backurs, and Virginia Vassilevska Williams. Tight hardness results for lcs and other sequence similarity measures. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 59–78. IEEE, 2015.

[AGW14] Amir Abboud, Fabrizio Grandoni, and Virginia Vassilevska Williams. Subcubic equivalences between graph centrality problems, APSP and diameter. In *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*, pages 1681–1697. SIAM, 2014.

[AHWW16] Amir Abboud, Thomas Dueholm Hansen, Virginia Vassilevska Williams, and Ryan Williams. Simulating branching programs with edit distance and friends: or: a polylog shaved is a lower bound made. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 375–388. ACM, 2016.

[AL96] Eric Allender and Klaus-Jörn Lange. $StUSPACE(\log n) \subseteq DSPACE(\log^2 n/ \log\log n)$. In *International Symposium on Algorithms and Computation*, pages 193–202. Springer, 1996.

[AR16] Udit Agarwal and Vijaya Ramachandran. Fine-grained complexity and conditional hardness for sparse graphs. *arXiv preprint arXiv:1611.07008*, 2016.

[AST94] Alok Aggarwal, Baruch Schieber, and Takeshi Tokuyama. Finding a minimum-weightk-link path in graphs with the concave monge property and applications. *Discrete & Computational Geometry*, 12(3):263–280, 1994.

[AU79] Alfred V Aho and Jeffrey D Ullman. Universality of data retrieval languages. In *Proceedings of the 6th ACM SIGACT-SIGPLAN symposium on Principles of programming languages*, pages 110–119. ACM, 1979.

[AWW14] Amir Abboud, Virginia Vassilevska Williams, and Oren Weimann. Consequences of faster alignment of sequences. In *International Colloquium on Automata, Languages, and Programming*, pages 39–51. Springer, 2014.

[AWW16] Amir Abboud, Virginia Vassilevska Williams, and Joshua Wang. Approximation and fixed parameter subquadratic algorithms for radius and diameter in sparse graphs. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete Algorithms*, pages 377–391. SIAM, 2016.

[AWY15] Amir Abboud, Ryan Williams, and Huacheng Yu. More applications of the polynomial method to algorithm design. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 218–230. SIAM, 2015.

[BGL17] Karl Bringmann, Allan Grønlund, and Kasper Green Larsen. A dichotomy for regular expression membership testing. In *Foundations of Computer Science (FOCS), 2017 IEEE 58th Annual Symposium on*, pages 307–318. IEEE, 2017.

[BI15] Arturs Backurs and Piotr Indyk. Edit distance cannot be computed in strongly subquadratic time (unless SETH is false). In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 51–58. ACM, 2015.

[BI16] Arturs Backurs and Piotr Indyk. Which regular expression patterns are hard to match? In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*, pages 457–466. IEEE, 2016.

[BK15] Karl Bringmann and Marvin Künnemann. Quadratic conditional lower bounds for string problems and dynamic time warping. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 79–97. IEEE, 2015.

[BK17] Karl Bringmann and Marvin Künnemann. Improved approximation for Fréchet distance on c-packed curves matching conditional lower bounds. *International Journal of Computational Geometry & Applications*, 27(01n02):85–119, 2017.

[BM15] Karl Bringmann and Wolfgang Mulzer. Approximability of the discrete Fréchet distance. *Journal of Computational Geometry*, 7(2):46–76, 2015.

[Bod96] Hans L Bodlaender. A linear-time algorithm for finding tree-decompositions of small treewidth. *SIAM Journal on computing*, 25(6):1305–1317, 1996.

[Bri14] Karl Bringmann. Why walking the dog takes time: Fréchet distance has no strongly subquadratic algorithms unless SETH fails. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 661–670. IEEE, 2014.

[BRS+18] Arturs Backurs, Liam Roditty, Gilad Segal, Virginia Vassilevska Williams, and Nicole Wein. Towards tight approximation bounds for graph diameter and eccentricities. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 267–280. ACM, 2018.

[CW16] Timothy M Chan and Ryan Williams. Deterministic APSP, Orthogonal Vectors, and More: Quickly derandomizing Razborov-Smolensky. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1246–1255. SIAM, 2016.

[CWHL11] S.C. Chen, J.Y. Wu, G.S. Huang, and R.C.T. Lee. Finding a longest increasing subsequence on a galled tree. In *the 28th Workshop on Combinatorial Mathematics and Computation Theory, Penghu, Taiwan*, 2011.

[Fre75] Michael L Fredman. On computing the length of longest increasing subsequences. *Discrete Mathematics*, 11(1):29–35, 1975.

[GI19] Jiawei Gao and Russell Impagliazzo. The fine-grained complexity of strengthenings of first-order logic. *Electronic Colloquium on Computational Complexity (ECCC)*, 26:9, 2019.

[GIKW17] Jiawei Gao, Russell Impagliazzo, Antonina Kolokolova, and Ryan Williams. Completeness for first-order properties on sparse structures with algorithmic applications. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '17, pages 2162–2181, 2017.

[GLL12] Jerrold R Griggs, Wei-Tian Li, and Linyuan Lu. Diamond-free families. *Journal of Combinatorial Theory, Series A*, 119(2):310–322, 2012.

[GP89] Zvi Galil and Kunsoo Park. A linear-time algorithm for concave one-dimensional dynamic programming. 1989.

[HL87] Daniel S Hirschberg and Lawrence L Larmore. The least weight subsequence problem. *SIAM Journal on Computing*, 16(4):628–638, 1987.

[IP01] Russell Impagliazzo and Ramamohan Paturi. On the complexity of k-SAT. *Journal of Computer and System Sciences*, 62(2):367–375, 2001.

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512–530, 2001.

[KP81] Donald E Knuth and Michael F Plass. Breaking paragraphs into lines. *Software: Practice and Experience*, 11(11):1119–1184, 1981.

[KPS17] Marvin Künnemann, Ramamohan Paturi, and Stefan Schneider. On the Fine-Grained Complexity of One-Dimensional Dynamic Programming. In *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, volume 80 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1–21:15, 2017.

[Lib13] Leonid Libkin. *Elements of finite model theory*. Springer Science & Business Media, 2013.

[LjLW12] Guan-Yu Lin, Jia jie Liu, and Yue-Li Wang. Finding a longest increasing subsequence from the paths in a complete bipartite graph. 2012.

[LWW18] Andrea Lincoln, Virginia Vassilevska Williams, and Ryan Williams. Tight hardness for shortest cycles and paths in sparse graphs. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1236–1252. Society for Industrial and Applied Mathematics, 2018.

[RS86] Neil Robertson and P.D Seymour. Graph minors. ii. algorithmic aspects of tree-width. *Journal of Algorithms*, 7(3):309 – 322, 1986.

[RVW13] Liam Roditty and Virginia Vassilevska Williams. Fast approximation algorithms for the diameter and radius of sparse graphs. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 515–524. ACM, 2013.

[Sch98] Baruch Schieber. Computing a minimum weightk-link path in graphs with the concave monge property. *Journal of Algorithms*, 29(2):204–222, 1998.

[Sto74] Larry Joseph Stockmeyer. *The complexity of decision problems in automata theory and logic*. PhD thesis, Massachusetts Institute of Technology, 1974.

[Var82] Moshe Y Vardi. The complexity of relational query languages. In *Proceedings of the fourteenth annual ACM symposium on Theory of computing*, pages 137–146. ACM, 1982.

[Wil88] Robert Wilber. The concave least-weight subsequence problem revisited. *Journal of Algorithms*, 9(3):418–425, 1988.

[Wil05] Ryan Williams. A new algorithm for optimal 2-constraint satisfaction and its implications. *Theoretical Computer Science*, 348(2):357–365, 2005.

[Wil14] Ryan Williams. Faster decision of first-order graph properties. In *Proceedings of the Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*, page 80. ACM, 2014.

[WW10] Virginia Vassilevska Williams and Ryan Williams. Subcubic equivalences between path, matrix and triangle problems. In *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*, pages 645–654. IEEE, 2010.

[Yao80] F Frances Yao. Efficient dynamic programming using quadrangle inequalities. In *Proceedings of the twelfth annual ACM symposium on Theory of computing*, pages 429–435. ACM, 1980.

A Problem examples

We give a list of problems that can be considered as instances of LWSP or Path_P.

Trip Planning (LWSP version of Airplane Refueling)

On a DAG where vertices represent cities and edges are roads, we wish to find a path for a vehicle, along which we wish to find a sequence of cities where the vehicle can rest and add fuel so that the cost is minimized. The cost of traveling between cities x and y is defined by cost $c_{x,y}$. $c_{x,y}$ can be defined in multiple ways, e.g. $c_{x,y}$ is $cost(y)$ if $dist(x, y) \leq M$ and ∞ otherwise. $dist(x, y)$ is the distance between x, y that can be computed by the positions of x, y. M is the maximal distance the vehicle can travel without resting. $cost(y)$ is the cost for resting at position y.

Longest Subset Chain on graphs (LWSP version of Longest Subset Chain)

On a DAG where each vertex corresponds to a set, we want to find a longest chain in a path of the graph such that each set is a subset of its successor. Here $c_{x,y}$ is -1 if S_x is a subset of S_y, and ∞ otherwise.

Multi-currency Coin Change (LWSP version of Coin Change)

Consider there are two different currencies, so there are two sets of coins. We need to find a way to get value V_1 for currency #1 and value V_2 for currency #2, so that the total weight of coins is minimized. Each pair of values $v_1 \in \{0, \ldots, V_1\}$ and $v_2 \in \{0, \ldots, V_2\}$ can be considered as a vertex. We connect vertex (v_1, v_2) to (v_1', v_2') iff $v_1' = v_1 + 1$ or $v_2' = v_2' + 1$. The whole graph is a grid, and we wish to find a subsequence of a path from $(0, 0)$ to (V_1, V_2) so that the cost is minimized. The cost is defined by $C_{(v_1,v_2),(v_1',v_2)} = w_{1,v_1'-v_1}$ and $C_{(v_1,v_2),(v_1,v_2')} = w_{2,v_2'-v_2}$, where $w_{i,j}$ is the weight of a coin of value j from currency $\#i$.

Pretty Printing with alternative expressions (LWSP version of Pretty Printing)

The Pretty Printing problem is to break a paragraph into lines, so that each line have roughly the same length. If a line is too long or too short, then there is some cost depending on the line length. The goal of the problem is to minimize the cost.

For some text, it is hard to print prettily. For example, if there are long formulas in the text, then sometimes its line gets too wide, but if we move the formula into the next line, the original line has too few words. One solution for this issue is to use alternate wording for the sentence, to rephrase a part of a sentence to its synonym. These sentences have different lengths, and formulas in some of them will be displayed better than others. These different ways can be considered as different paths in a graph, and we wish to find one sentence that has the minimal Pretty Printing cost.

A Path_P instance

Say we have a set of words, and we want to find a word chain (a chain of words so that the last letter of the previous word is the same as the first letter of the next word) so that the first word and the last word satisfy some properties, e.g. they do not have similar meanings, they have the same length, they don't have the same letters on the same positions, etc. Each word corresponds to a vertex in the graph. For words that can be consecutive in a word chain, we add an edge to the words.

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il